Received 29/03/13

Necessary and Sufficient Conditions for the Validity of Jensen's Inequality

A. Guessab and G. Schmeisser

March 24, 2013

Abstract. We consider the d-dimensional Jensen inequality

$$T[\varphi(f_1, \ldots, f_d)] \geq \varphi(T[f_1], \ldots, T[f_d]) \qquad (*)$$

as it was established by McShane in 1937. Here T is a functional, φ is a convex function defined on a closed convex set $K \subset \mathbb{R}^d$ and $f_1, \ldots, f_d$ are from some linear space of functions. Our aim is to find necessary and sufficient conditions for the validity of $(*)$. In particular, we show that if we exclude three types of convex sets K, then Jensen's inequality holds for a sublinear functional T *if and only if* T is linear, positive and satisfies $T[1] = 1$. Furthermore, for each of the excluded types of convex sets, we present nonlinear, sublinear functionals T for which Jensen's inequality holds. Thus the conditions on K are optimal. Our contributions generalize or complete several known results.

Mathematics Subject Classification (2010). Primary 26Dxx; 39Bxx; 39B62; Secondary 26A51; 26B25.

Keywords. Jensen's inequality, convex functions, sublinear functionals, positive functionals.

1. Introduction

Let I be an interval. In 1906 Jensen [6], one of the founders of the theory of convex functions [3, pp. 70], defined a function $\varphi : I \to \mathbb{R}$ as being *convex* if

$$\varphi\left(\frac{x+y}{2}\right) \leq \frac{\varphi(x) + \varphi(y)}{2} \tag{1.1}$$

for all $x, y \in I$. As a consequence, he obtained that, if a convex function $\varphi : I \to \mathbb{R}$ is in addition continuous, then

$$\varphi\left(\frac{\sum_{\nu=1}^{n} a_\nu x_\nu}{\sum_{\nu=1}^{n} a_\nu}\right) \leq \frac{\sum_{\nu=1}^{n} a_\nu \varphi(x_\nu)}{\sum_{\nu=1}^{n} a_\nu} \tag{1.2}$$

for all $x_1, \dots, x_n \in I$ and all positive $a_1, \dots, a_n$. He mentioned that under the additional assumption that φ'' exists, this inequality was already obtained by Hölder [4] in 1889. Furthermore, by a limiting process, Jensen extended (1.2) to integrals. He showed that

$$\varphi\left(\frac{\int_0^1 a(x)f(x)dx}{\int_0^1 a(x)dx}\right) \le \frac{\int_0^1 a(x)\varphi(f(x))dx}{\int_0^1 a(x)dx} \tag{1.3}$$

holding for integrable functions $a : [0,1] \to (0,\infty)$ and $f : [0,1] \to I$.

In classical textbooks, inequalities (1.2) and (1.3) are named after Jensen. However, in the more modern literature, the convexity of φ is *defined* as the validity of (1.2) for $n = 2$ and all admissible a_1, a_2, x_1, x_2. Then a convex function is automatically continuous on the interior of I.

Nowadays a Jensen's inequality is understood as a generalization of (1.3) due to McShane [8]. In its one-dimensional form it can be interpreted as the inequality (1.3) with the role of the weighted integral being taken by a linear functional. For a precise statement of that inequality in the d-dimensional case and for the results of this article, we first introduce some notation and definitions.

Let Ω be a nonempty set. Let $\mathcal{F}$ be any linear space of functions $f : \Omega \to \mathbb{R}$ that contains the real constants. Denote by $\mathcal{F}_0$ a subspace of bounded functions that contains the constants. Furthermore we denote by $\mathcal{F}_1$ a linear space of functions $f : \Omega \to \mathbb{R}$ that contains the constants and has the additional property that $f \in \mathcal{F}_1$ implies $|f| \in \mathcal{F}_1$. If Ω is a compact set, then $C(\Omega)$, the class of all functions that are continuous on Ω, qualifies as any of the spaces $\mathcal{F}$, $\mathcal{F}_0$ and $\mathcal{F}_1$. Note that definitions and statements valid for $\mathcal{F}$ being any space as introduced before, will also be valid for $\mathcal{F}_0$ and $\mathcal{F}_1$.

By $\boldsymbol{f} \in \mathcal{F}^d$, we mean that

$$\boldsymbol{f} = (f_1, \dots, f_d) \qquad (f_j \in \mathcal{F} \text{ for } j = 1, \dots, d).$$

Generally, we use bold-faced letters for d-tuples and, if not specified otherwise, the corresponding *normal* letter with an index j, say, denotes the jth component.

