to appear in Colloq. Math.

REMARKS ON THE COMPARISON OF WEIGHTED QUASI-ARITHMETIC MEANS

GYULA MAKSA AND ZSOLT PÁLES

Abstract. In this paper, we present comparison theorems for the weighted quasi-arithmetic means and for weighted Bajraktarević means without supposing in advance that the weights are the same.

1. Introduction

Throughout this paper $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and I will denote the sets of all positive integers, integers, rational numbers, real numbers, and a nonvoid open subinterval of $\mathbb{R}$, respectively. Let $\varphi : I \to \mathbb{R}$ be a continuous and strictly monotonic function, $2 \le n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in]0,1[$ such that $\sum_{k=1}^n \lambda_k = 1$. The function $M_{\varphi,\lambda}$ defined on I^n by

$$M_{\varphi;\lambda}(x_1, \dots, x_n) = \varphi^{-1}\left(\sum_{k=1}^n \lambda_k \varphi(x_k)\right)$$

is called a weighted quasi-arithmetic mean with generating function φ and weights $\lambda_1, \dots, \lambda_n$. In this note we discuss the following comparison problem for these means: What properties have to be imposed upon the continuous and strictly monotonic functions $\varphi, \psi : I \to \mathbb{R}$ and the weights $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n$ in order that the inequality

$$M_{\varphi;\lambda}(x_1, \dots, x_n) \le M_{\psi;\mu}(x_1, \dots, x_n)$$

be satisfied for all $(x_1, \dots, x_n) \in I^n$? This problem was raised and discussed in the monograph of Hardy–Littlewood–Pólya [**?**], supposing in advance that the weights are the same (see p. 66 in the second edition). However, as it is shown in this paper, this *a priori* condition can be omitted.

In the last section of the paper, we consider the analogous but more general problem for Bajraktarević means. However, to derive the necessity of the conditions, we impose additional regularity properties of the generating functions.

2. Comparison of weighted quasi-arithmetic means

Our first main result is the following comparison theorem.

Theorem 1. *Let $\varphi, \psi : I \to \mathbb{R}$ be continuous and strictly monotonic functions, $2 \le n \in \mathbb{N}$, and let $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in]0,1[$ with $\sum_{k=1}^n \lambda_k = \sum_{k=1}^n \mu_k = 1$. Then the inequality*

$$\varphi^{-1}\left(\sum_{k=1}^n \lambda_k \varphi(x_k)\right) \le \psi^{-1}\left(\sum_{k=1}^n \mu_k \psi(x_k)\right) \tag{1}$$

holds for all $(x_1, \dots, x_n) \in I^n$ if, and only if, $\psi \circ \varphi^{-1}$ is convex (concave) if ψ is increasing (decreasing) and $\lambda_k = \mu_k$ for all $k \in \{1, \dots, n\}$.

Date: October 31, 2009.

2000 *Mathematics Subject Classification.* Primary: 39B62.

Key words and phrases. Weighted quasi-arithmetic mean, weighted Bajraktarević mean, comparison, (λ, μ)-convexity, $\lambda-$Wright convexity.

This research has been supported by the Hungarian Scientific Research Fund (OTKA) Grants K-62316, NK-68040.

Proof. First we prove the necessity. We may (and do) suppose that ψ is strictly increasing and (**??**) holds for all $(x_1, \dots, x_n) \in I^n$.

Reduction to (λ, μ)*-convexity.* Let

$$J := \varphi(I), \qquad f := \psi \circ \varphi^{-1}, \qquad k \in \{1, \dots, n\} \text{ be fixed} \quad \text{and} \quad \lambda := \lambda_k, \quad \mu := \mu_k.$$

Then $\emptyset \neq J \subset \mathbb{R}$ is an open interval, $f : J \to \mathbb{R}$ is a continuous and strictly monotonic function. With the substitutions

$$x_k := \varphi^{-1}(x), \qquad x_\ell := \varphi^{-1}(y) \quad \text{for} \quad \ell \in \{1, \dots, n\} \setminus \{k\}, \qquad \text{where} \quad x, y \in J,$$

inequality (**??**) implies that

$$f(\lambda x + (1-\lambda)y) \le \mu f(x) + (1-\mu) f(y) \qquad (x, y \in J), \tag{2}$$

that is, f is a (λ, μ)-convex function on J (see Kuhn [**?**]).

