Mathematical
Inequalities
& Applications
Volume 3, Number 2 (2000), 169–176

NONCONVEX FUNCTIONS AND SEPARATION BY POWER MEANS

ZSOLT PÁLES

To the memory
of my friend and colleague,
Professor György Szabó

(*communicated by Z. Daróczy*)

Abstract. In this note we show that, for a nonconvex function defined on a real interval, there exists a point where this function behaves like a strictly concave function. Due to this result, global convexity can be characterized as pointwise convexity everywhere. As an application, a necessary and sufficient condition for the separability of quasiarithmetic means with power means is obtained.

1. Introduction

Let I be a (proper) real interval throughout this paper. A function $f : I \to \mathbb{R}$ is called *convex* (on the interval I) if

$$f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y) \tag{1}$$

for all $x, y \in I$ and $t \in [0,1]$. If (1) holds with the strict inequality sign "$<$" for $x \neq y$ and $t \in]0,1[$, then f is said to be *strictly convex* (on I). If the above properties are valid for $(-f)$ instead of f, then f is called *concave* and *strictly concave*, respectively. (See the book of Roberts and Varberg [4] for basic facts about convex functions.)

If f is C^2, i.e., it is twice continuously differentiable, then a well know necessary and sufficient condition for its convexity is the nonnegativity of the second derivative f''. Therefore, if f is nonconvex, then, for some interior point p of I, $f''(p) < 0$. Thus, f is strictly concave in a neighbourhood of p. That is, the nonconvexity of a C^2 function always yields its strict concavity in a neighbourhood of a point in I. Trivial examples show, that such a statement cannot be obtained without assuming f to be in C^2.

The aim of this note is to introduce the notion of *pointwise convexity* and show that if an (upper semicontinuous) function is nonconvex on I, then there exists a point where it is strictly concave. It easily follows from this statement that convexity on I is equivalent to being convex at each point of I.

Mathematics subject classification (1991): Primary 26A51, 26B25.

Key words and phrases: Convexity, quasiarithmetic means, power means, separation.

This research has been supported in part by the Hungarian National Research Science Foundation (OTKA) Grant T-016846 and by the High Educational Research and Development Fund (FKFP) Grant 0310/1997.

The precise notion is the following. A function $f: I \to \mathbb{R}$ is called *convex at an interior point* $p \in I$ if there exists a positive δ such that

$$f(p) \leqslant \frac{y-p}{y-x} f(x) + \frac{p-x}{y-x} f(y) \tag{2}$$

for all $x \in]p-\delta, p[$, $y \in]p, p+\delta[$. If (2) holds with "$<$" sign, then f is said to be *strictly convex at* $p \in I$. If, instead, $(-f)$ admits the above properties, then f is called *concave at* p and *strictly concave at* p, respectively.

Clearly, convexity on I always yields convexity at each interior point of I. However, the converse of this statement is not true in general. If $f: \mathbb{R} \to \mathbb{R}$ is defined as $f(0) = 0$ and $f(x) = 1$ for $x \neq 0$, then one can check that f is convex at each point, but it is not convex on $\mathbb{R}$.[1] Thus, some additional conditions on f should be imposed to obtain that pointwise convexity everywhere yields global convexity. It will turn out that the upper semicontinuity of f is sufficient for this aim. This result is proved in the next section and also some conditions that are equivalent to pointwise convexity are offered.

In the last section of this paper, we apply the results on the nonconvexity of functions to the following problem: Given two quasiarithmetic means, find necessary and sufficient conditions such that there exists a power mean separating these means. As a consequence, the known characterization of homogeneous means follows easily.

2. Main results

In our first result, we formulate several properties that are equivalent to pointwise convexity.

