An Araki-Lieb-Thirring inequality for geometrically concave and geometrically convex functions

Koenraad M.R. Audenaert

Department of Mathematics, Royal Holloway, University of London, Egham TW20 0EX, United Kingdom

Abstract

For positive definite matrices $A$ and $B$, the Araki-Lieb-Thirring inequality amounts to an eigenvalue log-submajorisation relation for fractional powers

$$\lambda(A^t B^t) \prec_{w(\log)} \lambda^t(AB), \quad 0 < t \leq 1,$$

while for $t \geq 1$, the reversed inequality holds. In this paper I generalise this inequality, replacing the fractional powers $x^t$ by a larger class of functions. Namely, a continuous, non-negative, geometrically concave function $f$ with domain $\text{dom}(f) = [0, x_0)$ for some positive $x_0$ (possibly infinity) satisfies

$$\lambda(f(A)f(B)) \prec_{w(\log)} f^2(\lambda^{1/2}(AB)),$$

for all positive semidefinite $A$ and $B$ with spectrum in $\text{dom}(f)$, if and only if $0 \leq xf'(x) \leq f(x)$ for all $x \in \text{dom}(f)$. The reversed inequality holds for continuous, non-negative, geometrically convex functions if and only if they satisfy $xf'(x) \geq f(x)$ for all $x \in \text{dom}(f)$. As an application I derive a complementary inequality to the Golden-Thompson inequality.

Key words: Log-majorisation, positive semidefinite matrix, matrix inequality

1991 MSC: 15A60

1 Introduction

The Araki-Lieb-Thirring (ALT) inequality \cite{217} states that for $0 < t \leq 1$ and positive definite matrices $A$ and $B$, the eigenvalues of $A^t B^t$ are log-submajorised by the eigenvalues of $(AB)^t$. For $n \times n$ matrices $X$ and $Y$ with

Email address: koenraad.audenaert@rhul.ac.uk (Koenraad M.R. Audenaert).
positive spectrum, the log-submajorisation relation $\lambda(X) \prec_w \lambda(Y)$ means that for all $k = 1, \ldots, n$, the following holds:

$$\prod_{j=1}^{k} \lambda_j(X) \leq \prod_{j=1}^{k} \lambda_j(Y).$$

This is equivalent to weak majorisation of the logarithms of the spectra

$$\log \lambda(X) \prec_w \log \lambda(Y).$$

Here and elsewhere, I adhere to the convention to sort eigenvalues in non-increasing order; that is, $\lambda_1(X) \geq \lambda_2(X) \geq \ldots \geq \lambda_n(X)$.

With this notation, the ALT inequality can be written as

$$\lambda(A^t B^t) \prec_w \lambda^t(AB), \quad 0 < t \leq 1. \tag{1}$$

For positive scalars $a$ and $b$, (1) reduces to the equality $a^t b^t = (ab)^t$.

One can ask whether similar inequalities hold for other functions than the fractional powers $x^t$. One possibility is to consider functions that satisfy

$$\lambda(f(A)f(B)) \prec_w \lambda^t(f(AB)).$$

As the scalar case reduces to $f(a)f(b) \leq f(ab)$ these functions must be super-multiplicative. Another possibility, and the one pursued here, is to consider functions satisfying

$$\lambda(f(A)f(B)) \prec \lambda^{2t} \left( \sqrt[2t]{\lambda(AB)} \right). \tag{2}$$

Here, the scalar case reduces to $f(a)f(b) \leq f^2(\sqrt{ab})$, for all $a, b > 0$. Functions satisfying this requirement are called geometrically concave (see Definition 1 below). In this paper I completely characterise the class of geometrically concave functions that satisfy (2) for all positive definite matrices $A$ and $B$.

