

# KLEIN'S TRACE INEQUALITY AND SUPERQUADRATIC TRACE FUNCTIONS

MOHSEN KIAN and MOHAMMAD W. ALOMARI

**ABSTRACT.** We show that if  $f$  is a non-negative superquadratic function, then  $A \mapsto \text{Tr}f(A)$  is a superquadratic function on the matrix algebra. In particular,

$$\text{Tr} f\left(\frac{A+B}{2}\right) + \text{Tr} f\left(\left|\frac{A-B}{2}\right|\right) \leq \frac{\text{Tr} f(A) + \text{Tr} f(B)}{2}$$

holds for all positive matrices  $A, B$ . In addition, we present a Klein's inequality for superquadratic functions as

$$\text{Tr}[f(A) - f(B) - (A - B)f'(B)] \geq \text{Tr}[f(|A - B|)]$$

for all positive matrices  $A, B$ . It gives in particular an improvement of the Klein's inequality for non-negative convex function. As a consequence, some variants of the Jensen trace inequality for superquadratic functions have been presented.

## 1. INTRODUCTION AND PRELIMINARIES

In study of quantum mechanical systems, there are many famous concepts which are related to the trace function  $A \mapsto \text{Tr}(A)$ . The well-known relative entropy of a density matrix  $\rho$  (a positive matrix of trace one) with respect of another density matrix  $\sigma$  is defined by

$$S(\rho|\sigma) = \text{Tr}(\rho \log \rho) - \text{Tr}(\rho \log \sigma).$$

More generally, for a proper (continuous) real function  $f$ , the study of the mapping  $A \mapsto \text{Tr}(f(A))$  is important.

The main subject of this paper, is to study this mapping for a class of real functions, the superquadratic functions. It is known that if  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous convex (monotone increasing) function, then the trace function  $A \mapsto \text{Tr}(f(A))$  is a convex (monotone increasing) function, see [14, 17]. In Section 2, we present this result for superquadratic functions.

For all Hermitian  $n \times n$  matrices  $A$  and  $B$  and all differentiable convex functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  with derivative  $f'$ , the well known Klein inequality reads as

$$\text{Tr} [f(A) - f(B) - (A - B)f'(B)] \geq 0. \quad (1.1)$$

---

2010 *Mathematics Subject Classification.* Primary: 47A56, 15A45 Secondary: 15A18, 15A42.

*Key words and phrases.* Klein's trace inequality, superquadratic trace function, majorization.

With  $f(t) = t \log t$  ( $t > 0$ ), this gives

$$S(A|B) = \text{Tr } A (\log A - \log B) \geq \text{Tr } (A - B)$$

for positive matrices  $A, B$ . If  $A$  and  $B$  are density matrices, then  $S(A, B) \geq 0$ . This is a classical application of the Klein inequality. See [7, 18]. To see a collection of trace inequalities the reader can refer to [8, 9, 10, 13, 19, 20] and references therein.

In Section 3, we present a Klein trace inequality for superquadrtic functions. We show that our result improves previous results in the case of non-negative functions. In-addition, some applications of our results present counterpart to some known trace inequalities. We give some examples to clarify our results.

Let  $\mathcal{B}(\mathcal{H})$  be the  $C^*$ -algebra of all bounded linear operators defined on a complex Hilbert space  $(\mathcal{H}; \langle \cdot, \cdot \rangle)$  with the identity operator  $I$ . When  $\dim \mathcal{H} = n$ , we identify  $\mathcal{B}(\mathcal{H})$  with the algebra  $\mathbb{M}_n$  of  $n$ -by- $n$  complex matrices. We denote by  $\mathbb{H}_n$  the real subspace of Hermitian matrices and by  $\mathbb{M}_n^+$  the cone of positive (semidefinite) matrices. The identity matrix of any size will be denoted by  $I$ .

Every Hermitian matrix  $A \in \mathbb{H}_n$  enjoys the spectral decomposition  $A = \sum_{j=1}^n \lambda_j P_j$ , where  $\lambda_j$ 's are eigenvalues of  $A$  and  $P_j$ 's are projection matrices with  $\sum_{j=1}^n P_j = I$ . If  $f$  is a continuous real function which is defined on the set of eigenvalues of  $A$ , then  $f(A)$  is the matrix defined using the spectral decomposition by  $f(A) = \sum_{j=1}^n f(\lambda_j) P_j$ . The eigenvalues of  $f(A)$  are just  $f(\lambda_j)$ . Moreover, If  $U$  is a unitary matrix, then  $f(U^* A U) = U^* f(A) U$ .

For  $A = [a_{ij}] \in \mathbb{M}_n$  the canonical trace of  $A$  is denoted by  $\text{Tr} A$  and is defined to be  $\sum_{j=1}^n a_{ii}$ . The canonical trace is a unitary invariant mapping, say  $\text{Tr} U A U^* = \text{Tr} A$  for every unitary matrix  $U$ . So, when  $\lambda_1, \dots, \lambda_n$  are eigenvalues of  $A$  and  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  is an orthonormal set of corresponding eigenvectors in  $\mathbb{C}^n$ , then

$$\text{Tr} A = \sum_{j=1}^n \lambda_j(A) = \sum_{j=1}^n \langle A \mathbf{u}_j, \mathbf{u}_j \rangle \quad \text{and} \quad \text{Tr} f(A) = \sum_{j=1}^n f(\lambda_j(A)) = \sum_{j=1}^n f(\langle A \mathbf{u}_j, \mathbf{u}_j \rangle).$$

If  $\mathcal{H}$  is a separable Hilbert space with an orthonormal basis  $\{e_i\}_i$ , an operator  $A \in \mathcal{B}(\mathcal{H})$  is said to be a trace class operator if

$$\|A\|_1 = \sum_i \left\langle (A^* A)^{1/2} e_i, e_i \right\rangle,$$

is finite. In this case, the trace of  $A$  is defined by  $\text{Tr } (A) = \sum_i \langle A e_i, e_i \rangle$  and is independent of the choice of the orthonormal basis. When  $\mathcal{H}$  is finite-dimensional, every operator is trace class and this definition of trace of  $A$  coincides with the definition of the trace of a matrix.

