STUDIA MATHEMATICA 192 (1) (2009)

Characterization of convex functions

by

JACEK TABOR (Kraków) and JÓZEF TABOR (Rzeszów)

Abstract. There are many inequalities which in the class of continuous functions are equivalent to convexity (for example the Jensen inequality and the Hermite–Hadamard inequalities). We show that this is not a coincidence: every nontrivial linear inequality which is valid for all convex functions is valid only for convex functions.

1. Introduction. There are many inequalities valid for convex functions. Probably the most well-known ones are the Jensen inequality

$$f\left(\frac{x+y}{2}\right) \le \frac{f(x)+f(y)}{2} \qquad \text{for } x, y \in \mathbb{R},$$

and the Hermite–Hadamard inequalities

$$f\left(\frac{x+y}{2}\right) \le \frac{1}{y-x}\int_x^y f(z)\,dz \le \frac{f(x)+f(y)}{2} \qquad \text{for } x, y \in \mathbb{R},\ x<y.$$

In fact, in the class of continuous functions, each of the above inequalities is equivalent to convexity (see [NP, Chapter 1]; the same concerns Popoviciu's inequality [NP, Th. 1.18]).

It is usually easy to check whether a given linear inequality holds for all convex functions with domain in $\mathbb{R}$. Namely, it is enough to verify that inequality for the functions $x \mapsto |x-p|$ for all $p \in \mathbb{R}$ (see [NP, comments after Theorem 1.5.7]). As a consequence one can prove an even more widely applicable result, which is an easy corollary of Popoviciu's Theorem [NP, Th. 4.2.7].

POPOVICIU'S THEOREM. *Let ν, μ be finite positive Borel measures on $[a,b]$. Then*

$$\int_a^b f(x)\,d\nu(x) \le \int_a^b f(x)\,d\mu(x)$$

2000 *Mathematics Subject Classification*: 26B25, 39B62.
Key words and phrases: convex function, Jensen inequality, approximation.

DOI: 10.4064/sm192-1-3

for all continuous convex functions $f : [a,b] \to \mathbb{R}$, *if and only if* $\nu([a,b]) = \mu([a,b])$ *and*

$$\int_a^t (t-x)\, d\nu(x) \le \int_a^t d\mu(x),$$

$$\int_t^b (x-t)\, d\nu(x) \le \int_t^b (x-t)\, d\mu(x) \quad \textit{for } t \in [a,b].$$

In this paper we deal with a problem, to some extent, opposite. Namely, we prove that every nontrivial (linear-type) inequality which is valid for all convex functions, gives in fact a characterization of convexity in the class of continuous functions. In particular, as a direct consequence of Theorem 2 below we obtain the following result:

THEOREM. *Let* K *be a compact subset of* $\mathbb{R}^n$ *and let* ν *and* μ *be distinct finite Borel measures on* K. *Assume that*

$$\int_K f(x)\, d\nu(x) \le \int_K f(x)\, d\mu(x)$$

for every continuous convex real-valued function f *such that* $K \subset \mathrm{dom}(f)$ (*where* dom *denotes domain*). *Let* W *be a convex subset of a Banach space* E *and let* $h \in C(W,\mathbb{R})$ *be such that*

$$\int_K h(a(x))\, d\nu(x) \le \int_K h(a(x))\, d\mu(x)$$

for every affine function $a : \mathbb{R}^n \to E$ *such that* $a(K) \subset W$. *Then* h *is convex.*

For more information on convex functions we refer the reader to [Ku, NP, Ro].

2. Approximation. Let K be a compact convex subset of $\mathbb{R}^n$. We denote by $C(K,\mathbb{R})$ the Banach space of all continuous functions from K into $\mathbb{R}$ with the supremum norm. For a Lipschitz function $g \in C(K,\mathbb{R})$, we denote by $\mathrm{lip}(g)$ the smallest Lipschitz constant of g. Let $\mathrm{Aff}(K)$ denote the set of all affine functions from K into K, and $\mathrm{Aff}_\varepsilon(K)$ the subset of $\mathrm{Aff}(K)$ consisting of functions with Lipschitz constant less than or equal to ε. Given a set $B \subset C(K,\mathbb{R})$, we denote by $\mathrm{wedge}(B)$ the smallest wedge containing B, where by wedge we understand a closed convex and positively homogeneous set.

