FUNCTIONAL ANALYSIS AND OPERATOR THEORY
BANACH CENTER PUBLICATIONS, VOLUME 30
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 1994

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 $\operatorname{Tr} A \le \operatorname{Tr} B$, where A and B are certain Hermitian complex matrices and 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 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) $|\tau(A)| \le cr(A)$ for every A, where c is a constant and r denotes the spectral radius.

(iii) $A^2 = 0$ implies $\tau(A) = 0$.

(iv) $|\tau(A^k)| \le \tau(A^*A)^{k/2}$ 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 $P \vee Q$ of the projections P and Q

1991 *Mathematics Subject Classification*: 15A42, 15A90.
The paper is in final form and no version of it will be published elsewhere.
Research partially supported by OTKA 1900.

may be defined as the orthogonal projection onto the linear span of the subspaces corresponding to P and Q. Similarly, the meet $P \wedge Q$ 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 *subadditive* if $f(P \vee Q) \le f(P) + f(Q)$ for every $P, Q \in \mathcal{P}$.

(v) Up to a constant factor, 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

$$|\mathrm{Tr}(ABAB)| \le \mathrm{Tr}(A^*ABB^*) \tag{1}$$

of (iv) will also be used below.

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 $\kappa_1 \ge \ldots \ge \kappa_n$ and $\lambda_1 \ge \ldots \ge \lambda_n$, then $A \le B$ implies $\kappa_i \le \lambda_i$ for every $1 \le i \le n$. Recall that if $\sum_i \lambda_i p_i$ is the spectral decomposition of A and the real function f is defined on the spectrum $\mathrm{Spec}(A)$ of A then $f(A)$ is defined by $f(A) = \sum_i f(\lambda_i) p_i$.

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

$$\mathrm{Tr}\, f(A) \le \mathrm{Tr}\, f(B)\,.$$

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

$$F(A) = \mathrm{Tr}\, f(A)$$

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

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

$$\mathrm{Tr}\, f(B) \ge \sum_i f(\mathrm{Tr}\, Bp_i)\,. \tag{2}$$

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

$$\begin{aligned}\mathrm{Tr}\, f(B) &= \sum_j f(s_j)\,\mathrm{Tr}\, q_j = \sum_i \sum_j f(s_j)\,\mathrm{Tr}\, q_j p_i \\ &\ge \sum_i f\Big(\sum_j s_j\,\mathrm{Tr}\, q_j p_i\Big) = \sum_i f(\mathrm{Tr}\, Bp_i)\,.\end{aligned}$$

To prove the proposition we write $\sum_i \mu_i p_i$ for the spectral decomposition of the convex combination $A = \lambda B_1 + (1-\lambda) B_2$. Applying (2) twice we infer that

$$\begin{aligned}
&\lambda \operatorname{Tr} f(B_1) + (1-\lambda)\operatorname{Tr} f(B_2) \\
&\qquad \ge \lambda \sum_i f(\operatorname{Tr} B_1 p_i) + (1-\lambda) \sum_i f(\operatorname{Tr} B_2 p_i) \\
&\qquad \ge \sum_i f(\lambda \operatorname{Tr} B_1 p_i + (1-\lambda)\operatorname{Tr} B_2 p_i) = \sum_i f(\operatorname{Tr} A\, p_i) = \operatorname{Tr} f(A)\,,
\end{aligned}$$

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* $f_k,\ g_k\colon [\alpha,\beta] \to \mathbb{R}$ *are such that for some* $c_k \in \mathbb{R}$,

$$\sum_k c_k f_k(x) g_k(y) \ge 0 \qquad \textit{for every } x, y \in [\alpha,\beta]\,,$$

then

$$\sum_k c_k \operatorname{Tr} f_k(A) g_k(B) \ge 0$$

whenever A, B *are selfadjoint matrices with* $\operatorname{Spec}(A)$, $\operatorname{Spec}(B) \subset [\alpha,\beta]$.