For a vector  $\mathbf{x} = (x_1, \dots, x_n)$  in  $\mathbb{R}^n$ , let  $\mathbf{x}^\downarrow$  and  $\mathbf{x}^\uparrow$  denotes the vectors obtained by rearranging entries of  $\mathbf{x}$  in decreasing and increasing order, respectively, i.e.,  $x_1^\downarrow \geq \dots \geq x_n^\downarrow$

and  $x_1^\uparrow \leq \dots \leq x_n^\uparrow$ . A vector  $\mathbf{x} \in \mathbb{R}^n$  is said to be weakly majorised by  $\mathbf{y} \in \mathbb{R}^n$  and denoted by  $\mathbf{x} \prec_w \mathbf{y}$  if  $\sum_{j=1}^k x_j^\downarrow \leq \sum_{j=1}^k y_j^\downarrow$  holds for every  $k = 1, \dots, n$ . If in addition  $\sum_{j=1}^n x_j^\downarrow = \sum_{j=1}^n y_j^\downarrow$ , then  $\mathbf{x}$  is said to be majorised by  $\mathbf{y}$  and is denoted by  $\mathbf{x} \prec \mathbf{y}$ . The trace of a vector  $\mathbf{x} \in \mathbb{R}^n$  is defined to be the sum of its entries and is denoted using a same notation as a matrix by  $\text{Tr } \mathbf{x}$ .

A matrix  $P = [p_{ij}] \in \mathbb{M}_n$  is said to be doubly stochastic if all of its entries are non-negative and

$$\sum_{i=1}^n p_{ij} = 1 \quad \text{for all } j \quad \text{and} \quad \sum_{j=1}^n p_{ij} = 1 \quad \text{for all } i.$$

For all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  it is well-known that  $\mathbf{x} \prec \mathbf{y}$  if and only if there exists a doubly stochastic matrix  $P$  such that  $\mathbf{x} = P\mathbf{y}$ , see [6, Theorem II.1.10]. More results concerning majorization can be found in [6, 14].

A function  $f : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$  is called convex if

$$f(\alpha t + (1 - \alpha)s) \leq \alpha f(t) + (1 - \alpha)f(s), \quad (1.2)$$

for all points  $s, t \in J$  and all  $\alpha \in [0, 1]$ . If  $-f$  is convex then we say that  $f$  is concave. Moreover, if  $f$  is both convex and concave, then  $f$  is said to be affine.

Geometrically, for all  $x, y \in J$  with  $x \leq t \leq y$ , the two points  $(x, f(x))$  and  $(y, f(y))$  on the graph of  $f$  are on or below the chord joining the endpoints. In symbols, we write

$$f(t) \leq \frac{f(y) - f(x)}{y - x} (t - x) + f(x)$$

for any  $x \leq t \leq y$  and  $x, y \in J$ .

Equivalently, given a function  $f : J \rightarrow \mathbb{R}$ , we say that  $f$  admits a support line at  $s \in J$  if there exists a  $\lambda \in \mathbb{R}$  such that

$$f(t) \geq f(s) + \lambda(t - s) \quad (1.3)$$

for all  $t \in J$ . The set of all such  $\lambda$  is called the subdifferential of  $f$  at  $s$  and it is denoted by  $\partial f$ . Indeed, the subdifferential gives us the slopes of the supporting lines for the graph of  $f$  so that if  $f$  is convex, then  $\partial f(s) \neq \emptyset$  at all interior points of its domain.

From this point of view, Abramovich et al. [3] extended the above idea for what they called superquadratic functions. Namely, a function  $f : [0, \infty) \rightarrow \mathbb{R}$  is called superquadratic provided that for all  $s \geq 0$  there exists a constant  $C_s \in \mathbb{R}$  such that

$$f(t) \geq f(s) + C_s(t - s) + f(|t - s|) \quad (1.4)$$

for all  $t \geq 0$ . A function  $f$  is called subquadratic if  $-f$  is superquadratic. Thus, for a superquadratic function we require that  $f$  is above its tangent line plus a translation of  $f$  itself. If  $f$  is differentiable and satisfies  $f(0) = f'(0) = 0$ , then one can easily see that the constant  $C_s$  in the definition is necessarily  $f'(s)$ , see [2].

Prima facie, superquadraticity looks to be stronger than convexity, but if  $f$  takes negative values then it may be considered weaker. On the other hand, non-negative subquadratic functions does not need to be concave. In other words, there exist subquadratic function which are convex. This fact helps us first to improve some results for convex functions and second to present some counterpart results concerning convex functions. Some known examples of superquadratic functions are power functions. For every  $p \geq 2$ , the function  $f(t) = t^p$  is superquadratic as well as convex. If  $1 \leq p \leq 2$ , then  $f(t) = -t^p$  is superquadratic and concave. To see more examples of superquadratic and subquadratic functions and their properties, the reader can refer to [2, 3, 1, 4, 5]. Among others, Abramovich et al. [3] proved that the inequality

$$f\left(\int \varphi d\mu\right) \leq \int f(\varphi(s)) - f\left(\left|\varphi(s) - \int \varphi d\mu\right|\right) d\mu(s) \quad (1.5)$$

holds for all probability measures  $\mu$  and all nonnegative,  $\mu$ -integrable functions  $\varphi$  if and only if  $f$  is superquadratic.