Likewise, as inequality (1) holds in the reversed sense for $t \geq 1$, one may ask for which functions $f$ the reversed inequality holds for all positive definite matrices $A$ and $B$:

$$f^2 \left( \sqrt[2t]{\lambda(AB)} \right) \prec_w \lambda(f(A)f(B)). \tag{3}$$

Here the scalar case restricts the class of functions to those satisfying the relation $f^2(\sqrt{ab}) \leq f(a)f(b)$. Such functions are called geometrically convex.
I also completely characterise the class of geometrically convex functions that satisfy (3) for all positive definite matrices $A$ and $B$.

The concepts of geometric concavity and geometric convexity were first studied by Montel [8] and have recently received attention from the matrix community [3, 5].

**Definition 1** Let $I$ be the interval $I = [0, x_0)$, with $x_0 > 0$ (possibly infinite). A function $f : I \rightarrow [0, \infty)$ is geometrically concave if for all $x, y \in I$, $\sqrt{f(x)f(y)} \leq f(\sqrt{xy})$. It is geometrically convex if for all $x, y \in I$, $\sqrt{f(x)f(y)} \geq f(\sqrt{xy})$.

Equivalently, a function $f(x)$ is geometrically concave (convex) if and only if the associated function $F(y) := \log(f(e^y))$ is concave (convex).

The main results of this paper are summarised in the next section, the proofs of the main theorems (Theorems 1 and 2) are given in Section 3, and the paper concludes with a brief application in Section 4.

## 2 Main Results

To state the main results of this paper most succinctly, let us define two classes of functions.

**Definition 2** A continuous non-negative function $f$ with domain an interval $I = [0, x_0)$ is in class $A$ if and only if it is geometrically concave and its derivative $f'$ satisfies $0 \leq xf'(x) \leq f(x)$ for all $x \in I$ where the derivative exists.

**Definition 3** A continuous non-negative function $f$ with domain an interval $I = [0, x_0)$ is in class $B$ if and only if it is geometrically convex and its derivative $f'$ satisfies $xf'(x) \geq f(x)$ for all $x \in I$ where the derivative exists.

In terms of the associated function $F(y) = \log(f(e^y))$, $f \in A$ if and only if $F(y)$ is concave and $0 \leq F'(y) \leq 1$ for all $y$ where $F$ is differentiable, and $f \in B$ if and only if $F(y)$ is convex and $1 \leq F'(y)$ for all $y$ where $F$ is differentiable.

There is a simple one-to-one relationship between these two classes; essentially $f$ is in class $A$ if and only if its inverse function $f^{-1}$ is in class $B$. However, some care must be taken as $A$ contains the constant functions and also those functions that are constant on some interval.
Proposition 1 A function $f$ that is non-constant on the interval $[0, x_1] \subseteq [0, x_0)$ is in class $A$ if and only if the inverse of the restriction of $f$ to $[0, x_1]$ is in class $B$.

Proof. It is clear that a concave monotonous function $f$ is always invertible over the entire interval where it is not constant. We will henceforth identify the inverse of $f$ with the inverse of the restriction of $f$ on that interval.

If $F(y)$ is the associated function of $f(x)$ then the associated function of $f^{-1}$ is the inverse function of $F$, $F^{-1}$. Now $f$ is in class $A$ if and only if $F$ is concave, monotonous and $F' \leq 1$. This implies that the inverse function $G = F^{-1}$ is convex and satisfies $G' \geq 1$, which in turn implies that $G$ is the associated function of a function $g$ in class $B$. This shows that $f \in A$ implies $f^{-1} \in B$.

A similar argument reveals that the converse statement holds as well. $\square$

The main result of this paper is the following theorem:

Theorem 1 Let $f$ be a continuous non-negative function with domain an interval $I = [0, x_0), x_0 > 0$ (possibly infinite), then

$$\lambda(f(A)f(B)) \prec_w \log f^2 \left( \sqrt{\lambda(AB)} \right)$$

holds for all positive definite matrices $A$ and $B$ with spectrum in $I$ if and only if $f$ is in class $A$.

That the right-hand side of (4) is well-defined follows from the following lemma:

Lemma 1 If $A$ and $B$ are positive semidefinite matrices with eigenvalues in the interval $I = [a, b], 0 \leq a < b$, the positive square roots of the eigenvalues of $AB$ are in $I$ as well.

