BULLETIN OF THE POLISH
ACADEMY OF SCIENCES
MATHEMATICS
Vol. 55, No. 3, 2007

PROBABILITY THEORY AND STOCHASTIC PROCESSES

Some Remarks on Functionals with the Tensorization Property

by

Paweł WOLFF

Presented by Stanisław KWAPIEŃ

Summary. We investigate the subadditivity property (also known as the tensorization property) of φ-entropy functionals and their iterations. In particular we show that the only iterated φ-entropies with the tensorization property are iterated variances. This is a complement to the result due to Latała and Oleszkiewicz on characterization of the standard φ-entropies with the tensorization property.

1. Introduction. An important feature of some functional inequalities for probability measures is the *tensorization property* (sometimes called the product property): if the inequality holds for each measure $\mu_1, \mu_2, \ldots$ then it also holds for the product measure $\mu_1 \otimes \mu_2 \otimes \cdots$. In this paper we focus on the tensorization property of entropy-energy inequalities, well-known examples of which are the logarithmic Sobolev inequality and Poincaré inequality.

By the *φ-entropy functional* we mean the functional $E\varphi(Z) - \varphi(EZ)$. For $\varphi(x) = x \log x$ we get the classical entropy functional, for $\varphi(x) = x^2$ we get the variance, and for $\varphi(x) = x^p$, $p \in (1, 2]$, the so-called p-variance. The family of entropy-energy inequalities corresponding to the p-variance, which interpolate between the logarithmic Sobolev and Poincaré inequalities, was introduced by Beckner [1] in the context of Gaussian measure on $\mathbb{R}^n$ and Haar measure on the sphere S^{n-1}. A more abstract treatment of this family of inequalities (in the context of arbitrary probability measures) was given by Latała and Oleszkiewicz [3]. One of the results in that paper states that if $\varphi\colon (0, \infty) \to \mathbb{R}$ belongs to the class Φ, that is, φ is either affine or convex with

2000 *Mathematics Subject Classification*: Primary 60E15; Secondary 28A35, 52A40.

Key words and phrases: product measure, convex functional, entropy, non-negativity, tensorization property.

Research partially supported by MEiN Grant 1 PO3A 012 29.

$1/\varphi''$ concave, then the φ-entropy functional has the tensorization property, i.e. for any random variable Z defined on any product space $\Omega_1 \times \Omega_2$,

$$E\varphi(Z) - \varphi(EZ) \le E[(E_1\varphi(Z) - \varphi(E_1Z)) + (E_2\varphi(Z) - \varphi(E_2Z))],$$

or, equivalently,

$$\Psi_2(Z) = E\varphi(Z) - E_1\varphi(E_2Z) - E_2\varphi(E_1Z) + \varphi(EZ) \ge 0.$$

(The solution of a similar characterization problem, concerning hypercontractivity with some more general functionals instead of L_p norms, was given by Oleszkiewicz [6]). In fact, the paper [3] contains a rigorous proof only of the statement that if $\varphi \in \Phi$ then the φ-entropy functional

$$\Psi_1(Z) = E\varphi(Z) - \varphi(EZ) \text{ is convex.}$$

Later on, in [2] it was suggested that the convexity of Ψ_1 might not imply the non-negativity of Ψ_2 straightforwardly. Therefore in order to obtain the latter, a variational formula for Ψ_2 was used (established by Bobkov for some particular functions φ; see [4, Section 4]). However, this formula strongly relies on the analytic conditions that φ satisfies (namely, that $\varphi \in \Phi$).

In order to make the picture clear, we shall provide a direct argument that the convexity of Ψ_1 is equivalent to the non-negativity of Ψ_2 (Proposition 1). We also give the proof of the converse part of the characterization result (Theorem 1): if the φ-entropy has the tensorization property (in other words, φ belongs to the class C_2) then $\varphi \in \Phi$. Finally, Theorem 2 addresses the question posed at the end of [3], concerning a characterization of the higher "tensorization classes" C_n for $n > 2$.

2. Notation and definitions. Throughout the paper, d and n stand for positive integers, U denotes an open, convex subset of $\mathbb{R}^d$ and $\varphi\colon U \to \mathbb{R}$ is a continuous function. By $(\Omega, \mathcal{F}, P)$, $(\Omega_k, \mathcal{F}_k, P_k)$, etc. we shall denote probability spaces. In the case of the product space $(\Omega, \mathcal{F}, P) = \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, P_k)$, for $K \subset \{1, \dots, n\}$, E_K stands for the expectation with respect to the product measure $\bigotimes_{k\in K} P_k$. For $k \in \{1, \dots, n\}$ we shall write E_k instead of $E_{\{k\}}$.

For $V \subseteq \mathbb{R}^d$, when writing $Z\colon (\Omega, \mathcal{F}, P) \to V$, we mean that Z is a random variable taking values in $\mathbb{R}^d$ and $P(Z \in V) = 1$.

For fixed $U \subseteq \mathbb{R}^d$, $\varphi\colon U \to \mathbb{R}$ and fixed $(\Omega, \mathcal{F}, P) = \bigotimes_{k=1}^n (\Omega_k, \mathcal{F}_k, P_k)$ we shall consider the functional Ψ_n acting on random variables Z defined on $(\Omega, \mathcal{F}, P)$ with $P(Z \in V) = 1$ for some compact, convex set $V \subset U$, and defined by

$$\Psi_n(Z) = \sum_{K\subseteq\{1,\dots,n\}} (-1)^{|K|} E_{K^c}\varphi(E_K Z). \tag{1}$$

The definition of the main object we investigate in this paper originates in [3]:

Definition 1. We say that $\varphi \in C_n(U)$ iff the functional Ψ_n is non-negative for any $(\Omega, \mathcal{F}, P) = \bigotimes_{k=1}^{n} (\Omega_k, \mathcal{F}_k, P_k)$, i.e. for every compact, convex set $V \subset U$ and every $Z\colon (\Omega, \mathcal{F}, P) \to V$,

$$\Psi_n(Z) \geq 0.$$

Remark 1. It is obvious that $C_n(U)$ is a convex cone.

