Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 9, No. 2/2017

JENSEN'S INTEGRAL INEQUALITY IN LOCALLY CONVEX SPACES*

Libor Vesely†

Abstract

We prove generalizations of Jensen's integral inequality for proper convex functions defined on a finite- or infinite-dimensional convex set in a locally convex topological vector space.

MSC: 26D15, 26B25, 46N10, 52A41

keywords: Convex function, locally convex space, Jensen's integral inequality

1 Introduction

The well-known Jensen's (finite) inequality asserts that if C is a convex set (in a real vector space), $f\colon C \to (-\infty, \infty]$ a convex function, $x_1, \ldots, x_n$ points of C, and $\lambda_1, \ldots, \lambda_n$ positive numbers whose sum equals 1, then the point $\sum_{i=1}^n \lambda_i x_i$ belongs to C, and

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \le \sum_{i=1}^{n} \lambda_i f(x_i)\,.$$

This inequality can be reformulated in any of the following two forms:
as

$$f\left(\int_C x\,d\mu(x)\right) \le \int_C f\,d\mu \tag{1}$$

*Accepted for publication on April 5-th, 2017
†libor.vesely@unimi.it Dipartimento di Matematica, Università degli Studi di Milano, Via C. Saldini 50, 20133 Milano, Italy.

where μ is the probability measure $\mu := \sum_{i=1}^{n} \lambda_i \delta_{x_i}$ on C (here δ_y denotes the Dirac measure at y);

or as

$$f\left(\int_{\Omega} G\,d\mu\right) \le \int_{\Omega} (f \circ G)\,d\mu \tag{2}$$

where $\Omega = \{1, \dots, n\}$, the mapping $G\colon \Omega \to C$ is given by $G(i) = x_i$ $(1 \le i \le n)$, and μ is the probability measure on Ω, given by $\mu(\{i\}) = \lambda_i$ $(1 \le i \le n)$.

By *Jensen's integral inequality* one usually means an inequality of type (2) where C is an interval of reals and μ is a probability measure on some set Ω. It can be formulated as follows.

Jensen's integral inequality. *Let f be a real-valued convex function on an interval $I \subset \mathbf{R}$, $(\Omega, \mathcal{A}, \mu)$ a probability space, and $G\colon \Omega \to I$ a μ-integrable function. Then $\int_{\Omega} G\,d\mu$ belongs to I, the Lebesgue integral $\int_I (f \circ G)\,d\mu$ exists (finite or infinite), and (2) holds.*

See e.g. [3, Theorem 2, p. 181], [6, Theorem 1.8.1], [10, Theorem 3.3], or [11, Exercise 22, p. 192] for similar or slightly less general statements.

The aim of the present paper is to provide general versions of the inequalities (1) and (2) in which C is a convex set in a locally convex topological vector space. In literature, we have found only one known result in this direction: a multivalued version of Jensen's integral inequality proved in [4] implies, roughly speaking, validity of (2) when C is an open convex subset of a Banach space, and f is a continuous convex function (with appropriate assumptions on the mapping G).

Any of the two versions (1), (2) of Jensen's integral inequality can be easily deduced from the other one. We prefer to prove first the form (1) which is simpler and more convenient from the geometric point of view, and then derive the second form of the theorem. Our proofs use standard techniques of infinite-dimensional Convex Analysis.

2 Preliminaries

Throughout the paper, all topological vector spaces (t.v.s., for short) are real and, if not otherwise specified, $X = (X, \tau)$ is a Hausdorff locally convex t.v.s., and (C, Σ, μ) is a *convex probability space in* X, that is, a positive measure space such that $\mu(C) = 1$ and C is a nonempty convex subset of X. As usual, $\overline{C}$, $\mathrm{aff}(C)$, and $\mathrm{span}(C)$ denote the closure, the affine hull, and the linear span of C, respectively. The topological dual of X is denoted by X^*.

By $\mathrm{eco}(C)$ we denote the *evenly convex hull* of C, that is, the intersection of all open halfspaces containing C; and we say that the set C is *evenly convex* if $\mathrm{eco}(C) = C$ (cf. [2]). The Hahn-Banach separation theorem easily implies that $\mathrm{eco}(C) \subset \overline{C}$, and *if C is either closed or open then C is evenly convex.*

By $\mathrm{ri}(C)$ we denote the *relative interior* of C, that is, the interior of C in its affine hull. The *dimension* of C is the dimension of $\mathrm{aff}(C)$. It is a well-known fact that *every nonempty finite-dimensional convex set C has nonempty relative interior* (see, e.g., [1, Theorem 2.4.6]). This fact and an easy application of the Hahn-Banach separation theorem give that

(∗) *if C is finite-dimensional and $x \in \mathrm{aff}(C) \setminus \mathrm{ri}(C)$, then there exists $x^* \in X^*$ such that* $\sup x^*(C) \leq x^*(x)$ *and x^* is nonconstant on C.*

By $\mathcal{B}(C)$ we mean the σ-algebra of all relatively τ-Borel subsets of C.

