Real Analysis Exchange
Vol. (), , pp. 541–552

August M. Zapała, Department of Mathematics and Natural Sciences
Catholic University of Lublin, Aleje Racławickie 14, 20-950 Lublin, Poland.
e-mail: `august.zapala@kul.lublin.pl`

JENSEN'S INEQUALITY FOR CONDITIONAL EXPECTATIONS IN BANACH SPACES

Abstract

In this note we present a simple proof of the inequality $\Phi\left(E^{\mathcal{A}}\xi\right) \leq E^{\mathcal{A}}\Phi(\xi)$ *a.s.* for separable random elements $\xi \in L_1(\Omega, \mathcal{F}, P; X)$ in a Banach space X, where $E^{\mathcal{A}}(\cdot)$ denotes conditional expectation with respect to the σ-field $\mathcal{A} \subset \mathcal{F}$, and $\Phi : X \to \mathbb{R}$ is a convex functional satisfying certain additional assumptions which are less restrictive than known till now. Some consequences of the above result are also discussed; e.g., it is shown that if ξ is a Gaussian random element in X, then there exists a constant $0 < c < \infty$ such that for each σ-field $\mathcal{A}_0 \subset \mathcal{F}$ the family $\left\{\exp\{c\left\|E^{\mathcal{A}}\xi\right\|^2\} : \mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{F}\right\}$ is uniformly integrable.

1 Introduction.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{A} \subset \mathcal{F}$ be a sub-σ-field of the σ-field $\mathcal{F}$ and let X be a Banach space. Denote by $L_p(\mathcal{A}; X)$, $1 \leq p \leq \infty$, the space of equivalence classes of separable $\mathcal{A}$-measurable Borel random elements $\xi : \Omega \to X$ such that

$$\|\xi\|_p = \left\{\int_\Omega \|\xi\|^p \, dP\right\}^{1/p} < \infty \text{ for } 1 \leq p < \infty$$

and

$$\|\xi\|_\infty = \operatorname*{ess\,sup}_{\omega\in\Omega} \|\xi(\omega)\| < \infty \text{ for } p = \infty.$$

Key Words: conditional expectation, Jensen's inequality, convex function, Gaussian random element.
Mathematical Reviews subject classification: 26D07, 60E15, 28C05, 28C20, 46G12, 60B11
Received by the editors January 14, 1997

It is fairly well-known that there exists a continuous linear operator $E^{\mathcal{A}}$ acting from $L_1(\mathcal{F};X)$ to $L_1(\mathcal{A};X)$ such that $\int_A E^{\mathcal{A}}\xi\,dP = \int_A \xi\,dP$ for all $\xi \in L_1(\mathcal{F};X)$ and arbitrary $A \in \mathcal{A}$. The random element $E^{\mathcal{A}}\xi \in L_1(\mathcal{A};X)$ is called conditional expectation of $\xi \in L_1(\mathcal{F};X)$ with respect to the σ-field $\mathcal{A}$ and it is defined uniquely as an element of $L_1(\mathcal{A};X)$; i.e., as a random element in X uniquely up to sets $A \in \mathcal{A}$ of P-measure zero.

Conditional expectations of random elements in a Banach space possess similar properties to that of real valued random variables. The following interesting feature of this notion is worth mentioning here. If $T : X \to Y$ is a continuous linear operator acting on X into another Banach space Y, then $E^{\mathcal{A}}(T\xi) = T\left(E^{\mathcal{A}}\xi\right)$ *a.s.* for every $\xi \in L_1(\mathcal{F};X)$. In particular, if $x^* \in X^*$ and $\xi \in L_1(\mathcal{F};X)$, then

$$E^{\mathcal{A}}(x^*\xi) = x^*\left(E^{\mathcal{A}}\xi\right) \quad \text{in} \quad L_1(\mathcal{A};\mathbb{R}). \tag{1}$$

The aim of this note is to present a simple proof of Jensen′s inequality for conditional expectations in a Banach space. In various monographs and survey articles devoted to conditional expectations and martingales such a property is either merely mentioned, cf. Vahania, Tarieladze and Chobanyan (1985), Ch. II, § 4, or even quite omitted, as in Diestel and Uhl (1977), Metivier and Pellaumail (1980), or Woyczyński (1978). Vahania, Tarieladze and Chobanyan in Ch. II, § 4, of their monograph formulated without proof the following result. If $\xi \in L_1(\mathcal{F};X)$ and $\Phi : X \to \mathbb{R}$ is a continuous convex functional in a Banach space X such that $\Phi(\xi) \in L_1(\mathcal{F};X)$, then

$$\Phi\left(E^{\mathcal{A}}\xi\right) \le E^{\mathcal{A}}\Phi(\xi) \ \textit{a.s.} \tag{2}$$

We consider a convex functional Φ on arbitrary convex closed separable subset $K \subset X$ and we do not assume that Φ is continuous, but only lower or upper semi–continuous. It is shown in § 3 that under these conditions (2) remains true. Next the examples of convex semi–continuous but discontinuous functionals $\Phi : K \to \mathbb{R}$ are given. Moreover, a few applications of the conditional Jensen's inequality are presented.

2 Preliminary Result.

The auxiliary result below is a conditional analogue of the similar statement for the usual expectation in a Banach space. We include it here for future reference, but it may be also of independent interest.