Remark 2. By slight abuse of notation, we can also define the functional Ψ_n inductively, as iterations of the φ-entropy functional $E\varphi(Z) - \varphi(EZ)$, namely

$$\Psi_n(Z) = E_n \Psi_{n-1}(Z) - \Psi_{n-1}(E_n Z). \tag{2}$$

(By $\Psi_{n-1}(Z)$ we mean the application of Ψ_{n-1} conditionally with the nth product coordinate fixed, whereas in $\Psi_{n-1}(E_n Z)$ we consider $E_n Z$ as a random variable defined on the product of all probability spaces except the nth). Now, it can be seen that the non-negativity of Ψ_n is tightly connected with the convexity of Ψ_{n-1}. A precise statement appears in Proposition 1 (equivalence of (i) and (ii′)).

Remark 3. The functional Ψ_n can be extended to a functional $\widetilde{\Psi}_n$ acting on a larger class of random variables whose values are not restricted almost surely to some compact subset of U. However, some integrability assumptions should be added to ensure that the right hand side of (1) is well-defined. It would be natural to assume that φ is convex, $E|Z| < \infty$ ($|\cdot|$ stands for Euclidean norm in $\mathbb{R}^d$) and $E|\varphi(Z)| < \infty$. Then Jensen's inequality implies that for each $K \subseteq \{1, \ldots, n\}$,

$$aE_K Z + b \leq \varphi(E_K Z) \leq E_K \varphi(Z) \quad \text{a.s.}$$

for some $a, b \in \mathbb{R}$. Since the lower and upper bounds are integrable with respect to E_{K^c}, each term in the sum (1) is well-defined and finite. As we shall see, in the context of the classes $C_n(U)$, the assumption that φ is convex is not restrictive at all. Moreover, an easy truncation argument will show that the non-negativity of $\widetilde{\Psi}_n$ is a consequence of the non-negativity of Ψ_n (see Proposition 1, equivalence of (i) and (iii)).

Example 1. Jensen's inequality implies that $C_1(U)$ contains exactly the convex functions on U.

Example 2. The class $C_2((0,\infty))$ is exactly the class of functions φ for which the subadditive φ-entropies are widely considered. The most important examples are $\varphi(x) = x^p$ for $p \in (1, 2]$ and $\varphi(x) = x \log(x)$. In the introduction we mentioned that $\Phi \subseteq C_2((0,\infty))$. In fact, we shall show that these two classes are equal (see Theorem 1).

3. Properties of the classes C_n. We start with a proposition giving some equivalent variants of the definition of the class C_n. The discrete cubes $\{-1,1\}^n_\lambda$ considered below are the n-fold products of the two-point probability space $\{-1,1\}$ endowed with the measure $\lambda\delta_1+(1-\lambda)\delta_{-1}$; if λ is omitted then it means that we take $\lambda=1/2$.

PROPOSITION 1. *The following assertions are equivalent*:

(i) $\varphi\in C_n(U)$,
(ii) *for every random variable* $Z\colon \{-1,1\}^n\to U$ *we have* $\Psi_n(Z)\ge 0$,
(ii′) *for every pair of random variables* $Z_1,Z_2\colon\{-1,1\}^{n-1}\to U$,

$$\frac{1}{2}\Psi_{n-1}(Z_1)+\frac{1}{2}\Psi_{n-1}(Z_2)\ge\Psi_{n-1}\left(\frac{Z_1+Z_2}{2}\right),$$

(iii) φ *is convex and for every* $(\Omega,\mathcal{F},P)=\bigotimes_{k=1}^n(\Omega_k,\mathcal{F}_k,P_k)$ *and every random variable* $Z\colon(\Omega,\mathcal{F},P)\to U$ *such that* $E|Z|<\infty$ *and* $E|\varphi(Z)|<\infty$ *we have* $\widetilde{\Psi}_n(Z)\ge 0$.

In the proof we shall use the following lemmas:

LEMMA 1. *Let* V *be a compact, convex subset of* $\mathbb{R}^d$ *and* $(\Omega,\mathcal{F},P)=(\Omega_1,\mathcal{F}_1,P_1)\otimes(\Omega_2,\mathcal{F}_2,P_2)$ *be a product probability space. For every* Z: $(\Omega,\mathcal{F},P)\to V$ *and every* $\varepsilon>0$ *there exists* $\widetilde{Z}\colon(\Omega,\mathcal{F},P)\to V$ *such that*

$$\widetilde{Z}=\sum_{i=1}^{M}\sum_{j=1}^{N}a_{ij}1_{A_i\times B_j},$$

where $a_{ij}\in V$ *and* $(A_i)_{i=1}^M,(B_j)_{j=1}^N$ *are measurable, finite partitions of* $(\Omega_1,\mathcal{F}_1,P_1)$ *and* $(\Omega_2,\mathcal{F}_2,P_2)$ *(respectively), and* $P(|\widetilde{Z}-Z|\ge\varepsilon)<\varepsilon$.