By $\mathbb{R}_+$ and $\mathbb{R}_-$ we denote the set of non-negative and non-positive real numbers, respectively. A functional T on $\mathcal{F}$ is said to be

(i) *positively homogeneous* if $T[\lambda f] = \lambda T[f]$ for all $\lambda \in \mathbb{R}_+$ and $f \in \mathcal{F}$;
(ii) *subadditive* if $T[f_1 + f_2] \le T[f_1] + T[f_2]$ for all $f_1, f_2 \in \mathcal{F}$;
(iii) *sublinear* if it is positively homogeneous and subadditive;
(iv) *linear* if

$$T[\lambda_1 f_1 + \lambda_2 f_2] = \lambda_1 T[f_1] + \lambda_2 T[f_2]$$

for all $\lambda_1, \lambda_2 \in \mathbb{R}$ and $f_1, f_2 \in \mathcal{F}$;
(v) *positive* if $T[f] \ge 0$ for $f \in \mathcal{F}$ with $f \ge 0$;
(vi) *normalized* if $T[1] = 1$.

Furthermore, for $\boldsymbol{f} \in \mathcal{F}^d$ we define

$$\boldsymbol{T}[\boldsymbol{f}] := \big(T[f_1], \dots, T[f_d]\big).$$

With this terminology the generalization of Jensen's inequality due to McShane [8, Theorem 2] can be stated as follows.

Theorem 1.1 (McShane). *In the previous notation, let T be a linear, positive and normalized functional on $\mathcal{F}$. For $d \in \mathbb{N}$ let $K \subset \mathbb{R}^d$ be a closed convex set. Then*

$$T[\varphi \circ \boldsymbol{f}] \geq \varphi\left(\boldsymbol{T}[\boldsymbol{f}]\right) \tag{1.4}$$

for all $\boldsymbol{f} \in \mathcal{F}^d$ such that $\boldsymbol{f}(x) \in K$ for $x \in \Omega$ and all continuous convex functions $\varphi : K \to \mathbb{R}$ such that $\varphi \circ \boldsymbol{f} \in \mathcal{F}$.

In a preliminary theorem [8, Theorem 1], McShane showed that under the hypotheses on $\boldsymbol{f}$ and φ the existence of $\varphi(\boldsymbol{T}[\boldsymbol{f}])$ is guaranteed, that is, $\boldsymbol{T}[\boldsymbol{f}] \in K$.

The continuity of φ is needed since McShane uses the d-dimensional analogue of (1.1) for the definition of convexity.

For $d = 1$ Theorem 1.1 has gained much attention in books on probability (see, e.g., [1, p. 20, Theorem 3.9]), where it is established for $f = X$ being a random variable and $T = \mathcal{E}$ being a (mathematical) expectation.

In several recent papers the possibility of extending Jensen's inequality (1.4) to sublinear functionals has been studied in the setting of probability and measure theory. In the case $d = 1$ and $K = \mathbb{R}$, Hu [5] gave an extension to certain nonlinear expectations. In the case $d = 2$ and $K = \mathbb{R}^2$, Jia [7] showed that Jensen's inequality holds for a sublinear, monotone and constant preserving expectation $\mathcal{E}$ if and only if $\mathcal{E}$ is linear. For general d but with $K = \mathbb{R}^d_+$, Haase [2] established an inequality which amounts to an extension of (1.4) to sublinear functionals provided φ does not attain any positive value.

In this paper we consider a sublinear functional T and specify closed convex sets K such that McShane's hypotheses on T are not only sufficient but also necessary for the validity of (1.4); see Section 2. For all the other closed convex sets in $\mathbb{R}^d$, we present nonlinear, sublinear functionals such that (1.4) holds; see Section 5.

2. The main result

Our main result shows that if we exclude three types of convex sets K and restrict ourselves to the subspace $\mathcal{F}_0$ of $\mathcal{F}$, then Jensen's inequality holds for a sublinear functional T if and only if T is *linear*, *positive* and *normalized*. Note that for the necessity we do not assume that T is monotone and constant preserving.

The restriction to $\mathcal{F}_0$ is necessary for the following reason. The convex set K in Theorem 1.1 may be bounded. In this case Jensen's inequality involves only bounded functions and so its validity cannot have any consequences for the behavior of T on $\mathcal{F} \setminus \mathcal{F}_0$.

