Strongly subadditive functions

Koenraad Audenaert\textsuperscript{1}, Fumio Hiai\textsuperscript{2} and Dénes Petz\textsuperscript{3}

\textsuperscript{1} Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK
\textsuperscript{2} Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan
\textsuperscript{3} Alfréd Rényi Institute of Mathematics, H-1364 Budapest, POB 127, Hungary

Abstract

Let $f : \mathbb{R}^+ \to \mathbb{R}$. The subject is the trace inequality $\text{Tr} f(A) + \text{Tr} f(P_2 A P_2) \leq \text{Tr} f(P_{12} A P_{12}) + \text{Tr} f(P_{23} A P_{23})$, where $A$ is a positive operator, $P_1, P_2, P_3$ are orthogonal projections such that $P_1 + P_2 + P_3 = I$, $P_{12} = P_1 + P_2$ and $P_{23} = P_2 + P_3$. There are several examples of functions $f$ satisfying the inequality (called (SSA)) and the case of equality is described.

MSC (2000): Primary 47A63; Secondary 26A51, 45A90

Key words and phrases: strong subadditivity, operator monotone functions, operator concave functions, trace inequality

1 Introduction

Matrix monotone and matrix concave functions play important roles in several applications. Assume that $f : \mathbb{R}^+ \to \mathbb{R}$ is a continuous function. It is matrix monotone if $0 \leq A \leq B$ implies $f(A) \leq f(B)$ for every matrix $A$ and $B$. The function $f$ is called matrix concave if one of the following two equivalent conditions holds:

\[ f(\lambda A + (1 - \lambda)B) \geq \lambda f(A) + (1 - \lambda)f(B) \quad (1) \]

for every number $0 < \lambda < 1$ and for positive definite square matrices $A$ and $B$ (of the same size). In the other condition the number $\lambda$ is (heuristically) replaced by a matrix:

\[ f(CAC^* + DBD^*) \geq Cf(A)C^* + Df(B)D^* \quad (2) \]
if $CC^* + DD^* = I$, see the books [3, 8] about the details. It is surprising that a matrix monotone function is matrix concave.

Motivated by some applications we study the functions $f$ which are strongly subadditivite in the following sense. Let $P_1, P_2, P_3$ be orthogonal projections such that $P_1 + P_2 + P_3 = I$. Then

$$\text{Tr} f(\mathbf{A}) + \text{Tr} f(P_2 \mathbf{A} P_2) \leq \text{Tr} f(P_{12} \mathbf{A} P_{12}) + \text{Tr} f(P_{23} \mathbf{A} P_{23}),$$

where $P_{12} := P_1 + P_2$ and $P_{23} := P_2 + P_3$. The special case when $P_2 = 0$ could be called subadditivity. This holds for any concave function [7, Theorem 2.4].

The first example $f(x) = \log x$ appeared already [2], here we have several other examples and a sufficient condition. The strongly subadditive functions are concave in the sense of real variable and all known examples are matrix concave.

2 Motivation

The second quantization in quantum theory is mathematically a procedure which associates an operator on the Fock space $\mathcal{F}(\mathcal{H})$ to an operator on the Hilbert space $\mathcal{H}$ [4]. The simplest example is $\mathcal{H} = \mathbb{C}$, then $\mathcal{F}(\mathcal{H})$ is $l^2(\mathbb{Z}^+)$. To the number $\mu > 0$ (considered as a positive operator) we associate $\Gamma(\mu)$ defined as

$$\Gamma(\mu)\delta_n = \mu^n\delta_n \quad (n = 0, 1, 2, \ldots),$$

where $\delta_n$ are the standard basis vectors. $\Gamma$ can be extended to arbitrary finite dimension by the formula

$$\Gamma(H_1 \oplus H_2) = \Gamma(H_1) \otimes \Gamma(H_2).$$

In this way to any positive operator $H \in B(\mathcal{H})$ we have a positive operator $\Gamma(H) \in B(\mathcal{F}(\mathcal{H}))$. The construction of a statistical operator, analogue of the Gaussian distribution, is slightly more complicated. For a positive operator $A$ set

$$\alpha(A) = \frac{\Gamma(H)}{\text{Tr} \Gamma(H)}, \quad \text{where} \quad H = A(I + A)^{-1}.$$ 