Proof. We take any $\varepsilon>0$ and any finite covering of V by (open) balls $U_i=B(a_i,\varepsilon)$ $(i=1,\ldots,L)$ such that $a_i\in V$. Then we take disjoint and measurable (with respect to $\mathcal{F}_1\otimes\mathcal{F}_2$) sets $C_i=Z^{-1}(U_i\setminus\bigcup_{j<i}U_j)$. Now we shall represent each C_i as a union of finitely many measurable product sets $A\times B$ in such a way that the measure of the symmetric difference of this union and C_i is small. Since $P_1\otimes P_2$ is the product measure, we can find countably many sets $A_{i,j}\in\mathcal{F}_1$ and $B_{i,j}\in\mathcal{F}_2$ $(j=1,2,\ldots)$ such that $C_i\subseteq\bigcup_{j=1}^\infty(A_{i,j}\times B_{i,j})$ and

$$P(C_i)+\varepsilon/L^2>\sum_{j=1}^{\infty}P_1(A_{i,j})P_2(B_{i,j}).$$

If we take m_i such that the tail of the above series for $j>m_i$ is less than ε/L^2 and put $\widetilde{C}_i=\bigcup_{j=1}^{m_i}(A_{i,j}\times B_{i,j})$, then

$$P(C_i \setminus \widetilde{C}_i) \le P\Big(\bigcup_{j>m_i}(A_{i,j}\times B_{i,j})\Big) < \varepsilon/L^2, \tag{3}$$

$$P(\widetilde{C}_i \setminus C_i) \le P\Big(\bigcup_{j=1}^{\infty}(A_{i,j}\times B_{i,j})\Big) - P(C_i) < \varepsilon/L^2. \tag{4}$$

We set

$$D_i = \widetilde{C}_i \setminus \bigcup_{i'\ne i}\widetilde{C}_{i'} \quad \text{for } i=1,\dots,L.$$

Obviously, the D_i are pairwise disjoint and each of them is a finite union of measurable product sets. Putting $D_0 = \Omega \setminus \sum_{i=1}^{L} D_i$ (which is also a finite union of product sets) and choosing an arbitrary $a_0 \in V$, we see that $\widetilde{Z} = \sum_{i=0}^{L} a_i 1_{D_i}$ has the desired form (to see this, take a joint subdivision of Ω_1 and Ω_2 generated by all (finitely many) product sets from $D_0, D_1, \dots, D_L$).

To finish the proof we show that $P(|\widetilde{Z}-Z|\ge\varepsilon)<\varepsilon$. For each i we have

$$\begin{aligned}\{|\widetilde{Z}-Z|\ge\varepsilon\}\cap C_i \subseteq C_i\setminus D_i &= (C_i\setminus\widetilde{C}_i)\cup\bigcup_{i'\ne i}(C_i\cap\widetilde{C}_{i'})\\ &\subseteq (C_i\setminus\widetilde{C}_i)\cup\bigcup_{i'\ne i}(\widetilde{C}_{i'}\setminus C_{i'}),\end{aligned}$$

since $C_i\cap\widetilde{C}_{i'} = (\widetilde{C}_{i'}\setminus C_{i'})\cap C_i \subseteq \widetilde{C}_{i'}\setminus C_{i'}$. Therefore for each $i=1,\dots,L$, (3) and (4) yield

$$P(\{|\widetilde{Z}-Z|\ge\varepsilon\}\cap C_i) \le P(C_i\setminus\widetilde{C}_i) + \sum_{i'\ne i} P(\widetilde{C}_{i'}\setminus C_{i'}) < \varepsilon/L. \ ■$$

LEMMA 2. *Let V be a compact, convex subset of $\mathbb{R}^d$ and $\varphi\colon V\to\mathbb{R}$ be a continuous function. If the sequence $Z_k\colon(\Omega_1,\mathcal{F}_1,P_1)\otimes(\Omega_2,\mathcal{F}_2,P_2)\to V$ converges in probability to Z then $E_1\varphi(E_2Z_k)\to E_1\varphi(E_2Z)$.*

Proof. Let $R>0$ satisfy $V\subseteq B(0,R)$. We take any $\varepsilon>0$ and k such that $P(|Z_k-Z|\ge\varepsilon)<\varepsilon$. Consider the measurable sets $A=\{|Z_k-Z|\ge\varepsilon\}\subseteq\Omega_1\times\Omega_2$ and $A_{\omega_1}=\{\omega_2\colon(\omega_1,\omega_2)\in A\}\subseteq\Omega_2$ for each $\omega_1\in\Omega_1$. By Fubini's theorem we get

$$\varepsilon > E1_A = \int_{\Omega_1} P_2(A_{\omega_1})\,P_1(d\omega_1) \ge \sqrt{\varepsilon}\,P_1(B),$$

where $B=\{\omega_1\colon P_2(A_{\omega_1})\ge\sqrt{\varepsilon}\}$ is a measurable subset of Ω_1, which yields $P_1(B)<\sqrt{\varepsilon}$. Now we write

$$\begin{aligned}|E_1\varphi(E_2Z_k)-E_1\varphi(E_2Z)| \le &\int_B |\varphi(E_2Z_k(\omega_1,\cdot)-\varphi(E_2Z(\omega_1,\cdot)|\,P_1(d\omega_1)\\ &+\int_{\Omega_1\setminus B}|\varphi(E_2Z_k(\omega_1,\cdot)-\varphi(E_2Z(\omega_1,\cdot)|\,P_1(d\omega_1),\end{aligned}$$

and the first term on the right hand side can be estimated by $2P_1(B)\sup_V|\varphi| < 2\sqrt{\varepsilon}\sup_V|\varphi|$. For $\omega_1 \notin B$,

$$\begin{aligned}|E_2Z_k(\omega_1,\cdot) - E_2Z(\omega_1,\cdot)| &\le E_2(\|Z_k\|_\infty + \|Z\|_\infty)1_{A_{\omega_1}} + \varepsilon E_2 1_{\Omega_2\setminus A_{\omega_1}} \\ &< 2R\sqrt{\varepsilon} + \varepsilon,\end{aligned}$$