Theorem 2.1. *Let T be a sublinear functional on $\mathcal{F}_0$. For $d \in \mathbb{N}$, let $K \subset \mathbb{R}^d$ be a closed convex set with the following properties:*

(i) *K is not a subset of $\mathbb{R}^d_+$ or $\mathbb{R}^d_-$;*

(ii) *K is not a subset of a line $\{\boldsymbol{a}+t\boldsymbol{w} : t \in \mathbb{R}\}$, where $\boldsymbol{a}, \boldsymbol{w} \in \mathbb{R}^d$, $\boldsymbol{w} \neq 0$ and the non-zero components of $\boldsymbol{w}$ are all of the same sign.*

Then, in order that

$$T[\varphi \circ \boldsymbol{f}] \geq \varphi\left(\boldsymbol{T}[\boldsymbol{f}]\right)$$

holds for all $\boldsymbol{f} \in \mathcal{F}_0^d$ such that $\boldsymbol{f}(x) \in K$ for $x \in \Omega$ and all continuous convex functions $\varphi : K \to \mathbb{R}$ such that $\varphi \circ \boldsymbol{f} \in \mathcal{F}_0$, it is necessary and sufficient that T is linear, positive and normalized.

Note that property (ii) excludes $d = 1$.

Remark 2.2. One can state an alternative version of Theorem 2.1 that extends to $\mathcal{F}$ by excluding bounded sets K in requiring that K contains a straight line $\{\boldsymbol{a}+t\boldsymbol{w} : t \in \mathbb{R}\}$, where $\boldsymbol{a}, \boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{w}$ has two non-zero components of different signs.[1]

3. Some lemmas

Before proving Theorem 2.1, we want to state some useful auxiliary results.

Lemma 3.1. *Let T be a sublinear functional on $\mathcal{F}$ with the additional property that $T[\beta] = \beta$ for all $\beta \in \mathbb{R}$. Then*

$$T[\alpha f + \beta] \geq \alpha T[f] + \beta \tag{3.1}$$

for all $\alpha, \beta \in \mathbb{R}$ and all $f \in \mathcal{F}$. Equality occurs when $\alpha \geq 0$.

Proof. The subadditivity of T and the preservation of constants yield

$$T[f+\beta] \leq T[f] + T[\beta] = T[f] + \beta$$

and

$$T[f] = T[f+\beta-\beta] \leq T[f+\beta] + T[-\beta] = T[f+\beta] - \beta.$$

Combining both inequalities, we obtain

$$T[f+\beta] = T[f] + \beta. \tag{3.2}$$

Next, if $\lambda \geq 0$, then, by the positive homogeneity, $T[\lambda f] = \lambda T[f]$. Furthermore

$$0 = T[\lambda f - \lambda f] \leq T[\lambda f] + T[-\lambda f] = \lambda T[f] + T[-\lambda f],$$

which shows that $T[-\lambda f] \geq -\lambda T[f]$. Hence, for any $\alpha \in \mathbb{R}$, we have

$$T[\alpha f] \geq \alpha T[f] \tag{3.3}$$

with equality when $\alpha \geq 0$. From (3.2) and (3.3), the conclusion of the lemma follows immediately. □

As a test for the linearity of a sublinear functional, we want to use the following known criterion; see, e.g., [7, Proposition 2.1].

Lemma 3.2. *Let T be a sublinear functional on $\mathcal{F}$. Then T is linear if and only if $T[f] + T[-f] \leq 0$ for all $f \in \mathcal{F}$.*

---

[1] This addresses a question asked by one of the referees.

We shall also use the following lemma, which can be verified by basic linear algebra.

Lemma 3.3. *Let K be a convex subset of $\mathbb{R}^d$. If K contains three points that are not collinear, then it contains a line segment*

$$\{\boldsymbol{a} + t\boldsymbol{w} \,:\, t \in [0,1]\}, \tag{3.4}$$

where $\boldsymbol{a}, \boldsymbol{w} \in \mathbb{R}^d$ and $\boldsymbol{w}$ has two non-zero components of different signs.