In particular, if $A = \lambda$, then

$$\alpha(\lambda)\delta_n = \frac{1}{1 + \lambda} \left( \frac{\lambda}{1 + \lambda} \right)^n \delta_n.$$ 

The von Neumann entropy

$$S(\alpha(A)) := -\text{Tr} \alpha(A) \log \alpha(A)$$

equals to $\text{Tr} \kappa(A)$, where $\kappa(x) := -x \log x + (x + 1) \log(x + 1)$ [4, 5].

2
From the formula
\[ \log x = \int_0^\infty \frac{1}{1+t} - \frac{1}{x+t} \, dt. \]
we get
\[ \kappa(x) = -\int_0^\infty \frac{x}{1+t} - \frac{x+1}{x+t} \, dt + \int_1^\infty \frac{x+1}{1+t} - \frac{x+1}{x+t+1} \, dt. \]
Since both integrands are matrix concave, the integrals are matrix concave, too.

\[ \kappa'(x) = \log \left( 1 + \frac{1}{x} \right) > 0 \]
and \( \kappa \) is monotone. Hence \( \kappa(x) \geq \kappa(0) = 0 \). The positivity together with matrix concavity implies matrix monotonicity, [6].

Let \( \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \) be a finite dimensional Hilbert space and let
\[ A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{12}^* & A_{22} & A_{23} \\ A_{13}^* & A_{23}^* & A_{33} \end{bmatrix}, \]
be a positive invertible operator and set
\[ B = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}, \quad C = \begin{bmatrix} A_{22} & A_{23} \\ A_{23}^* & A_{33} \end{bmatrix}. \]

The strong subadditivity of the von Neumann entropy,
\[ S(\alpha(A)) + S(\alpha(A_{22})) \leq S(\alpha(B)) + S(\alpha(C)), \]
has the equivalent form
\[ \text{Tr} \kappa(A) + \text{Tr} \kappa(A_{22}) \leq \text{Tr} \kappa(B) + \text{Tr} \kappa(C). \] (4)
The case of equality is studied in the paper [5] and the general properties of entropy are in the book [8].

**Proposition 2.1** The equality
\[ \text{Tr} \kappa(A) + \text{Tr} \kappa(A_{22}) \leq \text{Tr} \kappa(B) + \text{Tr} \kappa(C), \]
in the strong subadditivity holds if and only if \( A \) has the form
\[ A = \begin{bmatrix} A_{11} & [a & 0] & 0 \\ \begin{bmatrix} a^* \\ 0 \end{bmatrix} & [c & 0] & [0] \\ 0 & [0 & d] & [b] \\ 0 & 0 & A_{33} \end{bmatrix} = \begin{bmatrix} A_{11} & a \\ a^* & c \end{bmatrix} \begin{bmatrix} 0 \\ [d & b] \end{bmatrix} \]
(5)
where the parameters \( a, b, c, d \) (and 0) are matrices.
Note that the matrix \( c \) or \( d \) in the theorem can be \( 0 \times 0 \).

We are interested in the (differentiable) functions \( f \) such that the inequality

\[
\text{Tr} f(A) + \text{Tr} f(A_{22}) \leq \text{Tr} f(B) + \text{Tr} f(C)
\]

(\text{SSA})

holds. We call this \textit{strong subadditivity} for the function \( f \). The strong subadditivity holds for the function \( \kappa \). Another equivalent formulation of the strong subadditivity is [3].

### 3 Particular examples

**Example 3.1** If

\[
A = \begin{bmatrix}
a & 0 & d \\
0 & b & 0 \\
d^* & 0 & c
\end{bmatrix}
\]

(6)

is a numerical matrix, then it is an exercise to show that (SSA) holds for this kind of \( A \) if and only if \( f \) is a concave function.