and the uniform continuity of f yields

$$|E_1\varphi(E_2Z_k) - E_1\varphi(E_2Z)| < 2\sqrt{\varepsilon}\sup_V|\varphi| + \delta(2R\sqrt{\varepsilon}+\varepsilon) \to 0 \quad \text{as } \varepsilon \to 0,$$

where $\delta(\varepsilon)$ is the modulus of continuity of φ. ∎

Proof of Proposition 1. The implications (i)⇒(ii), (iii)⇒(i), (iii)⇒(ii) and (ii)⇔(ii′) are obvious. The proof of the implication (i)⇒(iii) is postponed until after the proof of Proposition 2. Now we prove (ii)⇒(i). It suffices to show that for any fixed compact, convex $V \subset U$ and any fixed $(\Omega_k, \mathcal{F}_k, P_k)$ (for $k = 1,\dots,n-1$),

$$\text{(5)}\qquad \begin{aligned}&\Psi_n(Z) \ge 0 \text{ for every } Z\colon \bigotimes_{k=1}^{n-1}(\Omega_k,\mathcal{F}_k,P_k)\otimes\{-1,1\} \to V \\ &\Rightarrow\ \Psi_n(Z)\ge 0 \text{ for every } (\Omega_n,\mathcal{F}_n,P_n) \text{ and } Z\colon \bigotimes_{k=1}^{n}(\Omega_k,\mathcal{F}_k,P_k) \to V,\end{aligned}$$

which means that the convexity of Ψ_{n-1} (even just 1/2-convexity) implies the non-negativity of Ψ_n. Applying this argument n times we get (i).

First note that the implication (5) holds for $(\Omega_n,\mathcal{F}_n,P_n) = \{-1,1\}_\lambda$ with $\lambda \in (0,1)$. Indeed, the hypothesis of (5) states that for any pair of random variables $Z_1, Z_2\colon \bigotimes_{k=1}^{n-1}(\Omega_k,\mathcal{F}_k,P_k) \to V$,

$$\text{(6)}\qquad \lambda\Psi_{n-1}(Z_1) + (1-\lambda)\Psi_{n-1}(Z_2) \ge \Psi_{n-1}(\lambda Z_1 + (1-\lambda)Z_2)$$

for $\lambda = 1/2$, hence also for any $\lambda = j_i2^{-i}$ $(0 < j_i < 2^i)$. Letting $\lambda_i \to \lambda$ we get (6) for any $\lambda \in [0,1]$, because $X_i := \lambda_iZ_1 + (1-\lambda_i)Z_2 \to \lambda Z_1 + (1-\lambda)Z_2 =: X$ a.s., so $E_KX_i \to E_KX$ a.s. (the sequence (X_i) is bounded a.s.) and also $E_{K^c}\varphi(E_KX_i) \to E_{K^c}\varphi(E_KX)$ (φ is continuous and bounded on V).

Now we show that $(\Omega_n,\mathcal{F}_n,P_n)$ can be an arbitrary probability space. Fix any $Z\colon \bigotimes_{k=1}^{n}(\Omega_k,\mathcal{F}_k,P_k) \to V$. Lemma 1 implies that for any $\varepsilon > 0$ we may take $\widetilde{Z}\colon \bigotimes_{k=1}^{n}(\Omega_k,\mathcal{F}_k,P_k) \to V$ such that $P(|\widetilde{Z} - Z| \ge \varepsilon) < \varepsilon$ and

$$\widetilde{Z}(\omega',\omega_n) = \sum_{j=1}^{N}\widetilde{Z}_j(\omega')1_{B_j}(\omega_n),$$

where $\widetilde{Z}_j\colon \bigotimes_{k=1}^{n-1}(\Omega_k,\mathcal{F}_k,P_k) \to V$, $\omega' \in \prod_{k=1}^{n-1}\Omega_k$, $(B_j)_{j=1}^N$ is a finite, measurable partition of $(\Omega_n,\mathcal{F}_n,P_n)$, and $\omega_n \in \Omega_n$. Then applying (6) $N-1$

times we get

$$E_n\Psi_{n-1}(\widetilde{Z}) = \sum_{j=1}^{N} P_n(B_j)\Psi_{n-1}(\widetilde{Z}_j) \ge \Psi_{n-1}\Big(\sum_{j=1}^{N} P_n(B_j)\Psi_{n-1}(\widetilde{Z}_j)\Big)$$
$$= \Psi_{n-1}(E_n\widetilde{Z}),$$

hence, due to (2), $\Psi_n(\widetilde{Z}) \ge 0$. Lemma 2 implies that $|E_{K^c}\varphi(E_K\widetilde{Z}) - E_{K^c}\varphi(E_K Z)|$ is small for each $K \subseteq \{1,\dots,n\}$, hence letting $\varepsilon \to 0$ we obtain $\Psi_n(Z) \ge 0$. ■