Proof. The three non-collinear points span a non-degenerate triangle $\Delta \subset K$. Let $\boldsymbol{a}$ be the center of gravity of Δ. Knowing that the inradius of Δ is positive, we conclude that there exist two linearly independent vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$ and a $\delta > 0$ such that $\boldsymbol{a} + \lambda\boldsymbol{u} + \mu\boldsymbol{v} \in K$ whenever $\lambda^2 + \mu^2 \leq \delta^2$. As a consequence of the linear independence of $\boldsymbol{u}$ and $\boldsymbol{v}$, there exist indices k and ℓ with $1 \leq k < \ell \leq d$ such that

$$\det \begin{pmatrix} u_k & v_k \\ u_\ell & v_\ell \end{pmatrix} \neq 0\,.$$

Hence the linear system

$$\begin{pmatrix} u_k & v_k \\ u_\ell & v_\ell \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

has a unique solution. Defining

$$\delta' := \frac{\delta}{\sqrt{\alpha^2 + \beta^2}} \qquad \text{and} \qquad \boldsymbol{w} := \alpha\delta'\boldsymbol{u} + \beta\delta'\boldsymbol{v},$$

we see that $w_k = \delta'$, $w_\ell = -\delta'$ and $\boldsymbol{a} + t\boldsymbol{w} \in K$ for $t \in [0,1]$. □

4. Proof of Theorem 2.1

The sufficiency follows from McShane's Theorem 1.1.

We now turn to the necessity. First we want to show that the validity of (1.4) entails that T preserves constants.

Take any $c \in \mathbb{R}$ and define $\varphi(\boldsymbol{y}) \equiv c$ for $\boldsymbol{y} \in K$, which is a convex function. In this case Jensen's inequality implies that

$$T[c] \geq c \qquad (c \in \mathbb{R}). \tag{4.1}$$

Next, by its property (i), K contains a point $\boldsymbol{b}$ which has a positive and a negative coordinate, say $b_k > 0$ and $b_\ell < 0$. The functions $\varphi(\boldsymbol{y}) := y_k - b_k$ and $\boldsymbol{f} := \boldsymbol{b}$ are admissible for Jensen's inequality. Since $T[\varphi \circ \boldsymbol{f}] = 0$ and

$$\varphi\left(\boldsymbol{T}[\boldsymbol{f}]\right) = T[b_k] - b_k\,,$$

Jensen's inequality implies that $T[b_k] \leq b_k$. Together with (4.1) we obtain $T[b_k] = b_k$. Since T is positively homogeneous, we see that $T[\lambda b_k] = \lambda b_k$ for every $\lambda \geq 0$. Hence $T[\beta] = \beta$ for all $\beta \geq 0$.

Analogously, consider $\varphi(\boldsymbol{y}) := y_\ell - b_\ell$ and $\boldsymbol{f} := \boldsymbol{b}$. Then by the same arguments $T[b_\ell] = b_\ell$, which implies that $T[\gamma] = \gamma$ for all $\gamma \leq 0$. Altogether, we have shown that T preserves constants. In particular, T is normalized.

Next we show that T is linear. Property (ii) in conjunction with Lemma 3.3 ensures that K contains a line segment (3.4) with a vector $\boldsymbol{w}$ that has two non-zero components of different signs, say $w_k > 0$ and $w_\ell < 0$. Now take any f from $\mathcal{F}_0$. Then for a sufficiently small $\varepsilon > 0$ and an appropriate $\gamma \in \mathbb{R}$, we have $\varepsilon f(x) + \gamma \in [0, 1]$ for all $x \in \Omega$, and so

$$\boldsymbol{f}(x) := \boldsymbol{a} + (\varepsilon f(x) + \gamma)\boldsymbol{w} \in K.$$

Defining

$$\varphi(\boldsymbol{y}) := \frac{y_k - a_k}{w_k} - \frac{y_\ell - a_\ell}{w_\ell},$$

our choices of $\boldsymbol{f}$ and φ are as required for Jensen's inequality (1.4). Now, the left-hand side of (1.4) yields

$$T[\varphi \circ \boldsymbol{f}] = T[\varepsilon f + \gamma - \varepsilon f - \gamma] = 0.$$

For the right-hand side, we find by employing Lemma 3.1 that

$$\begin{aligned}\varphi\left(\boldsymbol{T}[\boldsymbol{f}]\right) &= \frac{T[a_k + (\varepsilon f + \gamma)w_k] - a_k}{w_k} - \frac{T[a_\ell + (\varepsilon f + \gamma)w_\ell] - a_\ell}{w_\ell} \\ &= \frac{T[\varepsilon f w_k]}{w_k} + \gamma - \frac{T[\varepsilon f w_\ell]}{w_\ell} - \gamma = \varepsilon T[f] + \varepsilon T[-f]\,.\end{aligned}$$

Hence the validity of (1.4) gives $T[f] + T[-f] \leq 0$. Thus Lemma 3.2 ensures that T is linear.

