A SURVEY OF CERTAIN TRACE INEQUALITIES

DÉNES PETZ

Department of Mathematics, Faculty of Chemical Engineering
Technical University Budapest
Sztoczek u. 2, H ép II. 25, H-1521 Budapest XI, Hungary

This paper concerns inequalities like $\text{Tr} A \leq \text{Tr} B$, where $A$ and $B$ are certain Hermitian complex matrices and $\text{Tr}$ stands for the trace. In most cases $A$ and $B$ will be exponential or logarithmic expressions of some other matrices. Due to the interest of the author in quantum statistical mechanics, the possible applications of the trace inequalities will be commented from time to time. Several inequalities treated below have been established in the context of Hilbert space operators or operator algebras. Notwithstanding these extensions our discussion will be limited to matrices.

1. The trace of matrices. Before discussing trace inequalities we consider characterizations of the trace functional. Below $M_n$ will denote the algebra of $n \times n$ complex matrices and $M_n^{sa}$ will stand for the Hermitian part. We consider $\text{Tr}$ as a linear functional on $M_n$. It is well known that each of the following properties characterizes the trace functional up to a constant factor among the linear functionals on $M_n$.

   (i) $\tau(AB - BA) = 0$ for every $A$ and $B$.
   (ii) $|\tau(A)| \leq cr(A)$ for every $A$, where $c$ is a constant and $r$ denotes the spectral radius.
   (iii) $A^2 = 0$ implies $\tau(A) = 0$.
   (iv) $|\tau(A^k)| \leq \tau(A^{*}A)^{k/2}$ for some $k \in \mathbb{N}$ and for every $A$.

The selfadjoint idempotent matrices in $M_n$, called projections, correspond to subspaces of the linear space $\mathbb{C}^n$. So the join $P \lor Q$ of the projections $P$ and $Q$
may be defined as the orthogonal projection onto the linear span of the subspaces corresponding to $P$ and $Q$. Similarly, the meet $P \cap Q$ projects onto the intersection of these subspaces. With the operations $\vee$ and $\wedge$ the set $\mathcal{P}$ of projections in $M_n$ becomes a lattice which plays an important role in quantum mechanics. A function $f : \mathcal{P} \to \mathbb{R}^+$ is called \textit{subadditive} if $f(P \vee Q) \leq f(P) + f(Q)$ for every $P, Q \in \mathcal{P}$.

(v) Up to a constant factor, $\text{Tr}$ is the only linear functional which is subadditive when restricted to $\mathcal{P}$.

The last two characterizations of the trace were found in [25] and treated there in the more general context of $C^*$-algebras. The consequence

(1) $|\text{Tr}(ABAB)| \leq \text{Tr}(A^*ABB^*)$

of (iv) will also be used below.

\section*{2. Inequalities to warm up.} In this section we consider some trace inequalities that are obtained by diagonalization of matrices or by simple considerations about their eigenvalues. For example, the first proposition is based on the following fact. If $A$ and $B$ are selfadjoint matrices with eigenvalues $\kappa_1 \geq \ldots \geq \kappa_n$ and $\lambda_1 \geq \ldots \geq \lambda_n$, then $A \leq B$ implies $\kappa_i \leq \lambda_i$ for every $1 \leq i \leq n$. Recall that if $\sum_s \lambda_s p_s$ is the spectral decomposition of $A$ and the real function $f$ is defined on the spectrum $\text{Spec}(A)$ of $A$ then $f(A)$ is defined by $f(A) = \sum_s f(\lambda_s)p_s$.

**Proposition 1.** Let $A$ and $B$ be selfadjoint matrices and let $f : \mathbb{R} \to \mathbb{R}$ be increasing. Then $A \leq B$ implies

$$\text{Tr} f(A) \leq \text{Tr} f(B).$$

**Proposition 2.** Let $f : [\alpha, \beta] \to \mathbb{R}$ be convex. Then the functional

$$F(A) = \text{Tr} f(A)$$

is convex on the set $\{A \in M_n^{sa} : \text{Spec}(A) \subset [\alpha, \beta]\}$.