Proof. Let $A = \sum \lambda_i p_i$ and $B = \sum \mu_j q_j$ be the spectral decompositions. Then

$$\begin{aligned}
\sum_k c_k \operatorname{Tr} f_k(A) g_k(B) &= \sum_k \sum_{i,j} c_k \operatorname{Tr} p_i f_k(A) g_k(B) q_j \\
&= \sum_{i,j} \operatorname{Tr} p_i q_i \sum_k c_k f_k(\lambda_i) g_k(\mu_j) \ge 0
\end{aligned}$$

by the hypothesis.

In particular, if f is convex then

$$f(x) - f(y) - (x-y) f'(y) \ge 0 \tag{3}$$

and

$$\operatorname{Tr} f(A) \ge \operatorname{Tr} f(B) + \operatorname{Tr}(A-B) f'(B)\,. \tag{4}$$

For the choice $f(t) = -\eta(t) = t \log t$ we obtain

$$S(A,B) \equiv \operatorname{Tr} A(\log A - \log B) \ge \operatorname{Tr}(A-B) \tag{5}$$

for B strictly positive and A nonnegative. The left-hand side is called *relative entropy*. If A and B are density matrices, i.e. $\operatorname{Tr} A = \operatorname{Tr} B = 1$, then $S(A,B) \ge 0$. This is a classical application of the Klein inequality (cf. [27]). The stronger estimate

$$-\eta(x) + \eta(y) + (x-y)\eta'(y) \ge \frac{1}{2}(x-y)^2$$

allows another use of the Klein inequality. Namely,

$$S(A,B) \ge \frac{1}{2} \operatorname{Tr}(A-B)^2 + \operatorname{Tr}(A) - \operatorname{Tr}(B)\,, \tag{6}$$

which was obtained in [31].

From the inequality $1 + \log x \le x$ $(x > 0)$ one obtains

$$t^{-1}(a - a^{1-t}b^t) \le a(\log a - \log b) \le t^{-1}(a^{1+t}b^{-t} - a)$$

for $a, b, t > 0$. If T and S are nonnegative invertible matrices then Proposition 3 gives

$$\begin{aligned} t^{-1}\operatorname{Tr}(S - S^{1-t}T^t) &\le \operatorname{Tr} S(\log S - \log T) \\ &\le t^{-1}\operatorname{Tr}(S^{1+t}T^{-t} - S)\,, \end{aligned} \tag{7}$$

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

$$\operatorname{Tr} e^A e^B \ge \operatorname{Tr} e^{A+B}\,. \tag{8}$$

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 ([1]).

We also need the inequality

$$|\operatorname{Tr} X^{2k}| \le \operatorname{Tr}(XX^*)^k\,, \tag{9}$$

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

$$\operatorname{Tr} VHVH \le \operatorname{Tr} V^2H^2\,, \tag{10}$$

which is a particular case of the inequality

$$\operatorname{Tr} YHY^*g(H) \le \operatorname{Tr} YY^*Hg(H)\,, \tag{11}$$

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 $A, B \in M_n$,*

$$\operatorname{Tr} e^{(A+A^*)/2}e^{(B+B^*)/2} \ge |\operatorname{Tr} e^{A+B}|\,.$$

([1]) *Editorial note*: See also the application on p. 370 in this volume.

P r o o f. Substituting $X = AB$ into (9) we have

$$\mathrm{Tr}((ABB^*A^*)^s) \ge |\mathrm{Tr}((AB)^{2s})|\,,$$

where the left-hand side is nothing else but $\mathrm{Tr}(BB^*A^*A)^s$. Setting $s = 2^{k-1}$ with a positive integer k and using (9) gives

$$\begin{aligned}|\mathrm{Tr}(AB)^{2^k}| &\le \mathrm{Tr}(BB^*A^*A)^{2^{k-1}} \le |\mathrm{Tr}((BB^*A^*A)^2)^{2^{k-2}}|\\ &\le \mathrm{Tr}((BB^*A^*A)(BB^*A^*A)^*)^{2^{k-2}}\\ &= \mathrm{Tr}((A^*A)^2(BB^*)^2)^{2^{k-2}}\,.\end{aligned}$$

By a repeated application of this argument we easily infer that

$$\mathrm{Tr}((A^*A)^{2^{k-1}}(BB^*)^{2^{k-1}}) \ge |\mathrm{Tr}((AB)^{2^k})|\,.$$