Given a function $f\colon E \to \overline{\mathbf{R}} := [-\infty, \infty]$ and some $\alpha \in \overline{\mathbf{R}}$, we shall use an intuitively clear simplified notation, illustrated by the following particular case:

$$[f < \alpha] := \{x \in E : f(x) < \alpha\}.$$

Pettis integral. Let (Ω, Σ, μ) be any probability space. A mapping $G\colon \Omega \to X$ is said to be *Pettis integrable* (or τ-Pettis integrable) if for every $x^* \in X^*$ the composition $x^* \circ G$ is Legesgue μ-integrable and there exists $z \in X$ such that

$$\int_\Omega (x^* \circ G)\, d\mu = x^*(z), \quad x^* \in X^*.$$

Then one defines the *Pettis integral* $\int_\Omega G\, d\mu := z$. Since the elements of X^* separate points of X, the Pettis integral is unique if it exists.

Push-forward (or image) of a measure. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, (M, Σ) a measurable space, and $G\colon \Omega \to M$ an $\mathcal{A}$-to-Σ measurable mapping. Then we can define a measure ν on Σ by

$$\nu(A) := \mu(G^{-1}(A)), \quad A \in \Sigma.$$

This measure is called the *push-forward* (or *image*) of μ by G, and it is often denoted by $G\#\mu$. Given a Σ-measurable function $f\colon M \to \overline{\mathbf{R}}$, the Lebesgue integral $\int_\Gamma f\, d\nu$ exists (finite or infinite) if and only if the Lebesgue integral $\int_\Omega (f \circ G)\, d\mu$ exists; moreover, the two integrals coincide in this case.

3 Barycenter of a probability measure

Let (C, Ω, μ) be a convex probability space in X such that $\mathcal{B}(C) \subset \Sigma$. The *barycenter* x_μ of μ is defined as the Pettis integral of the identity mapping $I\colon C \to X$, $I(y) = y$,

$$x_\mu := \int_C y\, d\mu(y) = \int_C I\, d\mu.$$

In other words, the barycenter x_μ, if it exists, is the point of X such that

$$x^*(x_\mu) = \int_C x^*\, d\mu, \quad x^* \in X^*.$$

If the barycenter x_μ exists, and $x^* \in X^* \setminus \{0\}$ is such that $x^*(x_\mu) \geq s := \sup x^*(C)$, then $x^*(x_\mu) = s$, and μ is concentrated on the set $[x^* = s] \cap C$. (Indeed, since $\int_C \{x^*(x_\mu) - x^*\}\, d\mu = x^*(x_\mu) - \int_C x^*\, d\mu = 0$ and the integrand in braces is nonnegative, the latter has to be μ-almost everywhere null.)

Corollary 1. *If the barycenter x_μ exists, then $x_\mu \in \mathrm{eco}(C)$ $(\subset \overline{C})$.*

Proof. If $x_\mu \notin \mathrm{eco}(C)$, there exist $x^* \in X^* \setminus \{0\}$ and $\alpha \in \mathbf{R}$ such that $C \subset [x^* < \alpha]$ and $x^*(x_\mu) \geq \alpha$. But this contradicts Observation 3. □

We shall need the following formally stronger version of a result due to Rubin and Wesler [9], for which the same proof works.

Theorem 1. *If C is a finite-dimensional convex set and the barycenter x_μ exists, then μ is concentrated on a convex set $D \subset C$ for which $x_\mu \in \mathrm{ri}(D)$.*

Proof. Clearly, we can (and do) suppose that $C \subset \mathbf{R}^d$. Let $m \geq 0$ be the smallest dimension of an affine set $L \subset \mathbf{R}^d$ such that μ is concentrated on $C \cap L$. Fix such a set L, and denote $D := C \cap L$. Proceeding by contradiction, assume that $x_\mu \notin \mathrm{ri}(D)$. Then necessarily $m \geq 1$, and $x_\mu \subset \overline{D} \subset \mathrm{aff}(D)$ by Corollary 1. By $(*)$, there exists $\xi \in (\mathbf{R}^d)^*$ such that $\xi|_D$ is nonconstant, and $\xi(x_\mu) \geq s := \sup \xi(D)$. By Observation 3, μ is concentrated on $D \cap [\xi = s] = C \cap (L \cap [\xi = s])$. Since the set in parentheses is an affine proper subset of L, this contradicts the choice of L. □