**Proof.** First we note that for a pairwise orthogonal family $(p_i)$ of minimal projections with $\sum p_i = I$ we have

(2) $\text{Tr} f(B) \geq \sum_i f(\text{Tr} Bp_i).$

Indeed, using the convexity of $f$ we deduce (2) as follows. Let $\sum_j s_j q_j$ be the spectral decomposition of $B$. Then

$$\text{Tr} f(B) = \sum_j f(s_j) \text{Tr} q_j = \sum_i \sum_j f(s_j) \text{Tr} q_j p_i \geq \sum_i f(\sum_j s_j \text{Tr} q_j p_i) = \sum_i f(\text{Tr} Bp_i).$$

To prove the proposition we write $\sum \mu_i p_i$ for the spectral decomposition of the convex combination $A = \lambda B_1 + (1 - \lambda)B_2$. Applying (2) twice we infer that
\[
\lambda \text{Tr} f(B_1) + (1 - \lambda) \text{Tr} f(B_2) \\
\geq \lambda \sum_i f(\text{Tr} B_1 p_i) + (1 - \lambda) \sum_i f(\text{Tr} B_2 p_i) \\
\geq \sum_i f(\lambda \text{Tr} B_1 p_i + (1 - \lambda) \text{Tr} B_2 p_i) = \sum_i f(\text{Tr} A p_i) = \text{Tr} f(A),
\]
which is the convexity of the functional \( F \).

Some particular cases of the next simple and useful observation are sometimes called Klein inequalities.

**Proposition 3.** If \( f_k, g_k : [\alpha, \beta] \to \mathbb{R} \) are such that for some \( c_k \in \mathbb{R} \),
\[
\sum_k c_k f_k(x) g_k(y) \geq 0 \quad \text{for every } x, y \in [\alpha, \beta],
\]
then
\[
\sum_k c_k \text{Tr} f_k(A) g_k(B) \geq 0
\]
whenever \( A, B \) are selfadjoint matrices with \( \text{Spec}(A), \text{Spec}(B) \subset [\alpha, \beta] \).

**Proof.** Let \( A = \sum \lambda_i p_i \) and \( B = \sum \mu_j q_j \) be the spectral decompositions. Then
\[
\sum_k c_k \text{Tr} f_k(A) g_k(B) = \sum_k \sum_{i,j} c_k \text{Tr} p_i f_k(A) g_k(B) q_j \\
= \sum_{i,j} \text{Tr} p_i q_j \sum_k c_k f_k(\lambda_i) g_k(\mu_j) \geq 0
\]
by the hypothesis.

In particular, if \( f \) is convex then
\[
f(x) - f(y) - (x - y)f'(y) \geq 0
\]
and
\[
\text{Tr} f(A) \geq \text{Tr} f(B) + \text{Tr}(A - B)f'(B).
\]
For the choice \( f(t) = -\eta(t) = t \log t \) we obtain
\[
S(A, B) \equiv \text{Tr} A(\log A - \log B) \geq \text{Tr}(A - B)
\]
for \( B \) strictly positive and \( A \) nonnegative. The left-hand side is called relative entropy. If \( A \) and \( B \) are density matrices, i.e. \( \text{Tr} A = \text{Tr} B = 1 \), then \( S(A, B) \geq 0 \). This is a classical application of the Klein inequality (cf. [27]). The stronger estimate
\[
-\eta(x) + \eta(y) + (x - y)\eta'(y) \geq \frac{1}{2}(x - y)^2
\]
allows another use of the Klein inequality. Namely,
\[
S(A, B) \geq \frac{1}{2} \text{Tr}(A - B)^2 + \text{Tr}(A) - \text{Tr}(B),
\]
which was obtained in [31].
From the inequality $1 + \log x \leq x$ ($x > 0$) one obtains
\[ t^{-1}(a - a^{-t}b) \leq a(\log a - \log b) \leq t^{-1}(a^{1+t}b^{-t} - a) \]
for $a, b, t > 0$. If $T$ and $S$ are nonnegative invertible matrices then Proposition 3 gives
\[ t^{-1} \text{Tr}(S - S^{-1}T^t) \leq \text{Tr}(\log S - \log T) \leq t^{-1} \text{Tr}(S^{1+t}T^{-t} - S), \]
which provides a lower as well as an upper estimate for the relative entropy [29].