Now replace A by $\exp(2^{-k}A)$ and B by $\exp(2^{-k}B)$:

$$\mathrm{Tr}((e^{2^{-k}A^*}e^{2^{-k}A})^{2^k/2}(e^{2^{-k}B}e^{2^{-k}B^*})^{2^k/2}) \ge |\mathrm{Tr}((e^{2^{-k}A}e^{2^{-k}B})^{2^k})|\,.$$

The obvious continuity of 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*

$$|\mathrm{Tr}\, e^{A+\mathrm{i}B}| \le \mathrm{Tr}\, e^A\,.$$

The relative entropy of nonnegative matrices defined by (5) is related to the functional $B \mapsto \log \mathrm{Tr}\, e^{A+B}$ by the Legendre transform. Namely, $B \mapsto \log \mathrm{Tr}\, e^{A+B}$ 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*

$$\log \mathrm{Tr}\, e^{A+\log Y} = \max\{\mathrm{Tr}\, XA - S(X,Y) : X \textit{ is positive},\ \mathrm{Tr}\, X = 1\}\,. \tag{12}$$

On the other hand, if X is positive with $\mathrm{Tr}\, X = 1$ *and B is Hermitian, then*

$$S(X, e^B) = \max\{\mathrm{Tr}\, XA - \log \mathrm{Tr}\, e^{A+B} : A \textit{ is Hermitian}\}\,. \tag{13}$$

P r o o f. Define

$$F(X) = \mathrm{Tr}\, XA - S(X,Y)$$

for nonnegative X with $\mathrm{Tr}\, X = 1$. When $P_1, \ldots, P_n$ are projections of rank one with $\sum_{i=1}^n P_i = 1$, we write

$$F\Big(\sum_{i=1}^n \lambda_i P_i\Big) = \sum_{i=1}^n (\lambda_i \,\mathrm{Tr}\, P_i A + \lambda_i \,\mathrm{Tr}\, P_i \log Y - \lambda_i \log \lambda_i)\,,$$

where $\lambda_i \geq 0$, $\sum_{i=1}^n \lambda_i = 1$. Since

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

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

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

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

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

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

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

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

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

$$\frac{d}{dt} \operatorname{Tr} e^{\log X + tS}\Big|_{t=0} = \operatorname{Tr} XS\,.$$

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

Let us make the following definition:

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

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

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

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 $X = e^{A+B} / \operatorname{Tr} e^{A+B}$ for Hermitian A and B we have

$$\begin{aligned}\log \operatorname{Tr} e^A e^B &\geq \operatorname{Tr} XA - S_{\mathrm{co}}(X, e^B)\\ &\geq \operatorname{Tr} XA - S(X, e^B) = \log \operatorname{Tr} e^{A+B}\end{aligned}$$

by (15), which further shows that $\operatorname{Tr} e^{A+B} = \operatorname{Tr} e^A e^B$ holds if and only if $AB = BA$. This derivation of the Golden–Thompson inequality as well as characterization 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

$$\mathrm{Tr}(X^{1/2}YX^{1/2})^q \le \mathrm{Tr}\, X^{q/2}Y^qX^{q/2}$$

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

$$\mathrm{Tr}(X^{1/2}YX^{1/2})^{rp} \le \mathrm{Tr}(X^{r/2}Y^rX^{r/2})^p \tag{16}$$

for every $r \ge 1$ and $p > 0$. This implies that the function

$$p \mapsto \mathrm{Tr}(e^{pB/2}e^{pA}e^{pB/2})^{1/p} \tag{17}$$

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

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

$$\mathrm{Tr}\, e^{A+B} \le \mathrm{Tr}(e^{pB/2}e^{pA}e^{pB/2})^{1/p}\,. \tag{18}$$

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

$$\mathrm{Tr}\, e^{A+B+C} \le \mathrm{Tr}\, e^Ae^Be^C$$

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:

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

For $A \in M_n^{\mathrm{sa}}$ let $\mathcal{T}_{\exp A} : M_n \to M_n$ be defined by

$$\mathcal{T}_{\exp A}(K) = \int_0^\infty (t + \exp A)^{-1}K(t + \exp A)^{-1}\, dt\,. \tag{19}$$