As a matrix extension of (1.5), Kian [15] showed that if  $f : [0, \infty) \rightarrow \mathbb{R}$  is a continuous superquadratic function, then

$$f(\langle A\mathbf{u}, \mathbf{u} \rangle) \leq \langle f(A)\mathbf{u}, \mathbf{u} \rangle - \langle f(|A - \langle A\mathbf{u}, \mathbf{u} \rangle|)\mathbf{u}, \mathbf{u} \rangle \quad (1.6)$$

holds for every positive matrix  $A \in \mathbb{M}_n^+$  and every unit vector  $\mathbf{u} \in \mathbb{C}^n$ . More generally, it has been shown in [16] that if  $\Phi : \mathbb{M}_n \rightarrow \mathbb{M}_m$  is a unital positive linear map, then

$$f(\langle \Phi(A)\mathbf{u}, \mathbf{u} \rangle) \leq \langle \Phi(f(A))\mathbf{u}, \mathbf{u} \rangle - \langle \Phi(f(|A - \langle \Phi(A)\mathbf{u}, \mathbf{u} \rangle|))\mathbf{u}, \mathbf{u} \rangle \quad (1.7)$$

holds for every positive matrix  $A \in \mathbb{M}_n^+$  and every unit vector  $\mathbf{u} \in \mathbb{C}^n$ .

## 2. SUPERQUADRATIC TRACE FUNCTIONS

It is known that if  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous convex function, then the trace function  $A \mapsto \text{Tr}[f(A)]$  is a convex function on  $\mathbb{M}_n$ . In this section, we present this fact for superquadratic functions. We need some lemmas. Note that if  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  is a vector and  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a real function, we denote the vector  $(f(x_1), \dots, f(x_n))$  by  $f(\mathbf{x})$ .

**Lemma 2.1.** [6] *For  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*

- (i) *If  $\mathbf{x} \prec \mathbf{y}$ , then  $|\mathbf{x}| \prec |\mathbf{y}|$ , where  $|\mathbf{x}| = (|x_1|, \dots, |x_n|)$ .*
- (ii)  *$\mathbf{x} \prec \mathbf{y}$  if and only if  $\text{Tr}f(\mathbf{x}) \leq \text{Tr}f(\mathbf{y})$  for every convex function  $f$ .*

**Lemma 2.2.** *Assume that  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$  and  $f : [0, \infty) \rightarrow \mathbb{R}$  is a superquadratic function. If  $\mathbf{x} \prec \mathbf{y}$ , then there exists a doubly stochastic matrix  $P$  such that  $\text{Tr}f(\mathbf{x}) \leq \text{Tr}f(\mathbf{y}) - \text{Tr}PF$ , where  $F = [f(|x_i - y_j|)]$ .*

*Proof.* For  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ , if  $\mathbf{x} \prec \mathbf{y}$ , then there exists a doubly stochastic matrix  $P = [p_{ij}]$  such that  $\mathbf{x} = P\mathbf{y}$ . Therefore,  $x_i = \sum_{j=1}^n p_{ij}y_j$  for every  $i = 1, \dots, n$  and  $\sum_{j=1}^n p_{ij} = 1$ . If  $f$  is a superquadratic function, then from (1.5) we conclude that the inequality

$$f(x_i) = f\left(\sum_{j=1}^n p_{ij}y_j\right) \leq \sum_{j=1}^n p_{ij}f(y_j) - \sum_{j=1}^n p_{ij}f\left(\left|y_j - \sum_{j=1}^n p_{ij}y_j\right|\right) \quad (2.1)$$

holds for every  $i = 1, \dots, n$ . Summing over  $i$ , we obtain

$$\text{Tr}f(\mathbf{x}) \leq \text{Tr}f(\mathbf{y}) - \sum_{i,j=1}^n p_{ij}f(|y_j - x_i|).$$

If we put  $F = [f(|x_i - y_j|)]$ , then  $\sum_{i,j=1}^n p_{ij}f(|y_j - x_i|) = \text{Tr}PF$ . This completes the proof.  $\square$

**Lemma 2.3.** [3] *Let  $f : [0, \infty) \rightarrow \mathbb{R}$  be a superquadratic function. Then*

- (i)  $f(0) \leq 0$ ;
- (ii) *If  $f$  is differentiable and  $f(0) = f'(0) = 0$ , then  $C_s = f'(s)$  in (1.4) for all  $s \geq 0$ ;*
- (iii) *If  $f$  is non-negative, then  $f(0) = f'(0) = 0$  and  $f$  is convex and increasing.*

**Theorem 2.4.** *Let  $f : [0, \infty) \rightarrow \mathbb{R}$  be a continuous superquadratic function. If  $f$  is non-negative, then the mapping  $A \mapsto \text{Tr}[f(A)]$  is a superquadratic function on  $\mathbb{M}_n^+$ . More generally, the inequality*

$$\text{Tr}f\left(\frac{A+B}{2}\right) + \text{Tr}f\left(\left|\frac{A-B}{2}\right|\right) \leq \frac{\text{Tr}f(A) + \text{Tr}f(B)}{2} - \text{Tr}[PG + QF] \quad (2.2)$$

holds for some doubly stochastic matrices  $P = [p_{ij}]$  and  $Q = [q_{ij}]$ , in which

$$G = \left[f\left(\frac{1}{2}||\lambda_i| - |\mu_j - \nu_j||\right)\right] \quad \text{and} \quad F = \left[f\left(\frac{1}{2}|\xi_i - \mu_j - \nu_j|\right)\right],$$

where  $\lambda_i, \xi_i, \mu_i$  and  $\nu_i$  are eigenvalues of  $A - B$ ,  $A + B$ ,  $A$  and  $B$ , respectively.

