INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

Why sums of $\log^2(x)$? What conjecture?

(A comment on a recent preprint by Borisov & al. [1], posted on RG)

Miroslav Šilhavý

Preprint No. 50-2015

PRAHA 2015

Why sums of $\log^2(x)$? What conjecture?

(A comment on a recent preprint by Borisov & al. [1], posted on RG)

Dedicated to memories of Jaroslav Fuka and Vlastimil Pták, who taught me Pick's theory

M. Šilhavý
Institute of Mathematics of the AV ČR
Žitná 25
115 67 Prague 1
Czech Republic

Abstract This note corrects some intentionally misleading statements about my work [4] (posted earlier on RG) in a recent preprint by Borisov, Neff, Sra & Thiel [1] (posted on the same platform). In the danger of a possible confusion of interested readers, and solely because of that, I decided, after much thought, to take the exception and reply in some detail.

Both [4] and [1] deal with the inequality of the form

$$f(x_1)+\ldots+f(x_n) \le f(y_1)+\ldots+f(y_n), \tag{$*$}$$

for a real valued function $f : (0,\infty) \to \mathbb{R}$, conditioned by the inequalities

$$S^k(x) \le S^k(y) \quad \text{for} \quad k = 1,\ldots,n-1, \quad S^n(x) = S^n(y)$$

with elementary symmetric functions S^k, see Section 1 for a precise and full formulation. In Borisov, Neff, Sra & Thiel [1] only the particular case

$$f(x) = \log^2(x), \tag{$**$}$$

is treated and $(*)$ is accordingly called the *sum of squared logarithms inequality* in that case by the authors. In contrast, in [4], Inequality $(*)$ is established for a large class $\mathscr{C}$ of functions f which includes $(**)$ as an easy particular case. That particular case is in no way preferred over the other members of $\mathscr{C}$, and therefore any fuss with fancy concocted names (see above and below) seems inappropriate.

To demonstrate the power of my approach, I give a short proof of $(*)$ in the Borisov, Neff, Sra & Thiel's special case $(**)$ by repeating [4; Example 2.3]. Not only that the proof is much shorter and more straightforward than the later one given by Borisov, Neff, Sra & Thiel [1]; it also indicates the right track to a generalization. It is therefore *conceptual,* a virtue that [1] does not seem to possess. [4; Example 2.3] requires no knowledge beyond the additivity of the logarithm on the real positive semi–axis and the elementary integration of fractions over the same semi–axis.

Key words Elementary symmetric functions · Bernstein's theorem · completely monotone functions · Stieltjes transform · matrix monotone functions · Pick class · Nevanlinna's representation

MSC 2010 26D20 · 30E05 · 74A05

Contents

1 Introduction

In a rather dramatic description, the authors [1] state the *sum of squared logarithms inequality* ("SSLI"), originally in the form of a 'Conjecture,' in widely self–cited extended papers, well documented in [4]. The 'Conjecture' reads, after an equivalent rephrasing,

is it true that for every positive integer n and every pair of n-tuples of positive numbers $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ we have

$$\log^2(x_1)+\ldots+\log^2(x_n) \le \log^2(y_1)+\ldots+\log^2(y_n) \tag{1}$$

whenever the elementary symmetric functions of n variables $S^1, \dots, S^n$ satisfy

$$S^k(x) \le S^k(y) \quad \text{for} \quad k = 1, \dots, n-1, \quad S^n(x) = S^n(y)?$$

Here

$$S^k(x) = \sum_{1 \le i_1 < \cdots < i_k \le n} x_{i_1} \cdots x_{i_n}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad 1 \le k \le n.$$

The authors take the care to refer at length to their proofs in the cases $n = 2, 3, 4$ (rendered obsolete by the subsequent research, that of mine inclusive), then of the publication of the conjecture on *MathOverflow,* and finally of a sketch of a proof for a general n by Lev Borisov on the same platform. In contrast, their citation of my work on the problem, the work which I will describe shortly, is restricted to a single sentence, p. 3.