Since

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

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

$$(\mathcal{T}_{\exp A}(K))_{ij} = K_{ij}/\mathrm{Lm}(e^{a_i}, e^{a_j})\,, \tag{20}$$

where $\mathrm{Lm}(x,y)$ stands for the so-called *logarithmic mean* defined by

$$\mathrm{Lm}(x,y)=\begin{cases}(x-y)/(\log x-\log y) & \text{if } x\neq y\,,\\ x & \text{if } x=y\,.\end{cases}$$

Note that if $K=K^*$ and $AK=KA$, then $\mathcal{T}_{\exp A}(K)=\exp(-A)K$.

THEOREM 9. *Let A, B and C be Hermitian matrices. Then*

$$\operatorname{Tr} e^{A+B+C}\le \operatorname{Tr}\mathcal{T}_{\exp(-A)}(e^B)e^C\,. \tag{21}$$

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*

$$\operatorname{Tr} e^{K+K^*}\ge \operatorname{Tr} e^K e^{K^*} \tag{22}$$

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 $\operatorname{Tr} e^{A+B}$ and $\operatorname{Tr} e^A e^B$ in the case $e^{A+B}\neq e^A e^B$. The logarithmic analogue would be a relation between $\operatorname{Tr}\log XY$ and $\operatorname{Tr}(\log X+\log Y)$ for positive matrices X and Y. This relation is well known, $\operatorname{Tr}\log XY=\operatorname{Tr}(\log X+\log Y)$, 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

$$\operatorname{Tr} X\log X^{1/2}YX^{1/2}=\operatorname{Tr} X\log XY \tag{23}$$

because $\operatorname{Tr} X(X^{1/2}YX^{1/2})^n=\operatorname{Tr} X(XY)^n$ for $n\ge 1$.

PROPOSITION 11. *Let X and Y be positive matrices. Then*

$$\operatorname{Tr} X\log Y^{1/2}XY^{1/2}\le \operatorname{Tr} X(\log X+\log Y)\le \operatorname{Tr} X\log XY\,.$$

Proof. The first inequality is a consequence of (13) in the case $\operatorname{Tr} X=1$, which it is sufficient to consider. Let B be Hermitian and $A=\log e^{-B/2}Xe^{-B/2}$. Then by (13) we have

$$\begin{aligned}\operatorname{Tr} X(\log X-B)&\ge \operatorname{Tr} XA-\log\operatorname{Tr} e^{A+B}\\ &\ge \operatorname{Tr} XA-\log\operatorname{Tr}(e^{B/2}e^Ae^{B/2})\\ &=\operatorname{Tr} X\log e^{-B/2}Xe^{-B/2}-\log\operatorname{Tr} X\\ &=\operatorname{Tr} X\log e^{-B/2}Xe^{-B/2}\,.\end{aligned}$$

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

$$S_{\mathrm{BS}}(X,Y) = \operatorname{Tr} X \log X^{1/2} Y^{-1} X^{1/2}$$

we see that the second inequality is the same as

$$S(X,Y) \le S_{\mathrm{BS}}(X,Y) \tag{24}$$

for positive matrices X and Y with $\operatorname{Tr} X = \operatorname{Tr} Y = 1$. (The quantity S_{BS} is related to the works [6, 12].) The proof of (24) was given in [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 $\mathcal{A}_m$ be the m-fold tensor product $\bigotimes_1^m M_n$ 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 $Y_m = \bigotimes_1^m Y$ and write E_{Y_m} for the conditional expectation from $\mathcal{A}_m$ onto $\{Y_m\}'$ with respect to the trace. (When $Z = \sum_i \lambda_i P_i$ is the spectral decomposition of a selfadjoint Z, then $E_Z(A) = \sum_i P_i A P_i$.)