*Proof.* For a Hermitian matrix  $X$ , assume that  $\lambda^\downarrow(X)$  and  $\lambda^\uparrow(X)$  are eigenvalues of  $X$  arranged in decreasing order and increasing order, respectively. Recall that [6] if  $A, B$  are Hermitian matrices, then

$$\lambda^\downarrow(A) - \lambda^\downarrow(B) \prec \lambda^\downarrow(A - B) \prec \lambda^\downarrow(A) - \lambda^\uparrow(B) \quad (2.3)$$

and

$$\lambda^\downarrow(A) + \lambda^\uparrow(B) \prec \lambda^\downarrow(A + B) \prec \lambda^\downarrow(A) + \lambda^\downarrow(B). \quad (2.4)$$

From (2.3) we have

$$\lambda^\uparrow(B - A) \prec \lambda^\downarrow(B) - \lambda^\downarrow(A) \quad (2.5)$$

and noting Lemma 2.1 this gives

$$\left| \lambda^\downarrow(A - B) \right| \prec \left| \lambda^\downarrow(B) - \lambda^\downarrow(A) \right|. \quad (2.6)$$

We assume that  $\mu_j$  and  $\nu_j$  ( $j = 1, \dots, n$ ) are eigenvalues of  $A$  and  $B$  respectively, arranged in decreasing order. If  $f$  is superquadratic, then it follows from (2.6) and Lemma 2.2 that

$$\begin{aligned} \text{Tr} f(|A - B|) &= \sum_{j=1}^n f(|\lambda_j(A - B)|) = \text{Tr} f\left(\left| \lambda^\downarrow(A - B) \right|\right) \\ &\leq \text{Tr} f\left(\left| \lambda^\downarrow(B) - \lambda^\downarrow(A) \right|\right) - \text{Tr} PG \quad (\text{by Lemma 2.2}) \\ &= \sum_{j=1}^n f(|\mu_j - \nu_j|) - \text{Tr} PG, \end{aligned}$$

for some doubly stochastic matrix  $P = [p_{ij}]$ , in which  $G = [f(|\lambda_i| - |\mu_j - \nu_j|)]$  and  $\lambda_i$ 's are eigenvalues of  $A - B$ . This implies that for every  $\alpha \geq 0$ , the inequality

$$\text{Tr} f(\alpha|A - B|) \leq \sum_{j=1}^n f(\alpha|\mu_j - \nu_j|) - \text{Tr} P_\alpha G_\alpha \quad (2.7)$$

holds for some doubly stochastic matrix  $P_\alpha$ , in which  $G_\alpha = [f(\alpha|\lambda_i| - |\mu_j - \nu_j|)]$  and  $\lambda_i$ 's are eigenvalues of  $(A - B)$ . Now suppose that  $\alpha \in [0, 1]$ . Another use of Lemma 2.2 together with (2.4) gives

$$\text{Tr} f\left(\lambda^\downarrow(\alpha A + (1 - \alpha)B)\right) \leq \text{Tr} f\left(\alpha \lambda^\downarrow(A) + (1 - \alpha) \lambda^\downarrow(B)\right) - \text{Tr} QF \quad (2.8)$$

for some doubly stochastic matrix  $Q$ , where  $F = [f(|\xi_i - \alpha \mu_j - (1 - \alpha) \nu_j|)]$  and  $\xi_i$ 's are eigenvalues of  $\alpha A + (1 - \alpha)B$ . Therefore

$$\begin{aligned} &\text{Tr} f(\alpha A + (1 - \alpha)B) \\ &= \sum_{j=1}^n f\left(\lambda_j^\downarrow(\alpha A + (1 - \alpha)B)\right) \\ &\leq \sum_{j=1}^n f(\alpha \mu_j + (1 - \alpha) \nu_j) - \text{Tr} QF \quad (\text{by (2.8)}) \\ &\leq \sum_{j=1}^n \{\alpha f(\mu_j) + (1 - \alpha) f(\nu_j) - \alpha f((1 - \alpha)|\mu_j - \nu_j|) - (1 - \alpha) f(\alpha|\mu_j - \nu_j|)\} - \text{Tr} QF \\ &\quad (\text{since } f \text{ is superquadratic}) \\ &= \alpha \text{Tr} f(A) + (1 - \alpha) \text{Tr} f(B) - \alpha \sum_{j=1}^n f((1 - \alpha)|\mu_j - \nu_j|) - (1 - \alpha) \sum_{j=1}^n f(\alpha|\mu_j - \nu_j|) - \text{Tr} QF \\ &\leq \alpha \text{Tr} f(A) + (1 - \alpha) \text{Tr} f(B) \\ &\quad - \alpha \text{Tr} f((1 - \alpha)|A - B|) - (1 - \alpha) \text{Tr} f(\alpha|A - B|) - \text{Tr}[(1 - \alpha)P_\alpha G_\alpha + \alpha P_{1-\alpha} G_{1-\alpha} + QF], \end{aligned}$$

where the last inequality follows from (2.7). In particular, with  $\alpha = 1/2$  this gives

$$\mathrm{Tr} f\left(\frac{A+B}{2}\right) + \mathrm{Tr} f\left(\left|\frac{A-B}{2}\right|\right) \leq \frac{\mathrm{Tr} f(A) + \mathrm{Tr} f(B)}{2} - \mathrm{Tr}[PG + QF]$$

for some doubly stochastic matrices  $P = [p_{ij}]$  and  $Q = [q_{ij}]$ , in which

$$G = \left[ f\left(\frac{1}{2}||\lambda_i| - |\mu_j - \nu_j||\right) \right] \quad \text{and} \quad F = \left[ f\left(\frac{1}{2}|\xi_i - \mu_j - \nu_j|\right) \right],$$

where  $\lambda_i$  and  $\xi_i$  are eigenvalues of  $A - B$  and  $A + B$ , respectively. Equivalently

$$\begin{aligned} & \mathrm{Tr} f\left(\frac{A+B}{2}\right) + \mathrm{Tr} f\left(\left|\frac{A-B}{2}\right|\right) \\ & \leq \frac{\mathrm{Tr} f(A) + \mathrm{Tr} f(B)}{2} - \sum_{i,j=1}^n \left( p_{ij} f\left(\frac{1}{2}||\lambda_i| - |\mu_j - \nu_j||\right) + q_{ij} f\left(\frac{1}{2}|\xi_i - \mu_j - \nu_j|\right) \right), \end{aligned}$$

from which we conclude that if  $f$  is non-negative, then  $A \mapsto \mathrm{Tr} f(A)$  is a superquadratic function. This completes the proof.  $\square$

In 2003, Hansen & Pedersen [12] proved a trace version of then Jensen inequality. They showed that if  $f : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$  is a continuous convex function, then

$$\mathrm{Tr} \left[ f\left(\sum_{i=1}^k C_i^* A_i C_i\right) \right] \leq \mathrm{Tr} \left[ \sum_{i=1}^k C_i^* f(A_i) C_i \right] \quad (2.9)$$

for every  $k$ -tuple of Hermitian matrices  $(A_1, \dots, A_k)$  in  $\mathbb{M}_n$  with spectra contained in  $J$  and every  $k$ -tuple  $(C_1, \dots, C_k)$  of matrices with  $\sum_{i=1}^k C_i^* C_i = I$ .