3. The Golden–Thompson inequality and its extensions. In statistical mechanics Golden [13] has proved that if $A$ and $B$ are Hermitian and nonnegative definite matrices then
\[ \text{Tr} e^A e^B \geq \text{Tr} e^{A+B}. \]
He observed that this inequality may be used to obtain lower bounds for the Helmholtz free-energy function by partitioning the hamiltonian. Independently, C. J. Thompson proved (8) for Hermitian $A$ and $B$ without the requirement of definiteness and applied the inequality to obtain an upper bound for the partition function of an antiferromagnetic chain [32]. Nowadays (8) is termed the Golden–Thompson inequality and it is a basic tool in quantum statistical mechanics.

The simplest proof of the Golden–Thompson inequality uses the following exponential product formula for matrices.

**Lemma 4.** For any complex $n \times n$ matrices $A$ and $B$,
\[ \lim_{s \to \infty} (e^{A/s}e^{B/s})^s = \lim_{s \to \infty} (e^{B/(2s)}e^{A/s}e^{B/(2s)})^s = e^{A+B}. \]

It is worthwhile to note that [10] contains interesting historical remarks concerning the origin of the previous lemma (I).

We also need the inequality
\[ |\text{Tr} X^{2k}| \leq \text{Tr}(XX^*)^k, \]
which appeared in characterization (iv) of the matrix trace. Note that if $k = 1$ and $X = VH$ with selfadjoint $V$ and $H$, then (9) reduces to
\[ \text{Tr} VH^2V^2 \leq \text{Tr} V^2H^2, \]
which is a particular case of the inequality
\[ \text{Tr} VH^*g(H) \leq \text{Tr} YY^*Hg(H), \]
which holds provided that $H$ is selfadjoint and $g : \mathbb{R} \to \mathbb{R}$ is increasing [21]. The next theorem together with its proof is taken from [10].

**Theorem 5.** For every $A, B \in M_n$,
\[ \text{Tr} e^{(A+A^*)/2}e^{(B+B^*)/2} \geq |\text{Tr} e^{A+B}|. \]

(1) Editorial note: See also the application on p. 370 in this volume.
Proof. Substituting $X = AB$ into (9) we have
\[
\text{Tr}((AB^*A^*)^k) \geq |\text{Tr}((AB)^{2k})|,
\]
where the left-hand side is nothing else but $\text{Tr}(BB^*A^*A)$. Setting $s = 2^{k-1}$ with a positive integer $k$ and using (9) gives
\[
|\text{Tr}(AB)^{2k}| \leq \text{Tr}(BB^*A^*A)^{2k-1} \leq |\text{Tr}((BB^*A^*A)^2)|^{2k-2}
\[
= \text{Tr}((A^*A)^2(BB^*A^*A)^{2k-2}).
\]
By a repeated application of this argument we easily infer that
\[
\text{Tr}((A^*A)^{2k-1}(BB^*)^{2k-1}) \geq |\text{Tr}((AB)^{2k})|.
\]
Now replace $A$ by $\exp(2^{-k}A)$ and $B$ by $\exp(2^{-k}B)$:
\[
\text{Tr}((e^{2^{-k}A}e^{2^{-k}A})^{2k/2}(e^{2^{-k}B}e^{2^{-k}B})^{2k/2}) \geq |\text{Tr}((e^{2^{-k}A}e^{2^{-k}B})^{2k})|.
\]
The obvious continuity of $\text{Tr}$ together with the exponential product formula (that is, Lemma 4) allows us to obtain the theorem.

Inequality (8) is an obvious consequence of the theorem coupled with the following

**Corollary 6.** If $A$ and $B$ are selfadjoint then
\[
|\text{Tr}e^{A+iB}| \leq \text{Tr}e^{A}.
\]

The relative entropy of nonnegative matrices defined by (5) is related to the functional $B \mapsto \log \text{Tr}e^{A+B}$ by the Legendre transform. Namely, $B \mapsto \log \text{Tr}e^{A+B}$ is the Legendre transform or the conjugate function of $X \mapsto S(X,Y)$ when $Y = e^B$ and vice versa. This was proved in [24] in the general setup of von Neumann algebras; here is an elementary proof from [16].

