Mathematical
Inequalities
& Applications
Volume 6, Number 4 (2003), 571–579

CONVEXITY ACCORDING TO MEANS

CONSTANTIN P. NICULESCU

(*communicated by J. Mahwin*)

Abstract. Given a function $f : I \to J$ and a pair of means M and N, on the intervals I and J respectively, we say that f is MN-convex provided that $f(M(x,y)) \leqslant N(f(x), f(y))$ for every x, $y \in I$. In this context, we prove the validity of all basic inequalities in Convex Function Theory, such as Jensen's Inequality and the Hermite-Hadamard Inequality.

1. Introduction

At the core of the notion of convexity is the comparison of means. By a *mean* (on an interval I) we understand any function $M : I \times I \to I$ which verifies the following two properties:

$$\begin{array}{lll} (M1) & \inf\{s,t\} \leqslant M(s,t) \leqslant \sup\{s,t\} & \text{(Intermediacy)} \\ (M2) & M(s,t) = M(t,s) & \text{(Symmetry)} \end{array}$$

for every pair (s,t) of elements of I. When I is one of the intervals $(0,\infty)$, $[0,\infty)$ or $(-\infty,\infty)$, it is usual to add a third property, precisely

$$(M3) \qquad M(tx,ty) = tM(x,y) \qquad \text{for all } t > 0 \quad \text{(Homogeneity)}$$

but this assumption is not really necessary in our paper. Instead, we shall restrict ourselves to the case of continuous means (i.e., continuous in both arguments).

Several examples of means (of strictly positive variables) are listed below.

Hölder's means (also called *power means*):

$$\begin{aligned} H_p(s,t) &= ((s^p + t^p)/2)^{1/p}, \quad \text{for} \quad p \neq 0 \\ G(s,t) &= H_0(s,t) = \lim_{p \to 0} H_p(s,t) = \sqrt{st}. \end{aligned}$$

Then $A = H_1$ is the *arithmetic mean* and G is the *geometric mean*. The mean H_{-1} is known as the *harmonic mean*.

Lehmer's means:

$$L_p(s,t) = (s^p + t^p)/(s^{p-1} + t^{p-1}).$$

Mathematics subject classification (2000): 26A51, 26D07, 26D15, 60A10.
Key words and phrases: convex function, mean, random variable.
Partially supported by CNCSIS Grand 143/2001.

Paper No. 06-53/8.00$

Note that $L_1 = A$, $L_{1/2} = G$ and $L_0 = H_{-1}$. These are the only means that are both Lehmer means and Hölder means.

Stolarsky's means:

$$S_p(s,t) = [(s^p - t^p)/(ps - pt)]^{1/(p-1)}, \quad p \neq 0, 1.$$

The limiting cases ($p = 0$ and $p = 1$) give the *logarithmic* and *identric* means, respectively. Thus

$$\begin{aligned} S_0(s,t) &= \lim_{p \to 0} S_p(s,t) = \frac{s-t}{\log s - \log t} = L(s,t) \\ S_1(s,t) &= \lim_{p \to 1} S_p(s,t) = \frac{1}{e}\left(\frac{t^t}{s^s}\right)^{1/(t-s)} = I(s,t). \end{aligned}$$

Notice that $S_2 = A$ and $S_{-1} = G$.

Of course, there are many other ways to introduce families of means. For example, from known means we can generate new ones. In this respect we recall here the *Gaussian compound* $M \otimes N$ of two means M and N, an iterative procedure which extends the celebrated *Gaussian arithmetic-geometric* mean iteration,

$$(A \otimes G)(s,t) = (\pi/2) \Big/ \int_0^{\pi/2} \left[s^2 \cos^2\theta + t^2 \sin^2\theta\right]^{-1/2} d\theta.$$

See the paper of B. C. Carlson [8] for details.

A comprehensive account of the entire topics of means is given in [6].