**Example 3.2** The strong subadditivity does not hold for the function \( f(t) = -1/t \). The following counterexample is due to Ando [1]: Let

\[
X \equiv A^{-1} := \begin{bmatrix}
4 & 8 & -2 \\
8 & 20 & 0 \\
-2 & 0 & 9
\end{bmatrix}.
\]

Then

\[
A = \begin{bmatrix}
\frac{45}{16} & \frac{9}{8} & \frac{5}{8} \\
\frac{9}{8} & \frac{17}{4} & \frac{1}{4} \\
\frac{5}{8} & \frac{1}{4} & \frac{1}{4}
\end{bmatrix}.
\]

We have

\[
\text{Tr} A^{-1} = 33, \quad \text{Tr} A_{22}^{-1} = 2, \quad \text{Tr} B^{-1} = \frac{212}{9}, \quad \text{Tr} C^{-1} = 12
\]

and (SSA) becomes

\[
33 + 2 \geq \frac{212}{9} + 12
\]

and this is not true.

**Example 3.3** It is elementary that the strong subadditivity holds for the function \( f(t) = -t^2 \). The equality holds if and only \( A_{13} = 0 \). \( \square \)

**Example 3.4** It was proved in [2] that the strong subadditivity holds for the function \( f(t) = \log t \) and the equality holds if and only if \( A_{13} = A_{12}A_{22}^{-1}A_{23} \).
We present an alternative approach. Now (SSA) is equivalent to
\[
\text{Det } A \cdot \text{Det } A_{22} \leq \text{Det } B \cdot \text{Det } C.
\]
Let
\[
\hat{A} := \text{Diag}(A_{11}^{-1/2}, A_{22}^{-1/2}, A_{33}^{-1/2})A \text{ Diag}(A_{11}^{-1/2}, A_{22}^{-1/2}, A_{33}^{-1/2}).
\]
Then (SSA) is equivalent to
\[
\text{Det } \hat{A} \leq \text{Det } \hat{B} \cdot \text{Det } \hat{C}.
\]
In other words, we may assume that the diagonal of \(A\) consists of \(I\)'s. Since
\[
\begin{bmatrix}
I & -\hat{A}_{12} & 0 \\
0 & I & 0 \\
0 & 0 & I
\end{bmatrix}
\begin{bmatrix}
I & \hat{A}_{12} & \hat{A}_{13} \\
\hat{A}_{12}^* & I & \hat{A}_{23}^* \\
\hat{A}_{13}^* & \hat{A}_{23}^* & I
\end{bmatrix}
\begin{bmatrix}
I & 0 & 0 \\
0 & I & -\hat{A}_{23} \\
0 & 0 & I
\end{bmatrix}
= 
\begin{bmatrix}
I - \hat{A}_{12}A_{12}^* & 0 & \hat{A}_{13} - \hat{A}_{12}\hat{A}_{23} \\
\hat{A}_{12}^* & I & 0 \\
\hat{A}_{13}^* & \hat{A}_{23}^* & I - \hat{A}_{23}\hat{A}_{23}
\end{bmatrix},
\]
equality holds in (SSA) if \(\hat{A}_{13} = \hat{A}_{12}\hat{A}_{23}\), equivalently \(A_{13} = A_{12}A_{22}^{-1}A_{23}\). This condition is sufficient for the equality. \(\square\)

**Example 3.5** Since
\[
\frac{1}{2}(A + \text{Diag}(1, 1, -1)A \text{ Diag}(1, 1, -1)) = 
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{12}^* & A_{22} & 0 \\
0 & 0 & A_{33}
\end{bmatrix}
\]
we get a majorization
\[
A \succ 
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{12}^* & A_{22} & 0 \\
0 & 0 & A_{33}
\end{bmatrix},
\]
that is, the eigenvalue vector \(\vec{\lambda}(A)\) majorizes that of
\[
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{12}^* & A_{22} & 0 \\
0 & 0 & A_{33}
\end{bmatrix}.
\]
For any concave function \(f\), this implies that \(f \circ \vec{\lambda}(A)\) is weakly majorized by the \(f \circ \lambda\) of
\[
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{12}^* & A_{22} & 0 \\
0 & 0 & A_{33}
\end{bmatrix}
\]
so that
\[
\text{Tr } f(A) \leq \text{Tr } f \left( 
\begin{bmatrix}
A_{11} & A_{12} & 0 \\
A_{12}^* & A_{22} & 0 \\
0 & 0 & A_{33}
\end{bmatrix} \right) = \text{Tr } f(B) + \text{Tr } f(A_{33}). \quad (7)
\]
Hence
\[ \text{Tr} f(A) + \text{Tr} f(A_{22}) \leq \text{Tr} f(B) + \text{Tr} f(A_{22}) + \text{Tr} f(A_{33}). \]
This says that if \( A_{23} = 0 \) (or \( A_{12} = 0 \)), then (SSA) holds for every concave function \( f \).