In the rest of this section, using the concept of superquadratic functions and Theorem 2.4, we present variants of (2.9) for superquadratic functions, which give in particular some refinements of the Hansen–Pedersen trace inequality (2.9) in the case of non-negative functions. Beside our results concerning (2.9), we give a conjecture as follows.

**Conjecture.** If  $f : [0, \infty) \rightarrow \mathbb{R}$  is a continuous superquadratic function, then

$$\mathrm{Tr} \left[ f\left(\sum_{i=1}^k C_i^* A_i C_i\right) \right] \leq \mathrm{Tr} \left[ \sum_{i=1}^k C_i^* f(A_i) C_i \right] - \mathrm{Tr} \left[ \sum_{i=1}^k C_i^* f\left(\left|A_i - \mathrm{Tr} \left[ \sum_{i=1}^k C_i^* A_i C_i \right]\right|\right) C_i \right] \quad (2.10)$$

for every  $k$ -tuple of positive matrices  $(A_1, \dots, A_k)$  in  $\mathbb{M}_n^+$  and every  $k$ -tuple  $(C_1, \dots, C_k)$  of matrices with  $\sum_{i=1}^k C_i^* C_i = I$ .

We now use Theorem 2.4 to present the first variant of (2.9) for superquadratic functions.

**Corollary 2.5.** Assume that  $f : [0, \infty) \rightarrow [0, \infty)$  is a continuous function. If  $f$  is superquadratic, then

$$\mathrm{Tr} f(C^* AC) + \mathrm{Tr} f(D^* AD) \leq \mathrm{Tr} [C^* f(A)C + D^* f(A)D] - \mathrm{Tr} [f(|DAC|) + f(|C^* AD^*|)] \quad (2.11)$$

for every positive matrix  $A \in \mathbb{M}_n^+$  and every isometry  $C$ , where  $D = \sqrt{1 - CC^*}$ .

*Proof.* To prove (2.11), we apply Theorem 2.4 and then employ a similar argument as in [11, Theorem 1.9]. Assume that  $A, B \in \mathbb{M}_n^+$ . If  $C \in \mathbb{M}_n$  and  $C^*C = I$ , then the block matrices  $U = \begin{bmatrix} C & D \\ 0 & -C^* \end{bmatrix}$  and  $V = \begin{bmatrix} C & -D \\ 0 & C^* \end{bmatrix}$  are unitary matrices in  $\mathbb{M}_{2n}$ , provided that  $D = (I - CC^*)^{1/2}$ . With  $\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$  we compute

$$\frac{U^*\tilde{A}U + V^*\tilde{A}V}{2} = (C^*AC) \oplus (DAD + CBC^*) \quad (2.12)$$

and

$$\left| \frac{U^*\tilde{A}U - V^*\tilde{A}V}{2} \right| = |DAC| \oplus |C^*AD|. \quad (2.13)$$

Now we use Theorem 2.4 to write

$$\begin{aligned} & \text{Tr} f(C^*AC) + \text{Tr} f(DAD + CBC^*) \\ &= \text{Tr} f\left(\frac{U^*\tilde{A}U + V^*\tilde{A}V}{2}\right) \quad (\text{by (2.12)}) \\ &\leq \text{Tr} \frac{f(U^*\tilde{A}U) + f(V^*\tilde{A}V)}{2} - \text{Tr} f\left(\left| \frac{U^*\tilde{A}U - V^*\tilde{A}V}{2} \right| \right) \quad (\text{by Theorem 2.4}) \\ &= \text{Tr} \frac{U^*f(\tilde{A})U + V^*f(\tilde{A})V}{2} - \text{Tr} f\left(\left| \frac{U^*\tilde{A}U - V^*\tilde{A}V}{2} \right| \right) \\ &= \text{Tr} [C^*f(A)C + Df(A)D + Cf(B)C^*] - \text{Tr} [f(|DAC|) + f(|C^*AD|)], \end{aligned}$$

where the last equality follows from (2.12) and (2.13). Putting  $B = 0$  and noting that  $f(0) \leq 0$ , this gives the desired inequality.  $\square$

We remark that a non-negative superquadratic function  $f$  is convex and satisfies  $f(0) = 0$ . If  $C^*C = I$ , then with  $D = \sqrt{I - CC^*}$  we have  $D^*D = I - CC^* \leq I$ . It follows from (2.9) that

$$\text{Tr} f(C^*AC) + \text{Tr} f(D^*AD) \leq \text{Tr} C^*f(A)C + \text{Tr} D^*f(A)D.$$

Therefore Corollary 2.5 gives a refinement of (2.9), when  $f$  is a non-negative superquadratic function.

To present the second variant of (2.9), we give the following version of (1.6) and (1.7). The proof is similar to those of [15, Theorem 2.1] and [16, Theorem 2.3]. We include the proof for the sake of completeness.

**Lemma 2.6.** *Let  $f : [0, \infty) \rightarrow \mathbb{R}$  be a continuous superquadratic function and  $\Phi : \mathbb{M}_n \rightarrow \mathbb{M}_m$  be a unital positive linear map. If  $\tau$  is an state on  $\mathbb{M}_m$ , then*

$$f(\tau(\Phi(A))) \leq \tau(\Phi(f(A))) - \tau(\Phi(f(|A - \tau(\Phi(A))|)))$$

for every positive matrix  $A$ .