Lemma. *Let $\xi \in L_1(\mathcal{F}; X)$ and let $A \subseteq X$ be a separable subset of a Banach space X, such that $\xi \in A$ a.s. Then, for an arbitrary σ-field $\mathcal{A} \subset \mathcal{F}$,*

$$E^{\mathcal{A}}\xi \in \overline{\operatorname{conv} A} \ a.s.,$$

where $\overline{\operatorname{conv} A}$ is the closed convex hull of the set A.

Proof. Let $K = \overline{\operatorname{conv} A}$ and observe that under our assumptions K is a separable subset of X. In fact, finite rational linear combinations of points taken from the separability set $A_s \subset A$ form a countable dense subset of K. Thus, assume that $\{x_1, x_2, ...\}$ is a denumerable dense subset of K. Let

$$B_{n,1} = \{x \in K : \|x - x_1\| < 1/n\},$$
$$B_{n,k} = \{x \in K : \|x - x_j\| \geq 1/n,\ j = 1 \dots k-1,\ \|x - x_k\| < 1/n\} \text{ for } k > 1,$$

and $f_n(x) = \sum_k x_k \mathcal{J}\{x \in B_{n,k}\}$, $n \geq 1$. Then $B_{n,k}$, $k \geq 1$ are mutually disjoint, $\bigcup_k B_{n,k} = K$ for each $n \geq 1$, and $\|f_n(x) - x\| < 1/n$, $x \in K$. Consequently, $\|f_n(\xi(\omega)) - \xi(\omega)\| \leq 1/n$ for all $\omega \in \Omega$, whereas $f_n(\xi(\omega)) = \sum_k x_k \mathcal{J}\{\xi(\omega) \in B_{n,k}\}$ are elementary random elements in X. Furthermore,

$$\|f_n(\xi(\omega))\| \leq \|f_n(\xi(\omega)) - \xi(\omega)\| + \|\xi(\omega)\| \leq 1/n + \|\xi(\omega)\|,$$

so that $f_n(\xi(\omega)) \in L_1(\mathcal{F}; X)$, and in addition $f_n(\xi(\omega)) \in K$ *a.s.*

Since $E^{\mathcal{A}} : L_1(\mathcal{F}; X) \to L_1(\mathcal{A}; X)$ is a bounded linear operator with norm 1 (cf. Vahania, Tarieladze and Chobanyan, Ch. II, § 4, Prop. 4.1, p. 108, or Diestel and Uhl, Ch. V, § 1, Th. 4, p. 123), we have

$$\begin{aligned}\left\|E^{\mathcal{A}} f_n(\xi) - E^{\mathcal{A}}\xi\right\|_1 \leq \|f_n(\xi) - \xi\|_1 &= \int_\Omega \|f_n(\xi) - \xi\| \, dP \\ &\leq 1/n \to 0 \text{ as } n \to \infty;\end{aligned}$$

i.e., $E^{\mathcal{A}} f_n(\xi) \to E^{\mathcal{A}}\xi$ in $L_1(\mathcal{A}; X)$. Selecting from $\left\{E^{\mathcal{A}} f_n(\xi)\right\}$ a suitable subsequence $\left\{E^{\mathcal{A}} f_{n'}(\xi)\right\}$ that is convergent with probability 1 to $E^{\mathcal{A}}\xi$, we infer that the relation $E^{\mathcal{A}} f_n(\xi) \in K$ *a.s.* implies $E^{\mathcal{A}}\xi = \lim_{n'} E^{\mathcal{A}} f_{n'}(\xi) \in K$ *a.s.* Therefore it suffices to prove our lemma for an elementary random element $\eta = \sum_k x_k \mathcal{J}_{A_k} \in L_1(\mathcal{F}; X)$ such that $x_k \in K$, $k \geq 1$, where $A_k \in \mathcal{F}$ are any disjoint random events, $\bigcup_k A_k = \Omega$.

To this end, let $\eta = \sum_{k \geq 1} x_k \mathcal{J}_{A_k} \in L_1(\mathcal{F}; X)$, where $x_k \in K$ and A_k are as above, and let $x_0 \in K$ be fixed arbitrarily. Put $A_0 = A_0^{(n)} = \Omega \setminus \bigcup_{k=1}^n A_k$ and notice that $\eta_n(\omega) = \sum_{k=0}^n x_k \mathcal{J}_{A_k}(\omega) \to \eta(\omega)$ as $n \to \infty$ for each $\omega \in \Omega$. Indeed, if $\omega \in \Omega$, then $\omega \in A_r$ for some $r = r(\omega) \geq 1$, and so

$$\|\eta_n(\omega) - \eta(\omega)\| = \left\| \sum_{k=0}^n x_k \mathcal{J}_{A_k}(\omega) - \sum_{k\geq 1} x_k \mathcal{J}_{A_k}(\omega) \right\| = \|x_r - x_r\| = 0$$

whenever $n \geq r$. Moreover,

$$E^{\mathcal{A}} \eta_n = \sum_{k=0}^n x_k E^{\mathcal{A}} \mathcal{J}_{A_k} = \sum_{k=0}^n x_k P^{\mathcal{A}}[A_k] \in K \ \textit{a.s.},$$

because the right side is a finite convex linear combination of $x_0, x_1, ..., x_n \in K$. Since $\eta \in L_1(\mathcal{F}; X)$, we conclude that

$$\|\eta\|_1 = \int_\Omega \left\| \sum_{k\geq 1} x_k \mathcal{J}_{A_k} \right\| dP = \sum_{k\geq 1} \|x_k\| \, P[A_k] < \infty.$$