The aim of this paper is to discuss how functions behave under the action of means, by considering a much more general concept of a convex function. Letting M and N be two means defined on the intervals I and J respectively, a function $F : I \to J$ will be called MN-*convex* provided that

$$F(M(\mathscr{F})) \leqslant N(F(\mathscr{F})) \quad \text{for every pair } \mathscr{F} \text{ of elements of } I. \tag{MN}$$

In the next section we shall describe the natural process of continuation of means from pairs of real numbers to random variables, a fact which can be seen as a *nonlinear theory of integration*. As a consequence we shall be able to extend (MN) to random variables and thus to obtain a far reaching generalization of *Jensen's Inequality.*

The idea to extend the theory of convexity in the form (MN) is not new. It goes back to people like J. Hadamard, G. H. Hardy and P. Montel [14], who considered it under certain degrees of generality. See Section 3, which is devoted to the case where M and N is one of the classical means A and G. Among the most recent contributions to the extended theory of convexity we should notice here the papers by J. Matkowski and J. Rätz [13], and D. Borwein, J. Borwein, G. Fee and R. Girgensohn [4].

In section 4 we discuss the connection of our results with the notion of comparative convexity (in the sense of G. Hardy, J. E. Littlewood and G. Pólya [10]).

2. The canonical continuation of a mean

In what follows we shall restrict ourselves to the case of continuous means and continuous convex functions.

Under the presence of continuity the inequality (MN) can be refined to allow weighted combinations: First, the dyadic combinations are defined as

$$\begin{aligned} M(x_1,x_2;1/2,1/2) &= M(x_1,x_2) \\ M(x_1,x_2;3/4,1/4) &= M(M(x_1,x_2),x_1) \\ M(x_1,x_2;1/4,3/4) &= M(M(x_1,x_2),x_2) \end{aligned}$$

and so on. Then, for $\lambda \in [0,1]$, we put

$$M(x_1,x_2;1-\lambda,\lambda) = \lim_{n\to\infty} M(x_1,x_2;1-d_n,d_n);$$

here $(d_n)_n$ is any sequence of dyadic numbers with $\lambda = \lim_{n\to\infty} d_n$.

The weighted combinations $M(x_1,\dots,x_n;\lambda_1,\dots,\lambda_n)$ of length $\geqslant 2$ can be defined in the same manner. For example,

$$M(x_1,x_2,x_3;\lambda_1,\lambda_2,\lambda_3) = M\left(M\left(x_1,x_2;\frac{\lambda_1}{1-\lambda_3},\frac{\lambda_2}{1-\lambda_3}\right),\, x_3;1-\lambda_3,\lambda_3\right).$$

EXAMPLE 1. If I is any interval and A denotes the arithmetic mean then

$$A(x_1,\dots,x_n;\lambda_1,\dots,\lambda_n) = \sum_{k=1}^{n} \lambda_k x_k.$$

If I is any subinterval of $(0,\infty)$ and G denotes the geometric mean then

$$G(x_1,\dots,x_n;\lambda_1,\dots,\lambda_n) = \prod_{k=1}^{n} x_k^{\lambda_k}.$$

We can bring together both examples above (as well as all Hölder's means) by considering the so called *quasi-arithmetic mean,*

$$\mathfrak{M}_\varphi\,(s,t) = \varphi^{-1}\left(\frac{1}{2}\,\varphi(s) + \frac{1}{2}\,\varphi(t)\right)$$

which is associated to a strictly monotone continuous mapping $\varphi : I \to \mathbb{R}$. For it,

$$\mathfrak{M}_\varphi(x_1,\dots,x_n;\lambda_1,\dots,\lambda_n) = \varphi^{-1}\left(\sum_{k=1}^{n} \lambda_k \varphi(x_k)\right).$$

An easy inductive argument leads us to the following result:

LEMMA 1. (The discrete form of Jensen's inequality). *Under the presence of continuity, for every MN-convex function $F : I \to J$,*

$$F(M(x_1,\dots,x_n;\lambda_1,\dots,\lambda_n)) \leqslant N((F(x_1),\dots,F(x_n);\lambda_1,\dots,\lambda_n)) \tag{J}$$

for every $x_1,\dots,x_n \in I$ and every $\lambda_1,\dots,\lambda_n \in [0,1]$ with $\sum_{k=1}^{n} \lambda_k = 1$.