*Proof.* If  $A$  is a positive matrix, then applying the functional calculus to (1.4) with  $t = A$  and then applying the positive linear functional  $\tau$  gives the inequality

$$\tau(f(A)) \geq f(s) + C_s(\tau(A) - s) + \tau(f(|A - s|))$$

for every  $s \geq 0$ . Put  $s = \tau(A)$  to obtain

$$\tau(f(A)) \geq f(\tau(A)) + \tau(f(|A - \tau(A)|)). \quad (2.14)$$

Now assume that  $\Phi : \mathbb{M}_n \rightarrow \mathbb{M}_m$  is a unital positive liner map. If  $\tau$  is an state on  $\mathbb{M}_m$ , then the mapping  $\psi_\tau : \mathbb{M}_n \rightarrow \mathbb{C}$  defined by  $\psi_\tau(X) = \tau(\Phi(X))$  is an state on  $\mathbb{M}_n$ . Applying (2.14) to  $\psi_\tau$  gives the desired inequality.  $\square$

The canonical trace is a positive linear functional on  $\mathbb{M}_n$ . If  $\tau(A) = 1/n\text{Tr}(A)$ , then Lemma 2.6 concludes the following result.

**Proposition 2.7.** *Let  $f : [0, \infty) \rightarrow \mathbb{R}$  be a continuous superquadratic function. If  $\Phi : \mathbb{M}_n \rightarrow \mathbb{M}_m$  is a unital positive linear map, then*

$$f\left(\frac{1}{n}\text{Tr}\Phi(A)\right) \leq \frac{1}{n}\text{Tr}\left[\Phi(f(A)) - \Phi\left(f\left(\left|A - \frac{1}{n}\text{Tr}\Phi(A)\right|\right)\right)\right]$$

for every positive matrix  $A \in \mathbb{M}_n^+$ .

In the next result, we present another variant of the Hansen-Pedersen trace inequality (2.9) for superquadratic functions. We need a well-known fact from matrix analysis.

**Lemma 2.8.** [6] *If  $A \in \mathbb{H}_n$  is a Hermitian matrix, then*

$$\sum_{j=1}^k \lambda_j(A) = \max \sum_{j=1}^k \langle A\mathbf{u}_j, \mathbf{u}_j \rangle, \quad (k = 1, \dots, n) \quad (2.15)$$

where the maximum is taken over all choices of orthonormal set of vectors  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ .

**Proposition 2.9.** *Let  $f : [0, \infty) \rightarrow \mathbb{R}$  be a continuous superquadratic function. If  $\Phi : \mathbb{M}_n \rightarrow \mathbb{M}_m$  is a unital positive linear map, then*

$$\text{Tr}f(\Phi(A)) \leq \text{Tr}\Phi(f(A)) - \min \left\{ \sum_{j=1}^n \langle \Phi(f(|A - \langle \Phi(A)\mathbf{u}_j, \mathbf{u}_j \rangle|))\mathbf{u}_j, \mathbf{u}_j \rangle \right\},$$

for every positive matrix  $A \in \mathbb{M}_n^+$ , where the minimum is taken over all choices of orthonormal system of vectors  $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ .

*Proof.* Assume that  $\lambda_1, \dots, \lambda_n$  are eigenvalues of  $\Phi(A)$  and  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  is orthonormal system of corresponding eigenvectors of  $\Phi(A)$ . Then

$$\begin{aligned} \text{Tr}f(\Phi(A)) &= \sum_{j=1}^n f(\lambda_j(\Phi(A))) \\ &= \sum_{j=1}^n f(\langle \Phi(A)\mathbf{u}_j, \mathbf{u}_j \rangle) \\ &\leq \sum_{j=1}^n [\langle \Phi(f(A))\mathbf{u}_j, \mathbf{u}_j \rangle - \langle \Phi(f(|A - \langle \Phi(A)\mathbf{u}_j, \mathbf{u}_j \rangle|))\mathbf{u}_j, \mathbf{u}_j \rangle] \quad (\text{by (1.7)}) \\ &\leq \text{Tr } \Phi(f(A)) - \sum_{j=1}^n \langle \Phi(f(|A - \langle \Phi(A)\mathbf{u}_j, \mathbf{u}_j \rangle|))\mathbf{u}_j, \mathbf{u}_j \rangle, \end{aligned}$$

in which the last inequality follows from Lemma 2.8. This completes the proof.  $\square$

### 3. KLEIN INEQUALITY

In this section, we present a Klein trace inequality for superquadratic functions. In particular, we show that if  $f$  is non-negative, a refinement of the Klein inequality (1.1) holds. The next lemma can be found in [6].

**Lemma 3.1.** [6] *If  $X, Y \in \mathbb{M}_n$  are Hermitian matrices, then the inequality*

$$\text{Tr}XY \leq \langle \lambda^\downarrow(X), \lambda^\downarrow(Y) \rangle \quad (3.1)$$

holds.

The main result of this section is the following Klein inequality for superquadratic functions.

**Theorem 3.2** (Klein's Inequality for superquadratic functions). *Assume that  $f : [0, \infty) \rightarrow \mathbb{R}$  is a differentiable superquadratic function with  $f(0) = f'(0) = 0$ . Then*

$$\text{Tr}[f(A) - f(B) - (A - B)f'(B)] \geq \min \left\{ \sum_{j=1}^n f(|x - y|); x \in \sigma(A), y \in \sigma(B) \right\} \quad (3.2)$$

for all  $A, B \in \mathbb{M}_n^+$  in which  $\sigma(A)$  is the set of eigenvalues of  $A$ . In particular, if  $f$  is non-negative, then

$$\text{Tr}[f(A) - f(B) - (A - B)f'(B)] \geq \text{Tr}f(|A - B|) \quad (3.3)$$

for all  $A, B \in \mathbb{M}_n^+$ .