The proof of the convexity of f*.* Let $p(u) = u^2 + (1-u)^2$, $u \in]0,1[$ and, for $t \in]0,1[$ define

$$a_0(t) := t \qquad \text{and} \qquad a_m(t) := p(a_{m-1}(t)) \quad \text{if} \quad m \in \mathbb{N}.$$

Then

$$p :]0,1[\to \left[\frac{1}{2}, 1\right[, \qquad a_{m+1}(t) \le a_m(t) \quad \text{if} \quad m \in \mathbb{N}, \qquad \text{and} \qquad \lim_{m\to\infty} a_m(t) = \frac{1}{2} \quad \text{for} \quad t \in]0,1[.$$

We show, by induction on m, that

$$f(a_m(\lambda)x + (1-a_m(\lambda))y) \le a_m(\mu) f(x) + (1-a_m(\mu)) f(y) \tag{3}$$

holds for all $x, y \in J$ and $0 \le m \in \mathbb{Z}$. Indeed, it is obvious that (**??**) holds for $m = 0$. Suppose that $0 < m \in \mathbb{Z}$ and (**??**) holds for $m-1$ instead of m, that is,

$$f(a_{m-1}(\lambda)x + (1-a_{m-1}(\lambda))y) \le a_{m-1}(\mu) f(x) + (1-a_{m-1}(\mu)) f(y) \qquad (x, y \in J). \tag{4}$$

Thus, by the definition of p and a_m, and by using (**??**) repeatedly, for all $x, y \in J$, we have that

$$\begin{aligned}
&f(a_m(\lambda)x + (1-a_m(\lambda))y) = f\big(p(a_{m-1}(\lambda))x + (1-p(a_{m-1}(\lambda)))y\big) \\
&\quad = f\big((a_{m-1}(\lambda)^2 + (1-a_{m-1}(\lambda))^2)x + 2a_{m-1}(\lambda)(1-a_{m-1}(\lambda))y\big) \\
&\quad = f\big(a_{m-1}(\lambda)(a_{m-1}(\lambda)x + (1-a_{m-1}(\lambda))y) + (1-a_{m-1}(\lambda))(a_{m-1}(\lambda)y + (1-a_{m-1}(\lambda))x)\big) \\
&\quad \le a_{m-1}(\mu) f\big(a_{m-1}(\lambda)x + (1-a_{m-1}(\lambda))y\big) + (1-a_{m-1}(\mu)) f\big(a_{m-1}(\lambda)y + (1-a_{m-1}(\lambda))x\big) \\
&\quad \le p(a_{m-1}(\mu)) f(x) + (1-p(a_{m-1}(\mu))) f(y) = a_m(\mu) f(x) + (1-a_m(\mu)) f(y)
\end{aligned}$$

which proves (**??**). Since f is continuous and $\lim_{m\to\infty} a_m(t) = \frac{1}{2}$ for all $t \in]0,1[$, thus (**??**) implies that f is Jensen-convex, and hence it is convex (see [**?**]).

The proof of the equality $\lambda = \mu$ $(\lambda_k = \mu_k,\ k \in \{1, \dots, n\})$. Since f is convex, it is absolutely continuous. Therefore, there exists a subset $A \subseteq J$ of full Lebesgue measure such that $f'(y)$ exists for all $y \in A$. Furthermore, f' is integrable on any compact interval and the Newton–Leibniz formula holds. Hence, if f' vanished on A then f would be constant, which is impossible. Thus there exists a point $\xi \in A$ such that $f'(\xi) \neq 0$. Now, define the function F on J by

$$F(x) := f(\lambda x + (1-\lambda)\xi) - \mu f(x) - (1-\mu) f(\xi).$$

Then it follows from (**??**) that $F(x) \le 0$. On the other hand $F(\xi) = 0$. This shows that F has a maximum at ξ. Thus

$$0 = F'(\xi) = \lambda f'(\xi) - \mu f'(\xi) = (\lambda - \mu) f'(\xi),$$

whence we obtain that $\lambda = \mu$. Therefore $\lambda_k = \mu_k$ for all $k \in \{1, \dots, n\}$.