The continuous version of Jensen's Inequality can be derived from the discrete case, by noticing that

$$M(x_1, \dots, x_n; \lambda_1, \dots, \lambda_n)$$

can be thought of as the mean $M(h;\mu)$ of the random variable

$$h : \{1, \dots, n\} \to I, \quad h(i) = x_i \quad \text{for} \quad i = 1, \dots, n$$

with respect to the probability measure

$$\mu : \mathscr{P}(\{1, \dots, n\}) \to [0, 1], \quad \mu(A) = \sum_{i \in A} \lambda_i.$$

If (X, Σ, μ) is an arbitrary probability field, it is still possible to define the mean $M(h;\mu)$ for certain real random variables $h \in L^1_{\mathbb{R}}(\mu)$ with values in I. In fact, letting $(\Sigma_a)_\alpha$ a upward directed net of finite subfields Σ whose union is Σ, the mathematical expectation $\mathscr{E}(F|\Sigma_a)$ of $F \in L^1(\mu)$ with respect to Σ_a gives rise to a positive contractive projection

$$P_\alpha : L^1(\mu) \to L^1(\mu|\Sigma_a), \quad P_\alpha(F) = \mathscr{E}(F|\Sigma_a)$$

and

$$\mathscr{E}(F|\Sigma_a) \to F \text{ in the norm topology of } L^1(\mu);$$

due to the Lebesgue theorem on dominated convergence.

A real random variable $h \in L^1_{\mathbb{R}}(\mu)$ (with values in I) will be called *M-intgrable* provided that the limit

$$M(h;\mu) = \lim_\alpha M(P_\alpha(h); \mu|\Sigma_a)$$

exists whenever $(\Sigma_a)_\alpha$ is a upward directed net of finite subfields Σ whose union is Σ.

For the quasi-arithmetic mean $\mathfrak{M}_\varphi$ (associated to a strictly monotone continuous mapping $\varphi : I \to \mathbb{R}$) and the probability field associated to the restriction of the Lebesgue measure to an interval $[s, t] \subset I$, the construction above yields

$$\mathfrak{M}_\varphi \left(id_{[s,t]}; \frac{1}{t-s}\, dx \right) = \varphi^{-1} \left(\frac{1}{t-s} \int_s^t \varphi(x)\, dx \right)$$

which coincides with the so called *integral φ-mean* of s and t, also denoted $Int_\varphi(s,t)$. As noticed M. E. Mayes [15], the class of all integral means equals the class of all differential means; recall that the *differential ψ-mean* of s and t (associated to a strictly monotone differentiable and convex mapping $\psi : I \to \mathbb{R}$) is given by the formula

$$D_\psi(s,t) = (\psi')^{-1} \left(\frac{\psi(t) - \psi(s)}{t-s} \right).$$

THEOREM 1. (The continuous form of Jensen's inequality). *Suppose that $F : I \to J$ is a continuous MN-convex function and (X, Σ, μ) is a probability field. Then*

$$F(M(h;\mu)) \leqslant N((F \circ h; \mu)) \tag{J}$$

for every $h \in L^1_{\mathbb{R}}(\mu)$ such that h is M-integrable and $F \circ h$ is N-integrable.

Proof. Pass to the limit in Lemma 1. □

COROLLARY 1. (The Hermite-Hadamard inequality). *Suppose that $F : I \to J$ is a continuous function. Then F is MN-convex if, and only if, for every $s < t$ in I and every Radon probability measure μ on $[s, t]$ we have the inequality*

$$F(M(s;t)) \leqslant N((F|[s,t];\mu)). \tag{HH}$$

Proof. The necessity follows from Theorem 1 (applied to $h = id_{[s,t]}$). The sufficiency represents the particular case where $\mu = (\varepsilon_s + \varepsilon_t)/2$. □

It is worth to mention the possibility to extend Theorem 1 beyond the case of probability measures. That can be done under the additional assumption of positive homogeneity (both for the means M and N, and the involved function F) following the model of Lebesgue theory, where formulae such as

$$\int_{\mathbb{R}} f(x)\,dx = \lim_{n\to\infty} \left[2n\left(\frac{1}{2n}\int_{-n}^{n} f(x)\,dx\right)\right]$$

hold.