**Proposition 7.** If $A$ is Hermitian and $Y$ is strictly positive, then
\[
\log \text{Tr}e^{A+iB} = \max\{\text{Tr}XA - S(X,Y) : X \text{ is positive}, \text{Tr}X = 1\}.
\]

On the other hand, if $X$ is positive with $\text{Tr}X = 1$ and $B$ is Hermitian, then
\[
S(X,e^B) = \max\{\text{Tr}XA - \log \text{Tr}e^{A+B} : A \text{ is Hermitian}\}.
\]

**Proof.** Define
\[
F(X) = \text{Tr}XA - S(X,Y)
\]
for nonnegative $X$ with $\text{Tr}X = 1$. When $P_1, \ldots, P_n$ are projections of rank one with $\sum_{i=1}^{n} P_i = 1$, we write
\[
F\left(\sum_{i=1}^{n} \lambda_i P_i\right) = \sum_{i=1}^{n} (\lambda_i \text{Tr}P_iA + \lambda_i \text{Tr}P_i\log Y - \lambda_i \log \lambda_i),
\]
$\lambda_i \geq 0$, $\sum_{i=1}^n \lambda_i = 1$. Since

$$\frac{\partial}{\partial \lambda_i} F\left( \sum_{i=1}^n \lambda_i P_i \right) \bigg|_{\lambda_i=0} = +\infty,$$

we see that $F(X)$ attains its maximum at a positive matrix $X_0$ with $\text{Tr} X_0 = 1$. Then for any Hermitian $S$ with $\text{Tr} S = 0$, we have

$$0 = \frac{d}{dt} F(X_0 + tS) \bigg|_{t=0} = \text{Tr} S (A + \log Y - \log X_0),$$

so that $A + \log Y - \log X_0 = cI$ with $c \in \mathbb{R}$. Therefore $X_0 = e^{A+\log Y} / \text{Tr} e^{A+\log Y}$ and $F(X_0) = \log \text{Tr} e^{A+\log Y}$ by simple computation.

We now prove (13). It follows from (12) that the functional $A \mapsto \log \text{Tr} e^{A+B}$ defined on the Hermitian matrices is convex. Let $A_0 = \log X - B$ and

$$G(A) = \text{Tr} X A - \log \text{Tr} e^{A+B},$$

which is concave on the Hermitian matrices. Then for any Hermitian $S$ we have

$$\frac{d}{dt} G(A_0 + tS) \bigg|_{t=0} = 0,$$

because $\text{Tr} X = 1$ and

$$\frac{d}{dt} \text{Tr} e^{\log X + tS} \bigg|_{t=0} = \text{Tr} X S.$$

Therefore $G$ has the maximum $G(A_0) = \text{Tr} X (\log X - B)$, which is the relative entropy of $X$ and $e^B$.

Let us make the following definition:

$$(14) \quad S_{\text{co}}(X,e^B) = \max \{ \text{Tr} X A - \log \text{Tr} e^{A+B} : A \text{ is Hermitian} \}.$$  

(The interested reader will find an explanation for the notation in [15].) It follows from the Golden–Thompson inequality that

$$(15) \quad S_{\text{co}}(X,e^B) \leq S(X,e^B).$$

This inequality may be proved within the theory of relative entropy. In fact, it is a particular case of the monotonicity of the relative entropy. In [22, 23] it was established that $[X,e^B] = 0$ is a necessary and sufficient condition for equality in (15).

Conversely, the Golden–Thompson inequality can be recovered from (15). Putting $X = e^{A+B} / \text{Tr} e^{A+B}$ for Hermitian $A$ and $B$ we have

$$\log \text{Tr} e^{A+B} \geq \text{Tr} X A - S_{\text{co}}(X,e^B) \geq \text{Tr} X A - S(X,e^B) = \log \text{Tr} e^{A+B}$$

by (15), which further shows that $\text{Tr} e^{A+B} = \text{Tr} e^A e^B$ holds if and only if $AB = BA$. This derivation of the Golden–Thompson inequality as well as characteriza-
tion of equality were performed in [24, Corollary 5] in the general setup of von Neumann algebras.

In the course of proving Theorem 5 the inequality

$$\text{Tr}(X^{1/2}YX^{1/2})^q \leq \text{Tr} X^{q/2}Y^q X^{q/2}$$

was obtained for \( q = 2^k \) and positive matrices \( X \) and \( Y \). According to Araki [5],

$$\text{Tr}(X^{1/2}YX^{1/2})^r \leq \text{Tr}(X^{r/2}Y^r X^{r/2})^p$$

for every \( r \geq 1 \) and \( p > 0 \). This implies that the function

$$p \mapsto \text{Tr}(e^{pB/2}e^{pA}e^{pB/2})^{1/p}$$

is increasing for \( p > 0 \). Its limit at \( p = 0 \) is \( \text{Tr} e^{A+B} \). Hence the next theorem is a strengthened variant of the Golden–Thompson inequality.

