Journal of Mathematical Inequalities
Volume 10, Number 1 (2016), 1–17
doi:10.7153/jmi-10-01

ON THE SUM OF SQUARED LOGARITHMS INEQUALITY AND RELATED INEQUALITIES

FOZI M. DANNAN, PATRIZIO NEFF AND CHRISTIAN THIEL

(*Communicated by G. Toader*)

Abstract. We consider the sum of squared logarithms inequality and investigate possible connections with the theory of majorization. We also discuss alternative sufficient conditions on two sets of vectors $a,b \in \mathbb{R}^n_+$ so that

$$\sum_{i=1}^{n} (\log a_i)^2 \leqslant \sum_{i=1}^{n} (\log b_i)^2 .$$

Generalizations of some inequalities from information theory are obtained, including a generalized information inequality and a generalized log sum inequality, which states for $a,b \in \mathbb{R}^n_+$ and $k_1,\ldots,k_n \in [0,\infty)$:

$$\sum_{i=1}^{n} a_i \log \prod_{s=1}^{m} \left(\frac{a_i}{b_i} + k_s \right) \geqslant \log \prod_{s=1}^{m} (1+k_s) .$$

1. Introduction – the sum of squared logarithms inequality

The *Sum of Squared Logarithms Inequality* (SSLI) was introduced in 2013 by Bîrsan, Neff and Lankeit [1], with the authors giving a proof for $n \in \{2,3\}$. Recently, Pompe and Neff [12] have shown the inequality for $n=4$, in which case it reads: Let $a_1,a_2,a_3,a_4,b_1,b_2,b_3,b_4 > 0$ be given positive numbers such that

$$\begin{aligned}
a_1+a_2+a_3+a_4 &\leqslant b_1+b_2+b_3+b_4 ,\\
a_1 a_2 + a_1 a_3 + a_2 a_3 + a_1 a_4 + a_2 a_4 + a_3 a_4 &\leqslant b_1 b_2 + b_1 b_3 + b_2 b_3 + b_1 b_4 + b_2 b_4 + b_3 b_4 ,\\
a_1 a_2 a_3 + a_1 a_2 a_4 + a_2 a_3 a_4 + a_1 a_3 a_4 &\leqslant b_1 b_2 b_3 + b_1 b_2 b_4 + b_2 b_3 b_4 + b_1 b_3 b_4 ,\\
a_1 a_2 a_3 a_4 &= b_1 b_2 b_3 b_4 .
\end{aligned}$$

Then

$$(\log a_1)^2 + (\log a_2)^2 + (\log a_3)^2 + (\log a_4)^2 \leqslant (\log b_1)^2 + (\log b_2)^2 + (\log b_3)^2 + (\log b_4)^2 .$$

The general form of this inequality can be conjectured as follows.

Mathematics subject classification (2010): 26D05, 26D07.
Keywords and phrases: Sum of squared logarithms inequality, exponential functions inequalities, log sum inequality, Gibbs' inequality, information inequality.

Paper JMI-10-01

DEFINITION 1.1. Let $x \in \mathbb{R}^n$. We denote by $e_k(x)$ the k-th *elementary symmetric polynomial*, i.e. the sum of all $\binom{n}{k}$ products of exactly k components of x, so that

$$e_k(x) := \sum_{i_1 < \ldots < i_k} x_{i_1} x_{i_2} \ldots x_{i_k} \qquad \text{for all } k \in \{1, \ldots, n\}\,;$$

note that $e_n(x) = x_1 \cdot x_2 \ldots \cdot x_n$.

CONJECTURE 1.2. (Sum of squared logarithms inequality) *Let $a, b \in \mathbb{R}^n_+$. If*

$$e_k(a) \leqslant e_k(b) \qquad \textit{for all } k \in \{1, \ldots, n-1\}$$

and $e_n(a) = e_n(b)$, then

$$\sum_{i=1}^n (\log a_i)^2 \;\leqslant\; \sum_{i=1}^n (\log b_i)^2\,. \tag{SSLI}$$

Alternatively, we can express the same statement as a minimization problem: Let $a \in \mathbb{R}^n_+$ be given and define

$$\mathscr{E}_a := \left\{ b \in \mathbb{R}^n_+ \mid e_k(a) \leqslant e_k(b) \quad \textit{for all } k \in \{1, \ldots, n-1\} \quad \textit{and} \quad e_n(a) = e_n(b) \right\}.$$

Then

$$\inf_{b \in \mathscr{E}_a} \left\{ \sum_{i=1}^n (\log b_i)^2 \right\} \;=\; \sum_{i=1}^n (\log a_i)^2\,.$$

The sum of squared logarithms inequality (SSLI) has important applications in matrix analysis and nonlinear elasticity theory [1, 7, 8, 9, 10, 11]. We notice that the previous conjecture for $n \geqslant 2$ puts conditions on the elementary symmetric polynomials of the numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ for obtaining (SSLI).

In this article we obtain (SSLI) and the so called sum of powered logarithms inequality (3.5) under some alternative conditions. We also introduce some new inequalities for the exponential functions. An extension of the log sum inequality is also obtained, which yields generalizations of the information inequality.

2. Preliminaries

The concept of *majorization* is of great importance to (SSLI) and related inequalities. In the following we state the basic definitions as well as some fundamental properties of majorization. For a larger survey we refer to Marshall, Olkin and Arnold [5].

There are various ways to define Majorization. Due to Hardy, Littlewood and Pólya [4] we can formulate the following theorem:

THEOREM 2.1. *Let $x, y \in \mathbb{R}^n_+$, then the following are equivalent:*

a) $\sum_{i=1}^n x_i^{\downarrow} = \sum_{i=1}^n y_i^{\downarrow}$ *and* $\sum_{i=1}^k x_i^{\downarrow} \leqslant \sum_{i=1}^k y_i^{\downarrow}$ *for all* $1 \leqslant k \leqslant n-1$.

b) $\sum_{i=1}^n f(x_i) \leqslant \sum_{i=1}^n f(y_i)$ *for all convex continous functions* $f : \mathbb{R} \to \mathbb{R}$.

If the expressions hold, x is said to be majorized by y, written $x \prec y$.

Similarly, the concept of the weak majorization can be fomulated (see Tomic [13] and Weyl [14]):

THEOREM 2.2. *Let $x,y \in \mathbb{R}^n_+$, then the following are equivalent:*

a) $\sum_{i=1}^k x_i^{\downarrow} \leqslant \sum_{i=1}^k y_i^{\downarrow}$ *for all* $1 \leqslant k \leqslant n$.

b) $\sum_{i=1}^n f(x_i) \leqslant \sum_{i=1}^n f(y_i)$ *for all convex monotone increasing continous functions* $f : \mathbb{R} \to \mathbb{R}$.

If the expressions hold, x is said to be weakly majorized by y, written $x \prec_w y$.

If x is weakly majorized by y, we state, in particular, that it is weakly majorized from below. Analogously we can also characterize weak majorization from above ($x \prec^w y$), were the inequality in a) holds with greater or equal for all vectors in descending order and equivalent die inequality in b) for all convex monotone decreasing functions.