Finally, we show that T is positive. Again we use that K contains a line segment (3.4). Clearly, there exists a vector $\boldsymbol{v} \in \mathbb{R}^d$ such that $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 1$, where the standard inner product has been used. Define

$$\varphi(\boldsymbol{y}) := \langle \boldsymbol{v}, \boldsymbol{y} - \boldsymbol{a} \rangle - \min\left\{\langle \boldsymbol{v}, \boldsymbol{y} - \boldsymbol{a} \rangle, 0\right\},$$

which is a convex function. Take any $f \in \mathcal{F}_0$ such that $f \geq 0$. Then, for a sufficiently small $\varepsilon > 0$, we have $\varepsilon f(x) \in [0, 1]$ for all $x \in \Omega$. Hence

$$\boldsymbol{f}(x) := \boldsymbol{a} + \varepsilon f(x)\boldsymbol{w} \in K \qquad \text{for all} \quad x \in \Omega,$$

and so $\boldsymbol{f}$ is admissible for Jensen's inequality. Since

$$\langle \boldsymbol{v}, \boldsymbol{f}(x) - \boldsymbol{a} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle \varepsilon f(x) = \varepsilon f(x),$$

we see that

$$\left(\varphi \circ \boldsymbol{f}\right)(x) = \varepsilon f(x) - \min\{\varepsilon f(x), 0\} = \varepsilon f(x).$$

Thus $\varphi \circ \boldsymbol{f} \in \mathcal{F}_0$ and

$$T[\varphi \circ \boldsymbol{f}] = T[\varepsilon f]. \tag{4.2}$$

Since we know already that T is linear and preserves constants, we find that

$$\langle \boldsymbol{v}, \boldsymbol{T}[\boldsymbol{f}] - \boldsymbol{a} \rangle = \langle \boldsymbol{v}, \boldsymbol{T}[\varepsilon f \boldsymbol{w}] \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle T[\varepsilon f] = T[\varepsilon f].$$

Therefore

$$\varphi\left(\boldsymbol{T}[\boldsymbol{f}]\right) = T[\varepsilon f] - \min\{T[\varepsilon f], 0\}. \tag{4.3}$$

With (4.2) and (4.3) the validity of Jensen's inequality yields

$$T[\varepsilon f] \geq T[\varepsilon f] - \min\{T[\varepsilon f], 0\},$$

or equivalently, $\min\{T[\varepsilon f], 0\} \geq 0$. This shows that $T[f] \geq 0$ and so T is positive. ☐

Remark 4.1. In the previous proof of the necessity, we needed Jensen's inequality for piecewise affine convex functions φ only. Hence its validity for these particular functions φ implies that it holds for all convex functions φ.

5. On the optimality of the conditions (i) and (ii) in Theorem 2.1

In this section we want to show that in Theorem 2.1, conditions (i) and (ii) cannot be abandoned. In fact, if K does not satisfy (i) or (ii), we shall see that there exist nonlinear, sublinear functionals T for which Jensen's inequality holds. In the following proposition, condition (i) is violated. The result may be compared with [2, Proposition 1.1] in which $K = \mathbb{R}^d_+$ and an opposite Jensen inequality is obtained with φ being replaced by a nonnegative concave function F. Since $-F$ is convex, an inequality of the form (1.4) can be deduced.

Proposition 5.1. *Let K be a closed convex subset of $\mathbb{R}^d_-$ or $\mathbb{R}^d_+$. Then there exists a nonlinear, sublinear functional T on $\mathcal{F}_1$ such that Jensen's inequality* (1.4) *holds for $\boldsymbol{f}$ and φ as specified in Theorem* 1.1 *with respect to the function space $\mathcal{F}_1$.*

Proof. Let L be any linear, positive and normalized functional on $\mathcal{F}_1$. Then, by McShane's Theorem 1.1

$$L[\varphi \circ \boldsymbol{f}] \geq \varphi(\boldsymbol{L}[\boldsymbol{f}]) \tag{5.1}$$

for the given K and φ and $\boldsymbol{f}$ as specified in Theorem 1.1 with $\mathcal{F}$ replaced by $\mathcal{F}_1$.