Given a measurable σ-field (X, Σ, μ), a function $h : X \to \mathbb{R}$ will be called M*-integrable* provided the limit

$$M(h;\mu) = \lim_{n\to\infty} \left[\mu(\Omega_n)\cdot M\left(h|\Omega_n; \frac{\mu|\Sigma\cap\Omega_n}{\mu(\Omega_n)}\right)\right] \tag{Int}$$

exists for every increasing sequence $(\Omega_n)_n$ of elements of Σ with $\cup\Omega_n = X$. Then

$$F(M(h;\mu)) \leqslant N((F\circ h;\mu)) \tag{J'}$$

for every $h \in L^1_{\mathbb{R}}(\mu)$ such that h is M-integrable and $F\circ h$ is N-integrable. Restricting (Int) to a suitable type of increasing sequences $(\Omega_n)_n$ with $\cup\Omega_n = X$, one can introduce a concept of M-integrability in the sense of principal value, a case for which (J ') still works.

3. Convexity associated to A and G

In this section we shall illustrate the concept of MN-convexity in the simplest case i.e., when M, $N \in \{A, G\}$.

Depending on which type of mean, arithmetic (A), or geometric (G), it is given on the domain and the codomain of definition, we can encounter one of the following four classes of functions:

AA − convex functions, the usual convex functions
AG − convex functions
GA − convex functions
GG − convex functions.

Notice that while A makes no restriction about the interval I where it applies (it is so because $x, y \in I, \lambda \in [0,1]$ implies $(1-\lambda)x + \lambda y \in I$), the use of G forces us to restrict to the subintervals J of $(0,\infty)$ in order to assure that

$$x, y \in J,\ \lambda \in [0,1] \Rightarrow x^{1-\lambda}y^{\lambda} \in J.$$

The *AG-convex functions* (usually known as log*-convex functions*) *are* those functions $F : I \to (0,\infty)$ for which

$$x, y \in I \quad \text{and} \quad \lambda \in [0,1] \Rightarrow F((1-\lambda)x + \lambda y) \leqslant F(x)^{1-\lambda}F(y)^{\lambda}, \qquad (AG)$$

i.e., for which $\log F$ is convex.

The class of all *GA-convex* functions is constituted by all functions $F : I \to \mathbb{R}$ (defined on subintervals of $(0,\infty)$) for which

$$x, y \in I \quad \text{and} \quad \lambda \in [0,1] \Rightarrow F(x^{1-\lambda}y^{\lambda}) \leqslant (1-\lambda)F(x) + \lambda f(y). \qquad (GA)$$

In the context of twice differentiable functions $F : I \to \mathbb{R}$, GA-convexity means $x^2F'' + xf' \geqslant 0$, so that all twice differentiable nondecreasing convex functions are also GA-convex.

The *GG-convex functions* (called in [16] *multiplicatively convex functions*) are those functions $F : I \to J$ (acting on subintervals of $(0,\infty)$) such that

$$x, y \in I \quad \text{and} \quad \lambda \in [0,1] \Rightarrow F(x^{1-\lambda}y^{\lambda}) \leqslant F(x)^{1-\lambda}F(y)^{\lambda}. \qquad (GG)$$

Due to the following form of the $AM - GM$ Inequality,

$$a, b \in (0,\infty),\ \lambda \in [0,1] \Rightarrow a^{1-\lambda}b^{\lambda} \leqslant (1-\lambda)a + \lambda b, \qquad (*)$$

every log-convex function is also convex. The most notable example of such a function is Euler's *gamma function,*

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt, \quad x > 0.$$

See H. Bohr and J. Mollerup [3]; their argument is recalled by E. Artin [2]. See also [16].