For $f \in C(K,\mathbb{R})$, we say that $f \in C^k$ if there exists an open neighbourhood U of K and $f_U \in C^k(U,\mathbb{R})$ such that $f_U|_K = f$.

Let $\mathrm{Conv}(K) \subset C(K,\mathbb{R})$ be the set of all convex functions on K.

Given $g \in C(K,\mathbb{R})$ and $\varepsilon > 0$ we put

$$\mathrm{Aff}_\varepsilon(g) := \{g \circ a : a \in \mathrm{Aff}_\varepsilon(K)\}.$$

Our aim in this section is to show that an arbitrary function from $C(K,\mathbb{R})$ can be approximated in the supremum norm by a sum (with nonnegative coefficients) of a convex function and affine modifications of a fixed nonconvex one.

THEOREM 1. *Let $K \subset \mathbb{R}^n$ be a compact convex set and let $g : K \to \mathbb{R}$ be a continuous function which is not convex. Let $\varepsilon > 0$. Then*

$$C(K,\mathbb{R}) = \mathrm{wedge}(\mathrm{Conv}(K) \cup \mathrm{Aff}_\varepsilon(g)).$$

We postpone the proof to the end of this section and precede it by a few auxiliary lemmas. We use the following notation. Let η be a mollifier, that is, a nonnegative function from $C^\infty(\mathbb{R}^n,\mathbb{R})$ with support in the unit ball $B(0,1)$ and such that $\int \eta = 1$. For $\delta > 0$ we put $\eta_\delta(x) := \eta(x/\delta)/\delta^n$.

Given a function f defined on a subset of $\mathbb{R}^n$ and $x \in \mathbb{R}^n$, let $T_x f$ the function $T_x f : y \mapsto f(y-x)$. Given $\varepsilon > 0$ we define the homothety at x by

$$H_x^\varepsilon(y) := x + \varepsilon(y-x) \quad \text{for } y \in \mathbb{R}^n.$$

LEMMA 1. *Let W be a convex compact subset of $\mathbb{R}^n$ with nonempty interior, let $r > 0$ and let $g \in C(W + B(0,r),\mathbb{R})$. For $\delta \in (0,r)$ define $g_\delta : W \to \mathbb{R}$ by the formula*

$$g_\delta(x) := \int\limits_{B(0,\delta)} \eta_\delta(y) g(x-y)\, dy \quad \textit{for } x \in W.$$

Then

(i) $g_\delta \in C^\infty(W,\mathbb{R})$;
(ii) $\lim_{\delta\to 0^+} g_\delta = g|_W$ *in* $C(W,\mathbb{R})$;
(iii) $g_\delta \in \mathrm{wedge}\{(T_a g)|_W : a \in B(0,\delta)\}$.

Proof. Since (i) and (ii) are well-known (see for example [Ev, Appendix C.4]), we sketch the proof of (iii). Given $\varepsilon > 0$, by uniform continuity of g on compact sets we find $\delta' > 0$ such that

$$|g(x-y) - g(x-y')| \le \varepsilon \quad \text{whenever } x \in W, y, y' \in B(0,\delta), \|y-y'\| \le \delta'.$$

Decompose $B(0,\delta)$ into a disjoint union of finitely many measurable subsets $\{Y_i\}$ with diameter less than δ'. Choose points $y_i \in Y_i$ arbitrarily and put

$$h := \sum_i \int\limits_{Y_i} \eta_\delta(y)\, dy \cdot (T_{y_i} g)|_W.$$

Clearly, $h \in \mathrm{wedge}\{(T_a g)|_W : a \in B(0,\delta)\}$. We finish the proof by showing that h approximates g_δ in the supremum norm. Indeed, for $x \in W$,

$$\begin{aligned}|g_\delta(x) - h(x)| &= \Big|\sum_i \int\limits_{Y_i} \eta_\delta(y) g(x-y)\,dy - \sum_i \int\limits_{Y_i} \eta_\delta(y) g(x-y_i)\,dy\Big| \\ &\le \sum_i \int\limits_{Y_i} \eta_\delta(y)|g(x-y) - g(x-y_i)|\,dy \\ &\le \varepsilon \sum_i \int\limits_{Y_i} \eta_\delta(y)\,dy = \varepsilon \int\limits_{B(0,\delta)} \eta_\delta(y)\,dy = \varepsilon. \ \blacksquare\end{aligned}$$

