A Lieb-Thirring inequality for singular values

Koenraad M.R. Audenaert

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

Abstract

Let $A$ and $B$ be positive semidefinite matrices. We investigate the conditions under which the Lieb-Thirring inequality can be extended to singular values. That is, for which values of $p$ does the majorisation $\sigma(B^pA^p) \prec_w \sigma((BA)^p)$ hold, and for which values its reversed inequality $\sigma(B^pA^p) \succ_w \sigma((BA)^p)$.

Key words: Inequalities, Lieb-Thirring inequality, Singular values, Majorisation

1991 MSC: 15A60

The famous Lieb-Thirring inequality [6] states that for positive semidefinite matrices $A$ and $B$, and $p \geq 1$, $\text{Tr}(AB)^p \leq \text{Tr}(A^pB^p)$, while for $0 < p \leq 1$ the inequality is reversed. Many generalisations of this inequality exist [2,7], one of the most notable being the Araki-Lieb-Thirring inequality [1]. For positive matrices $A$ and $B$, and any unitarily invariant norm $||| \cdot |||$, the following holds (see also Theorem IX.2.10 in [3]): $|||(BAB)^p||| \leq |||B^pA^pB^p|||$ when $p \geq 1$, and the reversed inequality when $0 < p \leq 1$. This inequality can be equivalently expressed as the weak majorisation relation between singular values $\sigma((BAB)^p) \prec_w \sigma(B^pA^pB^p)$. Here, $\sigma(X) \prec_w \sigma(Y)$ if and only if \( \sum_{j=1}^{k} \sigma_j(X) \leq \sum_{j=1}^{k} \sigma_j(Y) \), for $1 \leq k \leq d$, where $\sigma_j(X)$ denotes the $j$-th largest singular value of $X$.

In this paper we study the related question whether a majorisation relation exists between the singular values of the non-symmetric product $B^pA^p$ and $(BA)^p$. The latter expression is well-defined because the eigenvalues of a product of positive semidefinite matrices are real and non-negative. Our main result is the following:

**Theorem 1** Let $A, B \geq 0$ be $d \times d$ matrices. For $0 < p \leq 1/2$,

$$\sigma(B^pA^p) \prec_w \sigma((BA)^p).$$

Email address: koenraad.audenaert@rhul.ac.uk (Koenraad M.R. Audenaert).
In addition, if \( d = 2 \), the range of validity extends to \( 0 < p \leq 1 \).

For \( p \geq d - 1 \) and for \( p \in \mathbb{N}_0 \), the reversed inequality holds:

\[
\sigma(B^p A^p) \succ_w \sigma((BA)^p).
\] (2)

In the first half of the paper, we prove this Theorem for \( p \) satisfying the condition \( 1/p \in \mathbb{N}_0 \) or \( 1/p \geq d - 1 \) or \( p \in \mathbb{N}_0 \) or \( p \geq d - 1 \). We do so by chaining together two majorisations; in terms of the first inequality, (1), we chain together \( \sigma(A^p B^p) \prec_w \sigma^p(AB) \) and \( \sigma^p(AB) \prec_w \sigma((AB)^p) \). While the first majorisation indeed holds generally and is a straightforward consequence of the original Lieb-Thirring inequality, see Theorem 2, the second majorisation turns out to be subject to the rather surprising condition on \( p \) (Theorem 3). In the second half of this paper, we follow a different route and obtain validity of (1) for \( 0 < p \leq 1/2 \).

Henceforth, we abbreviate the term positive semidefinite as PSD.

The following Theorem is already well-known:

**Theorem 2** For \( A, B \) PSD, and \( 0 < p \leq 1 \),

\[
\sigma(A^p B^p) \prec_w \sigma^p(AB).
\]

For \( p \geq 1 \), the direction of the majorisation is reversed.

**Proof.** We only have to prove the statement for \( \sigma_1 \), i.e. the infinity norm \( \| \cdot \|_\infty \).

From that we can derive the full majorisation statement by using the well-known trick, due to Weyl, of replacing \( X \) by its antisymmetric tensor powers, as in [1].