Proof. We have $a \leq A, B \leq b$, which implies

$$a^2 \leq aA \leq A^{1/2}BA^{1/2} \leq bA \leq b^2.$$ 

Hence, $a^2 \leq \lambda_i(A^{1/2}BA^{1/2}) \leq b^2$, so that $a \leq \lambda_i^{1/2}(AB) \leq b$. $\square$

A simple consequence of Theorem 1 is that the reversed inequality holds if and only if $f$ is in class $B$.

Theorem 2 Let $g$ be a continuous non-negative function with domain an interval $I = [0, x_0), x_0 > 0$ (possibly infinite), then

$$g^2 \left( \sqrt{\lambda(XY)} \right) \prec_w \log \lambda(g(X)g(Y)).$$

(5)
holds for all positive definite matrices $X$ and $Y$ with spectrum in $I$ if and only if $g$ is in class $\mathcal{B}$.

3 Proofs

We now turn to the proofs of Theorems 1 and 2.

3.1 Proof of necessity

To show necessity of the conditions $f \in \mathcal{A}$ ($f \in \mathcal{B}$) I consider two special $2 \times 2$ matrices with eigenvalues $a$ and $b$, $0 \leq b < a$, such that $a \in \text{dom}(f)$ and $f$ is differentiable in $a$:

$$ A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad B = U \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} U^*, \quad \text{with } U = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. $$

We will consider values of $\theta$ close to 0. The largest eigenvalue of $AB$ can be calculated in a straight-forward fashion. The quantity $f^2(\sqrt{\lambda_1(AB)})$ can then be expanded in a power series of the variable $\theta$. To second order this yields

$$ f^2(\sqrt{\lambda_1(AB)}) = f(a)^2 - \frac{a(a - b)f(a)f'(a)}{a + b} \theta^2 + O(\theta^4). $$

In a similar way we also get

$$ \lambda_1(f(A)f(B)) = f(a)^2 - \frac{f(a)^2(f(a) - f(b))}{f(a) + f(b)} \theta^2 + O(\theta^4). $$

Hence, to satisfy inequality (1), the following must be satisfied for all $0 \leq b < a \in \text{dom}(f)$:

$$ \frac{a(a - b)f(a)f'(a)}{a + b} \leq \frac{f(a)^2(f(a) - f(b))}{f(a) + f(b)}. $$

In particular, take $b = 0$. As $f$ has to be geometrically concave, $f(0) = 0$. The condition then becomes

$$ af(a)f'(a) \leq f(a)^2, \forall a \in \text{dom}(f), $$

which reduces to the defining condition for $f \in \mathcal{A}$.

Necessity of the condition $xf'(x) \geq f(x)$ in Corollary 2 also follows immediately from this special pair of matrices.
Note that in the preceding proof we see why the domain of $f$ should include the point $x = 0$.

### 3.2 Proof of sufficiency for Theorem

Now I turn to proving sufficiency of the condition $f \in \mathcal{A}$. The main step consists in showing that the set of functions $f$ for which the inequality \( (4) \) holds is ‘geometrically convex’; that is, the set of associated functions $F$ for these $f$ is convex. To show this a number of preliminary propositions are needed.

**Lemma 2** Let $R_1$, $R_2$, $S_1$ and $S_2$ be positive semidefinite matrices such that $R_1$ commutes with $R_2$ and $S_1$ with $S_2$. Let $R = R_1^{1/2} R_2^{1/2}$ and $S = S_1^{1/2} S_2^{1/2}$. Then

$$\lambda_1((R^{1/2} SR^{1/2})^2) \leq \lambda_1(R_1^{1/2} S_1 R_1^{1/2} R_2^{1/2} S_2 R_2^{1/2}). \tag{6}$$

**Proof.** We will prove this by showing that the equality

$$\lambda_1(R_1^{1/2} S_1 R_1^{1/2} R_2^{1/2} S_2 R_2^{1/2}) = a$$

implies the inequality $\lambda_1((R^{1/2} SR^{1/2})^2) \leq a$.