Now suppose that $K \subset \mathbb{R}^d_-$. Then we define a functional T by

$$T[f] := L[2f + |f|] \qquad (f \in \mathcal{F}_1).$$

It has the following properties:

(i) $T[f] \geq L[f]$ for $f \in \mathcal{F}_1$;
(ii) $T[f] = L[f]$ for $f \in \mathcal{F}_1$, $f \leq 0$;
(iii) T is not linear;
(iv) T is sublinear.

As regards (i), we know that a positive linear functional is monotonically increasing and so the obvious inequality $2f+|f| \geq f$ implies that $L[2f+|f|] \geq L[f]$, which is (i). For (ii) we simply observe that $2f + |f| = f$ when $f \leq 0$. Since $T[-1] = -1$ but $-T[1] = -3$, we see that T cannot be linear. As regards (iv), it is obvious that T is positively homogeneous. Since

$$T[f_1+f_2] = L[2(f_1+f_2)+|f_1+f_2|] \leq L[2(f_1+f_2)+|f_1|+|f_2|] = T[f_1]+T[f_2],$$

we see that T is subadditive.

Now, by (ii) and the fact that $K \subset \mathbb{R}^d_-$, we have $\varphi(\boldsymbol{L}[\boldsymbol{f}]) = \varphi(\boldsymbol{T}[\boldsymbol{f}])$ while (i) gives $T[\varphi \circ \boldsymbol{f}] \ge L[\varphi \circ \boldsymbol{f}]$. Hence Jensen's inequality for T follows from (5.1).

Analogously, if $K \subset \mathbb{R}^d_+$, we define $T[f] := L[|f|]$ and see that (i), (iii) and (iv) hold while (ii) is valid for $f \ge 0$. Thus Jensen's inequality for T follows again from (5.1). □

The next proposition shows that if condition (ii) is violated, then there exits a class of sublinear functionals T, containing nonlinear ones, such that Jensen's inequality holds.

Proposition 5.2. *Let T be a sublinear functional on $\mathcal{F}$ with the additional properties:*

(i) $T[\lambda] = \lambda$ *for all* $\lambda \in \mathbb{R}$;
(ii) $f \le g \implies T[f] \le T[g] \qquad (f, g \in \mathcal{F})$.

Let K be a closed convex subset of a line

$$\Lambda := \{\boldsymbol{a} + t\boldsymbol{w} \,:\, t \in \mathbb{R}\},$$

where $\boldsymbol{a}, \boldsymbol{w} \in \mathbb{R}^d$, $\boldsymbol{w} \ne 0$ and the non-zero components of $\boldsymbol{w}$ are all of the same sign. Then Jensen's inequality (1.4) *holds for $\boldsymbol{f}$ and φ as specified in Theorem* 1.1.

Proof. We may assume that the components of $\boldsymbol{w}$ are all non-negative since otherwise we may simply replace $\boldsymbol{w}$ by $-\boldsymbol{w}$ without changing the line Λ.

Since K is closed and convex, there exists a closed interval $I \subset \mathbb{R}$ such that

$$K = \{\boldsymbol{a} + t\boldsymbol{w} \,:\, t \in I\}.$$

Let $\boldsymbol{f}$ be as specified in Theorem 1.1. The condition $\boldsymbol{f}(x) \in K$ for $x \in \Omega$ means that we can associate with each $x \in \Omega$ a $t \in I$, which we may call $g(x)$, such that

$$\boldsymbol{f}(x) = \boldsymbol{a} + g(x)\boldsymbol{w}. \tag{5.2}$$

This way, we have defined a function $g : \Omega \to I$. Suppose that $w_k > 0$. Then it follows from (5.2) that $g(x) = (f_k(x) - a_k)/w_k$ for $x \in \Omega$, which shows that $g \in \mathcal{F}$.