PROPOSITION 2. *$C_{n+1}(U) \subseteq C_n(U)$.*

Proof. Let $\varphi \in C_{n+1}(U)$. By Proposition 1 it is sufficient to show that $\Psi_n(Z) \ge 0$ for any Z defined on $\Omega = \{-1,1\}^n$ taking values in U. Define $\bar{Z}$ on the $(n+1)$-fold product $\{-1,1\}^n \times \Omega$ by

$$\bar{Z}(\varepsilon_1,\dots,\varepsilon_n,\bar{\varepsilon}) = Z(\varepsilon_1\bar{\varepsilon}_1,\dots,\varepsilon_n\bar{\varepsilon}_n),$$

where $\varepsilon_k \in \{-1,1\}$ and $\bar{\varepsilon} = (\bar{\varepsilon}_1,\dots,\bar{\varepsilon}_n) \in \Omega$. Since $\varphi \in C_{n+1}(U)$, we have $\Psi_{n+1}(\bar{Z}) = E_{n+1}\Psi_n(\bar{Z}) - \Psi_n(E_{n+1}\bar{Z}) \ge 0$. But $\Psi_n(\bar{Z}(\cdot,\bar{\varepsilon}))$ does not depend on the choice of $\bar{\varepsilon}$ and is equal to $\Psi_n(Z)$. Similarly $E_{n+1}\bar{Z}(\varepsilon_1,\dots,\varepsilon_n,\cdot)$ does not depend on ε_k and is equal to EZ, so we obtain $\Psi_{n+1}(\bar{Z}) = \Psi_n(Z)$. ■

Now we can finish the proof of Proposition 1:

Proof of Proposition 1, (i)⇒(iii). Fix any $\varphi \in C_n(U)$, $(\Omega,\mathcal{F},P) = \bigotimes_{k=1}^n(\Omega_k,\mathcal{F}_k,P_k)$ and $Z\colon (\Omega,\mathcal{F},P) \to U$ such that $E|Z| < \infty$ and $E|\varphi(Z)| < \infty$. Proposition 2 implies that $\varphi \in C_1(U)$, i.e. φ is convex. Take any increasing sequence of compact, convex subsets $V_i \subset U$ such that $\bigcup_i V_i = U$, and fix $v_0 \in V_1$. Then we define

$$Z_i = Z1_{Z\in V_i} + v_0 1_{Z\notin V_i},$$

which converges to Z a.s. We shall prove that

$$E_{K^c}\varphi(E_K Z_i) \to E_{K^c}\varphi(E_K Z), \tag{7}$$

which obviously implies that $\Psi_n(Z_i) \to \widetilde{\Psi}_n(Z)$. Since $|Z_i| \le |Z| + |v_0|$ and $E_K|Z| < \infty$ a.s., Lebesgue's dominated convergence theorem implies that $E_K Z_i \to E_K Z$ a.s. and by continuity of φ also $\varphi(E_K Z_i) \to \varphi(E_K Z)$ a.s. The convexity of φ yields

$$aE_K Z_i + b \le \varphi(E_K Z_i) \le E_K\varphi(Z_i)$$

for some $a,b \in \mathbb{R}$. Since $E_K\varphi(Z_i) \le E_K|\varphi(Z)| + \varphi(v_0)$ and $|aE_K Z_i + b| \le |a|(E_K|Z| + |v_0|) + |b|$ and both upper bounds are integrable with respect to E_{K^c}, Lebesgue's theorem applied once again gives $E_{K^c}\varphi(E_K Z_i) \to E_{K^c}\varphi(E_K Z)$. ■

From now on, we shall write Ψ_n, even if we really mean the extension $\widetilde{\Psi}_n$.

We should mention that e.g. in the case of the class $C_2((0,\infty))$ one may have $\Psi_2(Z) \geq 0$ not only for $Z > 0$ a.s., but also for Z having an atom at 0, as long as φ can be extended continuously to $[0,\infty)$ (cf. Example 2). Generally, we can state the following

REMARK 4. If $\varphi\colon U \to \mathbb{R}$ extends continuously to $\overline{\varphi}\colon \overline{U} \to \mathbb{R}$, then $\varphi \in C_n(U)$ implies that $\Psi_n(Z) \geq 0$ for every random variable Z defined on an n-fold product space and taking values in $\overline{U}$ and satisfying $E|Z| < \infty$ and $E|\overline{\varphi}(Z)| < \infty$. (More precisely, Ψ_n here is a natural extension of the functional (1).) Indeed, since $\varphi \in C_1(U)$, $\overline{\varphi}$ is also convex. Fixing $v_0 \in U$ and defining $Z_\varepsilon = Z1_{\{Z\notin\partial U\}} + ((1-\varepsilon)Z + \varepsilon v_0)1_{\{Z\in\partial U\}}$ for $\varepsilon \in (0,1)$ we obtain random variables Z_ε with values in U converging to Z a.s. The proof that $\Psi_n(Z_\varepsilon) \to \Psi_n(Z)$ as $\varepsilon \to 0$ is the same as in the case of (7).

THEOREM 1. *Let $U = (a,b) \subseteq \mathbb{R}$ be an open interval (possibly with $a = -\infty$ or $b = \infty$) and let $\varphi\colon U \to \mathbb{R}$ be a continuous function. Then $\varphi \in C_2(U)$ iff φ is an affine function or φ is twice differentiable with $\varphi'' > 0$ and $1/\varphi''$ is concave.*

Proof. The "if" part appears in [3] (in fact, for $a = 0$ and $b = \infty$, but it also works for any $a < b$). More precisely, it was proved there that Ψ_1 is convex. But this means that assertion (ii′) from Proposition 1 is satisfied, and so also is (i).