THEOREM 1. *Let $I \subset \mathbb{R}$ be a proper interval, p be an interior point of I, and $f: I \to \mathbb{R}$ be an arbitrary function. Then the following properties are equivalent:*

(i) *f is convex at the point p;*

(ii) *There exists a positive δ such that, for $x \in]p-\delta, p[$, $y \in]p, p+\delta[$,*

$$\frac{f(x)-f(p)}{x-p} \leqslant \frac{f(y)-f(p)}{y-p}; \tag{3}$$

(iii) *There exist a positive δ and a constant c such that, for $x \in]p-\delta, p+\delta[$,*

$$f(p) + c(x-p) \leqslant f(x); \tag{4}$$

(iv) *There exists a positive δ such that, for all $n \in \mathbb{N}$, $x_1, \dots, x_n \in]p-\delta, p+\delta[$, $t_1, \dots, t_n \in [0,1]$ with*

$$t_1 + \dots + t_n = 1, \qquad t_1x_1 + \dots + t_nx_n = p,$$

the following inequality is valid

$$f(p) \leqslant t_1 f(x_1) + \dots + t_n f(x_n). \tag{5}$$

[1] This counterexample is due to György Gát.

(v) *There exists a positive* δ *such that, for* $x \in]p-\delta,p[$, $y \in]p,p+\delta[$, $t \in [0,1]$ *with* $tx+(1-t)y = p$, *the following inequality is valid*

$$f(p) \leqslant tf(x) + (1-t)f(y). \tag{6}$$

Proof. The implication $(i) \Longrightarrow (ii)$ is obvious, because, on the domain indicated, (3) is equivalent to (2).

To verify $(ii) \Longrightarrow (iii)$, assume that (ii) is valid. Then, it follows from (3) that

$$c := \sup_{x\in]p-\delta,p[} \frac{f(x)-f(p)}{x-p} \leqslant \inf_{y\in]p,p+\delta[} \frac{f(y)-f(p)}{y-p}.$$

Hence, for $x \in]p-\delta,p[$, $y \in]p,p+\delta[$,

$$\frac{f(x)-f(p)}{x-p} \leqslant c \leqslant \frac{f(y)-f(p)}{y-p}.$$

Rearranging this inequality, we obtain (4).

Now assume that (iii) is valid. Choose $x_1,\ldots,x_n$ and $t_1,\ldots,x_n$ as indicated in (iv). Substituting $x = x_i$ into (4), multiplying the inequality by t_i and adding up the inequalities so obtained, we get (5). Thus, (iii) implies (iv).

Clearly, (v) is the $n = 2$ special case of (iv). Hence, (iv) yields (v).

To complete the proof of the theorem, assume that (v) holds. Let $x \in]p-\delta,p[$ and $y \in]p,p+\delta[$. Then $t = \dfrac{y-p}{y-x}$ satisfies the conditions of (v). Substituting this t into (6), we get (2). Thus we have proved that f is convex at p.

□

REMARK 1. Analogous properties can be formulated for pointwise strict convexity as well.

REMARK 2. It follows from property (ii) of the above theorem, that if f is convex at p, then, for the upper left and lower right Dini derivatives, we have

$$d^-f(p) := \limsup_{x\to 0-0} \frac{f(x)-f(p)}{x-p} \leqslant \liminf_{y\to 0+0} \frac{f(y)-f(p)}{y-p} =: d_+f(p).$$

Conversely, if $d^-f(p) < d_+f(p)$, then it is easy to see that p is strictly convex at p.

The following observation is also useful: A function is (strictly) convex on I, (strictly) convex at an interior point $p \in I$ if and only if, for $a,b \in \mathbb{R}$, the function $g(x) = f(x) + ax + b$ admits the corresponding properties. Therefore, we may subtract affine functions from f whenever it is necessary.

The main result of this section is contained in the following theorem. Its proof makes use of the ideas of the proof the main Lemma in [3] and Gy. Szabó.

THEOREM 2. *Let* $I \subset \mathbb{R}$ *be a proper interval and* $f : I \to \mathbb{R}$ *be an upper semicontinuous function that is nonconvex on* I. *Then there exists an interior point* p *of* I *such that* f *is strictly concave at* p.