Consider \( 0 < p \leq 1 \). By the Araki-Lieb-Thirring inequality for the infinity norm \( \| \cdot \|_\infty \), we have

\[
\| AB^2 A \|^p = \|(AB^2 A)^p\| \geq \| A^p B^{2p} A^p \|.
\]

Noting that \( \|XX^*\| = \|X\|^2 \), this gives, indeed,

\[
\| AB \|^p \geq \| A^p B^p \|.
\]

This inequality was first proven by Heinz (see Theorem IX.2.3 in [3]). For \( p \geq 1 \), the direction of the inequalities is reversed. \( \Box \)

For the second majorisation we need a lemma, which relates the question to a result by FitzGerald and Horn.
Lemma 1 Let \((\lambda_i)_i\) be a sequence of \(d\) non-negative numbers. The \(d \times d\) matrix \(C\) with entries
\[
C_{i,j} = \frac{1 - \lambda_i^\alpha \lambda_j^\alpha}{1 - \lambda_i \lambda_j},
\]
is PSD if \(\alpha \in \mathbb{N}_0\) or \(\alpha \geq d - 1\).

Proof. This expression can be represented in integral form as
\[
C_{i,j} = \alpha \int_{0}^{1} dt \,(t + (1 - t)\lambda_i \lambda_j)^{\alpha - 1}.
\]
Thus \(C\) is PSD if the integrand is. Since for \(0 \leq t \leq 1\) the matrix \((t + (1 - t)\lambda_i \lambda_j)_{i,j}\) is PSD and has non-negative entries, \(C\) being PSD follows from a Theorem of FitzGerald and Horn [5] that states that the \(q\)-th entrywise power of an entrywise non-negative PSD matrix is again PSD, provided either \(q \in \mathbb{N}_0\) or \(q \geq d - 2\). Here, \(q = \alpha - 1\), hence the condition is \(\alpha \in \mathbb{N}_0\) or \(\alpha \geq d - 1\). \(\square\)

Theorem 3 Let \(X\) be a \(d \times d\) matrix with non-negative real eigenvalues. For \(p\) in the range \(0 < p \leq 1\), the majorisation
\[
\sigma^p(X) \prec_w \sigma(X^p)
\]
holds, provided \(1/p \in \mathbb{N}_0\) or \(1/p \geq d - 1\).

For the range \(p \geq 1\), the direction of the majorisation is reversed, and the conditions for validity are \(p \in \mathbb{N}_0\) or \(p \geq d - 1\).

Proof. Consider the case \(0 < p \leq 1\) first.

Again, we consider the inequality \(\sigma^p_1(X) \leq \sigma_1(X^p)\), from which the majorisation of the Theorem follows by the Weyl trick.

An equivalent statement of the inequality is: \(\|X^p\| = 1\) implies \(\|X\| \leq 1\) (obtainable via rescaling \(X\)).

If we impose that \(X\) be diagonalisable, it has an eigenvalue decomposition \(X = S \Lambda S^{-1}\), where \(S\) is invertible, and \(\Lambda\) is diagonal, with diagonal entries \(\lambda_k \geq 0\). Then
\[
\|X^p\| = 1 \implies (X^p)^* (X^p) \leq 11
\implies S^{-*} \Lambda^p S^* S \Lambda^p S^{-1} \leq 11
\implies \Lambda^p S^* S \Lambda^p \leq S^* S.
\]
Let us introduce the matrix \(A = S^* S\), which of course is positive definite, by invertibility of \(S\). Thus the statement \(\|X^p\| = 1\) is equivalent with \(\Lambda^p A \Lambda^p \leq A\).
Likewise, the statement $||X|| = 1$ is equivalent with $\Lambda A \Lambda \leq A$. We therefore have to prove the implication

$$\Lambda^p A \Lambda^p \leq A \implies \Lambda A \Lambda \leq A.$$  \hspace{1cm} (3)

Now note that, since $\Lambda$ is diagonal, the condition $\Lambda^p A \Lambda^p \leq A$ can be written as

$$A' := A \circ (1 - \lambda_i^p \lambda_j^p)_{i,j=1} \geq 0,$$

where $\circ$ denotes the Hadamard product. Likewise, $\Lambda A \Lambda \leq A$ can be written as

$$A \circ (1 - \lambda_i \lambda_j)_{i,j=1} \geq 0.$$

In terms of $A'$, this reads

$$A' \circ C \geq 0,$$

with

$$C := \left( \frac{1 - \lambda_i \lambda_j}{1 - \lambda_i^p \lambda_j^p} \right)_{i,j=1}^d.$$

Thus, by Schur’s Theorem [4], the implication (3) would follow from non-negativity of the matrix $C$. Using Lemma [1], we find that a sufficient condition is $1/p \in \mathbb{N}_0$ or $1/p \geq d - 1$.