The following lemma, which shows elementary properties of the so-called *(weak) logarithmic majorization*, follows directly from the logarithmic laws and the monotonicity of the logarithm.

LEMMA 2.3. (Logarithmic majorization) *Let $x,y \in \mathbb{R}^n_+$. Then*

$$\begin{aligned} \log x \prec_w \log y \quad &\textit{if and only if} \quad x_1^{\downarrow} \cdot \ldots \cdot x_k^{\downarrow} \leqslant y_1^{\downarrow} \cdot \ldots \cdot y_k^{\downarrow} \quad \textit{for all } k \in \{1,\ldots,n\}, \\ \log x \prec^w \log y \quad &\textit{if and only if} \quad x_1^{\uparrow} \cdot \ldots \cdot x_k^{\uparrow} \geqslant y_1^{\uparrow} \cdot \ldots \cdot y_k^{\uparrow} \quad \textit{for all } k \in \{1,\ldots,n\}, \\ \log x \prec \log y \quad &\textit{if and only if} \quad \log x \prec_w \log y \quad \textit{and} \quad x_1 \cdot \ldots \cdot x_n = y_1 \cdot \ldots \cdot y_n, \end{aligned}$$

where we abbreviate $\log z := (\log z_1, \log z_2, \ldots, \log z_n)$ *for* $z \in \mathbb{R}^n_+$.

PROPOSITION 2.4. *Let $x,y \in \mathbb{R}^n_+$. Then*

$$\log x \prec_w \log y \quad \textit{implies} \quad x \prec_w y$$

and

$$x \prec^w y \quad \textit{implies} \quad \log x \prec^w \log y .$$

From Theorem 2.1 we can see, the mapping $\varphi : \mathbb{R}^n \to \mathbb{R}$ with $x \mapsto \sum_{i=1}^n f(x_i)$ where f is convex has the property

$$x \prec y \quad \text{implies} \quad \varphi(x) \leqslant \varphi(y) .$$

We call this property *Schur-convexity*. if $-\varphi$ is Schur-convex, so we call φ Schur-concave. Analogously to Theorem 2.1 and Theorem 2.2 we can generalize to a important property of Schur-convex functions:

PROPOSITION 2.5. *Let $\varphi : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be Schur-convex. If φ is also monotone increasing (resp. monotone decreasing), then*

$$\varphi(x) \leqslant \varphi(y) \qquad \textit{for all } x,y \in D \textit{ with } x \prec_w y \textit{ (resp. } x \prec^w y).$$

THEOREM 2.6. *The elementary symmetric polynomials* $e_k : \mathbb{R}^n \to \mathbb{R}$ *are Schur-concave and monotone increasing for all* $k \in \{1, \dots, n\}$ *and even strictly Schur-concave for* $k \geqslant 2$.

COROLLARY 2.7. *Let* $x, y \in \mathbb{R}^n_+$. *Then* $x \prec y$ *implies* $e_k(x) \geqslant e_k(y)$ *for all* $k \in \{1, \dots, n\}$. *However* $e_k(x) \geqslant e_k(y)$ *does not imply* $x \prec y$ *in general: For that consider* $a = (2,2,2)$ *and* $b = (1,1,1)$.

If x ist not only a permutation of y, the inequality between the e_k is even strict for $k \geqslant 2$. In inverse conclusion we can say:

COROLLARY 2.8. *If* $x \prec y$ *and* $e_k(x) = e_k(y)$ *for any* $k \in \{2, \dots, n\}$, *then* $x^{\downarrow} = y^{\downarrow}$ *and thus* $e_k(x) = e_k(y)$ *for every* $k \in \{2, \dots, n\}$.

3. Different conditions for SSLI and the sum of powered logarithms inequality

In the following theorems we give different conditions that guarantee the validity of (SSLI). We will use Chebyshev's sum inequality in certain cases, which states:

LEMMA 3.1. (Chebyshev's sum inequality) *If* $a_1 \leqslant a_2 \leqslant \dots \leqslant a_n$ *and* $b_1 \leqslant b_2 \leqslant \dots \leqslant b_n$ *are two monotone increasing sequences of real numbers, then*

$$\sum_{i=1}^{n} a_i b_i \geqslant \frac{1}{n} \left(\sum_{i=1}^{n} a_i \right) \left(\sum_{i=1}^{n} b_i \right) \geqslant \sum_{i=1}^{n} a_i b_{n+1-i} . \tag{3.1}$$

THEOREM 3.2. *Let* $a, b \in \mathbb{R}^n_+$ *and there exists a rearrangement of* a *and* b *that satisfy*

$$\frac{b_1}{a_1} \geqslant \frac{b_2}{a_2} \geqslant \dots \geqslant \frac{b_n}{a_n} \quad \text{and} \quad a_1 b_1 \geqslant a_2 b_2 \geqslant \dots \geqslant a_n b_n . \tag{3.2}$$

If we additionally assume one of the following two conditions

$$e_n(a) \leqslant e_n(b) \quad \text{and} \quad e_n(a) \cdot e_n(b) \geqslant 1 , \tag{3.2a}$$

or

$$e_n(a) \geqslant e_n(b) \quad \text{and} \quad e_n(a) \cdot e_n(b) \leqslant 1 , \tag{3.2b}$$

then we get the sum of squared logarithms inequality (SSLI)*:*

$$\sum_{i=1}^{n} (\log a_i)^2 \leqslant \sum_{i=1}^{n} (\log b_i)^2 .$$

Proof. First we assume Condition (3.2a):
Due to the monotonicity of the logarithm it follows from the assumptions that

$$\log \frac{b_1}{a_1} \geqslant \log \frac{b_2}{a_2} \geqslant \dots \geqslant \log \frac{b_n}{a_n} \quad \text{and} \quad \log(a_1 b_1) \geqslant \log(a_2 b_2) \geqslant \dots \geqslant \log(a_n b_n) .$$

Now we can estimate with Chebychev's inequality (3.1) using $\tilde{a}_{n+k-1} := \log \frac{b_i}{a_i}$ and $\tilde{b}_{n+k-1} := \log a_i b_i$:

$$\begin{aligned}
\sum_{i=1}^n (\log b_i)^2 - \sum_{i=1}^n (\log a_i)^2 &= \sum_{i=1}^n \left((\log b_i)^2 - (\log a_i)^2\right) \\
&= \sum_{i=1}^n (\log b_i - \log a_i)(\log b_i + \log a_i) \\
&= \sum_{i=1}^n \log \frac{b_i}{a_i} \cdot \log b_i a_i \\
&\overset{(3.1)}{\geqslant} \frac{1}{n}\left(\sum_{i=1}^n \log \frac{b_i}{a_i}\right)\left(\sum_{i=1}^n \log b_i a_i\right) \\
&= \frac{1}{n}\left(\log \underbrace{\prod_{i=1}^n \frac{b_i}{a_i}}_{e_n(b)/e_n(a)}\right)\left(\log \underbrace{\prod_{i=1}^n b_i a_i}_{e_n(a)\,e_n(b)}\right) \\
&\geqslant \frac{1}{n} \log 1 \log 1 = 0\,.
\end{aligned}$$

Finally $\sum_{i=1}^n (\log y_i)^2 - \sum_{i=1}^n (\log a_i)^2 \geqslant 0$ is equivalent to (SSLI).