4 Jensen's integral inequality

A convex function $f\colon C \to (-\infty, \infty]$ is called *proper* if the set $[f < \infty]$ is nonempty. The *epighraph* of f is the set

$$\mathrm{epi}(f) = \{(x, t) \in C \times \mathbf{R} : f(x) \leq t\}.$$

It is well-known that f is lower semicontinuous if and only if $\mathrm{epi}(f)$ is relatively closed in $C \times \mathbf{R}$. Recall that *a finite convex function on a finite-dimensional convex set C is continuous at any point of* $\mathrm{ri}(C)$ (see, e.g., [1, Theorem 2.1.12]).

We say that a closed hyperplane $H \subset X$ *strictly separates* two points $x, y \in X$ if the set $\{x, y\}$ is not contained in any of the two closed halfspaces determined by H.

It is an easy exercise to show the following fact. *Let H be a closed hyperplane in $X \times \mathbf{R}$ that strictly separates two "vertically situated points", that is, two points of $X \times \mathbf{R}$ of the form (x, t) and (x, s). Then H coincides with the graph of a continuous affine function $a \colon X \to \mathbf{R}$.* (The simple argument can be found, e.g., in the proof of [7, Proposition 3.15].)

We shall need the following simple separation-type lemma.

Lemma 1. *Let C be a nonempty convex set in a Hausdorff locally convex t.v.s. X. Let $f \colon C \to (-\infty, \infty]$ be a convex function. Let $z \in C$ and $t \in \mathbf{R}$ be such that $t < f(z) < \infty$. Assume that either f is lower semicontinuous, or the restriction $f|_{\mathrm{aff}([f<\infty])}$ is continuous at z. Then there exists a continuous affine function $a \colon X \to \mathbf{R}$ such that $a|_C \leq f$, and $t < a(z)$.*

Proof. Assume first that f is lower semicontinuous. Since the point (z, t) does not belong to $\mathrm{epi}(f)$, which is closed in $C \times \mathbf{R}$, there exist an open convex neighborhood V of z, and a real number $t' \in (t, f(z))$ such that the convex sets $V \times (-\infty, t')$ and $\mathrm{epi}(f)$ are disjoint. So, the two latter sets can be separated by a closed hyperplane $H \subset X \times \mathbf{R}$. Since H strictly separates the points (z, t) and $(z, f(z))$, it coincides with the graph of a continuous affine function $a \colon X \to \mathbf{R}$ such that $a \leq f$ on C, and $t < a(z)$.

Now, define $D := [f < \infty]$, and assume that $f|_{\mathrm{aff}(D)}$ is continuous at z. We can (and do) suppose that $z = 0$. Then $Y := \mathrm{aff}(D)$ is a subspace of X. Clearly, the point $(0, f(0) + 1)$ belongs to the interior of $\mathrm{epi}(f|_D)$ in $Y \times \mathbf{R}$. Fix an arbitrary real number $t' \in (t, f(0))$, and separate the point $(0, t')$ from $\mathrm{epi}(f|_D)$ by a closed hyperplane $H \subset Y \times \mathbf{R}$, which strictly separates the points $(0, t)$ and $(0, f(0) + 1)$. As above, this hyperplane coincides with the graph of a continuous affine function $\tilde{a} \colon Y \to \mathbf{R}$ such that $\tilde{a} \leq f|_D$ and $t < \tilde{a}(z)$. Now, any continuous affine extension $a \colon X \to \mathbf{R}$ of $\tilde{a}$ does the job. □

Now we are ready for the main results of the present paper.

Theorem 2 (Jensen's integral inequality, first form)**.** *Let X be a Hausdorff locally convex t.v.s., and (C, Σ, μ) a convex probability space in X, such that $\Sigma \supset \mathcal{B}(C)$ and the barycenter x_μ of μ exists (in X). Let $f \colon C \to (-\infty, \infty]$*

be a proper convex function. Assume that at least one *of the following two conditions holds.*

(I) C is finite-dimensional and f is Σ-measurable.

(II) C is evenly convex and f is lower semicontinuous.

Then:

(a) $x_\mu \in C$;

(b) the Lebesgue integral $\int_C f\,d\mu$ exists (finite or infinite);

(c) the Jensen's integral inequality $f(x_\mu) \le \int_C f\,d\mu$ holds.

Proof. (a) follows immediately from Theorem 1 and Corollary 1.