Lemma 12. *For every positive Z in $\mathcal{A}_m$ with* $\operatorname{Tr} Z = 1$,

$$S(Z, E_{Y_m}(Z)) \le n \log(m+1)\,.$$

Having the lemma at our disposal we obtain (24) from the chain

$$\begin{aligned} mS_{\mathrm{BS}}(X,Y) &= S_{\mathrm{BS}}(X_m, Y_m) \ge S_{\mathrm{BS}}(E_{Y_m}(X_m), Y_m) \\ &= S(E_{Y_m}(X_m), Y_m) = S(X_m, Y_m) - S(X_m, E_{Y_m}(X_m)) \\ &\ge S(X_m, Y_m) - n\log(m+1) = mS(X,Y) - n\log(m+1) \end{aligned}$$

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 [14].

For $0 \le \alpha \le 1$ the *α-power mean* of positive matrices X and Y is defined by

$$X \#_\alpha Y = X^{1/2}(X^{-1/2} Y X^{-1/2})^\alpha X^{1/2}\,.$$

This is the operator mean corresponding to an operator monotone function x^α, $x \ge 0$. In particular, $X \#_{1/2} Y = X \# Y$ is the geometric mean of X and Y which was introduced in [26].

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

(i) *If A and B are Hermitian, then*

$$\operatorname{Tr}\,(e^{pA} \#_\alpha e^{pB})^{1/p} \le \operatorname{Tr} e^{(1-\alpha)A + \alpha B} \tag{25}$$

for $0 \le \alpha \le 1$.

(ii) *If X and Y are positive, then*

$$\operatorname{Tr} X(\log X + \log Y) \le \frac{1}{p} \operatorname{Tr} X \log X^{p/2} Y^p X^{p/2}\,. \tag{26}$$

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:

$$\frac{d}{d\alpha}\operatorname{Tr}(e^{pA}\#_\alpha e^{pB})^{1/p}\bigg|_{\alpha=0}=\frac{1}{p}\operatorname{Tr}e^{A}\log e^{-pA/2}e^{pB}e^{-pA/2}\,.$$

THEOREM 13. *Let A and B be Hermitian and* $0\le\alpha\le 1$. *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* $0\le\alpha\le 1$, *then*

$$\lim_{s\to\infty}(e^{A/s}\#_\alpha e^{B/s})^{s}=e^{(1-\alpha)A+\alpha B}\,.$$

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 $\kappa_1^A\ge\ldots\ge\kappa_n^A$ and $\kappa_1^B\ge\ldots\ge\kappa_n^B$; then A is said to be *submajorized* by B, in notation $A\prec_{\mathrm{w}}B$, if

$$\sum_{i=1}^{k}\kappa_i^A\le\sum_{i=1}^{k}\kappa_i^B$$

for every $1\le k\le n$. If in addition $\operatorname{Tr}A=\operatorname{Tr}B$ 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 $A\prec_{\mathrm{w}}B$ implies that $\operatorname{Tr}f(A)\le\operatorname{Tr}f(B)$ 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*

$$\log(e^{tA/2}e^{tB}e^{tA/2})^{1/t} \prec_{\mathrm{w}} \log(e^{uA/2}e^{uB}e^{uA/2})^{1/u}$$

holds whenever $0 \le t \le u$.

Ando [3] deduced the next theorem from a norm inequality using the method of anti-symmetric tensor product.

THEOREM 17. *For Hermitian matrices A and B and for $0 < \alpha < 1$ the majorization relation*

$$\log(e^{pA}\#_\alpha e^{pB})^{1/p} \prec \log(e^{rA}\#_\alpha e^{rB})^{1/r}$$

holds whenever $0 \le r \le p$.

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

THEOREM 18. *For an arbitrary $n \times n$ matrix K,*

$$e^K e^{K^*} \prec_{\mathrm{w}} e^{K+K^*}.$$

References

[1] P. M. Alberti and A. Uhlmann, *Stochasticity and Partial Order. Doubly Stochastic Maps and Unitary Mixing*, Deutscher Verlag Wiss., Berlin, 1981.

[2] T. Ando, *Majorization, doubly stochastic matrices and comparison of eigenvalues*, Linear Algebra Appl. 118 (1989), 163–248.

[3] —, private communication, 1992.