Next we assume condition (3.2b).

We set $\tilde{a}, \tilde{b} \in \mathbb{R}_+^n$ with $\tilde{a}_k := \frac{1}{a_{n+1-k}}$ and $\tilde{b}_k := \frac{1}{b_{n+1-k}}$ for $k \in \{1, \ldots, n\}$, so that we have $(\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_n) = (a_n^{-1}, a_{n-1}^{-1}, \ldots, a_1^{-1})$ and $(\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_n) = (b_n^{-1}, b_{n-1}^{-1}, \ldots, b_1^{-1})$. Then

$$\frac{\tilde{b}_k}{\tilde{a}_k} = \frac{\frac{1}{b_{n+1-k}}}{\frac{1}{a_{n+1-k}}} = \frac{a_{n+1-k}}{b_{n+1-k}} \quad \text{and} \quad \tilde{a}_k \tilde{b}_k = \frac{1}{a_{n+1-k} b_{n+1-k}}\,.$$

Hence

$$\frac{\tilde{b}_1}{\tilde{a}_1} = \frac{a_n}{b_n} \geqslant \frac{\tilde{b}_2}{\tilde{a}_2} = \frac{a_{n-1}}{b_{n-1}} \geqslant \ldots \geqslant \frac{\tilde{b}_n}{\tilde{a}_n} \quad \text{and} \quad \tilde{a}_1 \tilde{b}_1 = \frac{1}{a_n b_n} \geqslant \tilde{a}_2 \tilde{b}_2 = \frac{1}{a_{n-1} b_{n-1}} \geqslant \ldots \geqslant \tilde{a}_n \tilde{b}_n\,.$$

Furthermore

$$e_n(\tilde{b}) = \frac{1}{e_n(b)} \geqslant \frac{1}{e_n(a)} = e_n(\tilde{a}) \quad \text{and} \quad e_n(\tilde{a}) e_n(\tilde{b}) = \frac{1}{e_n(a)} \cdot \frac{1}{e_n(b)} \geqslant 1\,,$$

thus $\tilde a$ and $\tilde b$ satisfy condition (3.2a). Therefore (SSLI) holds for $\tilde a$ and $\tilde b$, and we find

$$\begin{aligned}\sum_{i=1}^n (\log \tilde a_i)^2 \leqslant \sum_{i=1}^n (\log \tilde b_i)^2 &\Leftrightarrow \sum_{i=1}^n \left(\log \frac{1}{a_{n+1-i}}\right)^2 \leqslant \sum_{i=1}^n \left(\log \frac{1}{b_{n+1-i}}\right)^2 \\ &\Leftrightarrow \sum_{i=1}^n \left(\log \frac{1}{a_i}\right)^2 \leqslant \sum_{i=1}^n \left(\log \frac{1}{b_i}\right)^2 \\ &\Leftrightarrow \sum_{i=1}^n (-\log a_i)^2 \leqslant \sum_{i=1}^n (-\log b_i)^2 \\ &\Leftrightarrow \sum_{i=1}^n (\log a_i)^2 \leqslant \sum_{i=1}^n (\log b_i)^2 \quad \square \end{aligned} \tag{3.4}$$

EXAMPLE 3.3. Neither the conditions in Conjecture 1.2 are stronger than in Theorem 3.2 nor conversely:

i) With $a=(14,2,10)$ and $b=(20,2,7)$ we have $e_k(a) \leqslant e_k(b)$ for all $k \in \{1,\dots,n-1\}$ and $e_n(a)=e_n(b)$ but there is no rearrangement of a and b that satisfies $\frac{b_1}{a_1} \geqslant \frac{b_2}{a_2} \geqslant \frac{b_3}{a_3}$ and $a_1 b_1 \geqslant a_2 b_2 \geqslant a_3 b_3$.

ii) With $a=(6,5,7)$ and $b=(10,8,3)$ we have $\frac{10}{6} \geqslant \frac{8}{5} \geqslant \frac{3}{7}$ and $6\cdot 10 \geqslant 5\cdot 8 \geqslant 7\cdot 3$. Because of $e_n(a)=210$ and $e_n(b)=240$ we have $e_n(a) \leqslant e_n(b)$ and $e_n(a)\,e_n(b) \geqslant 1$ but not $e_n(a)=e_n(b)$.

iii) With $a=(2,2,2)$ and $b=(4,2,1)$ we have $e_k(a) \leqslant e_k(b)$ for all $k \in \{1,\dots,n-1\}$ and $e_n(a)=e_n(b)$. Moreover $\frac{4}{2} \geqslant \frac{2}{2} \geqslant 12$, $2\cdot 4 \geqslant 2\cdot 2 \geqslant 2\cdot 1$ and $e_n(a)\,e_n(b) \geqslant 1$.

THEOREM 3.4. (sum of powered logarithms inequality) *Let $a,b \in \mathbb{R}^n$ and $p \in \mathbb{R}$ with $a_i > 1$, $b_i > 1$ and $p < 0$. Assume $a \prec^w b$. Then*

$$\sum_{i=1}^n (\log a_i)^p \leqslant \sum_{i=1}^n (\log b_i)^p . \tag{3.5}$$

REMARK 3.5. In order for $(\log a_i)^p$ and $(\log b_i)^p$ to be well defined for all $p \in \mathbb{R}$, we must assume $a_i > 1$ and $b_i > 1$.

Proof. From $a \prec^w b$ with Proposition 2.4 we obtain

$$\sum_{i=1}^k \log a_i^{\uparrow} \geqslant \sum_{i=1}^k \log b_i^{\uparrow} \qquad \text{for all } k \in \{1,\dots,n\} .$$

Let $x,y \in \mathbb{R}^n_+$ with $x := \log a$, $y := \log b$ (therefore $x_i := \log a_i$ and $y_i := \log b_i$ for all $k \in \{1,\dots,n\}$) then $x \prec^w y$. We now consider the function $g : \mathbb{R}_+ \to \mathbb{R}$ with $g(z) = z^p$, then

$$g'(z) = \underbrace{p}_{<0} \cdot \underbrace{z^{p-1}}_{>0} < 0 \qquad \text{and} \qquad g''(z) = \underbrace{p\,(p-1)}_{>0} \cdot \underbrace{z^{p-2}}_{>0} > 0 .$$

For this reason g is monotone decreasing and convex. With Theorem 2.1 we obtain $\sum_{i=1}^n x_i^p \leqslant \sum_{i=1}^n y_i^p$. Resubstitution now directly yields the statement. □

REMARK 3.6. $a \prec b$ is a condition too weak and $a \prec b$ plus $e_n(a) = e_n(b)$ is too strong for (SSLI):

For $a = (3,2,2)$ and $b = (4,2,1)$ we get $a \prec b$ but $2.17 \approx \sum(\log a_i)^2 < \sum(\log b_i)^2 \approx 2.40$.