> "Meanwhile, inspired by this new approach, Miroslav Šilhavý (Czech Academy of Science) has found a solution [33] to classifying all functions W such that $e_1(x) \le e_1(y), \dots e_{n-1}(x) \le e_{n-1}(y)$, $e_n(x) = e_n(y)$ implies $W(x) \le W(y)$."⋆

In fact, my work–

- is by no means "inspired by this new approach," as it is based on a novel idea of inviting into play the machinery of Pick's functions, no trace of which occurs in "this new approach;"
- it involves no "classification" at all; the authors use this misnomer to avoid saying plainly that rather,
- it proves their 'conjecture' in a broadly but directly generalized form, since
- it describes the class $\mathscr{C}$ of all indefinitely differentiable functions $f : (0, \infty) \to \mathbb{R}$ which have the property that for every positive integer n and every pair of n-tuples of positive numbers $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ we have the following generalization of (1):

$$f(x_1)+\ldots+f(x_n) \le f(y_1)+\ldots+f(y_n) \tag{2}$$

whenever

$$S^k(x) \le S^k(y) \quad \text{for} \quad k = 1, \dots, n-1, \quad S^n(x) = S^n(y); \tag{3}$$

⋆ In addition to seeing the processes in my head and thus knowing what inspired me, there is an inacurracy in the authors' description. As a matter of fact, instead of being "inspired by this new approach," I started my work on the problem essentially by a comission by Neff. In a letter dated 29 May 2015, Neff asked me to disseminate his *MathOverflow* announcement of the conjecture among my friends, which I did. He failed to mention that the problem had already been solved by Borisov.

- a simple, easily accessible and often easily verifiable sufficient condition for f to belong to $\mathscr{C}$ is that the function $x \mapsto xf'(x)$, $x > 0$, admits an analytic extension $\varphi : \mathbb{C} \sim (-\infty, 0] \to \mathbb{C}$ with nonnegative imaginary part $\operatorname{Im} \varphi(z)$ for all $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$; (in the class of monotone functions f this condition is actually a *necessary and sufficient* for $f \in \mathscr{C}$);
- the function $f(x) = \log^2(x)$ clearly satisfies that sufficient condition: $xf'(x) = 2 \log x$ indeed has the required analytic extension; thus it satisfies (1);
- *however,* $\log^2(x)$ *is an "ordinary member" of the class* $\mathscr{C}$*, in no way preferred over others*; no mystery connected with its explicit form;
- using the above sufficient/necessary condition, the reader will have no difficulty to verify that, e.g., the following functions $f(x)$ belong to $\mathscr{C}$:

$$x^p, \quad -1 \le p \le 1,$$

  while

$$x^p, \quad p \ \text{ real}, \quad |p| > 1,$$

  do not belong to $\mathscr{C}$. The class $\mathscr{C}$ is large; further references to instances of $\mathscr{C}$ are found in [4; Examples 4.3].

The second, and last, place⋆⋆ where the authors [1] cite [4] is p. 15:

> "More generally, extensions of the SSLI to other matrix monotone functions may be obtained by building on [33]."

- What 'building on?' No need of it: My work contains *a complete and entirely explicit description of all functions from* $\mathscr{C}$, no further work, not to speak about 'building on,' is neither necessary nor possible;
- what 'more generally?' No special cases have been mentioned previously. In addition,
- the relationship with matrix monotone functions (described inaccurately and cryptically by the quoted sentence) is one of the main messages of my work, clearly stated there, and not the "discovery" in the first bullet on p. 15 of Borisov, Neff, Sra & Thiel [1], under which the citation of my work misleadingly appears.

2 The shortest proof of "SSLI," as described in [4]

Based on the ideas [4], which however, are not necessary for the present proof, we start from the following representation

$$\log^2 x = 2 \int_0^\infty \big(\log\left[(x+y)/(1+y)\right] - (\log x)/(1+y^2) \big) \, dy/y, \quad x > 0. \tag{4}$$

This is elementary to verify: the equality plainly holds for $x = 1$ and the differentiation with respect to x followed by a subsequent multiplication by $x/2$ leads to

⋆⋆ Compare with the frequency of citation of their works in my paper (altogether 16 in–text–citations and 7 cited papers of the authors). My absence in [1] occurs in a situation when the reference list in [1] contains a number of plainly irrelevant papers, some of them even not cited in the text (as is more or less frequent in writings of the second author of [1]).

$$\log x = -\int_0^\infty \left((x+y)^{-1} - y/(1+y^2) \right) dy$$

which is verified by an easy integration [2; Example 2, p. 27]. Let now n be a positive integer and $x \in \mathbb{R}^n$ have positive components. Then the repesentation (4) and the additivity of logarithm give

$$\begin{aligned}
\sum_{i=1}^n \log^2(x_i) &= 2\int_0^\infty \Big(\sum_{i=1}^n \log\left[(x_i+y)/(1+y)\right] - \sum_{i=1}^n (\log x_i)/(1+y^2) \Big)\, dy/y \\
&= 2\int_0^\infty \Big(\log \prod_{i=1}^n \left[(x_i+y)/(1+y)\right] - (\log(x_1 \cdots x_n))/(1+y^2) \Big)\, dy/y \\
&= 2\int_0^\infty \Big(\log\Big(\sum_{i=1}^n y^{n-i} S^i(x)/(1+y)^n \Big) - S^n(x)/(1+y^2) \Big)\, dy/y.
\end{aligned}$$

We see that the last expression is an increasing function of $S^1(x), \ldots, S^{n-1}(x)$ when $S^n(x)$ remains constant. Thus the set of inequalities

$$S^k(x) \le S^k(y) \quad \text{for} \quad k = 1, \ldots, n-1, \qquad S^n(x) = S^n(y)$$

implies

$$\log^2(x_1) + \ldots + \log^2(x_n) \le \log^2(y_1) + \ldots + \log^2(y_n). \quad \square$$

(Compare with the proof of the same assertion on pp. 3–11 in Borisov, Neff, Sra & Thiel [1].)