[4] H. Araki, *Golden–Thompson and Peierls–Bogoliubov inequalities for a general von Neumann algebra*, Comm. Math. Phys. 34 (1973), 167–178.

[5] —, *On an inequality of Lieb and Thirring*, Lett. Math. Phys. 19 (1990), 167–170.

[6] V. P. Belavkin and P. Staszewski, *C^*-algebraic generalization of relative entropy and entropy*, Ann. Inst. Henri Poincaré Sect. A 37 (1982), 51–58.

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

[8] M. Breitenbecker and H. R. Grümm, *Note on trace inequalities*, Comm. Math. Phys. 26 (1972), 276–279.

[9] J. E. Cohen, *Inequalities for matrix exponentials*, Linear Algebra Appl. 111 (1988), 25–28.

[10] J. E. Cohen, S. Friedland, T. Kato and F. P. Kelly, *Eigenvalue inequalities for products of matrix exponentials*, ibid. 45 (1982), 55–95.

[11] S. Friedland and W. So, *On the product of matrix exponentials*, ibid. 196 (1994), 193–205.

[12] J. I. Fujii and E. Kamei, *Relative operator entropy in noncommutative information theory*, Math. Japon. 34 (1989), 341–348.

[13] S. Golden, *Lower bounds for the Helmholtz function*, Phys. Rev. 137 (1965), B1127–B1128.

[14] F. Hiai, *Some remarks on the trace operator logarithm and relative entropy*, in: Quantum Probability and Related Topics VIII, World Sci., Singapore, to appear.

[15] F. Hiai and D. Petz, *The proper formula for relative entropy and its asymptotics in quantum probability*, Comm. Math. Phys. 143 (1991), 99–114.

[16] —, —, *The Golden–Thompson trace inequality is complemented*, Linear Algebra Appl. 181 (1993), 153–185.

[17] S. Klimek and A. Lesniewski, *A Golden–Thompson inequality in supersymmetric quantum mechanics*, Lett. Math. Phys. 21 (1991), 237–244.
[18] H. Kosaki, *An inequality of Araki–Lieb–Thirring* (*von Neumann algebra case*), Proc. Amer. Math. Soc. 114 (1992), 477–481.
[19] E. H. Lieb, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, Adv. in Math. 11 (1973), 267–288.
[20] A. W. Marshall and I. Olkin, *Inequalities*: *Theory of Majorization and Its Applications*, Academic Press, New York, 1979.
[21] —, —, *Inequalities for the trace function*, Aequationes Math. 29 (1985), 36–39.
[22] D. Petz, *Properties of quantum entropy*, in: Quantum Probability and Applications II, L. Accardi and W. von Waldenfels (eds.), Lecture Notes in Math. 1136, Springer, 1985, 428–441.
[23] —, *Sufficient subalgebras and the relative entropy of states on a von Neumann algebra*, Comm. Math. Phys. 105 (1986), 123–131.
[24] —, *A variational expression for the relative entropy*, ibid. 114 (1988), 345–348.
[25] D. Petz and J. Zemánek, *Characterizations of the trace*, Linear Algebra Appl. 111 (1988), 43–52.
[26] W. Pusz and S. L. Woronowicz, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phys. 8 (1975), 159–170.
[27] D. Ruelle, *Statistical Mechanics. Rigorous Results*, Benjamin, New York, 1969.
[28] M. B. Ruskai, *Inequalities for traces on von Neumann algebras*, Comm. Math. Phys. 26 (1972), 280–289.
[29] M. B. Ruskai and F. K. Stillinger, *Convexity inequalities for estimating free energy and relative entropy*, J. Phys. A 23 (1990), 2421–2437.
[30] W. So, *Equality cases in matrix exponential inequalities*, SIAM J. Matrix Anal. Appl. 13 (1992), 1154–1158.
[31] R. F. Streater, *Convergence of the quantum Boltzmann map*, Comm. Math. Phys. 98 (1985), 177–185.
[32] C. J. Thompson, *Inequality with applications in statistical mechanics*, J. Math. Phys. 6 (1965), 1812–1813.
[33] R. C. Thompson, *High, low and qualitative roads in linear algebra*, Linear Algebra Appl. 162–164 (1992), 23–64.