For $a = (4,4,4)$ and $b = (10,1,1)$ we get $a \prec b$ but $5.77 \approx \sum(\log a_i)^2 > \sum(\log b_i)^2 \approx 5.30$.

f we think $a \prec b$ implies $e_k(a) \geqslant e_k(b)$ and the only thing we have additionally to ensure for the use of Conjecture 1.2 to get (SSLI) is $e_n(a) = e_n(b)$ then remember: We have shown $a \prec b$ and $e_n(a) = e_n(b)$ implies $a^{\downarrow} = b^{\downarrow}$, therefore only two permutations of one vector satisfy this condition.

We have shown by Lemma 2.7 that $a \prec b$ implies $e_k(a) \geqslant e_k(b)$ (note the reverse inequality).

What about using $a \prec b$ and $e_n(a) = e_n(b)$ as sufficient requirements for (SSLI)?

We can easily show $a \prec b$ and $e_n(a) = e_n(b)$ imply the logarithmic majorization $\log a \prec \log b$. Since the mapping $t \mapsto t^2$ is convex, it follows from Theorem 2.1 that the inequality $\sum_{i=1}^n (\log a_i)^2 \leqslant \sum_{i=1}^n (\log b_i)^2$ holds. However, for $a \prec b$, we can apply Corollary 2.7 to find $e_k(a) \geqslant e_k(b)$, hence Conjecture 1.2 implies $\sum_{i=1}^n (\log a_i)^2 \geqslant \sum_{i=1}^n (\log b_i)^2$ and thus $\sum_{i=1}^n (\log a_i)^2 = \sum_{i=1}^n (\log b_i)^2$. This is not surprising: we already know (see Remark 2.8) that

$$a \prec b \quad \text{and} \quad e_n(a) = e_n(b) \quad \Rightarrow \quad a^{\downarrow} = b^{\downarrow}.$$

Thus the vectors a and $b \in \mathbb{R}_+^n$ are equal up to permutations.

4. Related inequalities

PROPOSITION 4.1. *Let* $x, y \in \mathbb{R}_+^n$ *and* $m \in \mathbb{R}_+$. *Assume additionally* $\log x \prec_w \log y$, *then*

$$\sum_{i=1}^n \mathrm{e}^{m x_i} \leqslant \sum_{i=1}^n \mathrm{e}^{m y_i}. \tag{4.1}$$

Proof. We set $\varphi : \mathbb{R}^n \to R$ with $\varphi(x) = \sum_{i=1}^n \mathrm{e}^{m x_i}$. With Proposition 2.4 we have $x \prec_w y$. Since $x \mapsto \mathrm{e}^{mx}$ is convex and monotone increasing, it follows directly from Theorem 2.2 that $\varphi(x) \leqslant \varphi(y)$. □

THEOREM 4.2. *If the real numbers* a, b, c, x, y, z *satisfy*

$$a + b + c = x + y + z = 0$$

and

$$a^2 + b^2 + c^2 = x^2 + y^2 + z^2 \neq 0,$$

then

$$e^{xy}+e^{yz}+e^{zx} \;<\; \left(e^{x^2}+e^{y^2}+e^{z^2}\right)\exp\left(-3\sqrt[3]{\frac{1}{4}a^2b^2c^2}\right) \tag{4.2}$$

and

$$e^{ab}+e^{bc}+e^{ca} \;<\; \left(e^{a^2}+e^{b^2}+e^{c^2}\right)\exp\left(-3\sqrt[3]{\frac{1}{4}x^2y^2z^2}\right). \tag{4.3}$$

Proof. For $\alpha,\beta,\gamma\in\mathbb{R}$ we obtain $(\alpha+\beta+\gamma)^2=\alpha^2+\beta^2+\gamma^2+2\alpha\beta+2\beta\gamma+2\gamma\alpha$, therefore under the conditions

$$\begin{aligned} ab+bc+ca &= \frac{1}{2}\left((a+b+c)^2-(a^2+b^2+c^2)\right) \\ &= \frac{1}{2}\left((x+y+z)^2-(x^2+y^2+z^2)\right) = xy+yz+zx \end{aligned}$$

and we can set

$$p \;:=\; ab+bc+ca \;=\; xy+yz+zx\,.$$

Furthermore

$$x^3+px-xyz = x^3+x^2y+xyz+x^2z-xyz = x^2(x+y+z) = x^2\cdot 0 = 0$$

and analogously

$$y^3+py-xyz \;=\; 0 \qquad \text{and} \qquad z^3+pz-xyz \;=\; 0\,.$$

Therefore the cubic equation $X^3+pX-xyz=0$ has exactly the three solutions $X\in\{x,y,z\}$.

Following Cardano's method (see Cardano [2]) the cubic equation $X^3+pX+q=0$ has exactly three real roots, if and only if

$$D := \left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3 < 0\,.$$

Therefore we obtain

$$\left(\frac{-xyz}{2}\right)^2+\left(\frac{p}{3}\right)^3 \;=\; \frac{x^2y^2z^2}{4}+\frac{p^3}{27} \;<\; 0\,,$$

thus $p^3 < -27\frac{x^2y^2z^2}{4}$ and therefore

$$p \;=\; xy+yz+zx \;<\; -3\sqrt[3]{\frac{x^2y^2z^2}{4}}\,. \tag{4.4}$$

Analogously we obtain from $X^3+pX-abc=0$

$$p \;=\; ab+bc+ca \;<\; -3\sqrt[3]{\frac{a^2b^2c^2}{4}}\,. \tag{4.5}$$

Now

$$p \quad = \quad p+x^2-x^2 \quad = \quad xy+yz+zx+x^2-x^2 \quad = \quad yz-x^2\,,$$

thus

$$yz = x^2+p$$

and analogously $xz = y^2+p$ and $xy = z^2+p$. According to (4.5) and the monotonicity of the exponential function we have

$$\begin{aligned} \mathrm{e}^{yz} &< \mathrm{e}^{x^2}\exp\left(-3\sqrt[3]{\frac{a^2b^2c^2}{4}}\right), \\ \mathrm{e}^{zx} &< \mathrm{e}^{y^2}\exp\left(-3\sqrt[3]{\frac{a^2b^2c^2}{4}}\right), \\ \mathrm{e}^{xy} &< \mathrm{e}^{z^2}\exp\left(-3\sqrt[3]{\frac{a^2b^2c^2}{4}}\right) \end{aligned}$$

and summing up we obtain (4.2). The proof of (4.3) proceeds analogously. □

THEOREM 4.3. *Let $I \subseteq \mathbb{R}$ and assume $f_1,\dots,f_n : I \to \mathbb{R}$ with*

$$\sum_{i=1}^n f_i(t) = 0 \quad \textit{and} \quad f_1(t) \leqslant f_2(t) \leqslant \ldots \leqslant f_n(t) \qquad \textit{for all } t \in I$$

and $g : I \to \mathbb{R}$ with

$$g(t) = \sum_{i=1}^n \mathrm{e}^{f_i(t)} \qquad \textit{for all } t \in I\,.$$

i) If $f_1'(t) \leqslant f_2'(t) \leqslant \ldots \leqslant f_n'(t)$ for all $t \in I$, then $g'(t) \geqslant 0$ for all $t \in I$, respectively g is monotone increasing.

ii) If $f_1'(t) \geqslant f_2'(t) \geqslant \ldots \geqslant f_n'(t)$ for all $t \in I$, then $g'(t) \leqslant 0$ for all $t \in I$, respectively g is monotone decreasing.

