Mediterr. J. Math. 13 (2016), 119–126
DOI 10.1007/s00009-014-0480-4
0378-620X/16/010119-8
published online October 16, 2014

Mediterranean Journal of Mathematics

CrossMark

A Simple Proof of the Jensen-Type Inequality of Fink and Jodeit

Marcela V. Mihai and Constantin P. Niculescu

Dedicated to Tudor Zamfirescu on the occasion of his 70th birthday

Abstract. We discuss the extension of Jensen's inequality to the framework of quasiconvex functions. Moreover, it is proved that our results work for a class of signed measures larger than the class of probability measures.

Mathematics Subject Classification. Primary 26A51; Secondary 26D15.

Keywords. Jensen's inequality, absolutely continuous function, convex function, quasiconvex function, signed measure, subdifferential.

From time to time, it is worth looking at old classical results. And surprises are not far away. The aim of this paper is to discuss the case of Jensen's inequality, a basic result in real analysis, known to characterize the convex functions. Trying to understand a result stated without proof by Fink and Jodeit [4], we discovered that Jensen's inequality actually works in a much more general framework related to quasiconvexity and the restriction to probability measures can be relaxed, allowing suitable signed measures. The details are given below.

We start by recalling that a real-valued function f defined on an interval I is called *quasiconvex* if

$$f\left((1-\lambda)x+\lambda y\right) \leq \max\left\{f(x), f(y)\right\}$$

for all $x, y \in I$ and $\lambda \in [0,1]$. Quasiconvexity is equivalent to the fact that all level sets $L_\lambda = \{x \in I : f(x) \leq \lambda\}$ are convex, whenever $\lambda \in \mathbb{R}$. Clearly, every convex function is also quasiconvex, but the converse fails. For example, every monotonic function is quasiconvex. The continuous quasiconvex functions have a nice monotonic behavior, first noticed by Martos [7]:

Lemma 1. *A continuous real-valued function f defined on an interval I is quasiconvex if and only if it is either monotonic or there exists an interior*

The second author was supported by a grant of the Romanian National Authority for Scientific Research, CNCS–UEFISCDI, Project Number PN-II-ID-PCE-2011-3-0257.

Birkhäuser

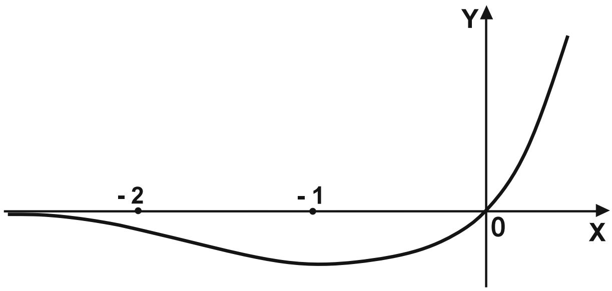

Figure 1. The graph of the function xe^x

point $c \in I$ *such that* f *is nonincreasing on* $(-\infty, c] \cap I$ *and nondecreasing on* $[c, \infty) \cap I$.

For details, see the book of Cambini and Martein [2], Theorem 2.5.2, p. 37. That book also contains a wealth of examples and applications.

Figure 1 shows the graph of the quasiconvex function xe^x. This function is decreasing on the interval $(-\infty, -1]$ and increasing on the interval $[-1, \infty)$. It is concave on $(-\infty, -2]$ and convex on $[-2, \infty)$. An important feature of this function that will be used in this paper is the fact that the tangent line to the graph at any point $x \geq -1$ is a support line.

Jensen's inequality is usually stated in the framework of positive measures of total mass 1 (that is, of probability measures). The main feature of a positive measure is the fact that the integral of a nonnegative function is a nonnegative number. Surprisingly, this property still works for some signed measures when restricted to suitable subcones of the cone of positive integrable functions. A simple example in the discrete case is offered by the following consequence of the Abel summation formula: If $(a_k)_{k=1}^n$ and $(b_k)_{k=1}^n$ are two families of real numbers such that

$$a_1 \geq a_2 \geq \cdots \geq a_n \geq 0 \quad \text{and} \quad \sum_{k=1}^{j} b_k \geq 0 \quad \text{for all } j \in \{1, 2, \ldots, n\},$$

then

$$\sum_{k=1}^{n} a_k b_k \geq 0.$$

Indeed, according to the Abel summation formula, we have

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n-1} \left[\left(\sum_{j=1}^{k} b_j \right) (a_k - a_{k+1}) \right] + \left(\sum_{j=1}^{n} b_j \right) a_n \geq 0.$$

This remark can be easily extended to the framework of Lebesgue integrability.