Proof. If f is nonconvex on I, then there exist $x_0, y_0 \in I$ and $t_0 \in [0,1]$ such that

$$f(t_0x_0 + (1-t_0)y_0) > t_0 f(x_0) + (1-t_0)f(y_0) \tag{7}$$

Without loss of generality, we may assume that $x_0 < y_0$, $0 < t_0 < 1$. Subtracting an affine function (a function of the form $x \mapsto ax+b$) from f, we may also assume that $f(x_0) = f(y_0) = 0$. Thus, (7) states, that the supremum of f on the interval $[x_0, y_0]$ is positive. By compactness of $[x_0, y_0]$ and upper semicontinuity of f, the set

$$H = \Big\{ u \in [x_0, y_0] \ : \ f(u) = \sup_{v \in [x_0, y_0]} f(v) \Big\}$$

is nonempty and compact. Let $p = \inf H$. Then p belongs to H and it is also an interior point of $[x_0, y_0]$, hence there exists a positive δ such that $]p-\delta, p+\delta[\subset]x_0, y_0[$.

The point p being the smallest element of H, we have

$$f(x) < f(p) \qquad \text{for all} \quad x \in]p-\delta, p[. \tag{8}$$

The point p being in H, we also have

$$f(y) \leqslant f(p) \qquad \text{for all} \quad y \in]p, p+\delta[. \tag{9}$$

Multiplying (8) and (9) by $\dfrac{y-p}{y-x}$ and $\dfrac{p-x}{y-x}$, respectively, and adding these two inequalities, we get

$$\frac{y-p}{y-x} f(x) + \frac{p-x}{y-x} f(y) < f(p)$$

for all $x \in]p-\delta, p[$, $y \in]p, p+\delta[$. Thus, f is strictly concave at p.

The proof of the theorem is complete.

□

As an immediate consequence, we obtain the following result.

COROLLARY 1. *Let $I \subset \mathbb{R}$ be a proper interval and $f : I \to \mathbb{R}$ be an upper semicontinuous function that is convex at each interior point of I. Then f is convex on I.*

Proof. If f were not convex on I, then, by Theorem 2, there would exist an interior point p of I where f would be strictly concave. Being convex at each interior point, we get an obvious contradiction. Hence f has to be a convex function on I.

□

3. Separation of quasiarithmetic means by power means

If $\phi : I \to \mathbb{R}$ is a strictly monotonic continuous function, then the $\mathscr{M}_\phi$-mean of the values $x_1, \dots, x_n \in I$ $(n \in \mathbb{N})$ is defined by

$$\mathscr{M}_\phi(x_1, \dots, x_n) := \phi^{-1}\Big(\frac{\phi(x_1) + \cdots + \phi(x_n)}{n} \Big).$$

The function $\mathscr{M}_\phi$ defined this way is called a *quasiarithmetic mean*. For properties and basic facts about quasiarithmetic means see the book of Hardy, Littlewood and Pólya [1].

For the comparison of quasiarithmetic means, the following result is well known (see [1]).

LEMMA 1. *Let I be a proper interval, $\phi, \psi : I \to \mathbb{R}$ be continuous strictly monotonic functions. Then the following conditions are equivalent:*

(i) *For all $x, y \in I$,*

$$\mathscr{M}_\phi(x,y) \leqslant \mathscr{M}_\psi(x,y);$$

(ii) *For all $n \in \mathbb{N}$, $x_1, \ldots, x_n \in I$,*

$$\mathscr{M}_\phi(x_1,\ldots,x_n) \leqslant \mathscr{M}_\psi(x_1,\ldots,x_n);$$

(iii) *If ϕ is increasing (decreasing), then $\phi \circ \psi^{-1}$ is concave (convex) on $\psi(I)$.*

If any of the equivalent conditions $(i)-(iii)$ is satisfied, then we say that $\mathscr{M}_\phi \leqslant \mathscr{M}_\psi$ on I.