Proof. The condition $\sum_{i=1}^n f_i(t) = 0$ for all $t \in I$ implies

$$\sum_{i=1}^n f_i'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^n f_i(t)\right) = 0 \qquad \text{for all } t \in I\,.$$

Assume $f_1'(t) \leqslant f_2'(t) \leqslant \ldots \leqslant f_n'(t)$ for all $t \in I$. From the monotonicity of the exponential function and Chebychev's inequality with $a_i := \mathrm{e}^{f_i(t)}$, $b_i := f_i'(t)$, we obtain

$$g'(t) = \sum_{i=1}^n f_i'(t)\cdot \mathrm{e}^{f_i(t)} \overset{(3.1)}{\geqslant} \frac{1}{n}\underbrace{\left(\sum_{i=1}^n f_i'(t)\right)}_{=0}\left(\sum_{i=1}^n \mathrm{e}^{f_i(t)}\right) = 0\,.$$

Now assume instead $f_1'(t) \geqslant f_2'(t) \geqslant \ldots \geqslant f_n'(t)$ for all $t \in I$. From the monotonicity of the exponential function and Chebychev's inequality with $a_i := \mathrm{e}^{f_i(t)}$, $b_{n+k-i} := f_i'(t)$, we obtain

$$g'(t) = \sum_{i=1}^{n} f_i'(t) \cdot \mathrm{e}^{f_i(t)} \overset{(3.1)}{\leqslant} \frac{1}{n} \underbrace{\left(\sum_{i=1}^{n} f_i'(t)\right)}_{=0} \left(\sum_{i=1}^{n} \mathrm{e}^{f_i(t)}\right) = 0. \qquad \square$$

EXAMPLE 4.4. Consider the functions $f_1, f_2, f_3 : \mathbb{R} \to \mathbb{R}$ with

$$f_1(x) = -x^2 + 1, \qquad f_2(x) = x - 1, \qquad f_3(x) = x^2 - x$$

and

$$f_1'(x) = -2x, \qquad f_2'(x) = 1, \qquad f_3'(x) = 2x - 1.$$

Then

$$f_1(x) + f_2(x) + f_3(x) = 0 \quad \text{for all } x \in \mathbb{R}.$$

Additionally

$$\begin{aligned} f_1(x) \leqslant f_2(x) \leqslant f_3(x) \quad &\text{and} \quad f_1'(x) \leqslant f_2'(x) \leqslant f_3'(x) \qquad \text{for all } x \in [1, \infty), \\ f_2(x) \leqslant f_3(x) \leqslant f_1(x) \quad &\text{and} \quad f_1'(x) \geqslant f_2'(x) \geqslant f_3'(x) \qquad \text{for all } x \in \left[\frac{1}{4}, 1\right]. \end{aligned}$$

Now we define $g : \mathbb{R} \to \mathbb{R}$ with

$$g(x) = \mathrm{e}^{-x^2+1} + \mathrm{e}^{-1+x} + \mathrm{e}^{x^2-x} = \mathrm{e}^{f_1(x)} + \mathrm{e}^{f_2(x)} + \mathrm{e}^{f_3(x)},$$

therefore we can conclude with Theorem 4.3: g is monotone increasing on $[1, \infty)$ and monotone decreasing on $[\frac{1}{4}, 1]$.

Now we generalize Theorem 4.3

THEOREM 4.5. *Let $I \subseteq \mathbb{R}$ and $f_1, \ldots, f_n : I \to \mathbb{R}$ with*

$$\sum_{i=1}^{n} f_i(t) = 0, \quad f_1(t) \leqslant f_2(t) \leqslant \ldots \leqslant f_n(t)$$

and

$$f_1'(t) \leqslant f_2'(t) \leqslant \ldots \leqslant f_n'(t) \quad \text{for all } t \in I$$

and assume $g, h : I \to \mathbb{R}$ with h positive and monotone increasing and

$$g_h(t) = \sum_{i=1}^{n} \mathrm{e}^{h(t) f_i(t)} \qquad \text{for all } t \in I.$$

Then $g_h'(t) \geqslant 0$ for all $t \in I$, which implies that g_h is monotone increasing.

Proof. The condition $\sum_{i=1}^{n} f_i(t) = 0$ for all $t \in I$ implies

$$\sum_{i=1}^{n} f_i'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{n} f_i(t)\right) = 0 \qquad \text{for all } t \in I\,.$$

With Chebychev's inequality using $a_i := f_i(t)$ resp. $a_i := f_i'(t)$ and $b_i := \mathrm{e}^{h(t) f_i(t)}$ we conclude

$$\begin{aligned} g_h'(t) &= \sum_{i=1}^{n} \left(h'(t) f_i(t) + h(t) f'(t)\right) \mathrm{e}^{h(t) f_i(t)} \\ &= h'(t)\left(\sum_{i=1}^{n} f_i(t) \cdot \, \mathrm{e}^{h(t) f_i(t)}\right) + h(t)\left(\sum_{i=1}^{n} f_i'(t) \cdot \, \mathrm{e}^{h(t) f_i(t)}\right) \qquad (4.6) \\ &\overset{(3.1)}{\geqslant} h'(t) \cdot \frac{1}{n} \underbrace{\left(\sum_{i=1}^{n} f_i(t)\right)}_{=0} \left(\sum_{i=1}^{n} \mathrm{e}^{h(t) f_i(t)}\right) + h(t) \underbrace{\left(\sum_{i=1}^{n} f_i'(t)\right)}_{=0} \left(\sum_{i=1}^{n} \mathrm{e}^{h(t) f_i(t)}\right) = 0\,. \quad \square \end{aligned}$$

THEOREM 4.6. *Let $I \subseteq \mathbb{R}$ and assume $f_1, \ldots, f_n : I \to \mathbb{R}$ with the properties*

$$\sum_{i=1}^{n} f_i(t) = 0\,, \quad f_1(t) \leqslant f_2(t) \leqslant \ldots \leqslant f_n(t) \quad \text{and} \quad f_1'(t) \leqslant f_2'(t) \leqslant \ldots \leqslant f_n'(t)$$

for all $t \in I$. Additionally $h : D \to \mathbb{R}$ is given monotone increasing and convex. In this regard $D \in \mathbb{R}$ provides $h(f_i(t))$ is well defined for all $i \in \{1, \ldots, n\}$ and all $t \in I$. We define furthermore $H : I \to \mathbb{R}$ with

$$H(t) \; = \; \sum_{i=1}^{n} \mathrm{e}^{h(f_i(t))}\,.$$

Then $H'(t) \geqslant 0$ for all $t \in I$ and H is monotone increasing.