The proof of the sufficiency is straightforward. □

Remark 2. We sketch some other possible ways of the proof of Theorem **??**. Kuhn [**?**], using transfinite tools, proved that if a function $f : J \to \mathbb{R}$ is (λ, λ)-convex for some $0 < \lambda < 1$, i.e.,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \qquad (x, y \in J)$$

then this inequality holds also for all $\lambda \in L \cap [0,1]$, where $L \subset \mathbb{R}$ is the smallest subfield containing λ. Consequently, f is Jensen-convex. The Jensen-convexity was proved in an elementary way by Daróczy-Páles [**?**] using the identity

$$\frac{x+y}{2} = \lambda\left(\lambda\frac{x+y}{2} + (1-\lambda)x\right) + (1-\lambda)\left(\lambda y + (1-\lambda)\frac{x+y}{2}\right).$$

It was observed by Matkowski-Pycia [**?**] that this identity implies the Jensen-convexity of (λ, μ)-convex functions, as well. The same was obtained by Kominek [**?**] using the Rodé' theorem ([**?**]). Thus, by the Bernstein-Doetsch theorem (see [**?**]), the (λ, μ)-convex functions bounded from above on an interval of positive length, are convex.

Another possibility to deduce the Jensen-convexity from the (λ, μ)-convexity is its connection to the λ-Wright-convexity. It is obvious that the (λ, μ)-convex functions are also λ-Wright-convex, i.e.,

$$f(\lambda x + (1-\lambda)y) + f(\lambda y + (1-\lambda)x) \leq f(x) + f(y) \qquad (x, y \in J). \tag{5}$$

(Indeed, interchanging x and y in (**??**) and adding the inequality so obtained to (**??**), we obtain (**??**)). Thus the results of Maksa–Nikodem–Páles [**?**] (the set of λ's for which (**??**) holds, is dense in $[0,1]$) or Kominek [**?**] (f has a limit at a point) or Olbryś [**?**] (f is measurable) or Olbryś [**?**] can be applied to prove that the (λ, μ)-convexity and a "weak" regularity property imply the convexity.

There are other proofs also for $\lambda = \mu$ when the (λ, μ)-convex function f is not constant. Some of them are based on a simplified version of the Rodé theorem ([**?**]) given by Kuhn [**?**]: If $f : J \to \mathbb{R}$ is (λ, μ)-convex then, for all $\tau \in J$, there exists a function $a : J \to \mathbb{R}$ satisfying

$$a(\lambda x + (1-\lambda)y) = \mu a(x) + (1-\mu)a(y), \qquad (x, y \in J) \tag{6}$$

such that $a \leq f$ and $a(\tau) = f(\tau)$. Supposing measurability, Kominek proved the equality $\lambda = \mu$ in [**?**]. Without using Rodé theorem and measurability, Matkowski and Pycia showed in [**?**] that, if λ and μ are not conjugate then every (λ, μ)-convex function is constant, while they are conjugate (i.e., they are both transcendental over $\mathbb{Q}$ or they are both algebraic over $\mathbb{Q}$ and have the same minimal polynomial) then there are non-constant additive solutions $f : \mathbb{R} \to \mathbb{R}$ of the equation

$$f(\lambda x + (1-\lambda)y) = \mu f(x) + (1-\mu)f(y). \qquad (x, y \in \mathbb{R})$$

The latter statement is also a simple consequence of Daróczy's results on (**??**) proved in [**?**].