We now show the converse implication. First assume that $\varphi \in C_2(U)\cap\mathcal{C}^2$. In this case we follow the idea of [3, Lemma 3]. Consider $F\colon U \times U \to \mathbb{R}$ defined by

$$F(x,y) = \frac{\varphi(x)+\varphi(y)}{2} - \varphi\left(\frac{x+y}{2}\right).$$

If a random variable $Z\colon \{-1,1\} \to U$ attains two values x and y then $\Psi_1(Z) = F(x,y)$. Therefore Proposition 1 ((i)⇒(ii′)) implies that F is convex. Since F is $\mathcal{C}^2$, D^2F is non-negative definite. Thus

$$\frac{\partial^2 F}{\partial x^2}(x,y) = \frac{1}{2}\,\varphi''(x) - \frac{1}{4}\,\varphi''\left(\frac{x+y}{2}\right) \geq 0.$$

Since $\varphi \in C_2(U) \subseteq C_1(U)$, we have $\varphi'' \geq 0$ and the above easily implies that if $\varphi''(x_0) = 0$ for some $x_0 \in U$, then also $\varphi''(x) = 0$ for $x \in ((a+x_0)/2, (b+x_0)/2)$. Applying this argument inductively we get $\varphi'' \equiv 0$, i.e. φ is affine. So further we assume $\varphi'' > 0$. The non-negativity of D^2F implies that

$$\frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2} \geq \frac{\partial^2 F}{\partial x \partial y}$$

and one easily checks that this is equivalent to the concavity of $1/\varphi''$ considered at the points x, y and $(x+y)/2$.

Now we show that the assumption $\varphi \in C_2(U)$ implies that $\varphi \in \mathcal{C}^2$. For $\varepsilon > 0$ let $U^\varepsilon = (a+\varepsilon, b-\varepsilon)$ and define $\varphi_\varepsilon\colon U^\varepsilon \to \mathbb{R}$ as the convolution $\varphi_\varepsilon = \varphi * \eta_\varepsilon$, where $\eta_\varepsilon \geq 0$ is a smooth approximation of δ_0 with $\operatorname{supp}(\eta_\varepsilon) \subseteq (-\varepsilon, \varepsilon)$. Since $C_2(U)$ is a convex cone, $\varphi_\varepsilon \in C_2(U^\varepsilon)$.

Since φ_ε is smooth, the first part of the proof implies that φ_ε is either affine, or has a strictly positive second derivative with $1/\varphi_\varepsilon''$ concave. Then it is easy to see that φ_ε'' is a convex function. Indeed, the affine case is obvious, and if $\varphi_\varepsilon'' > 0$ then the concavity of $1/\varphi_\varepsilon''$ considered at the points x, y and $(x+y)/2$ gives

$$\varphi_\varepsilon''\left(\frac{x+y}{2}\right) \leq \frac{2\varphi_\varepsilon''(x)\varphi_\varepsilon''(y)}{\varphi_\varepsilon''(x)+\varphi_\varepsilon''(y)} \leq \frac{\varphi_\varepsilon''(x)+\varphi_\varepsilon''(y)}{2}.$$

Therefore $\varphi_\varepsilon'' \geq 0$ and for some $x_0 \in \mathbb{R}$, φ_ε'' is non-increasing on $(-\infty, x_0] \cap U$ and non-decreasing on $[x_0, \infty) \cap U$, so φ_ε' is a non-decreasing, concave-convex function.

First we show that $\varphi \in \mathcal{C}^1$. Since $\varphi \in C_2(U) \subseteq C_1(U)$, φ is convex, so it is well-known that φ has a first derivative on a set $\mathcal{D}_\varphi$ with $\mathcal{ND}_\varphi = U \setminus \mathcal{D}_\varphi$ countable (so $\mathcal{ND}_\varphi$ is of zero Lebesgue measure and $\mathcal{D}_\varphi$ is dense in U). Moreover, φ' is continuous at all points of $\mathcal{D}_\varphi$ and φ is locally Lipschitz. Therefore Lebesgue's dominated convergence theorem yields

$$\text{(8)} \qquad \begin{aligned} \varphi_\varepsilon'(x) &= \lim_{h\to 0} \int \frac{\varphi(x-y+h)-\varphi(x-y)}{h}\, \eta_\varepsilon(y)\, dy \\ &= (\varphi' * \eta_\varepsilon)(x) \quad \text{ for } x \in U^\varepsilon \end{aligned}$$

(φ' is defined a.e.). Taking $\varepsilon \to 0$, by continuity of φ' in $\mathcal{D}_\varphi$,

$$\text{(9)} \qquad \lim_{\varepsilon\to 0} \varphi_\varepsilon'(x) = \varphi'(x) \quad \text{ for } x \in \mathcal{D}_\varphi.$$

Now fix any decreasing sequence $\varepsilon_k \to 0$ ($k = 0, 1, \ldots$) and think of ε_0 as small. Below we consider the φ_{ε_k} defined on one domain U^{ε_0}. The functions φ_{ε_k}' are non-decreasing and concave-convex and they pointwise converge on the dense set $U^{\varepsilon_0} \cap \mathcal{D}_\varphi$. This implies that they are also uniformly equicontinuous on any compact interval $[a_0, b_0] \subset U^{\varepsilon_0}$. Indeed, taking any $a_i, b_i \in U^{\varepsilon_0} \cap \mathcal{D}_\varphi$ ($i = 1, 2$) such that $a_1 < a_2 \leq a_0$ and $b_0 \leq b_1 < b_2$, we see that for sufficiently large k the Lipschitz constant of φ_{ε_k}' is less than

$$\max\left(\frac{\varphi'(a_2)-\varphi'(a_1)+1}{a_2-a_1}, \frac{\varphi'(b_2)-\varphi'(b_1)+1}{b_2-b_1}\right).$$

Therefore the Arzelà–Ascoli theorem implies that there exists a subsequence ε_{k_l} such that $\varphi_{\varepsilon_{k_l}}'$ converges uniformly on $[a_0, b_0]$ to some continuous function, which has to be the derivative of φ. Letting $\varepsilon_0 \to 0$ and $a_0 \to a$, $b_0 \to b$ we get $\varphi \in \mathcal{C}^1$. Moreover, φ' is also a non-decreasing, concave-convex function.