Proof. The condition $\sum_{i=1}^{n} f_i(t) = 0$ for all $t \in I$ implies

$$\sum_{i=1}^{n} f_i'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^{n} f_i(t)\right) = 0 \qquad \text{for all } t \in I\,.$$

Because h is monotone increasing $h'(x) \geqslant 0$ and $h(x) \leqslant h(y)$ for $x \leqslant y$, thus

$$h'(f_i(t)) \;\geqslant\; 0 \quad \text{and} \quad h(f_1(t)) \;\leqslant\; h(f_2(t)) \;\leqslant\; \ldots \;\leqslant\; h(f_n(t))$$

for all $t \in I$. By assumption h is convex and h' is monotone increasing, thus $h'(x) \leqslant h'(y)$ for $x \leqslant y$, therefore

$$h'(f_1(t)) \;\leqslant\; h'(f_2(t)) \;\leqslant\; \ldots \;\leqslant\; h'(f_n(t))$$

for all $t \in I$. If for real numbers a_1, a_2, b_1, b_2 the inequalities $0 < a_1 \leqslant a_2$ and $b_1 \leqslant b_2$ are satisfied, then $a_1 b_1 \leqslant a_2 b_2$. This applied iterated, we obtain

$$h'(f_1(t)) \cdot f_1'(t) \leqslant h'(f_2(t) \cdot f_2'(t) \leqslant \ldots \leqslant h'(f_n(t) \cdot f_n'(t)$$

for all $t \in I$. Finally we can easily show the original statement by applying Chebychev's inequality once with $a_i := h'(f_i(t) f_i'(t)$ and $b_i := \mathrm{e}^{h(f_i(t))}$, twice with $a_i := h'(f_i(t))$ and $b_i := f_i'(t)$. We obtain

$$\begin{aligned} H'(t) &= \sum_{i=1}^{n} (h \circ f_i)'(t)\, \mathrm{e}^{h(f_i(t))} = \sum_{i=1}^{n} h'(f_i(t))\, f_i'(t) \cdot \mathrm{e}^{h(f_i(t))} \\ &\overset{(3.1)}{\geqslant} \frac{1}{n} \left(\sum_{i=1}^{n} h'(f_i(t)) \cdot f_i'(t) \right) \left(\sum_{i=1}^{n} \mathrm{e}^{h(f_i(t))} \right) \\ &\overset{(3.1)}{\geqslant} \frac{1}{n^2} \left(\sum_{i=1}^{n} h'(f_i(t)) \right) \underbrace{\left(\sum_{i=1}^{n} f_i'(t) \right)}_{=0} \left(\sum_{i=1}^{n} \mathrm{e}^{h(f_i(t))} \right) = 0. \qquad \square \end{aligned} \tag{4.7}$$

The following theorem was proved by Bîrsan, Neff and Lankeit in [1]. Using Theorem 4.3 (respectively the generalizations Theorem 4.5 or Theorem 4.6) we can now show an alternative and otherwise very elementary proof:

THEOREM 4.7. *Let $a, b, c, x, y, z \in \mathbb{R}$ with*

$$a \geqslant b \geqslant c \quad \textit{and} \quad x \geqslant y \geqslant z. \tag{4.8}$$

Furthermore

$$a + b + c = x + y + z = 0 \quad \textit{and} \quad a^2 + b^2 + c^2 = x^2 + y^2 + z^2. \tag{4.9}$$

Then

$$\mathrm{e}^a + \mathrm{e}^b + \mathrm{e}^c \leqslant \mathrm{e}^x + \mathrm{e}^y + \mathrm{e}^z \tag{4.10}$$

if and only if $a \leqslant x$.

REMARK 4.8. Using Theorem 4.5 or Theorem 4.6 instead of Theorem 4.3 the following proof even allows to show the stronger statement unter the conditions of Theorem 4.7:

$$\mathrm{e}^{ma} + \mathrm{e}^{mb} + \mathrm{e}^{mc} \leqslant \mathrm{e}^{mx} + \mathrm{e}^{my} + \mathrm{e}^{mz} \qquad \text{if and only if} \qquad a \leqslant x \qquad \text{for all } m \in \mathbb{R}_+ .$$

Proof. Let us first fix a, b, c. Then we define for simplification $r \in \mathbb{R}_+$ with $r^2 := a^2 + b^2 + c^2$. From the conditions (4.8) and (4.9) we may uniquely determine y and z depending on $x \in \mathbb{R}_+$ (With (4.8) $x < 0$ implies $a + b + c < 0 + y + z \leqslant 0 + 0 + 0$; a contradiction to (4.9)). Since $z = -x - y$ we find $y^2 + (-x - y)^2 + x^2 - r^2 = 0$. Let x

and r be given, then we obtain a quadratic equation and we can solve with the quadratic formula to obtain

$$\begin{aligned} y^2+(-x-y)^2+x^2-r^2 &= 2y^2+2xy+2x^2-r^2=0 \\ \Leftrightarrow y^2+xy+x^2-\frac{1}{2}r^2 &= 0 \\ \Leftrightarrow y &= -\frac{1}{2}x\pm\sqrt{\frac{1}{4}x^2-x^2+\frac{1}{2}r^2} = -\frac{1}{2}x\pm\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2}. \end{aligned}$$

Inserting these two solutions into $z=-x-y$, we get

$$z = -\frac{1}{2}x\mp\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2}.$$

From these two possibilities $\pm$ for y, only the positive case, $\mp$ for z only the negative case remains to satisfy (4.8). Both equations have three real solutions, if and only if $-\frac{3}{4}x^2+\frac{1}{2}r^2\geqslant 0$. We must have $x\leqslant\sqrt{\frac{2}{3}r^2}$ for this. Moreover

$$\begin{aligned} x\geqslant y \quad&\Leftrightarrow\quad x\geqslant -\frac{1}{2}x+\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2} \quad\Leftrightarrow\quad \frac{3}{2}x\geqslant\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2} \\ &\Leftrightarrow\quad \frac{9}{4}x^2\geqslant -\frac{3}{4}x^2+\frac{1}{2}r^2 \quad\Leftrightarrow\quad 3x^2\geqslant\frac{1}{2}r^2 \quad\Leftrightarrow\quad x\geqslant\sqrt{\frac{1}{6}r^2}. \end{aligned}$$

With $D_r:=\left[\sqrt{\frac{1}{6}r^2},\sqrt{\frac{2}{3}r^2}\right]$ we obtain the differentiable functions $y,z:D_r\to\mathbb{R}$ and

$$y(x) = -\frac{1}{2}x+\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2} \quad\text{and}\quad z(x) = -\frac{1}{2}x-\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2} \tag{4.11}$$

as the unique solutions $(x,y(x),z(x))$ of (4.8) and (4.9). Because the given a,b,c satisfy these conditions, obviously $a\in D_r$, $b=y(a)$ and $c=y(b)$.