The study of the class of all multiplicatively convex functions can be easily reduced to that of all convex functions via a suitable change of variable and function [16]:

LEMMA 2. *Suppose that I is a subinterval of $(0,\infty)$ and $F : I \to (0,\infty)$ is a multiplicatively convex function. Then*

$$F = \log \circ F \circ \exp : \log(I) \to \mathbb{R}$$

is a convex function. Conversely, if J is an interval and $F : J \to \mathbb{R}$ is a convex function, then

$$F = \exp \circ F \circ \log : \exp(J) \to (0,\infty)$$

is a multiplicatively convex function.

An alternative characterization of multiplicative convexity of a function F is "$\log F(x)$ is a convex function of $\log x$". See [16], Lemma 2.1. Modulo this characterization, the class of all multiplicatively convex functions was first considered by P. Montel [14], in a beautiful paper discussing the analogues of the notion of convex function in n variables. However, the roots of the research in this area can be traced long time before him. Let us mention two such results here:

HADAMARD'S THREE CIRCLES THEOREM. *Let F be an analytical function in the annulus $a < |z| < b$. Then $\log M(r)$ is a convex function of $\log r$, where*

$$M(r) = \sup_{|z|=r} |F(z)|.$$

G. H. HARDY'S MEAN VALUE THEOREM. *Let F be an analytical function in the annulus $a < |z| < b$ and let $p \in [1, \infty)$. Then $\log M_p(r)$ is a convex function of $\log r$, where*

$$M_p(r) = \left(\frac{1}{2\pi} \int_0^{2\pi} |F(re^{i\theta})|^p \, d\theta \right)^{1/p}.$$

As $\lim_{n \to \infty} M_n(r) = M(r)$, Hardy's aforementioned result implies Hadamard's. As well known, Hadamard's result is instrumental in deriving the celebrating Riesz-Thorin Interpolation Theorem (see [10]).

The concept of multiplicative mean value and the multiplicative analogues of Jensen's Inequality and Hermite-Hadamard Inequality make the object of our paper [17].

4. The connection with relative convexity

The idea to introduce a notion of relative convexity appeared in the celebrated book of G. H. Hardy, J. E. Littlewood and G. Polya [10], p. 75: Suppose that $F, g : I \to J$ are two continuous functions and g is strictly monotone. Then F is said to be *convex with respect to* g (*abbreviated,* $g \triangleleft F$) if $F \circ g^{-1}$ is convex (in the usual sense) on the interval $g(I)$.

EXAMPLE 2. Under appropriate assumptions on the domain and the range of a function F, the following statements hold true:

i) F is convex if, and only if, $id \triangleleft F$;

ii) F is $\log$-convex if, and only if, $id \triangleleft \log F$;

iii) F is GG-convex if, and only if, $\log \triangleleft \log F$;

iv) F is GA-convex if, and only if, $\log \triangleleft F$.

As noticed G. T. Cargo [8], in the context of C^1-differentiable functions, F is convex with respect to an increasing function g if F'/g' is nondecreasing; in the context of C^2-differentiable functions, F is convex with respect to g if, and only if, $F''/F' \leqslant g''/g'$ (provided the two ratios exist).

The connection of relative convexity with the topic of our paper is expressed by the following Hermite-Hadamard type inequality:

PROPOSITION 1. *If $g \lhd F$, then*

$$F(Int_g(\{a,b\})) = F\left(g^{-1}\left(\frac{1}{b-a}\int_a^b g(x)\,dx\right)\right) \leqslant \frac{1}{b-a}\int_a^b F(x)\,dx.$$

for every $a<b$ in the domain of F and g.

COROLLARY 2. (*H. Alzer* [1]). *Supose that F is a strictly increasing continuous function such that $1/F^{-1}$ is convex. Then $1/x \lhd F$. As $I_{1/x}(\{a,b\})$ coincides with the logarithmic mean $L(a,b)$, it follows that*

$$F(L(a,b)) \leqslant \frac{1}{b-a}\int_a^b F(x)\,dx.$$

Many other interesting applications of Proposition 1 can be easily deduced via the paper [9], by N. Elezović and J. Pečarić.