Note that inequality (7) is written as
\[ \text{Tr} f(A) = \text{Tr} P f(A) P + \text{Tr} Q f(A) Q \leq \text{Tr} f(PAP + QAQ) \]
when \( P \) and \( Q \) are orthogonal projections and \( P + Q = I \). This is a special case of Jensen’s trace inequality for concave functions \[7, \text{Theorem 2.4}]. \]

\[\square\]

**Example 3.6** The representation
\[ y'(t) = \frac{\sin \pi t}{\pi} \int_0^\infty \frac{\lambda^{t-1}y}{\lambda + y} d\lambda \] (8)
is used to show that \( f(x) = x^t \) is operator monotone when \( 0 < t < 1 \). From this we obtain
\[ \int_0^x y^{t-1} dy = \frac{\sin \pi t}{\pi} \int_0^\infty \int_0^x \frac{\lambda^{t-1}}{\lambda + y} d\lambda dy \]
which gives
\[ x^t = \frac{t \sin \pi t}{\pi} \int_0^\infty \lambda^{t-1}(\log(x + \lambda) - \log \lambda) d\lambda. \]
So we have a similar formula to (8):
\[ x^t = \frac{t \sin \pi t}{\pi} \int_0^\infty \lambda^{t-1} \log \left(1 + \frac{x}{\lambda}\right) d\lambda. \] (9)

Since inequality (SSA) is true for the functions \( f_\lambda(x) := \log \left(1 + \frac{x}{\lambda}\right) \), by integration it follows for \( x^t \) when \( 0 < t < 1 \).

We analyze the condition for equality and use the decomposition \( \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \). For \( f_\lambda \) the equality condition is
\[ A_{13} = A_{12}(\lambda + A_{22})^{-1} A_{23}, \]
see Example 3.4. This holds for every \( \lambda > 0 \). If \( \lambda \to \infty \) in the relation
\[ \lambda A_{13} = A_{12} \left[ \lambda(\lambda + A_{22})^{-1} \right] A_{23}, \]
then we conclude \( A_{13} = 0 = A_{12} A_{23} \). The latter condition means that Rng \( A_{23} \subset \text{Ker} A_{12} \), or equivalently \( (\text{Ker} A_{12})^\perp \subset \text{Ker} A_{23}^4 \).

The linear combinations of the functions \( x \mapsto 1/(\lambda + x) \) form an algebra and due to the Stone-Weiersrass theorem \( A_{12} g(A_{22}) A_{23} = 0 \) for any continuous function \( g \).
We want to show that the equality implies the structure (5) of the operator $A$. We have $A_{23}: \mathcal{H}_3 \to \mathcal{H}_2$ and $A_{12}: \mathcal{H}_2 \to \mathcal{H}_1$. To show the structure (5), we have to find a subspace $H \subset \mathcal{H}_2$ such that

$$A_{23}H \subset H, \quad H^\perp \subset \text{Ker} \ A_{12}, \quad H \subset \text{Ker} \ A_{32},$$

or alternatively $(H^\perp =)K \subset \mathcal{H}_2$ should be an invariant subspace of $A_{22}$ such that

$$\text{Rng} \ A_{23} \subset K \subset \text{Ker} \ A_{12}.$$

Let

$$K := \left\{ \sum_i A_{22}^{n_i} A_{23} x_i : x_i \in \mathcal{H}_3, n_i \in \mathbb{Z}^+ \right\}$$