Thus

$$\begin{aligned}
\left\|E^{\mathcal{A}}\eta_n - E^{\mathcal{A}}\eta\right\|_1 \leq \|\eta_n - \eta\|_1 &= \int_\Omega \left\| \sum_{k=0}^n x_k \mathcal{J}_{A_k} - \sum_{k\geq 1} x_k \mathcal{J}_{A_k} \right\| dP \\
&= \sum_{k>n} \|x_0 - x_k\| \, P[A_k] \\
&\leq \|x_0\| \sum_{k>n} P[A_k] + \sum_{k>n} \|x_k\| \, P[A_k] \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Finally, choosing from $\left\{E^{\mathcal{A}}\eta_n\right\}$ an appropriate subsequence $\left\{E^{\mathcal{A}}\eta_{n'}\right\}$ convergent *a.s.* to $E^{\mathcal{A}}\eta$ we obtain $E^{\mathcal{A}}\eta = \lim_{n'} E^{\mathcal{A}}\eta_{n'} \in K$ *a.s.* □

3 The Main Theorem.

Let $K \subseteq X$ be a convex subset of a Banach space X. By analogy to the real case, a function $\Phi : K \to \mathbb{R}$ is called *convex*, if

$$\bigwedge_{x,y\in K} \ \bigwedge_{\substack{0\leq \alpha,\beta\in\mathbb{R} \\ \alpha+\beta=1}} \qquad \Phi\left(\alpha x + \beta y\right) \leq \alpha \Phi\left(x\right) + \beta \Phi\left(y\right).$$

Suppose now that $K \subseteq X$ is closed. The mapping $\Phi : K \to \mathbb{R}$ is said to be *upper semi-continuous*, if

$$\bigwedge_{x\in K} \ \bigwedge_{\substack{\{x_n\}\subset K \\ x_n \to x}} \qquad \limsup_n \Phi\left(x_n\right) \leq \Phi\left(x\right),$$

and it is called *lower semi-continuous*, if

$$\bigwedge_{x\in K} \ \bigwedge_{\substack{\{x_n\}\subset K \\ x_n \to x}} \qquad \Phi\left(x\right) \leq \liminf_n \Phi\left(x_n\right).$$

Theorem. *Let $K \subseteq X$ be a convex, closed, separable subset of a Banach space X with a nonempty interior $K^o \neq \emptyset$, such that $\xi \in K$ a.s., and let $\Phi : K \to \mathbb{R}$ be an upper or lower semi-continuous convex functional. If $\Phi(\xi) \in L_1(\mathcal{F};\mathbb{R})$, where $\xi \in L_1(\mathcal{F};X)$, then*

$$\Phi\left(E^{\mathcal{A}}\xi\right) \leq E^{\mathcal{A}}\Phi(\xi) \ \ a.s. \tag{3}$$

PROOF. Since ξ is a separable random element, we can restrict further arguments to a closed separable linear subspace $X_0 \subseteq X$. However, to simplify the notation we shall write X instead of X_0 and assume that X is separable. Consider the product space $X \times \mathbb{R}$ equipped with the usual Tychonov topology and put

$$D = \{(x,t) : \Phi(x) < t, x \in K\}, \text{ and } \overline{D} = \{(x,t) : \Phi(x) \leq t, x \in K\}.$$

Obviously, $X \times \mathbb{R}$ is a Banach space with norm $\|(x,t)\| = \|x\| + |t|$, and the sets D and $\overline{D}$ are convex in the product $X \times \mathbb{R}$. In fact, since K is convex and Φ is a convex functional, for $(x,t)\,,\,(y,s) \in D$ ($\overline{D}$ resp.) and $0 \leq \alpha, \beta \in \mathbb{R}$, $\alpha + \beta = 1$, we have $\alpha x + \beta y \in K$, and in addition

$$\Phi(\alpha x + \beta y) \leq \alpha\Phi(x) + \beta\Phi(y) < (\leq)\alpha t + \beta s;$$

i.e., $\alpha(x,t) + \beta(y,s) = (\alpha x + \beta y, \alpha t + \beta s) \in D$ ($\overline{D}$ resp.). Moreover, if Φ is upper semi-continuous, then D is open in the Tychonov topology of the product $K \times \mathbb{R}$. To see this, suppose $(x_n, t_n) \in D' = (K \times \mathbb{R}) \setminus D$; i.e., $x_n \in K$, $\Phi(x_n) \geq t_n$, and $(x_n, t_n) \to (x,t)$ in $X \times \mathbb{R}$. Then $\|x_n - x\| \to 0$ and $|t - t_n| \to 0$. Thus $x \in K$ for K is closed. According to upper semi-continuity of Φ,

$$\Phi(x) \geq \limsup_n \Phi(x_n) \geq \lim_n t_n = t;$$

so that $(x,t) \in D'$, which means that D' is closed in $K \times \mathbb{R}$ and *a fortiori* D is open. By analogy, under lower semi-continuity of Φ it can be easily shown that $\overline{D}$ is closed in $K \times \mathbb{R}$, and so in $X \times \mathbb{R}$ as well.