Lemma 2. *Let K be a compact convex subset of $\mathbb{R}^n$ with nonempty interior and let $a \in K$ and $r > 0$ be such that $B(a,r) \subset K$. Let $\varepsilon \in (0,1)$ and $g \in C(K,\mathbb{R})$. Set*

$$\overline{g} := g \circ H_a^\varepsilon.$$

Then

$$(T_x\overline{g})|_K \in \mathrm{Aff}_\varepsilon(g) \quad \text{for } x \in B(0,(1-\varepsilon)r/\varepsilon).$$

Proof. By the convexity of K,

$$H_b^\varepsilon(K) \subset K \quad \text{for } b \in K.$$

Pick $x \in B(0,(1-\varepsilon)r/\varepsilon)$. Then $\|\varepsilon x/(1-\varepsilon)\| < r$ and so

$$a + \frac{\varepsilon}{1-\varepsilon}\,x \in K.$$

Hence $g \circ H_{a+\varepsilon x/(1-\varepsilon)}$ is well-defined and the equality

$$T_x\overline{g} = g \circ H^\varepsilon_{a+\varepsilon x/(1-\varepsilon)}$$

completes the proof. ∎

Lemma 3. *Let K be a compact convex subset of $\mathbb{R}^n$ with nonempty interior. Let $\varepsilon \in (0,1)$ and suppose $g \in C(K,\mathbb{R})$ is not convex. Then there exists a C^∞ function $h \in \mathrm{wedge}(\mathrm{Aff}_\varepsilon(g))$ and $a \in \mathrm{int}\,K$ such that the function $K \ni x \mapsto D_a^2 h[x]$ attains a negative value.*

Proof. There exists a point $\overline{a} \in \mathrm{int}\,K$ such that g is not convex on any open convex neighbourhood of $\overline{a}$. Let $r > 0$ be such that $B(\overline{a},r) \subset K$. We define $\overline{g}$ by the formula

$$\overline{g} := g \circ H_{\overline{a}}^\varepsilon.$$

Since $H_{\overline{a}}^\varepsilon(B(\overline{a},r)) \subset B(\overline{a},r)$, $\overline{g}|_K$ is not convex. By a similar reasoning to that in the proof of Lemma 2 one can show that $K + B(0,(1-\varepsilon)r/\varepsilon) \subset \mathrm{dom}(\overline{g})$. In virtue of Lemma 2 we have

$$(T_x\overline{g})|_K \in \mathrm{Aff}_\varepsilon(g) \quad \text{for } x \in B(0,(1-\varepsilon)r/\varepsilon).$$

Making use of Lemma 1(iii) we obtain

$$\overline{g}_\delta|_K \in \mathrm{wedge}(\mathrm{Aff}_\varepsilon(g))$$

for sufficiently small δ.

By Lemma 1(ii) we know that $\lim_{\delta\to 0^+}\|\overline{g}_\delta|_K-\overline{g}|_K\|_{\sup}=0$. Hence $h:=\overline{g}_\delta|_K$ is not convex for some small δ. Since h is a C^∞ function, there exists an $a\in\operatorname{int}K$ such that the mapping $K\ni x\mapsto D_a^2h[x]$ is not nonnegative. ∎

Now we are ready to prove the main result of this section.