Using a standard continuity argument, we can now remove the restriction that $X$ be diagonalisable.

The case $p > 1$ is treated in a completely similar way, but relying instead on the non-negativity of the matrix

$$\left( \frac{1 - \lambda_i^p \lambda_j^p}{1 - \lambda_i \lambda_j} \right)_{i,j=1}^d.$$

For all other values of $p$ than the mentioned ones, the matrix $C$ encountered in the proof is in general no longer non-negative. Likewise, for these other values of $p$, counterexamples can be found to the inequality that we wanted to prove here, so the given conditions on $p$ are the best possible.

Combining Theorem 2 and Theorem 3 immediately proves Theorem 1 for $1/p \in \mathbb{N}_0$ or $1/p \geq d - 1$ or $p \in \mathbb{N}_0$ or $p \geq d - 1$.

\[ \star \star \star \]

To prove the remaining case covered by Theorem [1], we derive several equivalent forms of the inequalities (1) and (2). We again only need to treat the $\sigma_1$ case, as the full statement follows from it using the Weyl trick.
Consider first the case $0 < p \leq 1$. Then we need to consider

$$
\|B^p A^p\| \leq \|(BA)^p\|,
$$

(4)

since the largest singular value is just the operator norm.

As a first step, we reduce the expressions in such a way that only positive matrices appear with a fractional power.

By exploiting the relation $\|X\| = \|X^*X\|^{1/2}$, (4) is equivalent to

$$
\|A^p B^{2p} A^p\| \leq \|(AB)^p (BA)^p\|,
$$

which, by homogeneity of both sides, can be reformulated as

$$
\|(AB)^p (BA)^p\| \leq 1 \implies \|A^p B^{2p} A^p\| \leq 1,
$$

and, in terms of the PSD ordering,

$$
(AB)^p (BA)^p \leq 11 \implies A^p B^{2p} A^p \leq 11.
$$

(5)

**Lemma 2** For any $A > 0$ and $B \geq 0$, there exist diagonal $\Lambda \geq 0$ and invertible $S$ such that $A = SS^*$ and $AB = S\Lambda S^{-1}$, and, consequently, $B = S^{-*}\Lambda S^{-1}$.

*Proof.* Let $AB = T\Lambda T^{-1}$ be an eigenvalue decomposition of $AB$. Because $A$ and $B$ are PSD, the eigenvalues of $AB$ are non-negative, hence $\Lambda \geq 0$. Assuming that all eigenvalues of $AB$ are distinct, we show that $T^{-1}AT^{-*}$ is necessarily diagonal.

Indeed, from $AB = T\Lambda T^{-1}$ follows $T^{-1}AT^{-*} T^*BT = \Lambda$. The factors $X = T^{-1}AT^{-*}$ and $Y = T^*BT$ are positive definite, and positive semidefinite, respectively, since they are related to $A$ and $B$ by a $*$-conjugation. Now note that $\Lambda$ is diagonal and all its diagonal elements are distinct. This implies that $X$ and $Y$, both Hermitian, are themselves diagonal. This follows from taking the hermitian conjugate of $XY = \Lambda$, $YX = \Lambda$, and noting that the two equations taken together imply that $X$ and $Y$ commute and are therefore diagonalised by the same unitary conjugation. Then we see that the product $XY$ must also be diagonalised by that same unitary conjugation. However, $XY = \Lambda$ is already diagonal, so that $X$ and $Y$ must be diagonal too.

By a continuity argument, we see that there must exist a $T$ diagonalising both $AB$ (via a similarity) and $A$ (via a $*$-conjugation) even when the eigenvalues of $AB$ are not distinct.

The lemma now follows by putting $S = TX^{1/2}$. \(\Box\)
Using the Lemma, the left-hand side (lhs) of (5) can be rewritten as

\[(AB)^p(BA)^p = (SL^{-1})^p(S^{-*}ΛS^*)^p = SL^{-1}S^{-*}Λ^pS^*.\]

The condition \((AB)^p(BA)^p \leq 1\) then becomes

\[Λ^pS^{-1}S^{-*}Λ^p \leq S^{-1}S^{-*},\]

which turns into

\[Λ^pCA^p \leq C\]

on defining \(C = S^{-1}S^{-*} > 0\).