W.l.o.g. we can assume that the matrices $R_1$ and $S_1$ are invertible; then the equality indeed leads to the following sequence of implications:

\[
\begin{align*}
&\lambda_1(R_1^{1/2} S_1 R_1^{1/2} R_2^{1/2} S_2 R_2^{1/2}) = a \\
\implies & (R_1^{1/2} S_1 R_1^{1/2}) R_2^{1/2} S_2 R_2^{1/2} (R_1^{1/2} S_1 R_1^{1/2}) \leq a \\
\implies & R_2^{1/2} S_2 R_2^{1/2} \leq R_1^{-1/2} S_1^{-1} R_1^{-1/2} \\
\implies & S_1^{1/2} R_1^{1/2} R_2^{1/2} S_2 R_2^{1/2} R_1^{1/2} S_1^{1/2} = S_1^{1/2} R S_2 R S_1^{1/2} \leq a \\
\implies & \sigma_1(S_2^{1/2} R S_1^{1/2}) \leq \sqrt{a} \\
\implies & |\lambda_1(S_2^{1/2} R S_1^{1/2})| \leq \sqrt{a}.
\end{align*}
\]

The last implication is the simplest case of Weyl’s majorant theorem.

Now note that $S_2^{1/2} R S_1^{1/2}$ and $S_1^{1/2} R S_1^{1/2} = S_1^{1/4} S_2^{1/4} R S_1^{1/4} S_2^{1/4}$ have the same non-zero eigenvalues. Hence, $\lambda_1(S_1^{1/2} R S_1^{1/2}) = \lambda_1(R^{1/2} S R^{1/2}) \leq \sqrt{a}$, and the inequality $\lambda_1((R^{1/2} SR^{1/2})^2) \leq a$ follows. \( \square \)
Proposition 2 Under the conditions of Lemma 2

\[
\prod_{i=1}^{k} \lambda_i(R_i^{1/2} R_i^{1/2} S_i^{1/2} S_i^{1/2}) \leq \prod_{i=1}^{k} \lambda_i^{1/2}(R_i S_i) \lambda_i^{1/2}(S_i R_i). \tag{7}
\]

Proof. Since each side of (6) is the largest eigenvalue of a product of powers of matrices, we can use the well-known Weyl trick of replacing every matrix by its antisymmetric tensor power to boost the inequality to the log-submajorisation relation

\[
\prod_{i=1}^{k} \lambda_i((R_i^{1/2} S_i^{1/2})^2) \leq \prod_{i=1}^{k} \lambda_i(R_i^{1/2} R_i^{1/2} S_i^{1/2} S_i^{1/2}).
\]

Combining this with Lidskii’s inequality \(\prod_{i=1}^{k} \lambda_i(A B) \leq \prod_{i=1}^{k} \lambda_i(A) \lambda_i(B)\) (4, Corollary III.4.6), valid for positive definite \(A\) and \(B\), and then taking square roots yields inequality (7). \(\square\)

Inequality (7) can be interpreted as midpoint geometric convexity of the function

\[
p \mapsto f_k(p) = \prod_{i=1}^{k} \lambda_i(R_i^{p} R_i^{1-p} S_i^{p} S_i^{1-p});
\]

that is, \(f_k(1/2) \leq \sqrt{f_k(1)f_k(0)}\). We now use a standard argument (see e.g. the proof of Lemma IX.6.2 in [4]) to show that this actually implies geometric convexity in full generality, i.e. \(f_k(p) \leq f_k(1)^p f_k(0)^{1-p}\) for all \(p \in [0, 1]\).

Proposition 3 Under the conditions of Lemma 2 and for all \(p \in [0, 1]\),

\[
\prod_{i=1}^{k} \lambda_i(R_i^{p} R_i^{1-p} S_i^{p} S_i^{1-p}) \leq \prod_{i=1}^{k} \lambda_i^{p}(R_i S_i) \lambda_i^{1-p}(S_i R_i). \tag{8}
\]

Proof. By Proposition 2 the inequality holds for \(p = 1/2\). It trivially holds for \(p = 0\) and \(p = 1\).