Proof of Theorem 1. Without loss of generality we may assume that $\varepsilon<1$. Since every convex set has nonempty interior in the affine space spanned by it, it is enough to consider the case when $\operatorname{int}K\neq\emptyset$. We put

$$\mathcal{F}:=\operatorname{wedge}(\operatorname{Conv}(K)\cup\operatorname{Aff}_\varepsilon(g)).$$

We are going to show that

$$f\circ a\in\mathcal{F}\quad\text{ for } f\in\mathcal{F},\ a\in\operatorname{Aff}(K),\ \operatorname{lip}(a)\le 1. \tag{1}$$

This follows from the fact that the above inclusion trivially holds for $f\in\operatorname{Conv}(K)$ and $f\in\operatorname{Aff}_\varepsilon(g)$, and consequently also for functions belonging to the wedge spanned over $\operatorname{Conv}(K)\cup\operatorname{Aff}_\varepsilon(g)$. Let $\mathcal{G}$ denote the set of functions of the form

$$f+\sum_{i=1}^n\alpha_if_i,$$

where $f\in C(K,\mathbb{R})$ is convex, $f_i\in\operatorname{Aff}_\varepsilon(g)$, $\alpha_i\ge 0$ and $n\in\mathbb{N}$. As one can easily check, a version of (1) holds for $\mathcal{G}$, and consequently also for $\mathcal{F}=\operatorname{cl}\mathcal{G}$.

By Lemma 3 there exists a function $h\in\mathcal{F}$ of class C^∞ and $p\in\operatorname{int}K$ such that the mapping $A:=x\mapsto D_p^2h[x]$ is not nonnegative. Clearly without loss of generality (we can shift the origin of the coordinate system to p) we may assume that $p=0$. Then $0\in\operatorname{int}K$.

Now the function $\overline{h}:K\ni x\mapsto h(x)-h(0)-D_0h[x]$ is also an element of $\mathcal{F}$, as $\overline{h}-h$ is affine. We have

$$\overline{h}(x)=D_0^2h[x]+o(\|x\|^2)=A(x)+o(\|x\|^2)\quad\text{ for } x\in K.$$

For $M\ge 1$ we define the function $h_M:K\to\mathbb{R}$ by the formula

$$h_M(x):=M^2\overline{h}(x/M)\quad\text{ for } x\in K.$$

Since $0\in K$ and K is convex, h_M is well-defined. Since we have $h_M(x)=M^2(\overline{h}\circ H_0^{1/M})(x)$ and $\operatorname{lip}(H_0^{1/M})=1/M$, by (1) we see that $h_M\in\mathcal{F}$. As h_M tends uniformly (as $M\to\infty$) to A, we conclude that $A\in\mathcal{F}$.

Because A attains a negative value, there exists an orthonormal basis $e_1,\dots,e_n$ of $\mathbb{R}^n$ such that $A(e_1)<0$. From now on we change the canonical base to the new one. Then for $\lambda:=A(e_1)$ we have

$$A(x_1,0,\dots,0)=\lambda x_1^2.$$

Consider the map

$$P_k(x_1,\dots,x_n)=(x_k,0,\dots,0)\quad\text{ for } k=1,\dots,n.$$

Since $0 \in \operatorname{int} K$ and K is bounded, there exists $\delta \in (0,1]$ such that

$$\delta \cdot P_k|_K \in \mathrm{Aff}_\varepsilon(K) \quad \text{for } k = 1, \ldots, n.$$

Consequently, by (1), $p_k := A \circ (\delta \cdot P_k|_K) \in \mathcal{F}$. Then

$$p_k(x_1, \ldots, x_n) = \lambda\delta^2 x_k^2 \quad \text{for } (x_1, \ldots, x_n) \in K.$$

Since $\lambda < 0$, we see that the function $P_k : K \in x = (x_1, \ldots, x_n) \mapsto -x_k^2$ is also an element of $\mathcal{F}$. Consequently, the function

$$x = (x_1, \ldots, x_n) \mapsto -(x_1^2 + \cdots + x_n^2) = -\|x\|^2$$

is an element of $\mathcal{F}$.

Now we are ready to show that $C(K, \mathbb{R}) \subset \mathcal{F}$. First consider the case of C^∞ functions. Let f be a C^∞ function on K (that is, a restriction to K of a C^∞ function on the neighbourhood of K). Clearly, there exists $M > 0$ such that the function

$$F : K \ni x \mapsto f(x) + M\|x\|^2$$

is convex, which implies that $F \in \mathcal{F}$. Since the function $K \ni x \mapsto -M\|x\|^2$ is also an element of $\mathcal{F}$, we deduce that

$$K \ni x \mapsto F(x) + (-M \cdot \|x\|^2) = f(x)$$

is also an element of $\mathcal{F}$.