To show (b), it suffices to show that f is minorized, on C, by a continuous affine function $a\colon X \to \mathbf{R}$, which clearly belongs to $L_1(\mu)$. This is immediate by Lemma 1 if (II) holds. In the case (I), take any point z belonging to the relative interior of $[f < \infty]$, and notice that $f|_{\mathrm{aff}([f<\infty])}$ is continuous at z. Hence Lemma 1 applies again.

(c) If $\gamma := \int_C f\,d\mu = \infty$, we are done. So let $\gamma \in \mathbf{R}$. Proceeding by contradiction, assume that $\gamma < f(x_\mu)$. If (II) holds, apply directly Lemma 1 to get a continuous affine function a on X such that $a \le f$ on C, and $\gamma < a(x_\mu) \le f(x_\mu)$; and define $D := C$. Now assume (I). Since γ is finite, μ is concentrated on the convex set $[f < \infty]$. By Theorem 1, there exists a convex set $D \subset [f < \infty]$ such that μ is concentrated on D and $x_\mu \in \mathrm{ri}(D)$. As in the proof of (b), $f|_{\mathrm{aff}(D)}$ is continuous at x_μ, and hence Lemma 1 (applied to $f|_D$) provides a continuous affine function a on X such that $a \le f$ on D, and $\gamma < a(x_\mu) \le f(x_\mu)$. The rest of the proof is common to both cases. By definition of x_μ, we have

$$\gamma < a(x_\mu) = \int_C a\,d\mu = \int_D a\,d\mu \le \int_D f\,d\mu = \gamma,$$

which is a contradiction that completes the proof. □

Theorem 3 (Jensen's integral inequality, second form)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, X a Hausdorff locally convex t.v.s., and (C, Σ) a convex measurable space such that $\Sigma \supset \mathcal{B}(C)$. Let $f\colon C \to (-\infty, \infty]$ be a proper convex function. Assume that* at least one *of the conditions (I),(II) from Theorem 2 is satisfied. Let $G\colon \Omega \to C$ be an $\mathcal{A}$-to-Σ measurable mapping such that the Pettis integral $u := \int_\Omega G\,d\mu$ exists in X. Then:*

(a) $u \in C$;

(b) the Lebesgue integral $\int_\Omega (f \circ G)\, d\mu$ *exists (finite or infinite);*

(c) the Jensen's integral inequality $f(u) \le \int_\Omega (f \circ G)\, d\mu$ *holds.*

Proof. Consider the push-forward measure $\nu := G\#\mu$, which is a probability measure defined on Σ. Then the corresponding barycenter $x_\nu = \int_C x\, d\nu(x) = u$ exists, and $\int_C f\, d\nu = \int_\Omega (f \circ G)\, d\mu$. Apply Theorem 2. □

5 Appendix on existence

Let us finish the paper with collecting some conditions that assure existence of barycenters or Pettis integrals.

Let (C, Σ, μ) be a convex probability space in the Euclidean space $\mathbf{R}^d$, such that $\Sigma \supset \mathcal{B}(C)$. Let $e_1^*, \dots, e_d^* \in (\mathbf{R}^d)^*$ denote the coefficient functionals. Then the following assertions are equivalent:

(i) the barycenter x_μ exists (and belongs to C);

(ii) $\|\cdot\| \in L_1(\mu)$ for some (every) norm $\|\cdot\|$ on $\mathbf{R}^d$;

(iii) $\xi \in L_1(\mu)$ for each $\xi \in (\mathbf{R}^d)^*$;

(iv) $e_i^* \in L_1(\mu)$ for each $1 \le i \le d$.

Proof. The implications $(ii) \Rightarrow (iii) \Rightarrow (vi)$ and $(i) \Rightarrow (iv)$ are obvious. Now assume (iv). Then clearly the point $(\int_C e_1^*\, d\mu, \dots, \int_C e_d^*\, d\mu) \in \mathbf{R}^d$ is the barycenter of μ. Moreover, the ℓ_1-norm $\|\cdot\|_1 := |e_1^*| + \cdots + |e_d^*|$ is clearly μ-integrable and equivalent to $\|\cdot\|$, which implies (ii). □

If $\mathbf{R}^d$ is replaced by a general Hausdorff locally convex t.v.s. X then, of course, existence of x_μ implies that every $x^* \in X^*$ is μ-integrable. *The converse is not true in general.* Indeed, by [5, Example 3.5] there exists a function $f\colon (0,1] \to c_0$ which is not Pettis integrable with respect the Lebesgue measure λ on $(0,1]$, while $x^* \circ f \in L_1(\lambda)$ for each $x^* \in (c_0)^*$; then the push-forward measure $\mu := f\#\lambda$ is a probability measure on c_0 which has no barycenter, and every element of $(c_0)^*$ is μ-integrable.