Lemma 2. *Suppose that $f : [a, b] \to \mathbb{R}$ is a nonnegative absolutely continuous function and $g : [a, b] \to \mathbb{R}$ is an integrable function. Then,*

$$\int_a^b f(x)g(x)\mathrm{d}x \geq 0,$$

in each of the following two cases:

(i) *f is decreasing and $\int_a^x g(t)\mathrm{d}t \geq 0$ for all $x \in [a, b]$; or,*

(ii) *f is increasing and $\int_x^b g(t)\mathrm{d}t \geq 0$ for all $x \in [a, b]$.*

Proof. (i). Indeed,

$$\begin{aligned}\int_a^b f(x)g(x)\mathrm{d}x &= \int_a^b f(x)d\left(\int_a^x g(t)\mathrm{d}t\right)\\ &= \left[f(x)\int_a^x g(t)\mathrm{d}t\right]\Bigg|_{x=a}^{x=b} - \int_a^b f'(x)\left(\int_a^x g(t)\mathrm{d}t\right)\mathrm{d}x\\ &= f(b)\int_a^b g(t)\mathrm{d}t + \int_a^b (-f'(x))\left(\int_a^x g(t)\mathrm{d}t\right)\mathrm{d}x \geq 0,\end{aligned}$$

as a sum of nonnegative numbers. The integration by parts for absolutely continuous functions is motivated by Theorem 18.19, p. 287, in the monograph of Hewitt and Stromberg [5].

The case (ii) follows in a similar manner. □

By combining Lemma 1 and Lemma 2, we arrive at the following result:

Theorem 1. *Suppose that $f : [a, b] \to \mathbb{R}$ is a nonnegative absolutely continuous quasiconvex function and $g : [a, b] \to \mathbb{R}$ is an integrable function such that*

$$\int_a^x g(t)\mathrm{d}t \geq 0 \text{ and } \int_x^b g(t)\mathrm{d}t \geq 0 \quad \text{for every } x \in [a, b].$$

Then,

$$\int_a^b f(x)g(x)\mathrm{d}x \geq 0.$$

Corollary 1. *Under the hypotheses of Theorem 1 for the function g,*

$$\int_a^b f(x)g(x)\mathrm{d}x \geq 0$$

for every nonnegative continuous convex function $f : [a, b] \to \mathbb{R}$.

Proof. The absolute continuity of continuous convex functions $f : [a, b] \to \mathbb{R}$ is proved in [8], Proposition 1.6.1, p. 37.

□

Corollary 2. *Under the hypotheses of Theorem 1,*

$$\int_a^x (x-t)g(t)\mathrm{d}t \geq 0 \text{ and } \int_x^b (t-x)g(t)\mathrm{d}t \geq 0 \quad \text{for every } x \in [a, b].$$

A straightforward computation shows that the hypotheses of Theorem 1 are fulfilled by the function $g(x) = x^2 - \frac{1}{6}$, for $x \in [-1, 1]$. Moreover,

$$\int_{-1}^{1} \left(x^2 - \frac{1}{6} \right) \mathrm{d}x = \frac{1}{3} > 0.$$

This fact makes the measure $\left(x^2 - \frac{1}{6}\right) \mathrm{d}x$ very special in the class of signed measures.

Definition 1. (*Niculescu and Persson* [8]*, p. 179*) A signed Borel measure μ on an interval I is called a *Steffensen–Popoviciu measure* if $\mu(I) > 0$ and

$$\int_I f(x) \mathrm{d}\mu(x) \geq 0$$

for every nonnegative continuous convex function $f : I \to \mathbb{R}$.

According to Corollary 1, an example of such measure on the interval $[-1, 1]$ is $\left(x^2 - \frac{1}{6}\right) \mathrm{d}x$. Using the pushing-forward technique of constructing image measures, one can indicate examples of Steffensen–Popoviciu measures that have an arbitrarily given compact interval as support.