By Lemma 1, the class of C^∞ functions is dense in $C(K, \mathbb{R})$. This completes the proof of Theorem 1. ∎

Remark 1. The assertion of Theorem 1 can be reformulated in the following way. Every continuous function $h : K \to \mathbb{R}$ can be uniformly approximated by functions of the form

$$f + \sum_{i=1}^{n} \alpha_i f_i,$$

where f is a continuous convex function, $f_i \in \mathrm{Aff}_\varepsilon(g)$, $\alpha_i \geq 0$ and $n \in \mathbb{N}$.

3. Convex inequalities. Now we are ready to proceed to our main subject of interest, that is, to inequalities valid for convex functions.

Theorem 2. *Let K be a compact convex subset of $\mathbb{R}^n$. Let ν and μ be distinct finite positive Borel measures in K. Assume that*

$$\int_K f\, d\nu \leq \int_K f\, d\mu$$

for every continuous convex function $f : K \to \mathbb{R}$. Let E be a Banach space. Let $\varepsilon > 0$. Let $W \subset E$ be a convex set and let $h \in C(W, \mathbb{R})$ be such that

$$\int_K (h \circ a)\, d\nu \leq \int_K (h \circ a)\, d\mu \tag{2}$$

for every one-dimensional affine function $a : \mathbb{R}^n \to E$ such that $a(K) \subset W$ and $\mathrm{lip}(a) < \varepsilon$. *Then h is convex.*

Proof. Suppose that h is not convex. Then there exist $w_0, w_1 \in W$ such that h is not convex on the interval $[w_0, w_1]$. We define $\overline{h} : [0,1] \to \mathbb{R}$ by

$$\overline{h}(t) = h(w_0 + t(w_1 - w_0)) \quad \text{for } t \in [0,1].$$

Obviously $\overline{h}$ is a continuous function which is not convex, and therefore we can find $0 < t_1 < t_2 < 1$ such that $\overline{h}|_{[t_1,t_2]}$ is not convex. Let

$$\overline{t} := \inf\{t \in [t_1, t_2] : \overline{h}|_{[t_1,t]} \text{ is not convex}\}.$$

Since $h|_{[t_1,\overline{t}]}$ is convex and $h|_{[t_1,t]}$ is not convex for any $t > \overline{t}$, it follows that $\overline{h}$ is not convex on a neighbourhood of $\overline{t}$.

Now we choose an affine map $i : K \to [0,1]$ with $\overline{t} \in \mathrm{int}(i(K))$ and $\mathrm{lip}(i) \le 1/\|w_1 - w_0\|$ and define

$$a_0(x) := w_0 + i(x)(w_1 - w_0) \quad \text{for } x \in K.$$

Obviously $\mathrm{lip}(a_0) \le 1$. Let

$$\mathcal{G} := \Big\{ f \in C(K,\mathbb{R}) : \int_K f \, d\nu \le \int_K f \, d\mu \Big\}.$$

Obviously $\mathcal{G}$ is a wedge which contains all convex functions.

We put

$$g := h \circ a_0 \in C(K,\mathbb{R}).$$

Clearly g is not convex. Let $a : K \to K$ be an affine function with $\mathrm{lip}(a) \le \varepsilon$. Then

$$g \circ a = h \circ a_0 \circ a.$$

As $a_0 \circ a : K \to W$ is affine with $\mathrm{lip}(a_0 \circ a) \le \varepsilon$, by (2) we see that $g \circ a \in \mathcal{G}$. This means that $\mathrm{Aff}_\varepsilon(g) \subset \mathcal{G}$.