The power means defined below play an essential role in the class of quasiarithmetic means. For $p \in \mathbb{R}$, $x_1, \ldots, x_n > 0$, define

$$M_p(x_1,\ldots,x_n) := \begin{cases} \left(\dfrac{x_1^p + \cdots + x_n^p}{n}\right)^{1/p} & \text{if} \quad p \neq 0, \\ \sqrt[n]{x_1 \cdots x_n} & \text{if} \quad p = 0. \end{cases}$$

Then M_1, M_0, M_{-1} are the arithmetic, geometric and harmonic means, respectively. The means M_p $(p \in \mathbb{R})$ are called *power means*. Taking the function

$$\phi_p(t) := \begin{cases} t^p & \text{if} \quad p \neq 0, \\ \ln t & \text{if} \quad p = 0, \end{cases} \tag{10}$$

one can easily see that $M_p = \mathscr{M}_{\phi_p}$, that is, power means are quasiarithmetic means as well.

For the sake of convenience, we also introduce

$$M_p(x_1,\ldots,x_n) := \begin{cases} \max(x_1,\ldots,x_n) & \text{if} \quad p = \infty, \\ \min(x_1,\ldots,x_n) & \text{if} \quad p = -\infty. \end{cases}$$

The means M_∞ and $M_{-\infty}$ are, however, not quasiarithmetic means, because, for $x, y \in I$ with $x \neq y$, $\min(x,y) < \mathscr{M}_\phi(x,y) < \max(x,y)$ for all generating function ϕ.

The basic facts about power means are summarized in the following lemma (see [1] for the details).

LEMMA 2. *The means M_p ($p \in [-\infty, \infty]$) have the following properties:*

(i) *For all $p \in [-\infty, \infty]$, M_p is homogeneous, that is*

$$M_p(tx_1,\ldots,tx_n) = tM_p(x_1,\ldots,x_n) \qquad \textit{for all } n \in \mathbb{N},\ t, x_1, \ldots, x_n > 0;$$

(*ii*) *If* $-\infty \leqslant p \leqslant q \leqslant \infty$, *then* $M_p \leqslant M_q$ *on* $\mathbb{R}_+$, *that is,*

$$M_p(x_1,\dots,x_n) \leqslant M_q(x_1,\dots,x_n) \qquad \text{for all } n \in \mathbb{N},\ x_1,\dots,x_n > 0;$$

(*iii*) *For all fixed* $n \in \mathbb{N}$, $x_1,\dots,x_n > 0$, *the function* $p \mapsto M_p(x_1,\dots,x_n)$ *is continuous on* $[-\infty,\infty]$.

Now, using the characterization of nonconvexity from the previous section, we deduce a condition for the noncomparability of quasiarithmetic means and power means.

THEOREM 3. *Let* I *be a proper subinterval of the positive reals* $\mathbb{R}_+$, *let* $\phi : I \to \mathbb{R}$ *be continuous strictly monotonic function, and let* p *be a nonzero real number. Then, if* $\mathscr{M}_\phi \not\leqslant M_p$ *on* I, *then there exist an interior point* $u \in I$ *and a positive* δ *such that, for* $0 < \varepsilon < \delta$,

$$\mathscr{M}_\phi\big(u(1-\varepsilon)^{1/p}, u(1+\varepsilon)^{1/p}\big) > u = M_p\big(u(1-\varepsilon)^{1/p}, u(1+\varepsilon)^{1/p}\big). \tag{11}$$

Proof. The mean M_p is a quasiarithmetic mean generated by the function ϕ_p defined in (10). Therefore, by Lemma 1, if $\mathscr{M}_\phi \not\leqslant M_p$, then $\phi \circ \phi_p^{-1}$ is nonconcave (resp. nonconvex) on $\phi_p(I)$ if ϕ is increasing (resp. decreasing). Assume that ϕ is increasing. The proof in the other case is analogous. Then, by Theorem 2 of the previous section, $\phi \circ \phi_p^{-1}$ is strictly convex at an interior point $v \in \phi_p(I)$, that is, there exists a positive η such that

$$\phi(v^{1/p}) < \frac{y-v}{y-x}\phi(x^{1/p}) + \frac{v-x}{y-x}\phi(y^{1/p}) \tag{12}$$

for all $x \in]v-\eta, v[$, $y \in]v, v+\eta[$. Put

$$x := x(\varepsilon) := v(1-\varepsilon) \qquad \text{and} \qquad y := y(\varepsilon) := v(1+\varepsilon).$$