*Proof.* First we prove (3.3). Suppose that  $\lambda_j$  and  $\mu_j$  ( $j = 1, \dots, n$ ) are eigenvalues of  $A$  and  $B$ , respectively, arranged in decreasing order. If  $f$  is non-negative, then  $f'$  is a

monotone increasing function by Lemma 2.3 and so  $f'(\mu_j)$  ( $j = 1, \dots, n$ ) are eigenvalues of  $f(B)$  arranged in decreasing order. Hence

$$\begin{aligned} \text{Tr}(A - B)f'(B) &= \text{Tr } Af'(B) - \text{Tr } Bf'(B) \\ &= \text{Tr } Af'(B) - \sum_{j=1}^n \mu_j f'(\mu_j) \\ &\leq \sum_{j=1}^n \lambda_j f'(\mu_j) - \sum_{j=1}^n \mu_j f'(\mu_j) \quad \text{by (3.1)} \\ &= \sum_{j=1}^n (\lambda_j - \mu_j) f'(\mu_j). \end{aligned}$$

Moreover, it follows from proof of Theorem 2.4 that

$$\text{Tr } f(|A - B|) \leq \sum_{j=1}^n f(|\lambda_j - \mu_j|). \quad (3.4)$$

Note that if a superquadratic function  $f$  is differentiable on  $(0, \infty)$  and  $f(0) = f'(0) = 0$ , then Lemma 2.3 implies that

$$f(t) \geq f(s) + f'(s)(t - s) + f(|t - s|)$$

for all  $s, t \geq 0$ . This gives

$$f(\lambda_j) \geq f(\mu_j) + f'(\mu_j)(\lambda_j - \mu_j) + f(|\lambda_j - \mu_j|) \quad (j = 1, \dots, n)$$

and so

$$\sum_{j=1}^n f(\lambda_j) \geq \sum_{j=1}^n f(\mu_j) + \sum_{j=1}^n f'(\mu_j)(\lambda_j - \mu_j) + \sum_{j=1}^n f(|\lambda_j - \mu_j|), \quad (3.5)$$

which proves (3.3). In general case, when  $f$  is not assumed to be non-negative, we suppose that  $\lambda_j$  ( $j = 1, \dots, n$ ) are eigenvalues of  $A$  arranged in decreasing order and  $\mu_j$  ( $j = 1, \dots, n$ ) are eigenvalues of  $B$ , arranged in such a way that  $f'(\mu_1) \geq \dots \geq f'(\mu_n)$ . By a same argument as in the first part of the proof, this guarantees the inequality  $\text{Tr}(A - B)f'(B) \leq \sum_{j=1}^n (\lambda_j - \mu_j) f'(\mu_j)$ . It follows from (3.5) that

$$\text{Tr } f(A) \geq \text{Tr } f(B) + \text{Tr}(A - B)f'(B) + \sum_{j=1}^n f(|\lambda_j - \mu_j|),$$

from which we get (3.2).  $\square$

When the superquadratic function  $f$  is non-negative, then Theorem 3.2 gives a refinement of the Klein's inequality (1.1) for convex functions. Indeed, if  $f \geq 0$ , then

$$0 \leq \text{Tr}[f(A) - f(B) - (A - B)f'(B) - f(|A - B|)] \leq \text{Tr}[f(A) - f(B) - (A - B)f'(B)].$$

**Example 3.3.** The function  $f(t) = t^p$  is superquadratic for every  $p \geq 2$ . Theorem 3.2 gives

$$0 \leq \text{Tr}[A^p - B^p - p(A - B)B^{p-1} - |A - B|^p] \leq \text{Tr}[A^p - B^p - p(A - B)B^{p-1}]$$

for all  $A, B \in \mathbb{M}_n^+$  and every  $p \geq 2$ .

As a simple example, assume that  $p = 3$  and consider the positive matrices

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then

$$\text{Tr}[A^p - B^p - p(A - B)B^{p-1}] = 20 \quad \text{and} \quad \text{Tr}|A - B|^p \simeq 14.15.$$

On the other hand, if  $f \geq 0$  is a convex function and  $-f$  is a superquadratic function, then Theorem 3.2 provides an upper bound for the Klein's Inequality. Applying Theorem 3.2 to the superquadratic function  $-f$  we obtain

$$\text{Tr}[f(A) - f(B) - (A - B)f'(B)] \leq \max \left\{ \sum_{j=1}^n f(|x - y|); x \in \sigma(A), y \in \sigma(B) \right\} \quad (3.6)$$

for all  $A, B \in \mathbb{M}_n^+$ , while the left side is positive due to the Klein's Inequality for the convex function  $f$ .

**Example 3.4.** If  $1 \leq p \leq 2$ , then the function  $f(t) = t^p$  is convex and  $-f(t) = -t^p$  is superquadratic. It follows from (3.6) that

$$\text{Tr}[A^p - B^p - p(A - B)B^{p-1}] \leq \max \left\{ \sum_{j=1}^n |x - y|^p; x \in \sigma(A), y \in \sigma(B) \right\},$$

for all  $A, B \in \mathbb{M}_n^+$  and every  $1 \leq p \leq 2$ .

To see a simple example, let  $p = 3/2$  and consider the two matrices in Example 3.3.

Then

$$\text{Tr}[A^p - B^p - p(A - B)B^{p-1}] \simeq 3.36 \quad \text{and} \quad \max \left\{ \sum_{j=1}^n |x - y|^p; x \in \sigma(A), y \in \sigma(B) \right\} \simeq 6.19.$$

If  $f$  is a continuous convex function, then  $f(\langle A\mathbf{u}, \mathbf{u} \rangle) \leq \langle f(A)\mathbf{u}, \mathbf{u} \rangle$  for every unit vector  $\mathbf{u} \in \mathbb{C}^n$ , see [11]. If  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  is an orthonormal basis of  $\mathbb{C}^n$ , then

$$\begin{aligned} \sum_{j=1}^n f(\langle A\mathbf{u}_j, \mathbf{u}_j \rangle) &\leq \sum_{j=1}^n \langle f(A)\mathbf{u}_j, \mathbf{u}_j \rangle \\ &\leq \sum_{j=1}^n \lambda_j(f(A)) \quad \text{by (2.15)} \\ &= \text{Tr}f(A). \end{aligned}$$

In other words,

$$\sum_{j=1}^n f(\langle A\mathbf{u}_j, \mathbf{u}_j \rangle) \leq \text{Tr} f(A). \quad (3.7)$$

for every orthonormal basis  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  of  $\mathbb{C}^n$ . Inequality (3.7) is known as the *Peierls inequality*. The equality holds in (3.7) when  $\mathbf{u}_i$ 's are eigenvectors of  $A$ .