3 The general case

The proof in Section 2 opens the way to a treatment of the general case, the main result of [4].

Theorem.

(i) *If $f : (0, \infty) \to \mathbb{R}$ admits the representation*

$$\begin{aligned} f(x) &= a + bx + c\log x + d/x \\ &\quad + \int_{(0,\infty)} \left(\log\left[(x+y)/(1+y)\right] - \log x/(1+y^2) \right) d\mu(y) \end{aligned} \tag{5}$$

for all $x > 0$, where μ is a nonnegative measure on $(0, \infty)$ with $\int_{(0,\infty)} y/(1+y^2)\, d\mu(y) < \infty$ and $a, c \in \mathbb{R}$ are constants, with $b \ge 0$ and $d \ge 0$, then f belongs to the class $\mathscr{C}$, i.e., satisfies (2) *for any positive integer n and x and y n-tuples satisfying* (3).

(ii) *If f is monotone, then f belongs to $\mathscr{C}$ if and only if it admits the representation* (5).

Remarks.

(i) Treating integrals with respect to a nonnegative measure as a sort of generalized linear combinations with nonnegative coefficients, we see that (5) synthesizes the function f as a combination of the following "elementary members" of $\mathscr{C}$

$$f(x) = \pm 1, \quad f(x) = x, \quad f(x) = \pm\log x, \quad f(x) = 1/x, \tag{6}$$

$$f_y(x) = \log\left[(x+y)/(1+y)\right] - (\log x)/(1+y^2) \tag{7}$$

where $y > 0$ is a parameter. Indeed, using the additivity of the logarithm, it is elementary to verify that each member in (6) and (7) belongs to $\mathscr{C}$. Hence in view of the linearity of the basic inequality (2), so also does the combination of these with nonnegative coefficients, i.e., the function f. (Here we repeat the proof in Section 2). This completes the proof that the representation (5) implies (2).

(ii) The proof of the converse implication is less elementary as it is based on Bernstein's and Widder's theory of completely monotone functions.

(iii) There is a strong geometric idea behind (5). Namely, the set $\mathscr{C}$ is a convex cone: if $f_1, f_2 \in \mathscr{C}$ then $t_1 f_1 + t_2 f_2 \in \mathscr{C}$ for any $t_1 \geq 0$, $t_2 \geq 0$. A broad generalization of Carathéodory's theorem on extremal points of compact convex subsets of $\mathbb{R}^n$, viz., the Krein–Milman's theorem [3; Theorem 3.21], asserts that each point of a compact convex subset K of a "decent" topological vectorspace belongs to the closure of the convex envelope of the set of all extreme points of K. Thus any point of K is the limit of a (generalized) sequence of convex combinations of the extreme points of K. In the limit the finite convex combinations change into an integral with respect to a probability measure. This result is essentially the celebrated integral representation of Choquet. In the version for convex cones (as contrasted to convex compact sets) the restriction that the total mass of the measure be 1 disappears, and one is left with representations of the type (5). Thus one possible proof of the existence of a measure as in (5) (not yet tested) is to show that the elements of the list (6) and (7) are the extreme points of $\mathscr{C}$. (As already mentioned, different devices are used in [4] to prove (5).)

Acknowledgment The research was supported by the RVO: 67985840.

4 References

1 Borisov, L.; Neff, P.; Sra, S.; Thiel, Ch.: *The sum of squared logarithms inequality in arbitrary dimensions* (2015) Available at: `http://www.researchgate.net/publication/281084078`

2 Donoghue, Jr, W. F.: *Monotone matrix functions and analytic continuation* Berlin, Springer (1974)

3 Rudin, W.: *Functional analysis* New York, McGraw-Hill (1973)

4 Šilhavý, M.: *A functional inequality related to analytic continuation* (2015) Preprint No. 37-2015, Institute of Mathematics of the AV ČR Available at: `http://www.math.cas.cz/fichier/preprints/IM2015062310272944.pdf`, `https://www.researchgate.net/publication/281106921`