Now Theorem 1 shows that $\mathcal{G} = C(K,\mathbb{R})$, and consequently

$$\int_K f \, d\nu \le \int_K f \, d\mu$$

for every $f \in C(K,\mathbb{R})$. Putting $-f$ in place of f we obtain

$$\int_K f \, d\nu = \int_K f \, d\mu \quad \text{for } f \in C(K,\mathbb{R}),$$

which trivially implies that the measures ν and μ are equal. ∎

As a trivial consequence, the Jensen inequality and the Hermite–Hadamard inequalities in the class of continuous functions imply convexity. We provide the proof for the Hermite inequality in $\mathbb{R}$ (other proofs are similar).

Let δ_a denote the unit atom measure concentrated at a.

Corollary 1. *Let W be a convex subset of a Banach space E and let $g : W \to \mathbb{R}$ be a continuous function such that*

$$g\left(\frac{x+y}{2}\right) \le \int_0^1 g(x+t(y-x))\,dt \qquad \text{for } x, y \in W. \tag{3}$$

Then g is convex.

Proof. Let

$$K = [0,1], \qquad \nu = \delta_{1/2}, \qquad \mu = \lambda_1|_K,$$

where $\lambda_1|_K$ denotes the one-dimensional Lebesgue measure restricted to K. Obviously $\nu \ne \mu$.

By the Hermite inequality for every convex function $f : K = [0,1] \to \mathbb{R}$ we obtain

$$\int_K f\,d\nu = f(1/2) \le \int_0^1 f(s)\,ds = \int_K f\,d\mu.$$

Now, by (3), for every affine function $a : K \to W$ such that $a(K) \subset W$ we have

$$\begin{aligned}\int_K g \circ a\,d\nu = g(a(1/2)) &\le \int_0^1 g(a(0) + t(a(1) - a(0)))\,dt \\ &= \int_0^1 g(a(t))\,dt = \int_K g \circ a\,d\mu.\end{aligned}$$

Consequently, by Theorem 2, g is convex. ∎

In a similar manner one can prove that every continuous t-Wright convex function is convex. Recall that a function $f : V \to \mathbb{R}$, where V is convex, is called *t-Wright convex* (where $t \in (0,1)$) if

$$f(tx + (1-t)y) + f((1-t)x + ty) \le f(x) + f(y) \qquad \text{for all } x, y \in V.$$

Then we take $\nu = \delta_{t_0} + \delta_{1-t_0}$, $\mu = \delta_0 + \delta_1$.

The above mentioned result is well-known [MNP]. We present it to point out that Theorem 2 can be applied as a useful tool in the theory of convex functions.

Remark 2. One can ask if the space of affine transformations in Theorem 2 can be replaced by a smaller one. We show that the space of affine similarities is not sufficient.

Consider the inequality

$$f(0) \le \frac{1}{2\pi} \int_{S(0,1)} f(x)\,dS(x),$$

which is clearly satisfied for all subharmonic (and consequently also convex) functions f on $\mathbb{R}^2$.

Let $g(x_1, x_2) = x_1^2 - x_2^2$. One can easily check that $g \circ a$ satisfies the above inequality (in fact even equality) for every affine similarity a (this is because g is harmonic). However, clearly g is not convex.

PROBLEM 1. Are Theorems 1 and 2 valid in infinite-dimensional Banach spaces?

Acknowledgements. We would like to thank the referee for valuable remarks.

References

[Ev] L. Evans, *Partial Differential Equations*, Amer. Math. Soc., 1998.

[Ku] M. Kuczma, *An Introduction to the Theory of Functional Equations and Inequalities*, PWN, 1985.

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

[NP] C. Niculescu and L.-E. Persson, *Convex Functions and Their Applications*, CMS Books in Math. 23, Springer, 2006.

[Ro] R. Rockafellar, *Convex Analysis*, Princeton Math. Ser. 28, Princeton Univ. Press, 1970.

Jacek Tabor
Institute of Mathematics
Jagiellonian University
Łojasiewicza 6
30-348 Kraków, Poland
E-mail: jacek.tabor@im.uj.edu.pl

Józef Tabor
Institute of Mathematics
University of Rzeszów
Rejtana 16A
35-310 Rzeszów, Poland
E-mail: tabor@univ.rzeszow.pl

Received May 10, 2007
Revised version December 19, 2008 (6156)