Without loss of generality in the sequel we may and do assume that $\mathbb{O} = (\theta, 0) \in D$, where θ denotes the zero vector in X. For if $\theta \notin K^o$, but $K^o \ni x_0 \neq \theta$, then putting $\Phi_0(x) = \Phi(x + x_0)$ we see that $\Phi_0 : K - x_0 \to \mathbb{R}$ is upper (resp. lower) semi-continuous convex functional on $K - x_0$, $\theta \in (K - x_0)^o = K^o - x_0$ and $\xi - x_0 \in K - x_0$ *a.s.* Moreover, the inequality $\Phi_0\left(E^{\mathcal{A}}(\xi - x_0)\right) \leq E^{\mathcal{A}}\Phi_0(\xi - x_0)$ *a.s.* implies that

$$\begin{aligned}\Phi\left(E^{\mathcal{A}}(\xi - x_0) + x_0\right) &= \Phi_0\left(E^{\mathcal{A}}(\xi - x_0)\right) \leq E^{\mathcal{A}}\Phi_0(\xi - x_0) \\ &= E^{\mathcal{A}}\Phi\left((\xi - x_0) + x_0\right) \ \ a.s.\end{aligned}$$

Next, if $\Phi(\theta) \geq 0$, instead of Φ we consider $\Psi(x) = \Phi(x) - C$ with a suitably chosen constant C, $0 \leq \Phi(\theta) < C < \infty$. The mapping $\Psi : X \to \mathbb{R}$ is then obviously upper (resp. lower) semi-continuous convex functional, and whenever we prove the inequality $\Psi\left(E^{\mathcal{A}}\xi\right) \leq E^{\mathcal{A}}\Psi(\xi)$ *a.s.*, then (3) will follow automatically. Thus, from now on let $\mathbb{O} \in D$.

Clearly, $D = G \cap (K \times \mathbb{R})$, where G is an open set in $X \times \mathbb{R}$ such that $\mathbb{O} \in G$. Since $\theta \in K^{o}$ and $\Phi(\theta) < 0$, we have $\mathbb{O} = (\theta, 0) \in G \cap (K^{o} \times \mathbb{R}) \subset D$, which means that the interior D^{o} of D treated as a subset of $X \times \mathbb{R}$ is non-empty.

It is known that each continuous linear functional in $X_1 \times ... \times X_n$ is of the form $(x_1^*, ..., x_n^*)(x_1, ..., x_n) = x_1^*(x_1) + ... + x_n^*(x_n)$, where $x_i^* \in X_i^*$, $1 \leq i \leq n$; furthermore $\|(x_1^*, ..., x_n^*)\| = \|x_1^*\| \vee ... \vee \|x_n^*\|$ (see e.g. Alexiewicz (1969), Th. 10.3, Ch. III, p. 152). Suppose now that Φ is upper semi-continuous and observe that $(x, \Phi(x)) \notin D$ for every $x \in K$. Therefore, on the basis of Th. 8.10, Ch. III, p. 141 in Alexiewicz (cf. also Kantorovich and Akilov (1984), Th. 5, Ch. III, § 2, p. 107) we have

$$\bigwedge_{x \in K} \quad \bigvee_{\substack{z_x^* \in X^* \\ \alpha_x \in \mathbb{R}}} \quad z_x^*(y) + \alpha_x t \leq 1 \leq z_x^*(x) + \alpha_x \Phi(x) \tag{4}$$

for all $(y, t) \in D$. In particular, (4) is valid for each $x \in K^{o}$. If $x \in K \setminus K^{o} \neq \emptyset$, we select (z_x^*, α_x) in a special way. Namely, recall that

$$(x, t) \in \operatorname{cl} D \quad \Leftrightarrow \quad \bigvee_{\{(x_n, t_n)\} \subset D} \quad x_n \to x \ \wedge \ t_n \to t \ ,$$

where $\operatorname{cl} D$ stands for the closure of D in $X \times \mathbb{R}$. Moreover,

$$\limsup_{x_n \to x} \Phi(x_n) \leq \lim_{n \to \infty} t_n = t$$

for an arbitrary sequence of points $(x_n, t_n) \in D$, $n \geq 1$ such that $x_n \to x$ and $t_n \to t$. Thus, denoting $\mathcal{E}_x = \{\{x_n\} \subset K : x_n \to x\}$ we have in fact

$$\sup_{\{x_n\} \in \mathcal{E}_x} \left[\limsup_{x_n \to x} \Phi(x_n)\right] = \kappa_x \leq t \ ,$$

where $t = \inf\left\{g \in \mathbb{R} : \bigvee_{\{(x_n, t_n)\} \subset D} \ \{x_n\} \in \mathcal{E}_x \ , \ \lim t_n = g \text{ exists}\right\}$.

Otherwise, if $t < \kappa_x \leq \Phi(x) < \infty$, then $(x, t) \notin \operatorname{cl} D$. Hence, on account of Th. 8.11 and Corollary 8.12, Ch. III, § 8, pp. 142-143 of Alexiewicz (see also

Kantorovich and Akilov, Ch. III, § 2, Th. 6, p. 107 or Yosida (1978), Ch. IV, § 6, Th. 3′, p. 109) we obtain

$$\bigwedge_{x\in K\setminus K^{o}} \bigvee_{0<C_x<\infty} \bigvee_{\substack{z_x^*\in X^*\\ \alpha_x\in\mathbb{R}}} z_x^*(y)+\alpha_x t \le 1 < z_x^*(x)+\alpha_x\big(\Phi(x)-C_x\big), \tag{5}$$

because $(x, \Phi(x) - C_x) \notin \operatorname{cl} D$ for sufficiently large $C_x > \Phi(x) - \kappa_x$. Using (4) and (5) we shall characterize the set $\overline{D}$.