Then there exists a positive δ such that, for $0 < \varepsilon < \delta$, $x = x(\varepsilon)$ and $y = y(\varepsilon)$ are in the domain of (12). Substituting these values into (12), we get, for $0 < \varepsilon < \delta$,

$$\phi(v^{1/p}) < \frac{\phi(v^{1/p}(1-\varepsilon)^{1/p}) + \phi(v^{1/p}(1+\varepsilon)^{1/p})}{2}.$$

Let $u = v^{1/p} \in I$. Then, applying the inverse of ϕ to both sides of the above inequality, we arrive at (11).

□

REMARK 3. It is completely similar to prove that if $M_p \not\leqslant \mathscr{M}_\phi$ on I, then there exist an interior point $u \in I$ and a positive δ such that, for $0 < \varepsilon < \delta$,

$$M_p\big(u(1-\varepsilon)^{1/p}, u(1+\varepsilon)^{1/p}\big) = u > \mathscr{M}_\phi\big(u(1-\varepsilon)^{1/p}, u(1+\varepsilon)^{1/p}\big).$$

The main result of this section is the following theorem which characterizes the situation when two quasiarithmetic means can be separated by a power mean.

THEOREM 4. *Let I and J be proper subintervals of the positive reals $\mathbb{R}_+$ and let $\phi : I \to \mathbb{R}$ and $\psi : J \to \mathbb{R}$ be strictly monotonic continuous functions. Then in order that there exist a real number p such that*

$$\mathscr{M}_\phi \leqslant M_p \quad \text{on } I \qquad \text{and} \qquad M_p \leqslant \mathscr{M}_\psi \quad \text{on } J,$$

it is necessary and sufficient that, for all $x, y \in J$ and $t > 0$ with $tx, ty \in I$, the following inequality be valid:

$$\mathscr{M}_\phi(tx, ty) \leqslant t\mathscr{M}_\psi(x, y). \tag{13}$$

Proof. The necessity of condition (13) follows easily. If p exists such that M_p separates $\mathscr{M}_\phi$ and $\mathscr{M}_\psi$, then, for $x, y \in J$, $t > 0$, with $tx, ty \in I$, we have

$$\mathscr{M}_\phi(tx, ty) \leqslant M_p(tx, ty) = tM_p(x, y) \leqslant t\mathscr{M}_\psi(x, y),$$

which yields (13).

To prove the sufficiency, define the following sets:

$$A := \{p \in [-\infty, \infty] : \mathscr{M}_\phi \leqslant M_p \text{ on } I\}, \ B := \{p \in [-\infty, \infty] : M_p \leqslant \mathscr{M}_\psi \text{ on } J\}.$$

Then, by the properties of power means (listed in Lemma 1 above), A and B are closed intervals of the extended real numbers and

$$-\infty \notin A, \quad \infty \in A, \qquad \text{and} \qquad -\infty \in B, \quad \infty \notin B.$$

Therefore, there exist extended real numbers $a > -\infty$ and $b < \infty$ such that $A = [a, \infty]$, $B = [-\infty, b]$.

To prove the sufficiency of condition (13), we have to show that the intersection of A and B is not empty if (13) is valid. Assume, on the contrary, that $A \cap B = \emptyset$. Then $b < a$, and hence there exists a *nonzero real number* p such that $b < p < a$.