**Theorem 8.** The function \( \text{Tr}(e^{pB/2}e^{pA}e^{pB/2})^{1/p} \) is increasing in \( p \in (0, \infty) \) for Hermitian matrices \( A \) and \( B \). Its limit at \( p = 0 \) is \( \text{Tr} e^{A+B} \). In particular, for any \( p > 0 \),

$$\text{Tr} e^{A+B} \leq \text{Tr}(e^{pB/2}e^{pA}e^{pB/2})^{1/p}.$$  

It was proved by Friedland and So that the function (17) is either strictly monotone or constant [11]. The latter case corresponds to the commutativity of \( A \) and \( B \).

The formal generalization

$$\text{Tr} e^{A+B+C} \leq \text{Tr} e^{A} e^{B} e^{C}$$

of the Golden–Thompson inequality is false. However, if two of the three matrices commute then the inequality holds obviously. A nontrivial extension of the Golden–Thompson inequality to three operators is due to Lieb [19]. Before stating this extension we introduce some positive operators on the space \( M_n \) of matrices, which becomes a Hilbert space when endowed with the Hilbert–Schmidt scalar product:

$$\langle A, B \rangle = \text{Tr} AB^*.$$  

For \( A \in M_n \) let \( \mathcal{T}_{\exp A} : M_n \to M_n \) be defined by

$$\mathcal{T}_{\exp A}(K) = \int_0^\infty (t + \exp A)^{-1} K (t + \exp A)^{-1} dt.$$  

Since

$$\langle \mathcal{T}_{\exp A}(K), K \rangle = \int_0^\infty \text{Tr}(t + \exp A)^{-1} K (t + \exp A)^{-1} K^* dt$$

is nonnegative, the operator \( \mathcal{T}_{\exp A} \) is positive (definite). In a basis in which \( A \equiv \text{Diag}(a_1, \ldots, a_n) \) one can compute \( \mathcal{T}_{\exp A} \) explicitly. Namely,

$$(\mathcal{T}_{\exp A}(K))_{ij} = K_{ij}/\text{Lm}(e^{a_i}, e^{a_j}).$$
where \( \text{Lm}(x, y) \) stands for the so-called logarithmic mean defined by
\[
\text{Lm}(x, y) = \begin{cases} 
(x - y)/(\log x - \log y) & \text{if } x \neq y, \\
x & \text{if } x = y.
\end{cases}
\]
Note that if \( K = K^* \) and \( AK = KA \), then \( T_{\exp A}(K) = \exp(-A)K \).

**Theorem 9.** Let \( A, B \) and \( C \) be Hermitian matrices. Then
\[
\text{Tr} e^{A^2} e^B + C \leq e^{\text{Tr} e^{-A} (e^B)} e^{C}.
\]

Extensions of the Golden–Thompson inequality to infinite dimensions have extensive literature [8, 18, 4, 17, 28]. The review [33] contains several interesting results on the exponential function of matrices.

The next theorem is due to Bernstein except of the case of equality which was added by So [7, 30]. Although it contains an exponential trace inequality it does not concern selfadjoint matrices and the direction of the inequality is opposite to that of the Golden–Thompson inequality.

**Theorem 10.** Let \( K \) be an arbitrary \( n \times n \) matrix. Then
\[
\text{Tr} e^{K^2} \geq \text{Tr} e^{K^*} e^{K^*}
\]
and equality holds if and only if \( K \) is normal.

### 4. Logarithmic inequalities

The Golden–Thompson inequality is remarkable because it establishes a relation between \( \text{Tr} e^{A+B} \) and \( \text{Tr} e^A e^B \). The logarithmic analogue would be a relation between \( \text{Tr} \log XY \) and \( \text{Tr}(\log X + \log Y) \) for positive matrices \( X \) and \( Y \). This relation is well known, \( \text{Tr} \log XY = \text{Tr}(\log X + \log Y) \), due to the multiplicativity of the determinant. However, a slight modification leads to a logarithmic trace inequality. Note that for positive (invertible) matrices \( X \) and \( Y \), one can define \( \log XY \) by analytic functional calculus or by power series and get the equality
\[
\text{Tr} X \log X^{1/2} Y X^{1/2} = \text{Tr} X \log XY
\]
because \( \text{Tr} X(Y^{1/2} X^{1/2})^n = \text{Tr} X(Y)^n \) for \( n \geq 1 \).