Let $(y, s) \in \overline{D}$, so that $y \in K$ and $\Phi(y) \le s$. Then $\Phi(y) < s + 1/n$ for $n \ge 1$, and thus $(y, s + 1/n) \in D$. Hence $z_x^*(y) + \alpha_x (s + 1/n) \le 1 \le z_x^*(x) + \alpha_x \Phi(x)$ for each $n \ge 1$ and $x \in K$. Consequently, for $x \in K$

$$z_x^*(y) + \alpha_x s \le 1 \le z_x^*(x) + \alpha_x \Phi(x).$$

On the other hand, if $(y, s) \in (K\times\mathbb{R})\setminus\overline{D}$, then $s < \Phi(y)$, and so $s = \Phi(y) - \varepsilon$ for some $\varepsilon > 0$. We shall demonstrate the inequality $z_y^*(y) + \alpha_y s > 1$.

Notice first that $\alpha_x < 0$ for each $x \in K$. Indeed, from (4) it follows that

$$z_x^*(x) + \alpha_x t \le 1 \le z_x^*(x) + \alpha_x \Phi(x),$$

whenever $t > \Phi(x)$, $x \in K^{o}$. Hence $\alpha_x \le 0$. Furthermore, the equality $\alpha_x = 0$ implies that $z_x^*(x) = 1$. But if $x \in K^{o}$, then $x + rx \in K^{o}$ for some $r > 0$, and so for appropriately chosen $t_r > \Phi(x + rx)$, in view of (4) we have

$$1 + r = z_x^*(x + rx) + 0 \cdot t_r \le 1,$$

which leads to a contradiction. If $x \in K \setminus K^{o}$, then taking $t > \Phi(x)$; i.e., $(x, t) \in D$, we get $z_x^*(x) + \alpha_x t \le 1 < z_x^*(x) + \alpha_x (\Phi(x) - C_x)$, and this is possible only when $\alpha_x < 0$ as well. Therefore,

$$z_y^*(y) + \alpha_y s = z_y^*(y) + \alpha_y \Phi(y) - \alpha_y \varepsilon > z_y^*(y) + \alpha_y \Phi(y) \ge 1.$$

In other words, we have obtained the following relation

$$(y,t)\in\overline{D} \quad \Leftrightarrow \quad \bigwedge_{x\in K} (z_x^*(y)+\alpha_x t\le 1)\wedge(y\in K).$$

Observe next that each pair (z_x^*, α_x) is a continuous linear functional on $X \times \mathbb{R}$, thus the family $\{(y, t) \in K \times \mathbb{R} : z_x^*(y) + \alpha_x t = (z_x^*, \alpha_x)(y, t) > 1\}$, $x \in K$, forms an open covering of the set $(K \times \mathbb{R}) \setminus \overline{D}$. Since $X \times \mathbb{R}$ is a separable metric space, by virtue of the well-known Lindelöf theorem there can be found at most denumerable set $Q \subset X^* \times \mathbb{R}$ such that

$$(K\times\mathbb{R})\setminus\overline{D} = \bigcup_{(z^*,\alpha)\in Q} \{(y,t)\in K\times\mathbb{R} : z^*(y)+\alpha t>1\}.$$

Hence it follows that

$$(y,t) \in \overline{D} \quad \Leftrightarrow \quad \bigwedge_{(z^*,\alpha)\in Q} \quad (z^*(y) + \alpha t \leq 1) \wedge (y \in K), \tag{6}$$

in particular

$$\bigwedge_{(z^*,\alpha)\in Q} \quad z^*(y) + \alpha\Phi(y) \leq 1 \quad \text{for} \quad y \in K. \tag{7}$$

If Φ is lower semi-continuous and $\overline{D}$ is closed, then $(x, \Phi(x)-\lambda) \in (K\times\mathbb{R})\setminus\overline{D}$ for an arbitrary $x \in K$ and $\lambda > 0$. Thus, in view of Th. 3′, Ch. IV, § 6, p. 109 of Yosida (1978) (cf. also Kantorovich and Akilov, Ch. III, § 2, Th. 6, p. 107 or Alexiewicz, Ch. III, § 8, Th. 8.11 and Corollary 8.12, pp. 142-143),

$$\bigwedge_{x\in K} \quad \bigwedge_{\lambda>0} \quad \bigvee_{\substack{z^*_{x,\lambda}\in X^* \\ \alpha_{x,\lambda}\in\mathbb{R}}} \quad z^*_{x,\lambda}(y) + \alpha_{x,\lambda} t \leq 1 < z^*_{x,\lambda}(x) + \alpha_{x,\lambda}(\Phi(x) - \lambda) \tag{8}$$