Corollary 2 is related to Lemma 4.1.3 in [8], p. 179 (see also [3]), which shows that a signed Borel measure μ is a Steffensen-Popoviciu measure on an interval $[a, b]$ if and only if $\mu([a, b]) > 0$ and

$$\int_a^x (x - t) \mathrm{d}\mu(t) \geq 0 \quad \text{and} \quad \int_x^b (t - x) g(t) \mathrm{d}t \geq 0 \quad \text{for every } x \in [a, b]. \quad (SP)$$

If $g : [a, b] \to \mathbb{R}$ is an integrable function such that

$$\int_a^b g(x) \mathrm{d}x = 1,$$

we define the *barycenter* of the absolutely continuous measure $g(x)\mathrm{d}x$ as its moment of first order,

$$\beta_{g(x)\mathrm{d}x} = \int_a^b x g(x) \mathrm{d}x.$$

The barycenter belongs to $[a, b]$ when $g(x)\mathrm{d}x$ is a Steffensen–Popoviciu measure. This follows from (SP), by taking $x = a$ and $x = b$.

Alternatively, the barycenter can be characterized as the unique solution $\beta_{g(x)\mathrm{d}x}$ of the following equation involving the class of affine functions on $[a, b]$:

$$A\beta_{g(x)\mathrm{d}x} + B = \int_a^b (Ax + B) g(x) \mathrm{d}x, \text{ for every } A, B \in \mathbb{R}.$$

Theorem 1 easily yields the Jensen-type inequality stated by Fink and Jodeit [4]:

Theorem 2. *Suppose that $g : [a, b] \to \mathbb{R}$ is an integrable function that verifies the hypotheses of Theorem 1 and also the condition $\int_a^b g(x)\mathrm{d}x = 1$. Then,*

$$f\left(\beta_{g(x)\mathrm{d}x}\right) \leq \int_a^b f(x)g(x)\mathrm{d}x,$$

for every continuous convex function $f : [a, b] \to \mathbb{R}$.

Proof. Since f is the uniform limit of a sequence of convex polygonal functions, we may assume that f itself is of this particular type. This assures that the subdifferential $\partial f(x)$, of f at any point $x \in [a, b]$, is nonempty. Let $\lambda \in \partial f(\beta_{g(x)\mathrm{d}x})$. Then,

$$f(x) \geq f\left(\beta_{g(x)\mathrm{d}x}\right) + \lambda\left(x - \beta_{g(x)\mathrm{d}x}\right) \quad \text{for every } x \in [a, b].$$

and taking into account that $g(x)\mathrm{d}x$ is a Steffensen–Popoviciu measure, we conclude that

$$\begin{aligned}\int_a^b f(x)g(x)\mathrm{d}x &\geq \int_a^b \left(f\left(\beta_{g(x)\mathrm{d}x}\right) + \lambda\left(x - \beta_{g(x)\mathrm{d}x}\right)\right) g(x)\mathrm{d}x \\ &= f\left(\beta_{g(x)\mathrm{d}x}\right) + \lambda \int_a^b \left(x - \beta_{g(x)\mathrm{d}x}\right) g(x)\mathrm{d}x = f\left(\beta_{g(x)\mathrm{d}x}\right).\end{aligned}$$

□

An inspection of the argument of Theorem 2 reveals that a similar result works for open intervals. Indeed, in this case, convexity implies continuity and also the nonemptiness of the subdifferential at any point. See [8], Theorem 1.3.3, p. 21, and Lemma 1.5.1, p. 30.

Theorem 3. *Suppose that g is a real-valued integrable function defined on an open interval I such that*

$$\int_I g(x)\mathrm{d}x = 1$$

$$\beta_{g(x)\mathrm{d}x} = \int_I xg(x)\mathrm{d}x \in I$$

$$g(x)\mathrm{d}x \text{ is a Steffensen–Popoviciu measure on } I,$$

Then,

$$f\left(\beta_{g(x)\mathrm{d}x}\right) \leq \int_I f(x)g(x)\mathrm{d}x$$

for every convex function $f : I \to \mathbb{R}$ with the property that $fg \in L^1(I)$.