3. Comparison of weighted Bajraktarević means

A possible generalization of quasi-arithmetic means was introduced in 1963 by Bajraktarević in [**?**], [**?**]: Given a continuous and strictly monotonic function $\varphi : I \to \mathbb{R}$ and a positive function $\omega : I \to \mathbb{R}_+$, $2 \leq n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in]0,1[$ with $\sum_{k=1}^n \lambda_k = 1$, the mean $M_{\varphi,\omega,\lambda}$ is defined on I^n by

$$M_{\varphi,\omega;\lambda}(x_1, \dots, x_n) = \varphi^{-1}\left(\frac{\sum_{k=1}^n \lambda_k \omega(x_k)\varphi(x_k)}{\sum_{k=1}^n \lambda_k \omega(x_k)}\right).$$

In this section we consider the following general comparison problem: What properties have to be imposed upon the continuous and strictly monotonic functions $\varphi, \psi : I \to \mathbb{R}$, the positive function $\omega : I \to \mathbb{R}_+$, and the weights $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n$ in order that the inequality

$$M_{\varphi,\omega;\lambda}(x_1, \dots, x_n) \leq M_{\psi,\omega;\mu}(x_1, \dots, x_n) \tag{7}$$

be valid for all $(x_1, \dots, x_n) \in I^n$?

In the case $\lambda_k = \mu_k$ ($k \in \{1, \dots, n\}$), the necessary and sufficient conditions for inequality (**??**) assuming its validity for all $n \in \mathbb{N}$, were established by Daróczy and Losonczi in [**?**] and [**?**]. Here, we investigate this inequality with fixed number of variables $n \geq 2$ and, *a priori*, we do not assume that the condition $\lambda_k = \mu_k$ ($k \in \{1, \dots, n\}$) holds.

In the particular case when ω is a constant function, the above comparison problem obviously reduces to the problem dealt with in the previous section. Unfortunately, the approach used in that particular case cannot be followed for the setting when ω is nonconstant. Thus, to derive the necessity of the conditions, we need additional regularity assumptions on the functions $\varphi, \psi : I \to \mathbb{R}$ and $\omega : I \to \mathbb{R}_+$. Furthermore, in the proof of the necessity below, we prove first that $\mu_k = \lambda_k$ for all k and the convexity (concavity) of the function $f := \psi \circ \varphi^{-1}$ is deduced in the next step (while, in the proof of Theorem **??**, the convexity was obtained in the first step).

Theorem 3. *Let $\varphi, \psi : I \to \mathbb{R}$ be continuous and strictly monotonic functions, $\omega : I \to \mathbb{R}_+$ be a positive function, $2 \le n \in \mathbb{N}$, and let $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in]0,1[$ with $\sum_{k=1}^n \lambda_k = \sum_{k=1}^n \mu_k = 1$. Assume the following regularity condition: There exists a point $p \in I$ such that ω is continuous at p and the function $f := \psi \circ \varphi^{-1}$ is differentiable at $\varphi(p)$ with $f'(\varphi(p)) \neq 0$. Then the inequality*

$$\varphi^{-1}\left(\frac{\sum_{k=1}^n \lambda_k \omega(x_k)\varphi(x_k)}{\sum_{k=1}^n \lambda_k \omega(x_k)}\right) \le \psi^{-1}\left(\frac{\sum_{k=1}^n \mu_k \omega(x_k)\psi(x_k)}{\sum_{k=1}^n \mu_k \omega(x_k)}\right) \tag{8}$$

holds for all $(x_1, \dots, x_n) \in I^n$ if, and only if, f is convex (concave) if ψ is increasing (decreasing) and $\lambda_k = \mu_k$ for all $k \in \{1, \dots, n\}$.

Proof. To prove the necessity of the conditions, we suppose that ψ is strictly increasing and (**??**) holds for all $(x_1, \dots, x_n) \in I^n$. Let

$$J := \varphi(I), \qquad w := \omega \circ \varphi^{-1}, \qquad k \in \{1, \dots, n\} \text{ be fixed} \quad \text{and} \quad \lambda := \lambda_k, \quad \mu := \mu_k.$$