for all $(y,t) \in \overline{D}$, in particular $z^*_{x,\lambda}(y) + \alpha_{x,\lambda}\Phi(y) \leq 1$ whenever $x, y \in K$ and $\lambda > 0$. Putting $y = x$ in the last inequality we infer from (8) that $\alpha_{x,\lambda} < 0$ for $x \in K$ and $\lambda > 0$. Denote by Z the set of all the pairs $(z^*_{x,\lambda}, \alpha_{x,\lambda})$ indexed by $x \in K$ and $\lambda > 0$, for which (8) holds. Suppose that $(y,s) \in (K\times\mathbb{R})\setminus\overline{D}$; i.e., $y \in K$ and $s < \Phi(y)$. Then $s = \Phi(y) - \varepsilon$ for some $\varepsilon > 0$, and so taking $\lambda \leq \varepsilon$ we see that for $(z^*_{y,\lambda}, \alpha_{y,\lambda}) \in Z$,

$$1 < z^*_{y,\lambda}(y) + \alpha_{y,\lambda}(\Phi(y) - \lambda) \leq z^*_{y,\lambda}(y) + \alpha_{y,\lambda} s.$$

Hence

$$(y,t) \in \overline{D} \quad \Leftrightarrow \quad \bigwedge_{(z^*_{x,\lambda},\alpha_{x,\lambda})\in Z} \quad (z^*_{x,\lambda}(y) + \alpha_{x,\lambda} t \leq 1) \wedge (y \in K).$$

Applying again the Lindelöf theorem to the family of sets

$$\{(y,t) \in K\times\mathbb{R} : z^*_{x,\lambda}(y) + \alpha_{x,\lambda} t > 1\}, \quad (z^*_{x,\lambda}, \alpha_{x,\lambda}) \in Z,$$

we obtain (6) and (7) as well as previously.

Having established (7) we argue as follows. Substitute $y = \xi(\omega)$ in (7) and next evaluate conditional expectations of both sides of the obtained thus inequality with respect to the σ-field $\mathcal{A}$. Then for all $(z^*,\alpha) \in Q$ we get $E^{\mathcal{A}}z^*(\xi) + \alpha E^{\mathcal{A}}\Phi(\xi) \leq 1$ *a.s.*. By (1), $z^*E^{\mathcal{A}}\xi) + \alpha E^{\mathcal{A}}\Phi(\xi) \leq 1$ *a.s.*, for all $(z^*,\alpha) \in Q$; so that on the basis of (6) and the Lemma, $(E^{\mathcal{A}}\xi, E^{\mathcal{A}}\Phi(\xi)) \in \overline{D}$ with probability 1. By definition of $\overline{D}$ we have $\Phi(E^{\mathcal{A}}\xi) \leq E^{\mathcal{A}}\Phi(\xi)$ *a.s.*. □

Remark. It is an open question whether for an arbitrary (infinite dimensional) normed linear space X there exists a convex semi-continuous functional $\Phi : X \to \mathbb{R}$ which is not continuous. The possible candidate for such a convex mapping may be $x \to |x'(x)|$, where x' is any discontinuous linear functional on X (see Alexiewicz, Ch. III, § 8, Prop. 8.17, p. 145). However, as will be seen below, if $\xi(\omega) \in K$ for all $\omega \in \Omega$, where K is a proper closed convex subset of X and Φ need not be defined when $x \notin K$, then the relevant examples can be given even for finite dimensional spaces.

Example 1. Let $X = \mathbb{R}$. Then it is enough to take $K = \langle 0, \infty) \subset \mathbb{R}$ and $\Phi(t) = t^2$ for $t > 0$, accomplished with $\Phi(0) = c > 0$.

Example 2. A slightly more involved example may be constructed as follows. Let $X = \mathbb{R}^2$ and $K = \{x \in X : \|x\| \leq 1\}$, where $\|\cdot\|$ is the usual Euclidean norm in the plane. Define $\Phi(x) = \|x\|^2$ for $\|x\| < 1$ and $\Phi(x) = c > 1$ on the unit sphere $S = \{x \in X : \|x\| = 1\}$, except for a set $\{x_i, i \in I\} \subset S$, which does not have any accumulation points; moreover, for $i \in I$ put $\Phi(x_i) = c_i > c$. The same effect on the boundary of K can be achieved in a more general way by taking an arbitrary upper semi-continuous function $\varphi : \langle 0, 2\pi) \to \mathbb{R}$, such that $\varphi(t) \geq 1$ and $\limsup_{t \nearrow 2\pi} \varphi(t) \leq \varphi(0)$. Using the one-to-one correspondence $x \longleftrightarrow \exp\{it\}$, where $x \in S$, $t \in \langle 0, 2\pi)$, we define $\Phi(x) = \varphi(t)$ for $x = \exp\{it\} \in S$, $t \in \langle 0, 2\pi)$. It can be easily verified that Φ is convex and upper semi-continuous on K, but not continuous.