Let \(s, t \in [0, 1]\) be given. Applying Proposition 2 with the matrices \(R_1, S_1, R_2\) and \(S_2\) replaced by \(R_i^{s} R_i^{1-s}, S_i^{s} S_i^{1-s}\) and \(S_i^{s} S_i^{1-s}\) respectively, yields the inequality

\[
\prod_{i=1}^{k} \lambda_i(R_i^{(s+t)/2} R_i^{1-(s+t)/2} S_i^{(s+t)/2} S_i^{1-(s+t)/2})
\leq \prod_{i=1}^{k} \lambda_i^{1/2}(R_i^{s} R_i^{1-s} S_i^{s} S_i^{1-s}) \lambda_i^{1/2}(R_i^{s} R_i^{1-s} S_i^{s} S_i^{1-s}).
\]

Now assume that the inequality (8) holds for the values \(p = s\) and \(p = t\). Thus
\[
\prod_{i=1}^{k} \lambda_i^{1/2}(R_1^i R_2^{1-t} S_1^i S_2^{1-t}) \lambda_i^{1/2}(R_1^s R_2^{1-s} S_1^s S_2^{1-s})
\leq \prod_{i=1}^{k} \lambda_i^{(1-t)/2}(R_1 S_1) \lambda_i^{(1-s)/2}(S_2 R_2)
\leq \prod_{i=1}^{k} \lambda_i^{(s+t)/2}(R_1 S_1) \lambda_i^{1-(s+t)/2}(S_2 R_2).
\]

In other words, the assumption that (8) holds for the values \(p = s\) and \(p = t\) implies that it also holds for their midpoint \(p = (s + t)/2\).

Using induction this shows that (8) holds for all dyadic rational values of \(p\) (i.e. rationals of the form \(k/2^n\), with \(k\) and \(n\) integers such that \(k \leq 2^n\)).

Invoking continuity and the fact that the dyadic rationals are dense in \([0, 1]\), this finally implies that (8) holds for all real values of \(p\) in \([0, 1]\).

\[\blacksquare\]

We are now ready to prove our first intermediate result: convexity of the set of associated functions \(F\) for which the inequality (4) holds.

**Proposition 4** Let \(f_1(x)\) and \(f_2(x)\) be two continuous, non-negative functions with domain an interval \(I\) of the non-negative reals, and for which (4) holds for all positive semidefinite \(A\) and \(B\) with spectrum in \(I\). Let \(p \in [0, 1]\) and let \(f(x) = f_1^p(x) f_2^{1-p}(x)\). Then (4) holds for \(f\) too.

**Proof.** Let us fix the matrices \(A\) and \(B\) and let \(R_i = f_i(A)\) and \(S_i = f_i(B)\), \(i = 1, 2\). These matrices \(R_i\) and \(S_i\) clearly satisfy the conditions of Proposition 3 (positivity and commutativity). Hence

\[
\prod_{i=1}^{k} \lambda_i(f(A) f(B)) = \prod_{i=1}^{k} \lambda_i(f_1^p(A) f_2^{1-p}(A) f_1^p(B) f_2^{1-p}(B))
\leq \prod_{i=1}^{k} \lambda_i^p(f_1(A) f_1(B)) \lambda_i^{1-p}(f_2(A) f_2(B)).
\]

By the assumption that \(f_1\) and \(f_2\) satisfy inequality (4), this implies

\[
\prod_{i=1}^{k} \lambda_i(f(A) f(B)) \leq \prod_{i=1}^{k} f_1^{2p}(\lambda_i^{1/2}(AB)) f_2^{2(1-p)}(\lambda_i^{1/2}(AB))
\leq \prod_{i=1}^{k} f^2(\lambda_i^{1/2}(AB)),
\]