We present a variant of the Peierls inequality in the case when  $f$  is a superquadratic function. It gives in particular a refinement of the Peierls inequality if  $f$  is non-negative.

**Proposition 3.5.** *Assume that  $f$  is a superquadratic function. If  $A \in \mathbb{M}_n^+$ , then*

$$\sum_{j=1}^n f(\langle A\mathbf{u}_j, \mathbf{u}_j \rangle) + \sum_{j=1}^n \langle f(|A - \langle A\mathbf{u}_j, \mathbf{u}_j \rangle|) \mathbf{u}_j, \mathbf{u}_j \rangle \leq \text{Tr} f(A) \quad (3.8)$$

for every orthonormal basis  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  of  $\mathbb{C}^n$ . Equality holds if  $f$  is non-negative and  $\mathbf{u}_i$ 's are eigenvectors of  $A$ .

*Proof.* Let  $f$  be a superquadratic function. We apply the Jensen's operator inequality (1.6) and then we use (2.15). This gives (3.8).

If  $f$  is non-negative, then

$$\sum_{j=1}^n f(\langle A\mathbf{u}_j, \mathbf{u}_j \rangle) \leq \sum_{j=1}^n f(\langle A\mathbf{u}_j, \mathbf{u}_j \rangle) + \sum_{j=1}^n \langle f(|A - \langle A\mathbf{u}_j, \mathbf{u}_j \rangle|) \mathbf{u}_j, \mathbf{u}_j \rangle \leq \text{Tr} f(A).$$

Hence, choosing  $\mathbf{u}_i$ 's to be the eigenvectors of  $A$ , gives the equality in (3.7) and so in (3.8)  $\square$

## REFERENCES

- [1] S. Abramovich, S. Ivelić and J. Pečarić, *Improvement of Jensen-Steffensen's inequality for superquadratic functions*, Banach J. Math. Anal. **4** (1) (2010), 146–158.
- [2] S. Abramovich, G. Jameson and G. Sinnamon, *Inequalities for averages of convex and superquadratic functions*, J. Inequ. Pure Appl. Math. **5** (4) (2004), Article 91.
- [3] S. Abramovich, G. Jameson and G. Sinnamon, *Refining Jensen's inequality*, Bull. Math. Soc. Sci. Math. Roumanie, **47** (2004), 3–14.
- [4] M. W. Alomari, *Operator Popoviciu's inequality for superquadratic and convex functions of selfadjoint operators in Hilbert spaces*, Advan. Pure Appl. Math., accepted.
- [5] S. Banić, J. Pečarić and S. Varošanec, *Superquadratic functions and refinements of some classical inequalities*, J. Korean Math. Soc. **45** (2) (2008), 513–525.
- [6] R. Bhatia, *Matrix Analysis*, Springer-Verlag New York, 1997.
- [7] E. A. Carlen, *Trace inequalities and quantum entropy: An introductory course*, Book chapter, Contemporary Mathematics, 2010, DOI:10.1090/conm/529/10428.
- [8] E. A. Carlen and E. H. Lieb, *A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity*, Lett. Math. Phys. **83** (2008), 107–126.
- [9] S. Furuchi and M. Lin, *A matrix trace inequality and its application*, Linear Algebra Appl. **433** (2010), 1324–1328.

- [10] S. Furuchi, K. Kuriyama and K. Yanagi, *Trace inequalities for products of matrices*, Linear Algebra Appl. **430** (2009), 2271–2276.
- [11] T. Furuta, J. Mićić, J. Pečarić and Y. Seo, *Mond-Pečarić method in operator inequalities: Inequalities for bounded self-adjoint operators on a Hilbert space*, Element, Zagreb, 2005.
- [12] F. Hansen G.K. Pedersen, *Jensen's operator inequality*, Bull. London Math. Soc. **35** (2003), 553–564.
- [13] F. Hiai, *Concavity of certain matrix trace functions*, Taiwanese J. Math. **5**, no. 3 (2001), 535–554.
- [14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, 2013.
- [15] M. Kian, *Operator Jensen inequality for superquadratic functions*, Linear Algebra Appl. **456** (2014), 82–87.
- [16] M. Kian and S.S. Dragomir, *Inequalities involving superquadratic functions and operators*, *Mediterr. J. Math.* **11** (4) (2014), 1205–1214.
- [17] E. H. Lieb, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, *Advances in Math.* **11** (1973), 267–288.
- [18] D. Petz, *A survey of certain trace inequalities, functional analysis and operator theory center publications*, Vol. 30, Institute of Mathematics, Polish Academy od Sciences, Warszawa, 1994.
- [19] Kh. Shebrawi and H. Albadawi, *Trace inequalities for matrices*, Bull. Aust. Math. Soc. **87** (2013), 139–148.
- [20] Xin Min Yang, *A Matrix trace inequality*, J. Math. Anal. Appl. **263** (2001), 327–331.

MOHSEN KIAN: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BOJNORD, P.O. Box 1339, BOJNORD 94531, IRAN

*E-mail address:* `kian@ub.ac.ir`

MOHAMMAD W. ALOMARI: DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE AND INFORMATION TECHNOLOGY, IRBID NATIONAL UNIVERSITY, P.O. Box 2600, IRBID, P.C. 21110, JORDAN.

*E-mail address:* `mwomath@gmail.com`