Then $\emptyset \neq J \subset \mathbb{R}$ is an open interval. With the substitutions

$$x_k := \varphi^{-1}(x), \qquad x_\ell := \varphi^{-1}(y) \quad \text{for} \quad \ell \in \{1, \dots, n\} \setminus \{k\}, \qquad \text{where} \quad x, y \in J,$$

inequality (**??**) implies that

$$f\left(\frac{\lambda w(x)x + (1-\lambda)w(y)y}{\lambda w(x) + (1-\lambda)w(y)}\right) \le \frac{\mu w(x)f(x) + (1-\mu)w(y)f(y)}{\mu w(x) + (1-\mu)w(y)} \qquad (x, y \in J). \tag{9}$$

By our regularity assumption on f and ω, the function $w : J \to \mathbb{R}$ is positive and continuous at $\xi := \varphi(p) \in J$, furthermore, f is differentiable at ξ with $f'(\xi) \neq 0$. Now define the function F on J by

$$F(x) := f\left(\frac{\lambda w(x)x + (1-\lambda)w(\xi)\xi}{\lambda w(x) + (1-\lambda)w(\xi)}\right) - \frac{\mu w(x)f(x) + (1-\mu)w(\xi)f(\xi)}{\mu w(x) + (1-\mu)w(\xi)}.$$

First we prove that F is differentiable at ξ and $F'(\xi) = (\lambda - \mu)f'(\xi)$. We can write F in the form $F = f \circ g - h$, where the functions $g, h : J \to \mathbb{R}$ are defined by

$$g(x) := \frac{\lambda w(x)x + (1-\lambda)w(\xi)\xi}{\lambda w(x) + (1-\lambda)w(\xi)}, \qquad h(x) := \frac{\mu w(x)f(x) + (1-\mu)w(\xi)f(\xi)}{\mu w(x) + (1-\mu)w(\xi)}.$$

Observe that $g(\xi) = \xi$ and hence, by the continuity of w at ξ,

$$\lim_{x\to\xi} \frac{g(x) - g(\xi)}{x - \xi} = \lim_{x\to\xi} \frac{g(x) - \xi}{x - \xi} = \lim_{x\to\xi} \frac{\lambda w(x)}{\lambda w(x) + (1-\lambda)w(\xi)} = \lambda,$$

proving that g is differentiable at ξ and $g'(\xi) = \lambda$.

On the other hand, $h(\xi) = f(\xi)$, thus, by the continuity of w and the differentiability of f at ξ, we obtain

$$\lim_{x\to\xi} \frac{h(x) - h(\xi)}{x - \xi} = \lim_{x\to\xi} \frac{h(x) - f(\xi)}{x - \xi} = \lim_{x\to\xi} \frac{\mu w(x)}{\mu w(x) + (1-\mu)w(\xi)} \frac{f(x) - f(\xi)}{x - \xi} = \mu f'(\xi),$$

which shows that h is differentiable at ξ and $h'(\xi) = \mu f'(\xi)$.

Combining the above properties and using the chain rule, it follows that F is differentiable at ξ and $F'(\xi) = (\lambda - \mu)f'(\xi)$.

In view of inequality (**??**), we have that $F(x) \le 0$ for all $x \in I$. Thus $F(\xi) = 0$ implies that F has a maximum at the point ξ. Therefore,

$$0 = F'(\xi) = (\lambda - \mu)f'(\xi),$$

whence we obtain that $\lambda = \mu$ and hence $\lambda_k = \mu_k$ for all $k \in \{1, \dots, n\}$.

With the notation

$$\alpha(x,y) := \frac{\lambda w(x)}{\lambda w(x) + (1-\lambda) w(y)},$$

inequality (**??**) can be rewritten as

$$f\big(\alpha(x,y)x + (1-\alpha(x,y))y\big) \le \alpha(x,y) f(x) + (1-\alpha(x,y)) f(y) \qquad (x,y \in J).$$

Applying the characterization of convexity obtained in [**?**, Corollary 2.3], it follows that the function f must be convex. Thus, the proof of the necessity of the conditions is completed.

Conversely, assuming that f is convex (concave) if ψ is increasing (decreasing) and $\lambda_k = \mu_k$ for all $k \in \{1, \dots, n\}$, a standard argument shows that inequality (**??**) is valid. ☐

It is an open problem whether the regularity assumption from the theorem above can be removed so that the conclusion remains the same.