i.e. \(f\) satisfies inequality (4) as well. \(\blacksquare\)
We have already proven that membership of this class is a necessary condition for inequality (11) to hold. The set of associated functions $F$ for functions in class $A$ is the set of concave functions $F$ that satisfy $0 \leq F'(y) \leq 1$ for all $y$ in the domain of $F$ where $F$ is differentiable. This set is convex, as can be seen from the fact that, for $f \in A$, $F'$ is non-increasing and the range of $F'$ is $[0, 1]$. Hence, $F'$ is a convex combination of step functions $\Phi(b - y)$ (with $\Phi$ the Heaviside step function) and the constant functions 0 and 1:

$$F'(y) = r + s \int_{(-\infty, +\infty)} \Phi(b - y) d\mu(b),$$

where $r, s \geq 0$, $r + s \leq 1$, and $d\mu$ is a probability measure (normalised positive measure). Hence, such $F$ have the integral representation

$$F(y) = \alpha + ry + s \int_{(-\infty, +\infty)} \min(y, t) d\mu(t).$$

(9)

The additive constant $\alpha$ corresponds to multiplication of $f$ by $e^\alpha$, so we may assume that $\alpha = 0$. Since $r + s \leq 1$ it then follows that $f$ is in the geometric convex closure of $f(x) = 1$, $f(x) = x$ and $f(x) = \min(x, c)$ for $c \in I$ ($c = e^t$).

The next step of the proof is to show that inequality (11) holds for these extremal functions. For the functions $f(x) = 1$ and $f(x) = x$ this is of course trivial to prove. Hence let us consider the remaining function $f(x) = \min(x, c)$, with $c \in I$. As the constant $c$ can be absorbed in the matrices $A$ and $B$, we only need to check the function $f(x) = \min(x, 1)$. The action of this function on a matrix $A$ is to replace any eigenvalue of $A$ that is bigger than 1 by the value 1. I denote this matrix function by $\min(A, 1)$. For this function a stronger inequality can be proven than what is actually needed.

**Lemma 3** For $A, B \geq 0$, and for any $i$

$$\lambda_i(\min(A, 1) \min(B, 1)) \leq \min(\lambda_i(AB), 1).$$

*Proof.* Let $A_1 = \min(A, 1)$ and $B_1 = \min(B, 1)$. We have $A_1 \leq A$ and $B_1 \leq B$, so that, using Weyl monotonicity of the eigenvalues twice,

$$\lambda_i(A_1 B_1) = \lambda_i(A_1^{1/2} B_1 A_1^{1/2})$$
\begin{align*}
& \leq \lambda_i(A_1^{1/2} B A_1^{1/2}) \\
& = \lambda_i(B_1^{1/2} A_1 B_1^{1/2}) \\
& \leq \lambda_i(B_1^{1/2} A B_1^{1/2}) \\
& = \lambda_i(AB).
\end{align*}

This implies also that $\min(\lambda_i(A_1 B_1), 1) \leq \min(\lambda_i(AB), 1)$.
We also have $A_1 \leq I$ and $B_1 \leq I$, hence by Lemma 1
$$\min(\lambda_i(A_1 B_1), 1) = \lambda_i(A_1 B_1).$$
\[\Box\]

Since the inequality of this lemma implies the weaker log-submajorisation inequality
$$\lambda(\min(A, 1) \min(B, 1)) \prec_{w(\log)} \min(\lambda(AB), 1),$$
all extremal points of the class $A$ satisfy the inequality (4).

Finally, by Proposition 4 this implies that (4) holds for all functions in $A$, hence membership of $A$ is a sufficient condition. This ends the proof of Theorem 1.\[\Box\]

3.3 Proof of sufficiency for Theorem 2

Let $A = g(X)$ and $B = g(Y)$, with $g = f^{-1}$. Thus, $X = f(A)$ and $Y = f(B)$. Since $f$ is in $A$, $g$ is in $B$. Inequality (4) then gives
$$\lambda(XY) \prec_{w(\log)} f^2 \left( \sqrt{\lambda(g(X)g(Y))} \right). \tag{10}$$

The right-hand side features the function $w(x) = f^2(\sqrt{x})$. Because $f$ is geometrically concave, so is $w$. The inverse function $w^{-1}$ is given by $w^{-1}(y) = g^2(\sqrt{y})$. Therefore, $w^{-1}$ is geometrically convex. Furthermore, because $f'$ is non-negative, $w^{-1}$ is monotonously increasing.