**Proposition 11.** Let \( X \) and \( Y \) be positive matrices. Then
\[
\text{Tr} X \log Y^{1/2} X Y^{1/2} \leq \text{Tr} X(\log X + \log Y) \leq \text{Tr} X \log XY.
\]

**Proof.** The first inequality is a consequence of (13) in the case \( \text{Tr} X = 1 \), which it is sufficient to consider. Let \( B \) be Hermitian and \( A = \log e^{-B/2} X e^{-B/2} \). Then by (13) we have
\[
\text{Tr} X(\log X - B) \geq \text{Tr} X A - \log \text{Tr} e^{A+B} \\
\quad \geq \text{Tr} X A - \log \text{Tr}(e^{B/2} e^A e^{B/2}) \\
\quad = \text{Tr} X \log e^{-B/2} X e^{-B/2} - \log \text{Tr} X \\
= \text{Tr} X \log e^{-B/2} X e^{-B/2}.
\]
Hence the first stated inequality follows by letting \( B = -\log Y \).
The second inequality is deeper and its proof was given within relative entropy theory. Setting
\[ S_{BS}(X, Y) = \text{Tr} \, X \log X^{1/2} Y^{-1/2} Y^{1/2} \]
we see that the second inequality is the same as
\[ S(X, Y) \leq S_{BS}(X, Y) \]
for positive matrices \( X \) and \( Y \) with \( \text{Tr} \, X = \text{Tr} \, Y = 1 \). (The quantity \( S_{BS} \) is related to the works \cite{6, 12}.) The proof of \( 24 \) was given in \cite{15, 16} and applies some properties of the relative entropy quantities \( S \) and \( S_{BS} \). (Namely, monotonicity and additivity under tensor products.) The crucial part of the proof is a relative entropy estimate which is stated in the next lemma.

For each \( m \in \mathbb{N} \) let \( A_m \) be the \( m \)-fold tensor product \( \otimes_m M_n \) which is identified with the \( n^m \times n^m \) matrix algebra \( M_{n^m} \). For a positive matrix \( Y \) in \( M_n \) we set \( Y_m = \otimes_m Y \) and write \( E_Y \) for the conditional expectation from \( A_m \) onto \( \{ Y_m \} \) with respect to the trace. (When \( Z = \sum \lambda_i P_i \) is the spectral decomposition of a selfadjoint \( Z \), then \( E_Z(A) = \sum \lambda_i P_i A P_i \).

**Lemma 12.** For every positive \( Z \) in \( A_m \) with \( \text{Tr} \, Z = 1 \),
\[ S(Z, E_{Y_m}(Z)) \leq n \log(m + 1) \cdot \]

Having the lemma at our disposal we obtain \( 24 \) from the chain
\[ mS_{BS}(X, Y) = S_{BS}(X_m, Y_m) \geq S_{BS}(E_{Y_m}(X_m), Y_m) \]
\[ = S(E_{Y_m}(X_m), Y_m) = S(X_m, Y_m) - S(X_m, E_{Y_m}(X_m)) \]
\[ \geq S(X_m, Y_m) - n \log(m + 1) = mS(X, Y) - n \log(m + 1) \]
after dividing by \( m \) and letting \( m \to \infty \). (For details we refer to the original papers.)

We note that inequality \( 24 \) is extended to infinite dimensions in \cite{14}.

For \( 0 \leq \alpha \leq 1 \) the **\( \alpha \)-power mean** of positive matrices \( X \) and \( Y \) is defined by
\[ X \#_\alpha Y = X^{1/2}(X^{-1/2}YX^{-1/2})^\alpha X^{1/2} \cdot \]
This is the operator mean corresponding to an operator monotone function \( x^\alpha \), \( x \geq 0 \). In particular, \( X \#_{1/2} Y = X \# Y \) is the geometric mean of \( X \) and \( Y \) which was introduced in \cite{26}.