Example 3. A similar idea may be also adapted to the case of infinite dimensional spaces. Recall that a Banach space X is called *strictly convex*, if for an arbitrary $x, y \in S = \{x \in X : \|x\| = 1\}$ and $0 < \alpha, \beta \in \mathbb{R}$, $\alpha + \beta = 1$, we have $\|\alpha x + \beta y\| < 1$. Let $\psi : \langle 0, 1\rangle \to \mathbb{R}$ be a convex, continuous (bounded) real function and let X be a strictly convex Banach space. Put

$$\Phi(x) = \begin{cases} \psi(\|x\|) & \text{for} \quad \|x\| < 1 \ , \\ c > \psi(1) & \text{for} \quad \|x\| = 1 \ , \quad x \notin \{x_i, i \in I\} \ , \\ c_i > c & \text{for} \quad x = x_i \ , \quad i \in I \ , \end{cases}$$

where $\{x_i, i \in I\} \subset S = \{x \in X : \|x\| = 1\}$ does not have accumulation points. Then Φ is an upper semi-continuous, convex and discontinuous functional on the set $K = \{x \in X : \|x\| \leq 1\}$. If X is an arbitrary (infinite dimensional) Banach space, instead of the unit ball, we consider rather a compact convex set K contained in the closed unit ball. The convex continuous functional $\psi(\|x\|)$, $x \in K$, attains its supremum $c < \infty$ on K. Thus by the Krein-Milman theorem (see e.g. Yosida, Ch. XII, § 1, Th. and Corollary, pp. 362-363) the mapping $\Phi(x) = \psi(\|x\|)$ can be modified on the set $\{x_i, i \in I\} \subset K$ of all the extremal points of K by putting $\Phi(x_i) = c_i > c$, $i \in I$.

The next result is already known (see Diestel and Uhl, Ch. V, § 1, Th. 4, p. 123), but now it is a straightforward consequence of our theorem.

Corollary 1. *If $\xi \in L_p(\mathcal{F}; X)$, $1 < p \leq \infty$, then $E^{\mathcal{A}}\xi \in L_p(\mathcal{A}; X)$, and*

$$\left\|E^{\mathcal{A}}\xi\right\|_p \leq \|\xi\|_p\,; \tag{9}$$

i.e., $E^{\mathcal{A}}$ is a projection operator with norm 1 acting from $L_p(\mathcal{F}; X)$ to $L_p(\mathcal{A}; X)$.

PROOF. It can be easily verified that $\Phi(x) = \|x\|^p$, $x \in X$, $1 < p < \infty$, is a convex continuous functional on X. Hence, based on the above theorem, we have

$$\left\|E^{\mathcal{A}}\xi\right\|_p = (E\left\|E^{\mathcal{A}}\xi\right\|^p)^{1/p} \leq (E\left\{E^{\mathcal{A}}(\|\xi\|^p)\right\})^{1/p} = (E\|\xi\|^p)^{1/p} = \|\xi\|_p$$

whenever $\xi \in L_p(\mathcal{F}; X)$, $1 < p < \infty$. For $p = \infty$, (9) follows from the well-known equality $\lim\limits_{p\to\infty} \|\cdot\|_p = \|\cdot\|_\infty$ (cf. Yosida, Ch. I, § 2, Th. 1, p. 34, Kantorovich and Akilov, Ch. IV , § 3, p. 144, or Alexiewicz, Ch. IV , § 2, Th. 2.4, p. 219). The operator norm of $E^{\mathcal{A}}$ is equal to 1 in view of the property $E^{\mathcal{A}}\xi = \xi$ *a.s.* for $\xi \in L_p(\mathcal{A}; X)$. □

Corollary 2. *Let $K \subset X$ and $\Phi : K \to \mathbb{R}$ satisfy the assumptions of our theorem. Moreover, let $\{\mathcal{F}_t, t \in T \subset \mathbb{R}\} \subset \mathcal{F}$ be an increasing family of σ-fields and let $\{\xi_t, \mathcal{F}_t\}$ be a martingale such that $\xi_t \in K$ a.s. for all $t \in T$. Then $\{\Phi(\xi_t), \mathcal{F}_t\}$ is a submartingale.*

PROOF. The proof is immediate, because for $s < t$ we have

$$\Phi(\xi_s) = \Phi\left(E^{\mathcal{F}_s}\xi_t\right) \leq E^{\mathcal{F}_s}\Phi(\xi_t) \ a.s. \qquad \square$$

Using our theorem we can also derive an interesting result concerning Gaussian measures in a Banach space X. Recall that a (Borel separable) random element ξ in X is called *Gaussian*, if $x^*(\xi)$ is a real Gaussian random variable for each continuous linear functional $x^* \in X^*$, and it is called *Gaussian in the sense of Bernstein*, if for any two independent copies ξ_1, ξ_2 of ξ given on a common probability space, the random elements $\xi_1 + \xi_2$ and $\xi_1 - \xi_2$ are independent. In the described context these two definitions are equivalent.

Corollary 3. *If ξ is a Gaussian random element in X, then there exists a constant $0 < c < \infty$ such that for any fixed σ-field $\mathcal{A}_0 \subset \mathcal{F}$ the family of random elements $\left\{\exp\{c\left\|E^{\mathcal{A}}\xi\right\|^2 : \mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{F}\right\} \subset L_1(\mathcal{F}; X)$ is uniformly integrable.*

Proof. Observe that $t \to t^2$ and $s \to \exp\{cs\}$ are convex real functions. Therefore if $c > 0$, then for $x, y \in X$,

$$\exp\left\{c\|\alpha x + \beta y\|^2\right\} \le \exp\left\{c(\alpha\|x\| + \beta\|y\|)^2\right\} \le \exp\left\{c\left(\alpha\|x\|^2 + \beta\|y\|^2\right)\right\}$$
$$\le \alpha\exp\left\{c\|x\|^2\right\} + \beta\exp\left\{c\|y\|^2\right\}.$$

Consequently, $\Phi(x) = \exp\{c\|x\|^2\}$ for $c > 0$ is a convex continuous functional on X. Moreover, there exists a constant $0 < B < \infty$ such that $\exp\left\{B\|\xi\|^2\right\}$ is integrable (see e.g. Fernique (1970), Kuo (1975), Ch. III, § 3, Th. 3.1, p.159, Kwapień and Woyczyński (1992), Ch. III, § 7, pp. 54–55, or Zapała (1987)).