A monotonous convex function preserves the weak majorisation relation (4, Corollary II.3.4). Thus, a monotonous geometrically convex function preserves the log-submajorisation relation. Hence, when $w^{-1}$ is applied to both sides of (10) one obtains
$$w^{-1}(\lambda(XY)) \prec_{w(\log)} \lambda(g(X)g(Y)),$$
which is (5).\[\Box\]

4 Application

An interesting application concerns the function $f(x) = 1 - \exp(-x)$, which is in class $A$. A simple application of Theorem 1 leads to an inequality that is complementary to the famous Golden-Thompson inequality $\text{Tr} \exp(A + B) \leq \text{Tr} \exp(A) \exp(B)$ (where $A$ and $B$ are Hermitian).

In [6] (see also [1]) the inequality
$$\text{Tr}(\exp(pA) \# \exp(pB))^{2/p} \leq \text{Tr} \exp(A + B)$$
was proven, for every $p > 0$ and Hermitian $A$ and $B$. This is complementary to the Golden-Thompson inequality because it provides a lower bound on $\text{Tr} \exp(A + B)$. The bound obtained below is complementary in a different sense, as it provides an upper bound on $\text{Tr} e^{-A} e^{-B}$ (for positive $A$ and $B$).

**Theorem 3** For $A, B \geq 0$, and with $C = (A^{1/2} B A^{1/2})^{1/2}$,

$$\text{Tr}(e^{-A} e^{-B}) \leq \text{Tr}(e^{-A} + e^{-B}) + \text{Tr}(e^{-2C} - 2e^{-C}). \tag{11}$$

**Proof.** The inequality can be rewritten as $\text{Tr} f(A)f(B) \leq \text{Tr} f^2(C)$, with $f(x) = 1 - e^{-x}$. By Lemma 5 in [3] $f(x)$ is geometrically concave. Moreover, $f$ is in $\mathcal{A}$: obviously, $f' \geq 0$; secondly, $f'(x) = \exp(-x) \leq 1/(1 + x)$, so that $xf'(x) = x \exp(-x) \leq 1 - \exp(-x) = f(x)$.

Hence $f$ satisfies the conditions of Theorem 1. The inequality follows immediately from that theorem, as log-submajorisation implies weak majorisation, and majorisation of the trace, in particular. $\square$

**Acknowledgements**

Many thanks to F. Hiai for detailed comments on an earlier version of the manuscript. I am grateful for the hospitality of the University of Ulm, where this paper was completed.

**References**

[1] T. Ando and F. Hiai, “Log majorization and complementary Golden-Thompson inequalities”, Lin. Alg. Appl. **197**, 113–131 (1994).

[2] H. Araki, “On an inequality of Lieb and Thirring”, Lett. Math. Phys. **19**, 167–170 (1990).

[3] K.M.R. Audenaert, “Trace inequalities for completely monotone functions and Bernstein functions”, Linear Algebra Appl. In press (2012); see also eprint arXiv:1109.3057.

[4] R. Bhatia, *Matrix Analysis*, Springer, Heidelberg (1997).

[5] J.C. Bourin and F. Hiai, “Jensen and Minkowski type inequalities for operator means”, e-print arXiv:1106.2213 (2011).

[6] F. Hiai and D. Petz, “The Golden-Thompson trace inequality is complemented”, Linear Algebra Appl. **181**, 153–185 (1993).
[7] E. Lieb and W. Thirring, in E. Lieb, B. Simon and A. Wightman (eds), “Studies in Mathematical Physics”, pp. 301–302, Princeton Press (1976).

[8] P. Montel, “Sur les fonctions convexes et les fonctions sousharmoniques”, J. Math. Pures Appl. Neuvième Série, 7, 29–60 (1928).