Now we define for $m\in\mathbb{R}_+$ the function $\tilde{g}:D_r\to\mathbb{R}$ with $\tilde{g}(x)=\mathrm{e}^{my(x)}+\mathrm{e}^{mz(x)}$. We know $y(x)\geqslant z(x)$ and

$$y'(x) = \frac{1}{2}-\frac{\frac{3}{4}x}{\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2}} \leqslant \frac{1}{2}+\frac{\frac{3}{4}x}{\sqrt{-\frac{3}{4}x^2+\frac{1}{2}r^2}} = z'(x).$$

Thus, we can conclude with Theorem 4.5 or Theorem 4.6 (in both cases we can set $h(t):=mt$. For $m=1$ we can use directly Theorem 4.3) that $\tilde{g}$ is monotone increasing. Additionally $x\mapsto\mathrm{e}^{mx}$ is monotone increasing, so building the sum $g:\mathbb{R}\to\mathbb{R}$ with $g(x)=\mathrm{e}^{mx}+\mathrm{e}^{my(x)}+\mathrm{e}^{mz(x)}$ is also monotone increasing. Therefore

$$g(x)\geqslant g(a) \quad\text{if and only if}\quad x\geqslant a,$$

which is equivalent to the statement if we set $m=1$. $\square$

5. New Logarithmic inequalities in information theory

First we introduce Jensen's inequality (cf. Mitrinović, Pečarić, [6, eq. (2.1) p.191]) and the log sum inequality (cf. Cover, Thomas [3, 2.7 p.29]). Then we prove the information inequality which is in different notation known under the name Gibbs' inequality. The information inequality (cf. Cover, Thomas [3, 2.6 p.28]) is the most fundamental inequality in information theory. It asserts that the relative entropy between two probability distributions $p,q:\Omega\to[0,1]$, which is defined by

$$D(p\parallel q) := \sum_{x\in\Omega} p(x)\log\frac{p(x)}{q(x)} \tag{5.1}$$

(or, if p and q are probability measures on a finite set $\Omega=\{1,\dots,n\}$, by

$$D(p\parallel q) := \sum_{i=1}^{n} p_i\log\frac{p_i}{q_i}),$$

is nonnegative.

LEMMA 5.1. (Jensen's inequality) *Let $I\subseteq\mathbb{R}$ be an interval, $f:I\to\mathbb{R}$ be a convex function, $\lambda_1,\dots,\lambda_n$ positive numbers with $\lambda_1+\dots+\lambda_n=1$ and $x_1,\dots,x_n\in I$. Then*

$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right) \leqslant \sum_{i=1}^{n}\lambda_i f(x_i). \tag{5.2}$$

Note: With $n=2$ we have directly the definition of convexity of f.

With Jensen's inequality we can prove the so called log sum inequality:

LEMMA 5.2. (stronger log sum inequality) *Let $a_1,\dots,a_n,\ b_1,\dots,b_n\in\mathbb{R}_+$ and $k\geqslant 0$, then*

$$\sum_{i=1}^{n} a_i\log\left(\frac{a_i}{b_i}+k\right) \geqslant \left(\sum_{i=1}^{n} a_i\right)\log\left(\frac{1}{\sum_{i=1}^{n} b_i}\sum_{i=1}^{n} a_i+k\right). \tag{5.3}$$

We have equality if and only if $\frac{a_1}{b_1}=\frac{a_2}{b_2}=\dots=\frac{a_n}{b_n}$.

REMARK 5.3. For $k=0$ we get the log sum inequality:

$$\sum_{i=1}^{n} a_i\log\frac{a_i}{b_i} \geqslant \left(\sum_{i=1}^{n} a_i\right)\log\left(\frac{1}{\sum_{i=1}^{n} b_i}\sum_{i=1}^{n} a_i\right). \tag{5.4}$$

Proof. With $f(x) = x\log(x+k)$, $\lambda_i := b_i/\sum_{i=1}^n b_i$, $x_i := a_i/b_i$ and (5.2) we get

$$\begin{aligned}
&\sum_{i=1}^n a_i \log\left(\frac{a_i}{b_i}+k\right) = \left(\sum_{i=1}^n b_i\right) \frac{\sum_{i=1}^n b_i \frac{a_i}{b_i} \log\left(\frac{a_i}{b_i}+k\right)}{\sum_{i=1}^n b_i} = \sum_{i=1}^n \lambda_i x_i \log(x_i+k) \\
&= \sum_{i=1}^n \lambda_i f(x_i) \overset{(5.2)}{\geqslant} f\left(\sum_{i=1}^n \lambda_i x_i\right) = \left(\sum_{i=1}^n b_i\right)\left(\sum_{i=1}^n \lambda_i x_i\right) \log\left(\sum_{i=1}^n \lambda_i x_i + k\right) \\
&= \left(\sum_{i=1}^n b_i\right)\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} \log\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}+k\right)\right) = \left(\sum_{i=1}^n a_i\right) \log\left(\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}+k\right). \qquad (5.5)
\end{aligned}$$

If $c := \frac{a_1}{b_1} = \frac{a_2}{b_2} = \ldots = \frac{a_n}{b_n}$, we have $x_i = c$ and we get

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) = f\left(\sum_{i=1}^n \lambda_i c\right) = f(c) = \sum_{i=1}^n \lambda_i f(c) = \sum_{i=1}^n \lambda_i f(x_i).$$

If there is an $i \in \{1,\ldots,n-1\}$ with $\frac{a_i}{b_i} \neq \frac{a_{i+1}}{b_{i+1}}$ then we get

$$\sum_{i=1}^n \lambda_i f(x_i) > f\left(\sum_{i=1}^n \lambda_i x_i\right). \quad \square$$

PROPOSITION 5.4. (Gibbs' inequality / Information inequality) *Let P_n denote the set of probability measures on an n-element set, that is $P_n = \{p \in \mathbb{R}_+^n \mid \sum_{i=1}^n p_i = 1\}$. The following four expressions (the first three are named* Gibbs' inequality, *the last* Information inequality*) are equivalent and hold for all $a,b \in P_n$:*

$$\begin{aligned}
&i)\ \sup_{\xi\in P_n}\left\{\prod_{i=1}^n \xi_i^{a_i}\right\} = \prod_{i=1}^n a_i^{a_i}, \qquad ii)\ \inf_{\xi\in P_n}\left\{\sum_{i=1}^n a_i(-\log \xi_i)\right\} = \sum_{i=1}^n a_i(-\log a_i), \\
&iii)\ \sum_{i=1}^n a_i(-\log b_i) \geqslant \sum_{i=1}^n a_i(-\log a_i), \quad iv)\ D(a\parallel b) = \sum_{i=1}^n a_i \log\frac{a_i}{b_i} \geqslant 0. \qquad (5.6)
\end{aligned}$$