An example of a function g that fulfils the conditions of Theorem 3 is

$$g(x) = \begin{cases} \lambda e^{-x^2+1} & \text{if } |x| > 1 \\ \frac{6\lambda}{5}\left(x^2 - \frac{1}{6}\right) & \text{if } |x| \leq 1, \end{cases}$$

where the constant $\lambda > 0$ is chosen such that $\int_{\mathbb{R}} g(x)\mathrm{d}x = 1$. Using the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\mathrm{d}t,$$

one can show that the exact value of λ is

$$\lambda = \frac{1}{\frac{2}{5} + e\left(1 - \operatorname{erf}(1)\right)\sqrt{\pi}} = 0.863\,653\,206\ldots .$$

The barycenter of the measure $g(x)\mathrm{d}x$ is the origin. This example can be easily modified to provide examples of functions g on $\mathbb{R}$ for which $g(x)\mathrm{d}x$ has a prescribed barycenter.

Interestingly, the hypothesis concerning the convexity of the function f in Theorem 2 and Theorem 3 can be considerably relaxed using the concept of point of convexity, recently introduced by Niculescu and Rovenţa [9].

Definition 2. Given a real-valued continuous function f defined on an interval I, a point $a \in I$ is called a *point of convexity* of f relative to a neighborhood V of a (called neighborhood of convexity) if

$$f(a) \leq \sum_{k=1}^{n} \lambda_k f(x_k), \tag{J}$$

for every family of points $x_1, \ldots, x_n$ in V and every family of nonnegative weights $\lambda_1, \ldots, \lambda_n$ with $\sum_{k=1}^{n} \lambda_k = 1$ and $\sum_{k=1}^{n} \lambda_k x_k = a$.

Reversing the inequality (J), one obtains the notion of *point of concavity.*

Clearly, a continuous function $f : I \to \mathbb{R}$ is convex if and only if every point of I is a point of convexity relative to the whole domain. A simple sufficient condition for a point a to be a point of convexity is the nonemptiness of the subdifferential of f at a. Indeed, the condition $\lambda \in \partial f(a)$ means the existence of an affine function of the form $L(x) = f(a) + \lambda(x - a)$ such that

$$f(x) \geq L(x)$$

for all x in the domain of f. In this case,

$$f(a) = L(a) = L\left(\sum_{k=1}^{n} \lambda_k x_k\right) = \sum_{k=1}^{n} \lambda_k L(x_k) \leq \sum_{k=1}^{n} \lambda_k f(x_k),$$

for every family of points $x_1, \ldots, x_n$ in I and every family of nonnegative weights $\lambda_1, \ldots, \lambda_n$ with $\sum_{k=1}^{n} \lambda_k = 1$ and $\sum_{k=1}^{n} \lambda_k x_k = a$.

Thus, every point $x \geq -1$ is a point of convexity relative to the whole domain of the function xe^x. See Fig. 1. In the case of the arctangent function (which is convex on $(-\infty, 0]$ and concave on $[0, \infty)$), every point $x < 0$ is a point of convexity relative to the neighborhood $(-\infty, x^*]$, where $x^* > 0$ is the abscissa of the point where the tangent at x intersects again the graph. The phenomenon of the existence of convexity/concavity points is genuine for the class of continuous quasiconvex/quasiconcave functions.

The extension of Jensen's inequality to the case of points of convexity is as follows:

Theorem 4. *Suppose that $f : [a, b] \to \mathbb{R}$ is a continuous function and β is a point of convexity of f relative to the whole domain. Then,*

$$f(\beta) \leq \int_a^b f(x)\mathrm{d}\mu(x)$$

for every Borel probability measure μ on $[a, b]$ having the barycenter β.

Proof. The case of discrete probability measures is covered by Definition 2. In the general case, we should notice that every Borel probability measure μ on $[a, b]$ is the pointwise limit of a net of discrete Borel probability measures μ_α, each having the same barycenter as μ. See [8], Lemma 4.1.10, p. 183. □

Theorem 4 makes easy the computation of the extremum of certain functionals. For example, combining it with the technique of Dirac sequences (see [6], Chapater XI), one can prove that for every $\beta \geq -1$, the infimum of the functional

$$F(g) = \int_a^b xe^x g(x)\mathrm{d}x$$

over the convex set of all nonnegative integrable functions $g : [a, b] \to \mathbb{R}$ such that

$$\int_a^b g(x)\mathrm{d}x = 1 \quad \text{and} \quad \int_a^b xg(x)\mathrm{d}x = \beta$$

is βe^β.

We end our paper with a result that extends Theorem 4 to the framework of quasiconvex functions and signed measures. This is to be done by adapting the main ingredient in deriving Theorem 2 from Theorem 1: the fact that the sum between a convex function and a linear one is quasiconvex.