References

[1] M. Bajraktarević. Sur une généralisation des moyennes quasilinéaires. *Publ. Inst. Math. (Beograd) (N.S.)*, 3 (17):69–76, 1963.

[2] M. Bajraktarević. Über die Vergleichbarkeit der mit Gewichtsfunktionen gebildeten Mittelwerte. *Studia Sci. Math. Hungar.*, 4:3–8, 1969.

[3] F. Bernstein and G. Doetsch. Zur Theorie der konvexen Funktionen. *Math. Ann.*, 76(4):514–526, 1915.

[4] Z. Daróczy. Notwendige und hinreichende Bedingungen für die Existenz von nichtkonstanten Lösungen linearer Funktionalgleichungen. *Acta Sci. Math. (Szeged)*, 22:31–41, 1961.

[5] Z. Daróczy and L. Losonczi. Über den Vergleich von Mittelwerten. *Publ. Math. Debrecen*, 17:289–297 (1971), 1970.

[6] Z. Daróczy and Zs. Páles. Convexity with given infinite weight sequences. *Stochastica*, 11(1):5–12, 1987.

[7] Z. Daróczy and Zs. Páles. Gauss-composition of means and the solution of the Matkowski–Sutô problem. *Publ. Math. Debrecen*, 61(1-2):157–218, 2002.

[8] Z. Daróczy and Zs. Páles. A characterization of nonconvexity and its applications in the theory of quasi-arithmetic means. In C. Bandle, A. Gilányi, L. Losonczi, M. Plum, and Zs. Páles, editors, *Inequalities and Applications (Noszvaj, 2007)*, volume 157 of *International Series of Numerical Mathematics*, pages 251–260. Birkhäuser Verlag, 2008.

[9] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, Cambridge, 1934. (first edition), 1952 (second edition).

[10] Z. Kominek. On (a,b)-convex functions. *Arch. Math. (Basel)*, 58(1):64–69, 1992.

[11] Z. Kominek. A continuity result on t-Wright convex functions. *Publ. Math. Debrecen*, 63(1-2):213–219, 2003.

[12] M. Kuczma. *An Introduction to the Theory of Functional Equations and Inequalities*, volume 489 of *Prace Naukowe Uniwersytetu Śląskiego w Katowicach.* Państwowe Wydawnictwo Naukowe — Uniwersytet Śląski, Warszawa–Kraków–Katowice, 1985.

[13] N. Kuhn. A note on t-convex functions. In W. Walter, editor, *General Inequalities, 4 (Oberwolfach, 1983)*, volume 71 of *International Series of Numerical Mathematics*, pages 269–276. Birkhäuser, Basel, 1984.

[14] N. Kuhn. On the structure of (s,t)-convex functions. In *General inequalities, 5 (Oberwolfach, 1986)*, volume 80 of *Internat. Schriftenreihe Numer. Math.*, pages 161–174. Birkhäuser, Basel, 1987.

[15] L. Losonczi. Über den Vergleich von Mittelwerten die mit Gewichtsfunktionen gebildet sind. *Publ. Math. Debrecen*, 17:203–208 (1971), 1970.

[16] Gy. Maksa, K. Nikodem, and Zs. Páles. Results on t-Wright convexity. *C. R. Math. Rep. Acad. Sci. Canada*, 13(6):274–278, 1991.

[17] J. Matkowski and M. Pycia. On (α,a)-convex functions. *Arch. Math. (Basel)*, 64(2):132–138, 1995.

[18] A. Olbryś. On the measurability and the Baire property of t-Wright-convex functions. *Aequationes Math.*, 68(1-2):28–37, 2004.

[19] A. Olbryś. Some conditions implying the continuity of t-Wright convex functions. *Publ. Math. Debrecen*, 68(3-4):401–418, 2006.

[20] G. Rodé. Eine abstrakte Version des Satzes von Hahn–Banach. *Arch. Math. (Basel)*, 31:474–481, 1978.

Institute of Mathematics, University of Debrecen, H-4010 Debrecen, Pf. 12, Hungary
E-mail address: {maksa,pales}@math.klte.hu