Proof. First we prove the equality of the four expressions:

$$\begin{aligned}
&\sup_{b\in P_n}\left\{\prod_{i=1}^n b_i^{a_i}\right\} = \prod_{i=1}^n a_i^{a_i} \quad \Leftrightarrow \quad \inf_{b\in P_n}\left\{-\prod_{i=1}^n b_i^{a_i}\right\} = -\prod_{i=1}^n a_i^{a_i} \qquad (5.7) \\
\Leftrightarrow\quad & -\prod_{i=1}^n b_i^{a_i} \geqslant -\prod_{i=1}^n a_i^{a_i} \\
\Leftrightarrow\quad & -\log\prod_{i=1}^n b_i^{a_i} \geqslant -\log\prod_{i=1}^n a_i^{a_i} \quad \Leftrightarrow \quad -\sum_{i=1}^n \log b_i^{a_i} \geqslant -\sum_{i=1}^n \log a_i^{a_i} \\
\Leftrightarrow\quad & \sum_{i=1}^n a_i(-\log b_i) \geqslant \sum_{i=1}^n a_i(-\log a_i) \quad \Leftrightarrow \quad \sum_{i=1}^n a_i(\log a_i - \log b_i) \geqslant 0 \\
\Leftrightarrow\quad & \sum_{i=1}^n a_i \log\frac{a_i}{b_i} \geqslant 0 \quad \Leftrightarrow \quad D(a\parallel b) \geqslant 0.
\end{aligned}$$

Now let $a,b \in P_n$. With the stronger log sum inequality and because $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$ we get

$$\sum_{i=1}^{n} a_i \log\left(\frac{a_i}{b_i} + k\right) \geqslant \left(\sum_{i=1}^{n} a_i\right) \log\left(\frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^{n} a_i + k\right) = \log(1+k). \tag{5.8}$$

With $k = 0$ we obain inequality case iv) (information inequality). □

REMARK 5.5. Analogously we can denote inequality (5.8) as the *stronger information inequality*.

COROLLARY 5.6. (generalized log sum inequality) *Let* $a,b \in \mathbb{R}_+^n$ *and* $k_1,\ldots,k_n \in [0,\infty)$. *Then*

$$\sum_{i=1}^{n} a_i \log \prod_{s=1}^{m} \left(\frac{a_i}{b_i} + k_s\right) \geqslant \left(\sum_{i=1}^{n} a_i\right) \log \prod_{s=1}^{m} \left(\frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^{n} a_i + k_s\right). \tag{5.9}$$

Proof. With m numbers $k_1,\ldots,k_m \in [0,1)$ we obtain by m times building sums of the stronger log sum inequality (5.3)

$$\sum_{s=1}^{m}\left(\sum_{i=1}^{n} a_i \log\left(\frac{a_i}{b_i} + k_s\right)\right) \geqslant \sum_{s=1}^{m}\left(\sum_{i=1}^{n} a_i \log\left(\frac{1}{\sum_{i=1}^n b_i} \sum_{i=1}^{n} a_i + k_s\right)\right).$$

With use of distributivity and laws of logarithms we directly obtain the desired result. □

COROLLARY 5.7. (generalized information inequality) *Let* $a,b \in P_n$ *and* $k_1,\ldots,k_n \in [0,\infty)$. *Then*

$$\sum_{i=1}^{n} a_i \log \prod_{s=1}^{m} \left(\frac{a_i}{b_i} + k_s\right) \geqslant \log \prod_{s=1}^{m} (1+k_s). \tag{5.10}$$

Proof. We simplify (5.9) under the condition $a,b \in P_n$, therefore $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$. □

REFERENCES

[1] M. BÎRSAN, P. NEFF, AND J. LANKEIT, *Sum of squared logarithms–an inequality relating positive definite matrices and their matrix logarithm*, J. Inequal. Appl. 2013.

[2] G. CARDANO, *Ars Magna*, Petreius, 1545.

[3] T. COVER AND J. THOMAS, *Elements of information theory*, Wiley, 2012.

[4] G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge University Press, 1952.

[5] A. MARSHALL, I. OLKIN, AND B. ARNOLD, *Inequalities: Theory of majorization and its applications: Theory of majorization and its applications*, Springer Series in Statistics, Springer, 2010.

[6] D. S. MITRINOVIĆ, J. E. PEČARIĆ, AND A. M. FINK, *Classical and new inequalities in analysis*, Springer, 1993.

[7] P. NEFF, B. EIDEL, F. OSTERBRINK, AND R. MARTIN, *A Riemannian approach to strain measures in nonlinear elasticity*, Comptes Rendus Mécanique, **342**, 4 (2014), 254–257.
[8] P. NEFF, I. D. GHIBA, AND J. LANKEIT, *The exponentiated Hencky-logarithmic strain energy. Part I: Constitutive issues and rank–one convexity*, J. Elasticity (2014), 1–92.
[9] P. NEFF, J. LANKEIT, I.-D. GHIBA, R. MARTIN, AND D. STEIGMANN, *The exponentiated Hencky-logarithmic strain energy. Part II: Coercivity, planar polyconvexity and existence of minimizers*, Z. Angew. Math. Phys. (2015), 1–23.
[10] P. NEFF, J. LANKEIT, AND A. MADEO, *On Grioli's minimum property and its relation to Cauchy's polar decomposition*, Internat. J. Engrg. Sci. **80** (2014), 209–217.
[11] P. NEFF, Y. NAKATSUKASA, AND A. FISCHLE, *A logarithmic minimization property of the unitary polar factor in the spectral norm and the Frobenius matrix norm*, SIAM J. Matrix Anal. Appl. **35**, 3 (2014), 1132–1154.
[12] W. POMPE AND P. NEFF, *On the generalized sum of squared logarithms inequality*, J. Inequal. Appl. 2015 (1), 1–17.
[13] M. TOMIC, *Théoreme de Gauss relatif au centre de gravité et son application*, Bull. Soc. Math. Phys. Serbie **1** (1949), 31–40.
[14] H. WEYL, *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Natl. Acad. Sci. USA **35**, 7 (1949), 408.

(Received November 5, 2014)

Fozi M. Dannan
Arab International University
Syria
e-mail: fmdan@scs-net.org

Patrizio Neff
Head of Lehrstuhl für Nichtlineare Analysis und Modellierung
Fakultät für Mathematik, Universität Duisburg-Essen
Thea-Leymann Str. 9, 45127 Essen, Germany
e-mail: patrizio.neff@uni-due.de

Christian Thiel
Lehrstuhl für Nichtlineare Analysis und Modellierung
Fakultät für Mathematik, Universität Duisburg-Essen
Thea-Leymann Str. 9, 45127 Essen, Germany
e-mail: christian.thiel@uni-due.de