The following theorem collects several conditions under which the converse holds true. If Y is a Banach space, then w^* denotes the corresponding weak* topology $\sigma(Y^*, Y)$ on Y^*.

Theorem 4. *Let (C, Σ, μ) be a convex probability space in a Hausdorff locally convex t.v.s. (X, τ), such that $\Sigma \supset \mathcal{B}(C)$. Assume that*

$$x^* \in L_1(\mu) \text{ for every } x^* \in X^*,$$

and at least one *of the following conditions is satisfied.*

(a) C is finite-dimensional.

(b) C is relatively τ-compact.

(c) $(X, \tau) = (Y^*, w^*)$ for some Banach space Y.

(d) X is a reflexive Banach space.

(e) C is separable, and X is a Banach space not containing c_0.

Then μ admits a τ-barycenter.

Proof. (a) We can assume that $X = \text{span}(C)$, which is ismorphic to some $\mathbf{R}^d$. Apply Observation 5.

(b) follows directly from [8, Proposition 1.1].

(c) This is due to I.M. Gelfand (see [5, Proposition 3.3] for the reference). The linear operator $T: Y \to L_1(\mu)$, $Ty := y|_C$, has closed graph, and hence it is bounded. Then the functional $\phi(y) := \int_C y \, d\mu$ is an element of $Y^* = X$, and it is the barycenter of μ.

(d) By (c), μ admits a barycenter in the weak topology of X, and hence also in the norm topology.

(e) We can assume that $X = \overline{\text{span}}(C)$. In this case X is separable and does not contain c_0. Apply [5, Theorem 3.6]. □

Now, the push-forward argument from the proof of Theorem 3 gives immediately the following sufficient conditions for existence of Pettis integrals.

Corollary 2. *Let (C, Σ) be a convex measurable space in a Hausdorff locally convex t.v.s. (X, τ), such that $\Sigma \supset \mathcal{B}(C)$. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space, and $G: \Omega \to C$ an $\mathcal{A}$-to-Σ measurable mapping such that $x^* \circ G \in L_1(\mu)$ for every $x^* \in X^*$. If* at least one *of the conditions (a)-(e) from Theorem 4 is satisfied, then the τ-Pettis integral $\int_\Omega G \, d\mu$ exists (in X).*

Acknowledgement. The research was supported in part by GNAMPA - INdAM.

References

[1] J.M. Borwein, J.D. Vanderverff. *Convex Functions: Constructions, Characterizations and Counterexamples.* Encyclopedia of Mathematics and its Applications, Vol. 109, Cambridge University Press, Cambridge, 2010.

[2] V. Klee, E. Maluta, C. Zanco. Basic properties of evenly convex sets. *J. Convex Anal.* 14:137–148, 2007.

[3] M. Kuczma. *An Introduction to the Theory of Functional Equations and Inequalities. Cauchy's Equation and Jensen's Inequality.* Prace Naukowe Uniwersytetu Śląskiego w Katowicach [Scientific Publications of the University of Silesia], Vol. 489, Państwowe Wydawnictwo Naukowe (PWN), Warsaw, 1985.

[4] J. Matkowski, K. Nikodem. An integral Jensen inequality for convex multifunctions. *Results Math.* 26:348–353, 1994.

[5] K. Musial. Pettis integral. In *Handbook of Measure Theory, Vol. I, II.*, pp. 531–586, North-Holland, Amsterdam, 2002.

[6] C. Niculescu, L.-E. Persson. *Convex Functions and Their Applications. A Contemporary Approach.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Vol. 23, Springer, New York, 2006.

[7] R.R. Phelps. *Convex Functions, Monotone Operators and Differentiability. Second Edition.* Lecture Notes in Mathematics, Vol. 1364, Springer-Verlag, Berlin, 1993.

[8] R.R. Phelps. *Lectures on Choquet's Theorem. Second Edition.* Lecture Notes in Mathematics, Vol. 1757, Springer-Verlag, Berlin, 2001.

[9] H. Rubin, O. Wesler. A note on convexity in Euclidean n-space. *Proc. Amer. Math. Soc.* 9:522–523, 1958.

[10] W. Rudin. *Real and Complex Analysis. Second Edition.* McGraw-Hill Series in Higher Mathematics, McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1974.

[11] R.L. Wheeden, A. Zygmund. *Measure and Integral. An Introduction to Real Analysis.* Pure and Applied Mathematics, Vol. 43, Marcel Dekker, Inc., New York-Basel, 1977.