Relative convexity can clarify the meaning of certain cumbersome technical conditions. For example, Theorem 2.1 in [12] can be restated as follows:

THEOREM 2. *Let g be a positive increasing function on $[0,1]$, Φ a positive function of bounded variation on $[0,1]$, and $F:(0,\infty)\to\mathbb{R}$ positive, convex and differentiable such that $(1-x)F'(\lambda(1-x))$ is convex with respect to $\Phi(x)$ for every $\lambda>0$. Then*

$$\frac{\int_0^1 F(g(x))\,d\Phi(x)}{\int_0^1 d\Phi(x)} \leqslant F\left(\int_0^1 g(x)\,dx\right).$$

REFERENCES

[1] H. ALZER, *On an integral inequality*, Math. Rev. Anal. Numer. Th. Approx., **18** (1989), 101–103.

[2] E. ARTIN, *The Gamma Function*, Holt, Rinehart and Winston, New York, 1964. English translation of German original, *Einführung in die Theorie der Gammafunktion*, Teubner, 1931.

[3] H. BOHR AND J. MOLLERUP, *Laerebogi i mathematisk Analyse*, vol. III, pp. 149–164, Kopenhagen, 1922.

[4] D. BORWEIN, J. BORWEIN, G. FEE AND R. GIRGENSOHN, *Refined convexity and special cases of the Blaschke-Santalo inequality*, Math. Inequal. Appl., **4** (2001), 631–638.

[5] J. L. BRENNER AND B. C. CARLSON, *Homogeneous mean values: weights and asymptotics*, J. Math. Anal. Appl., **123** (1987), 265–280.

[6] P. S. BULLEN, D. S. MITRINOVIĆ AND P. M. VASIĆ, *Means and Their Inequalities*, D. Reidel Publishing Company, Dordrecht, 1988.

[7] G. T. CARGO, *Comparable means and generalized convexity*, J. Math. Anal. Appl., **12** (1965), 387–392.

[8] B. C. CARLSON, *Algorithms involving arithmetic and geometric means*, Amer. Math. Monthly, **78** (1971), 496–505.

[9] N. ELEZOVIĆ AND J. PEČARIĆ, *Differential and integral F-means and applications to digamma function*, Math. Inequal. Appl., **3** (2000), 189–196.

[10] G. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, Cambridge Mathematical Library, 2nd ed., 1952, Reprinted 1988.

[11] J. LAMPERTI, *On the isometries of certain function-spaces*, Pacific J. Math., **8** (1958), 459–466.

[12] S. M. MALAMUD, *Some complements to the Jensen and Chebyshev inequality and a problem of W. Walter*, Proc. Amer. Math. Soc., **129** (2001), 2671–2678.

[13] J. MATKOWSKI AND J. RÄTZ, *Convexity with respect to an arbitrary mean,* Int. Ser. Numer. Math., **123** (1997), 231–247.
[14] P. MONTEL, *Sur les functions convexes et les fonctions sousharmoniques,* Journal de Math., (9) **7** (1928), 29–60.
[15] M. E. MAYES, *Functions which parametrize means,* Amer. Math. Monthly, **90** (1983), 677–683.
[16] C. P. NICULESCU, *Convexity according to the geometric mean* , Math. Inequal. Appl., **3** (2000), 155–167.
[17] C. P. NICULESCU, *A multiplicative mean value and its applications.* In: *Inequality Theory and Applications* vol. 1, pp. 243–256, Nova Science Publishers (USA), 2001. (Y. J. Cho, S. S. Dragomir and J. Kim, Editors)
[18] A. W. ROBERTS AND D. E. VARBERG, *Convex functions,* Academic Press, New York and London, 1973.

(Received June 20, 2002)

Constantin P. Niculescu
University of Craiova
Department of Mathematics
Craiova 200585, Romania
e-mail: cniculescu@central.ucv.ro