be a set of finite sums. It is a subspace of $\mathcal{H}_2$. The property $\text{Rng} \ A_{23} \subset K$ and the invariance under $A_{22}$ are obvious. Since

$$A_{12} A_{22}^{n} A_{23} x = 0,$$

$K \subset \text{Ker} \ A_{12}$ also follows. □

4 Sufficient condition

**Theorem 4.1** Let $f: (0, +\infty) \to \mathbb{R}$ be a function such that $-f'$ is matrix monotone. Then the inequality (SSA) holds.

**Proof.** The idea of the previous example is followed. A matrix monotone function has the representation

$$a + bx + \int_0^\infty \left( \frac{\lambda}{\lambda^2 + 1} - \frac{1}{\lambda + x} \right) d\mu(\lambda),$$

where $b \geq 0$, see (V.49) in [3]. Therefore, we have the representation

$$f(t) = c - \int_1^t \left( a + bx + \int_0^\infty \left( \frac{\lambda}{\lambda^2 + 1} - \frac{1}{\lambda + x} \right) d\mu(\lambda) \right) dx.$$

By integration we have

$$f(t) = d - at - \frac{b}{2} t^2 + \int_0^\infty \left( \frac{\lambda}{\lambda^2 + 1} (1 - t) + \log \left( \frac{\lambda}{\lambda + 1} + \frac{t}{\lambda + 1} \right) \right) d\mu(\lambda).$$

The first quadratic part satisfies the (SSA) and we have to check the integral. Since $\log x$ is a strongly subadditive function, so is the integrand. The integration keeps the property. □

The previous theorem covers all known examples, but we can get new examples.
Example 4.1 By differentiation we can see that $f(x) = -(x + t) \log(x + t)$ with $t \geq 0$ satisfies (SSA). Similarly, $f(x) = -x^t$ satisfies (SSA) if $1 \leq t \leq 2$.

In some applications [9] the operator monotone functions

$$f_p(x) = p(1 - p) \frac{(x - 1)^2}{(x^p - 1)(x^{1-p} - 1)} \quad (0 < p < 1)$$

appear.

For $p = 1/2$ this is an (SSA) function. Up to a constant factor, the function is

$$(\sqrt{x} + 1)^2 = x + 2\sqrt{x} + 1$$

and all terms are known to be (SSA). The function $-f'_{1/2}$ is evidently matrix monotone.

Numerical computation shows that $-f'_p$ seems to be matrix monotone. □

Acknowledgements. This work was partially supported by the Hungarian Research Grant OTKA T068258 (D.P.) and Grant-in-Aid for Scientific Research (B)17340043 (F.H.) as well as by Hungary-Japan HAS-JSPS Joint Project (D.P. & F.H.). D.P. is also grateful to Professor Tsuyoshi Ando for communication and for his example.

References

[1] T. Ando, private communication, 2008.

[2] T. Ando and D. Petz, Gaussian Markov triplets approached by block matrices, Acta Sci. Math. (Szeged), 75(2009), 265–281.

[3] R. Bhatia, Matrix analysis, Springer, 1996.

[4] O. Bratteli and D.W. Robinson, Operator algebras and quantum statistical mechanics I, II, Springer, 1979, 1981.

[5] A. Jenčová, D. Petz and J. Pitrik, Markov triplets on CCR algebras, Acta Sci. Math. (Szeged) 76(2010), 27–50.

[6] F. Hansen and G.K. Pedersen, Jensen’s inequality for operators and Löwner’s theorem, Math. Ann. 258(1981/82), 229–241.

[7] F. Hansen and G. K. Pedersen, Jensen’s operator inequality, Bull. London Math. Soc. 35(2003), 553–564.

[8] D. Petz, Quantum Information Theory and Quantum Statistics, Springer-Verlag, Heidelberg, 2008.

[9] D. Petz and V.E. S. Szabó, From quasi-entropy to skew information, Int. J. Math. 20(2009), 1335–1345.