In general, the sum between a quasiconvex function and a linear function is not necessarily quasiconvex. See the case of the functions x^3 and $-3x$. On the other hand, there are differentiable quasiconvex functions $f : \mathbb{R} \to \mathbb{R}$ (such as xe^x) for which $f(x) - f'(c)x$ is still quasiconvex whenever c is a point such that $\partial f(c) \neq \emptyset$. Such functions verify the following version of Jensen's inequality.

Theorem 5. *Suppose that $g : [a, b] \to \mathbb{R}$ is an integrable function that verifies the conditions of Theorem 1 and $\int_a^b g(x)\mathrm{d}x = 1$. Then,*

$$f(\beta_{g(x)\mathrm{d}x}) \leq \int_a^b f(x)g(x)\mathrm{d}x,$$

for every quasiconvex function $f : [a, b] \to \mathbb{R}$ such that $\partial f\left(\beta_{g(x)\mathrm{d}x}\right)$ contains numbers λ with the property that $x \to f(x) - \lambda x$ is also quasiconvex.

Proof. By our hypotheses, for $\lambda \in \partial f\left(\beta_{g(x)\mathrm{d}x}\right)$, we have

$$f(x) \geq f(\beta_{g(x)\mathrm{d}x}) + \lambda(x - \beta_{g(x)\mathrm{d}x}) \quad \text{for every } x \in [a, b]$$

and the nonnegative function $f(x) - f(\beta_{g(x)\mathrm{d}x}) - \lambda(x - \beta_{g(x)\mathrm{d}x})$ is quasiconvex. By Theorem 1,

$$\begin{aligned} 0 &\leq \int_a^b \left[f(x) - f(\beta_{g(x)\mathrm{d}x}) - \lambda(x - \beta_{g(x)\mathrm{d}x})\right] g(x)\mathrm{d}x \\ &= \int_a^b f(x)g(x)\mathrm{d}x - f(\beta_{g(x)\mathrm{d}x}) \end{aligned}$$

and the proof is done. □

The argument of Theorem 5 also covers the case of robust quasiconvex functions, recently introduced by Barron, Goebel and Jensen [1].

Some of the results proved above (including Theorem 2 and Theorem 3) can be extended easily to the context of several variables. However, the characterization of Steffensen–Popoviciu measures in that context is still an open problem. Only few examples are known. See the paper of Niculescu and Spiridon [10].

References

[1] Barron, E.N., Goebel, R., Jensen, R.R.: Functions which are quasiconvex under linear perturbations. SIAM J. Optim **22**(3), 1089–1108 (2012)

[2] Cambini, A., Martein, L.: Generalized Convexity and Optimization. Theory and Applications, Lecture Notes in Economics and Mathematical Systems, vol. 616. Springer, Berlin (2009)

[3] Fink, A.M.: A best possible Hadamard inequality. Math. Inequal. Appl. **1**, 223–230 (1998)

[4] Fink, A.M., Jodeit, M.: On Chebyshev's other inequality. In: Inequalities in Statistics and Probability, Lecture Notes IMS, vol. 5, pp. 115–120. Institute of Mathematical Statistics, Haywood (1984)

[5] Hewitt, E., Stromberg, K.: Real and Abstract Analysis. Second printing corrected, Springer, Berlin, Heidelberg, New York (1969)

[6] Lang, S.: Undergraduate Analysis, 2nd edn. Springer, New York (1997)

[7] Martos, B.: Nonlinear Programming Theory and Methods. North-Holland, Amsterdam (1975)

[8] Niculescu, C.P., Persson L.-E.: Convex Functions and their Applications. A Contemporary Approach, CMS Books in Mathematics, vol. 23. Springer, New York (2006)

[9] Niculescu, C.P., Rovenţa, I.: Relative convexity and its applications. Aequat. Math. (2014). doi:10.1007/s00010-014-0319-x

[10] Niculescu, C.P., Spiridon, C.: New Jensen-type inequalities. J. Math. Anal. Appl **401**(1), 343–348 (2013)

Marcela V. Mihai and Constantin P. Niculescu
Department of Mathematics
University of Craiova
200585 Craiova
Romania
e-mail: cpniculescu@gmail.com; mmihai58@yahoo.com

Received: June 30, 2014.
Revised: September 10, 2014.
Accepted: October 4, 2014.