In the rest of this section we review some further results from \cite{16}. For each \( p > 0 \) the following statements (i) and (ii) are proved to be equivalent:

(i) If \( A \) and \( B \) are Hermitian, then
\[ \text{Tr} \left( e^{pA} \#_\alpha e^{pB} \right)^{1/p} \leq \text{Tr} \, e^{(1-\alpha)A + \alpha B} \]
for \( 0 \leq \alpha \leq 1 \).

(ii) If \( X \) and \( Y \) are positive, then
\[ \text{Tr} \, X (\log X + \log Y) \leq \frac{1}{p} \text{Tr} \, X (\log X^{p/2}Y^{p/2}) \cdot \]

(26)
Observe that the logarithmic inequality (26) extends the second inequality of Proposition 11. The equivalent exponential inequality (25) is opposite to that of Golden–Thompson. (25) and (26) are related by differentiation:
\[
\frac{d}{d\alpha} \text{Tr} \left( e^{pA} \# e^{pB} \right)^{1/p} \bigg|_{\alpha=0} = \frac{1}{p} \text{Tr} e^A \log e^{-pA/2} e^{pB} e^{-pA/2}.
\]

**Theorem 13.** Let \(A\) and \(B\) be Hermitian and \(0 \leq \alpha \leq 1\). Then the inequality (25) holds for every \(p > 0\). Moreover, the left-hand side converges to the right-hand side as \(p \to 0\).

This theorem shows some analogy with Theorem 8. While the convergence in Theorem 8 is based on Lemma 4 (called the Lie exponential formula), there is a somewhat similar formula with power means. We state it in the form of a lemma. Its proof does not differ essentially from the standard proof of the exponential product formula.

**Lemma 14.** If \(A\) and \(B\) are Hermitian and \(0 \leq \alpha \leq 1\), then
\[
\lim_{s \to \infty} (e^{A/s} \# e^{B/s})^s = e^{(1-\alpha)A + \alpha B}.
\]

**Theorem 15.** Let \(X\) and \(Y\) be nonnegative. Then the inequality (26) holds for every \(p > 0\). Moreover, the right-hand side converges to the left-hand side as \(p \to 0\).

It was conjectured in [16] that the limit appearing in the previous theorem is monotone. This follows from Theorem 17 below.

### 5. Majorization

Several inequalities for the trace can be strengthened in the form of submajorization. It turns out that in the case of trace inequalities discussed in Sections 3 and 4 the formulation by submajorization is very appropriate.

Let \(A\) and \(B\) be selfadjoint matrices with eigenvalues \(\kappa_1^A \geq \ldots \geq \kappa_n^A\) and \(\kappa_1^B \geq \ldots \geq \kappa_n^B\), then \(A\) is said to be submajorized by \(B\), in notation \(A \prec_w B\), if
\[
\sum_{i=1}^k \kappa_i^A \leq \sum_{i=1}^k \kappa_i^B
\]
for every \(1 \leq k \leq n\). If in addition \(\text{Tr}A = \text{Tr}B\) then \(A\) is said to be majorized by \(B\), in notation \(A \prec B\). Majorization and submajorization have an extensive literature; we only mention the main sources [2, 20]. In mathematical physics the same concept appears with different terminology and opposite notation [1]. (When \(D_1 \prec D_2\) holds for some density matrices then \(D_1\) is called more mixed than \(D_2\).)

It is well known that the relation \(A \prec_w B\) implies that \(\text{Tr} f(A) \leq \text{Tr} f(B)\) for every increasing convex function [20]. Therefore the following result is an extension of Theorem 8 as well as of the Golden–Thompson inequality [5, 11].
Theorem 16. For Hermitian matrices $A$ and $B$ the submajorization relation
\[ \log(e^{tA/2}e^{tB}e^{tA/2})^{1/t} \prec_w \log(e^{uA/2}e^{uB}e^{uA/2})^{1/u} \]
holds whenever $0 \leq t \leq u$.


Theorem 17. For Hermitian matrices $A$ and $B$ and for $0 < \alpha < 1$ the majorization relation
\[ \log(e^{\alpha A}e^{\alpha B})^{1/\alpha} \prec \log(e^{rA}e^{rB})^{1/r} \]
holds whenever $0 \leq r \leq p$.

Finally, here is the submajorization version of the Bernstein inequality [9].

Theorem 18. For an arbitrary $n \times n$ matrix $K$,
\[ e^{K}e^{K^*} \prec_w e^{K+K^*}. \]

References