The proof that $\varphi \in \mathcal{C}^2$ is similar. The equality (8) gives $\varphi_\varepsilon'' = (\varphi' * \eta_\varepsilon)'$ and (9) applied for φ' instead of φ (this is justified since φ' is a concave-convex function and all the facts concerning the derivative of φ' and the set $\mathcal{D}_{\varphi'}$ hold true as in the case of a convex function) yields

$$\varphi''_\varepsilon(x) = (\varphi' * \eta_\varepsilon)'(x) \to \varphi''(x) \quad \text{for } x \in \mathcal{D}_{\varphi'}.$$

Now using the fact that φ''_ε is convex, a similar argument shows that the convex functions $\varphi''_{\varepsilon_k}$ are uniformly equicontinuous on compact intervals. As a consequence, some subsequence $\varphi''_{\varepsilon_{k_l}}$ is uniformly convergent on compact intervals to some continuous function, which has to be the derivative of φ'. ∎

THEOREM 2. *Let $U \subseteq \mathbb{R}^d$ be an open, convex set. Then for all $n \geq 3$,*

$$C_n(U) = \{\varphi\colon U \to \mathbb{R} \mid \varphi(x) = Q(x) + v^*(x) + c\},$$

where Q is a non-negative definite quadratic form on $\mathbb{R}^d$, v is a linear functional on $\mathbb{R}^d$ and $c \in \mathbb{R}$.

Proof. The inclusion $\supseteq$ is easy. Since the expectation commutes with v^*, we can assume $\varphi(x) = Q(x)$. Moreover, we can take $U = \mathbb{R}^d$, because if $\varphi \in C_n(U)$ and $U' \subseteq U$ then $\varphi_{|U'} \in C_n(U')$.

We show that if $\varphi(x) = Q(x)$ is a quadratic form then

$$\Psi_n(Z) = \Psi_n(Z - E_n Z). \tag{10}$$

Indeed, denote by $Q(x,y)$ the bilinear form associated with $Q(x)$; then (2) yields

$$\begin{aligned}
&\Psi_n(Z - E_nZ) = E_n\Psi_{n-1}(Z - E_nZ) - \Psi_{n-1}(0)\\
&= E_n \sum_{K\subseteq\{1,\ldots,n-1\}} (-1)^{|K|} E_{K^c} Q(E_K(Z - E_nZ))\\
&= \sum_{K\subseteq\{1,\ldots,n-1\}} (-1)^{|K|} E_{K^c}E_n(Q(E_KZ) - 2Q(E_KZ, E_{K\cup\{n\}}Z) + Q(E_{K\cup\{n\}}Z))\\
&= \sum_{K\subseteq\{1,\ldots,n-1\}} (-1)^{|K|} E_{K^c}(E_nQ(E_KZ) - 2Q(E_nE_KZ, E_{K\cup\{n\}}Z)\\
&\qquad\qquad + Q(E_{K\cup\{n\}}Z))\\
&= \sum_{K\subseteq\{1,\ldots,n-1\}} (-1)^{|K|} (E_{K^c\cup\{n\}}Q(E_KZ) - Q(E_{K\cup\{n\}}Z)) = \Psi_n(Z).
\end{aligned}$$

Now, by induction on n, we prove that $\Psi_n \geq 0$, i.e. $Q \in C_n(\mathbb{R}^d)$. Obviously, $\Psi_1 \geq 0$. Then the formulas (10) and (2) imply that

$$\Psi_n(Z) = \Psi_n(Z - E_nZ) = E_n\Psi_{n-1}(Z - E_nZ) - \Psi_{n-1}(0) \geq 0,$$

since by the induction hypothesis $\Psi_{n-1}(Z - E_nZ) \geq 0$ a.s.

The inclusion $\subseteq$ is more tricky. First, Proposition 2 allows us to consider the case $n = 3$ only. The argument presented below is due to K. Oleszkiewicz and is reproduced here with his kind permission. (The author's argument was a bit more complicated and was not so general—it worked e.g. for $U = (0,\infty) \subseteq \mathbb{R}$ but not for finite intervals).

First, assume that $\varphi \in C_3(U)$ is ($\mathcal{C}^\infty$) smooth. We define $X\colon \{-1,1\}^3 \to \mathbb{R}$ by

$$X(\varepsilon_1,\varepsilon_2,\varepsilon_3) = \begin{cases} 3 & \text{if } |\varepsilon_1+\varepsilon_2+\varepsilon_3| = 3, \\ -1 & \text{otherwise.} \end{cases}$$

Fix $a \in U$ and $v \in \mathbb{R}^d$. For $\varepsilon \in \mathbb{R}$, we define $Z_\varepsilon = a + v\varepsilon X$. If $|\varepsilon|$ is sufficiently small, Z_ε has values in U. The hypothesis implies that $\Psi_3(Z_\varepsilon) \geq 0$. On the other hand, if we put $f(x) = \varphi(a+vx)$ for x from some open interval containing 0, we obtain

$$\Psi_3(Z_\varepsilon) = \sum_{K\subseteq\{1,2,3\}} (-1)^{|K|} E_{K^c} f(\varepsilon E_K X) = \frac{1}{4} f(3\varepsilon) - \frac{3}{2} f(\varepsilon) + 2f(0) - \frac{3}{4} f(-\varepsilon). \tag{11}$$

Notice that the right hand side vanishes if we take 1, x or x^2 as $f(x)$, and is equal to 6 for $f(x) = x^3$. Since f is smooth, applying Taylor's expansion $f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \frac{1}{6}f'''(0)x^3 + o(x^3)$ to (11) we obtain

$$\lim_{\varepsilon\to 0} \frac{\Psi_3(Z_\varepsilon)}{\varepsilon^3} = f'''(0).$$

Since $\Psi_3(Z_\varepsilon)/\varepsilon^3 \geq 0$ for $\varepsilon > 0$ and $\Psi_3(Z_\varepsilon)/\varepsilon^3 \leq 0$ for $\varepsilon < 0$, we obtain $f'''(0) = 0$, hence $D^3_{v,v,v}\varphi(a) = 0$ for any $v \in \mathbb{R}^d$ and $a \in U$, so $D^3\varphi \equiv 0$. An elementary reasoning shows that φ is of the desired form—we leave the details to the reader. (A similar result dealing with functions on an infinite-dimensional vector space was given e.g. in [5]. That result says that if a function restricted to any line is a one-variable polynomial of degree at most k, then the whole function is a polynomial of degree at most k.)