Similarly, the condition of the right-hand side (rhs) of (5), \(A^pB^{2p}A^p \leq 1\), can be rewritten as \(B^{2p} \leq A^{-2p}\), or

\[(S^{-*}ΛS^{-1})^{2p} \leq (SS^*)^{-2p} = (S^{-*}S^{-1})^{2p}.\]  \(6\)

Using the polar decomposition, we can put \(S^{-*} = UC^{1/2}\), where \(U\) is a unitary matrix. Then the condition of the rhs becomes \((UC^{1/2}ΛC^{1/2}U^*)^{2p} \leq (UCU^*)^{2p}\), or

\[(C^{1/2}ΛC^{1/2})^{2p} \leq C^{2p}.\]  \(7\)

Thus, implication (5) is equivalent to

\[Λ^pCA^p \leq C \implies (C^{1/2}ΛC^{1/2})^{2p} \leq C^{2p},\]  \(8\)

for \(0 \leq p \leq 1\), and \(C > 0\), \(Λ \geq 0\).

On left- and right-multiplying both sides of the lhs of \(8\) with \(C^{1/2}\), we get

\[(C^{1/2}Λ^pC^{1/2})^2 \leq C^2 \implies (C^{1/2}ΛC^{1/2})^{2p} \leq C^{2p}.\]

By putting \(A = C^{1/2}\) and \(B = Λ^p\), this becomes

\[(ABA)^2 \leq A^4 \implies (AB^{1/p}A)^{2p} \leq A^{4p}.\]

In this equivalent form, it is now easy to prove (1) for \(p \leq 1/2\).

Proof of Theorem 1 for \(0 \leq p \leq 1/2\): By operator monotonicity of the square root, \((ABA)^2 \leq A^4\) implies \(ABA \leq A^2\). Dividing out \(A\) on both sides, this is equivalent with \(B \leq 11\). This implies \(B^{1/p} \leq 11\), for all \(p > 0\), and thus \(AB^{1/p}A \leq A^2\). Since \(0 < p \leq 1/2\), operator monotonicity of the \(2p\)-th power finally implies \((AB^{1/p}A)^{2p} \leq A^{4p}\). \(\Box\)
For $d > 2$ and $1/2 < p < 1$, we have found counterexamples. To narrow down the search for counterexamples, we semi-intelligently chose a random positive diagonal $d \times d$ matrix $D$ and a random $d$-dimensional vector $\psi$ to construct $A$ and $B$ matrices:

$$A^2 = \left( \frac{\psi_k \overline{\psi}_l}{1 - D_{kk} D_{ll}} \right)_{i,j=1}^d$$

$$B = \| A^{-1} D A^2 D A^{-1} \|^{-1/2} D.$$  

The condition $(ABA)^2 \leq A^4$ is equivalent with $\| A^{-1} B A^2 B A^{-1} \| \leq 1$ and is thus satisfied by construction. However, with high probability $A$ and $B$ are found that violate $(AB^{1/p} A)^{2p} \leq A^{4p}$. As the violations are extremely small, all calculations have to be done in high-precision arithmetic (we used 60 digits of precision). This numerical procedure yielded counterexamples for $d = 3$ and $p$ between 0.89 and 1.

In a similar way counterexamples can be found in the regime $d > 2$ and $p > 1$. For $p \geq 1$, we find by a similar reasoning that the reversed inequality of (4) is equivalent to the converse of (5), and therefore to the converse implication

$$A^p C A^p \leq C \iff (C^{1/2} \Lambda C^{1/2})^{2p} \leq C^{2p}. \quad (9)$$

For $d = 3$ we have found counterexamples up to $p = 1.25$, but no higher. It is therefore imaginable that the second majorisation inequality in Theorem 1 could be valid under more general conditions, e.g. for $p \geq 2$ perhaps. For the time being, this problem is still open.

References

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

[2] K.M.R. Audenaert, “On the Araki-Lieb-Thirring inequality”, Int. J. of Information and Systems Sciences, 4(1), 78–83 (2008). See also eprint arXiv:math.FA/0701129.

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

[4] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge (1985).

---

1 A Mathematica notebook with these calculations is available from the author on request.
[5] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge (1991).

[6] E. Lieb and W. Thirring, in *Studies in Mathematical Physics* (Eds. E. Lieb, B. Simon and A. Wightman), 301–302, Princeton Press (1976).

[7] B-Y. Wang and F. Zhang, “Trace and Eigenvalue Inequalities for Ordinary and Hadamard Products of Positive Semidefinite Hermitian Matrices”, SIAM J. Matrix Anal. Appl. 16, 1173–1183 (1995).