Choose $0 < c < B$ and fix $\varepsilon > 0$ arbitrarily. Applying the above theorem we get $\exp\left\{c\left\|E^{\mathcal{A}}\xi\right\|^2\right\} \le E^{\mathcal{A}}\exp\left\{c\|\xi\|^2\right\}$ *a.s.* Hence, if $P[A] < (\varepsilon/M)^p$ for $A \in \mathcal{A}_0$, where $M = \left(E\exp\left\{cq\|\xi\|^2\right\}\right)^{1/q} < \infty$, $cq \le B$, $p, q > 1$, $1/p + 1/q = 1$, then by Hölder's inequality

$$\int_A \exp\left\{c\left\|E^{\mathcal{A}}\xi\right\|^2\right\} dP \le \int_A E^{\mathcal{A}}\exp\left\{c\|\xi\|^2\right\} dP = \int_A 1\cdot\exp\left\{c\|\xi\|^2\right\} dP$$
$$\le (P[A])^{1/p}\left(\int_\Omega \exp\left\{cq\|\xi\|^2\right\} dP\right)^{1/q} < \varepsilon,$$

and $E\exp\left\{c\left\|E^{\mathcal{A}}\xi\right\|^2\right\} \le E\left\{E^{\mathcal{A}}\exp\{c\|\xi\|^2\}\right\} = E\exp\left\{c\|\xi\|^2\right\} \le M < \infty$ uniformly with respect to σ-fields $\mathcal{A}$, $\mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{F}$. By Th.19, Ch. II, p. 22 in Dellacherie and Meyer (1978), the family $\left\{\exp\{c\left\|E^{\mathcal{A}}\xi\right\|^2 : \mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{F}\right\} \subset L_1(\mathcal{F}; X)$ is uniformly integrable. $\square$

Remark. 1. If $\{\mathcal{F}_t, t \in T \subset \mathbb{R}\} \subset \mathcal{F}$ is an increasing family of σ-fields and ξ is a Gaussian random element in X, then according to Corollary 2 for a fixed $c > 0$ sufficiently close to zero $\left\{\exp\{c\left\|E^{\mathcal{F}_t}\xi\right\|^2 : t \in T\right\}$ is a submartingale, thus its equi-integrability follows from a more general theorem concerning stopped martingales (cf. Metivier and Pellaumail, Ch. 4, § 8.3, Prop. 2, p. 96-97).

2. The result applied above concerning exponential integrability of Gaussian random vectors in Banach spaces was considerably improved by Talagrand (1984) as follows: if Y is a random vector with a symmetric Gaussian distribution in a (separable) Banach space X, then $E\exp\left\{c_0\|Y\|^2 - b\|Y\|\right\} < \infty$ for an arbitrary $b > 0$, where $c_0 = \left(2\sup_{\|x^*\|=1} E\left[x^*(Y)\right]^2\right)^{-1}$ (see Kwapień and Woyczyński, Ch. III, § 6, p. 89, cf. also Ch. II, § 6, Th. 2.6.1, pp. 52-53). Moreover, instead of the norm in a Banach space one can consider any pseudometric invariant under translations in a group (see Zapała (1987)).

References

[1] A. Alexiewicz, *Functional Analysis*, PWN, Warsaw, 1969 (in Polish).

[2] C. Dellacherie, P. A. Meyer P. A. *Probabilities and Potential*, North Holland, Amsterdam, New York 1978.

[3] J. Diestel, J. Uhl, *Vector Measures*, Providence 1977.

[4] X. Fernique, *Integrabilité des vecteurs Gaussiens*, C. R. Acad. Sci. Paris **270** (1970), 1698–1699.

[5] L. V. Kantorovich, G. P. Akilov, *Functional Analysis*, Nauka, Moscow 1984 (in Russian).

[6] H.-H. Kuo, *Gaussian Measures in Banach Spaces*, Springer-Verlag, Berlin, Heidelberg, New York 1975.

[7] S. Kwapień, W. A. Woyczyński, *Random Series and Stochastic Integrals: Single and Multiple*, Birkhäuser, Boston (1992).

[8] M. Metivier, J. Pellaumail, *Stochastic Integration*, Academic Press, New York (1980).

[9] M. Talagrand, *Sur l'intégrabilité des vecteurs Gaussiennes*, Z. Wahrscheinlichkeitstheorie verw. Geb. **68** (1984), 1–8.

[10] N. N. Vahania, V. I. Tarieladze, S. A. Chobanyan, *Probability Distributions in Banach Spaces*, Nauka, Moscow 1985 (in Russian) (English transl.: Reidel, Dordrecht 1987).

[11] W. A. Woyczyński, *Geometry and martingales in Banach spaces* II. *Advances in Probability and Related Topics*, No 4, p. 267–517. Marcel Dekker, New York, Basel (1978).

[12] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin, Heidelberg, New York (1978).

[13] A. M. Zapała, *On exponential integrability of a transferable pseudometric with respect to a Gaussian measure on a group*, Bull. Pol. Acad. Sci. **35** (1987), 597–600.