The general case (without assuming φ to be smooth) follows easily from the above. For $\varepsilon > 0$, we define

$$U^\varepsilon = \{x \in U\colon \overline{B}(x,\varepsilon) \subseteq U\}.$$

Clearly, U^ε is an open, convex subset of U. Define $\varphi_\varepsilon\colon U^\varepsilon \to \mathbb{R}$ as the convolution $\varphi_\varepsilon = \varphi * \eta_\varepsilon$, where $\eta_\varepsilon \geq 0$ is a smooth approximation of δ_0 with $\operatorname{supp}(\eta_\varepsilon) \subseteq B(0,\varepsilon)$. Since $C_3(U)$ is a convex cone, $\varphi_\varepsilon \in C_3(U^\varepsilon)$ and so φ_ε is a "quadratic function". Passing to the limit we conclude that so also is φ. ∎

The following proposition states what the "tensorization property" for the classes $C_n(U)$ means.

Proposition 3. *Let $\varphi \in C_{n+1}(U)$ $(n \geq 1)$. Let μ_k^0 and μ_k^1 for $k = 1,\ldots,n$ be probability measures. Then for any $Z\colon \bigotimes_{k=1}^n(\mu_k^0 \otimes \mu_k^1) \to U$ such* *that* $E|Z| < \infty$ *and* $E|\varphi(Z)| < \infty$ *we have*

$$\Psi_n(Z) \le E \sum_{A\subseteq\{1,\dots,n\}} \Psi_n^A(Z),$$

where $\Psi_n^A(Z)$ *means the functional* Ψ_n *applied to* Z *considered as a random variable defined on the product* $\bigotimes_{k=1}^n \mu_k^{I_A(k)}$ *with all coordinates* $\omega_k^{1-I_A(k)}$ *fixed.*

Proof. We shall prove that for $Z\colon (\mu_1^0\otimes\mu_1^1)\otimes\mu_2\otimes\cdots\otimes\mu_n \to U$ (satisfying appropriate integrability conditions) one has

$$\Psi_n(Z) \le E(\Psi_n^0(Z) + \Psi_n^1(Z)),$$

where $\Psi_n^0(Z)$ means Ψ_n applied to Z considered as a random variable defined on the product $\mu_1^0 \otimes \mu_2 \otimes \cdots \otimes \mu_n$ with ω_1^1 fixed (and similarly for $\Psi_n^1(Z)$). Labelling the product coordinates $\omega_1^0, \omega_1^1, \omega_2, \dots, \omega_n$ as $1^0, 1^1, 2, \dots, n$ respectively we have

$$\Psi_n(Z) = \sum_{\substack{K\subset\{1^0,1^1,2,\dots,n\}\\ |K\cap\{1^0,1^1\}|\neq 1}} (-1)^{|K|} E_{K^c}\varphi(E_K Z),$$

$$E\Psi_n^0(Z) = \sum_{K\subset\{1^0,2,\dots,n\}} (-1)^{|K|} E_{\{1^1\}\cup K^c}\varphi(E_K Z),$$

$$E\Psi_n^1(Z) = \sum_{K\subset\{1^1,2,\dots,n\}} (-1)^{|K|} E_{\{1^0\}\cup K^c}\varphi(E_K Z),$$

and we easily check that $E\Psi_n^0(Z) + E\Psi_n^1(Z) - \Psi_n(Z) = \Psi_{n+1}(Z)$.

Now observe that it suffices to apply the above argument recursively. ∎

Acknowledgments. I would like to thank Prof. Krzysztof Oleszkiewicz for introducing me to the subject and all the inspiring suggestions including his elegant arguments reproduced in the proof of Theorem 2. I am also indebted to Prof. Stanisław Kwapień for several comments which allowed me to improve the text of this paper.

References

[1] W. Beckner, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc. 105 (1989), 397–400.

[2] S. Boucheron, O. Bousquet, G. Lugosi and P. Massart, *Moment inequalities for functions of independent random variables*, Ann. Probab. 33 (2005), 514–560.

[3] R. Latała and K. Oleszkiewicz, *Between Sobolev and Poincaré*, in: Geometric Aspects of Functional Analysis, Lecture Notes in Math. 1745, Springer, Berlin, 2000, 147–168.

[4] M. Ledoux, *On Talagrand's deviation inequalities for product measures*, ESAIM Probab. Statist. 1 (1997), 63–87.

[5] S. Mazur und W. Orlicz, *Grundlegende Eigenschaften der polynomischen Operationen. Erste Mitteilung*, Studia Math. 5 (1934), 50–68.

[6] K. Oleszkiewicz, *Generalization of the hypercontraction method and the Khinchin–Kahane inequalities*, Master's Thesis, Warsaw Univ., 1994 (in Polish).

Paweł Wolff
Institute of Mathematics
Warsaw University
Banacha 2
02-097 Warszawa, Poland
E-mail: pwolff@mimuw.edu.pl

Received January 26, 2007;
received in final form June 14, 2007 (7582)