Then $p \notin A$ and $p \notin B$. Hence, by the previous theorem, there exist interior points $u \in I$, $v \in J$ and a positive δ, such that

$$\mathscr{M}_\phi\big(u(1-\varepsilon)^{1/p}, u(1+\varepsilon)^{1/p}\big) > u \tag{14}$$

and

$$\mathscr{M}_\psi\big(v(1-\varepsilon)^{1/p}, v(1+\varepsilon)^{1/p}\big) < v \tag{15}$$

for all $0 < \varepsilon < \delta$.

Now, applying (13) with $t = u/v$, $x = v(1-\varepsilon)^{1/p}$, $y = v(1+\varepsilon)^{1/p}$, (14) and (15) yield (for small positive ε)

$$\begin{aligned} u &< \mathscr{M}_\phi\big(u(1-\varepsilon)^{1/p}, u(1+\varepsilon)^{1/p}\big) \\ &= \mathscr{M}_\phi\Big(\frac{u}{v}v(1-\varepsilon)^{1/p}, \frac{u}{v}v(1+\varepsilon)^{1/p}\Big) \\ &\leqslant \frac{u}{v}\mathscr{M}_\psi\big(v(1-\varepsilon)^{1/p}, v(1+\varepsilon)^{1/p}\big) < \frac{u}{v}v = u. \end{aligned}$$

The contradiction obtained shows that the intersection of A and B cannot be empty.

The proof is complete. □

REMARK 4. Similarly to Lemma 1, the inequality (13) of Theorem 4, is equivalent to the following two conditions

(*i*) For all $n \in \mathbb{N}$, $x_1, \dots, x_n \in J$, $t > 0$ with $tx_1, \dots, tx_n \in I$, the following inequality holds:

$$\mathscr{M}_\phi(tx_1, \dots, x_n) \leqslant t\mathscr{M}_\psi(x_1, \dots, x_n);$$

(*ii*) If ϕ is increasing (decreasing), then, for all $t > 0$, the function $x \mapsto \phi\big(t\psi^{-1}(x)\big)$ is concave (convex) on $\psi(J \cap (I/t))$.

The proof is completely analogous to that of Lemma 1.

As a consequence of the above theorem, we obtain the characterization of homogeneous quasiarithmetic means (cf. [1]).

COROLLARY 2. *Let I be a proper subinterval of $\mathbb{R}_+$ and let $\phi : I \to \mathbb{R}$ be a strictly monotonic continuous function. Then $\mathscr{M}_\phi$ is homogeneous, that is*

$$\mathscr{M}_\phi(tx, ty) = t\mathscr{M}_\phi(x, y) \tag{16}$$

for all $x, y \in I$, $t > 0$ with $tx, ty \in I$ if and only if there exists a real number p such that

$$\mathscr{M}_\phi = M_p \qquad \text{on } I. \tag{17}$$

Proof. Clearly, if (17) holds, then $\mathscr{M}_\phi$ is homogeneous.

Conversely, if (16) is valid on the domain indicated, then condition (13) of the previous theorem is satisfied with $\psi = \phi$ and $J = I$. Therefore, there exists a power p such that $\mathscr{M}_\phi \leqslant M_p \leqslant \mathscr{M}_\phi$ holds on I. This yields (17) immediately. □

REFERENCES

[1] G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge Univ. Press, Cambridge – New York – London, 1934 (first edition), 1952 (second edition).

[2] M. KUCZMA, *An Introduction to the Theory of Functional Equations and Inequalities*, Państwowe Wydawnictwo Naukowe, Warszawa – Kraków – Katowice, 1985.

[3] ZS. PÁLES, *Hölder-type inequalities for quasiarithmetic means*, Acta Math. Hung., **47** (1986), 395-399.

[4] A. W. ROBERTS AND D. E. VARBERG, *Convex Functions*, Academic Press, New York – London, 1973.

(Received May 24, 1999)

Zsolt Páles
Institute of Mathematics and Informatics
Lajos Kossuth University
H-4010 Debrecen, Pf. 12
Hungary
e-mail: pales@math.klte.hu