Next we note that for φ as specified in Theorem 1.1, the function

$$\psi \,:\, t \longmapsto \varphi(\boldsymbol{a} + t\boldsymbol{v})$$

is a convex function defined on I and

$$(\varphi \circ \boldsymbol{f})(x) = \psi(g(x)).$$

Now suppose that $\psi(t) \ge \alpha t + \beta$ for all $t \in I$. Then

$$\psi \circ g \ge \alpha g + \beta. \tag{5.3}$$

Employing property (ii) and Lemma 3.1, we conclude that

$$T[\psi \circ g] \ge T[\alpha g + \beta] \ge \alpha T[g] + \beta. \tag{5.4}$$

Since $g(x) \in I$ for all $x \in \Omega$, we conclude by using properties (i) and (ii) of the functional T that $T[g] \in I$. Thus, from (5.3) and (5.4), it follows that

$$T[\varphi \circ \boldsymbol{f}] = T[\psi \circ g] \geq \sup_{\alpha,\beta\in\mathbb{R}} \left\{\alpha T[g] + \beta : \psi(T[g]) \geq \alpha T[g] + \beta\right\}. \quad (5.5)$$

Since ψ is a convex function, its graph may be represented as the envelope of all supporting lines, that is,

$$\psi(t) = \sup_{\alpha,\beta\in\mathbb{R}} \{\alpha t + \beta : \psi(t) \geq \alpha t + \beta\}.$$

Hence the right-hand side of (5.5) equals $\psi(T[g])$.

We have proved so far that

$$T[\varphi \circ \boldsymbol{f}] \geq \psi(T[g]) = \varphi(\boldsymbol{a} + T[g]\boldsymbol{w}).$$

Now, by Lemma 3.1 and our assumptions on the vector $\boldsymbol{w}$, there holds

$$a_j + T[g]w_j = T[a_j + gw_j] \qquad (j = 1, \ldots, d).$$

Hence

$$\varphi(\boldsymbol{a} + T[g]\boldsymbol{w}) = \varphi(\boldsymbol{T}[\boldsymbol{a} + g\boldsymbol{w}]) = \varphi(\boldsymbol{T}[\boldsymbol{f}]).$$

This completes the proof. □

Remark 5.3. An analysis of the previous proof reveals that instead of assuming that T is sublinear and satisfies (i) and (ii), it suffices to require that T satisfies (ii) and the conclusion of Lemma 3.1 holds. This variant of Proposition 5.2 generalizes [5, Theorem 4.1].

Remark 5.4. If $K = \mathbb{R}_+^d$ and φ is non-positive valued, the previous proof extends to arbitrary dimension $d \in \mathbb{N}$ by using supporting hyperplanes instead of supporting lines. This way, one obtains [2, Theorem 1.1].

As a consequence of Theorem 2.1 and Propositions 5.1 and 5.2, we arrive at the following conclusion: There exist nonlinear, sublinear functionals T for which Jensen's inequality as stated in Theorem 2.1 holds *if and only if* the closed convex set K is a subset of either R_+^d or $\mathbb{R}_-^d$ or of a line $\{\boldsymbol{a} + t\boldsymbol{w} : t \in \mathbb{R}\}$, where $\boldsymbol{a}, \boldsymbol{w} \in \mathbb{R}^d$, $\boldsymbol{w} \neq 0$ and the non-zero components of $\boldsymbol{w}$ are all of the same sign.

As emphasized by one of the referees, it may be interesting to give, for each of the three types of closed convex sets K, a complete description of all the sublinear functionals T which preserve the validity of Jensen's inequality.

Acknowledgment

We would like to thank the two referees for their useful proposals.

References

[1] H. Bauer, *Probability Theory*, Walter de Gruyter, Berlin, 1996.

[2] M. Haase, Convexity inequalities for positive operators, *Positivity* **11** (2007), 57–68.

[3] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1934.

[4] H. Hölder, Über einen Mittelwertsatz, *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen* (1889), 38–47.

[5] Y. Hu, On Jensen's inequality for g-expectation and for nonlinear expectation, *Arch. Math.* **85** (2005), 572–580.

[6] J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.* **30** (1906), 175–193.

[7] G. Y. Jia, The minimal sublinear expectations and their related properties, *Sci. China Ser. A* **52**(4) (2009), 785–793.

[8] E. J. McShane, Jensen's inequality, *Bull. Amer. Math. Soc.* **43** (1937), 521–527.

A. Guessab
Laboratoire de Mathématiques et de leurs Applications
UMR CNRS 4152
Université de Pau et des Pays de l'Adour
64000 Pau
France
e-mail: `allal.guessab@univ-pau.fr`

G. Schmeisser
Department of Mathematics
University of Erlangen-Nuremberg
91058 Erlangen
Germany
e-mail: `schmeisser@mi.uni-erlangen.de`