Journal of Nonlinear and Convex Analysis
Volume 1, Number 3, 2000, 233–244

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

CONDITIONAL EXPECTATION AND ERGODIC THEOREM FOR A POSITIVE INTEGRAND

MICHEL VALADIER

Abstract. An historical account of difficuties raised by the conditional expectation of an integrand is given. Then the Birkhoff ergodic therem for integrands, as initiated by C. Castaing and F. Ezzaki in 1992, is improved.

Contents

1. Introduction
2. Background on classical conditional expectation and difficulties raised by a parameter
3. Geometrical versus analytical points of view and short historical review
4. Disintegration and ergodic theorem for positive r.v.
5. Ergodic theorem for integrands
References

1. Introduction

The conditional expectation of random variables is a classical tool. When a parameter comes into play, that is when one handles an integrand, the situation is more involved. This has been studied in the seventies with a seminal paper by Bismut [6]. A related problem is raised by the conditional expectation of a multifunction. We will give an historical account in Section 3.

In 1992 C. Castaing and F. Ezzaki wrote a paper [10] giving several statements about the Birkhoff ergodic therem for integrands. Moreover, in the same time, C. Licht and G. Michaille [31–32] studying homogenization problems were also interested by an extension of the ergodic theorem (note that this started with papers of Dal Maso-Modica [15–16]; see my survey [46] for some comments about the difference between stationarity and ergodicity). The hypotheses of [10] are sometimes restrictive and some proofs can be made more direct. We will improve their results. A first result about the ergodic theorem for multifunctions is in Hess [21]. In the same direction we must mention researches by Choirat-Hess-Seri [12] and Korf-Wets [28–29].

Note that the strong law of large numbers for multifunctions was firstly studied by Hess [21] and Artstein-Hart [2], but received a general treatment in 1985: Hess [22] and Hiai [24]. For integrands see Attouch-Wets [4].

I thank C. Castaing and G. Michaille for their acute reading of some preliminary versions.

1991 *Mathematics Subject Classification.* 60–02, 60 F 15, 28 D 05.
Key words and phrases. ergodicity, Birkhoff's theorem, integrands, epi-convergence.

2. Background on classical conditional expectation and difficulties raised by a parameter

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{I}$ a sub-σ-algebra of $\mathcal{F}$. The conditional expectation operator $\mathbf{E}^{\mathcal{I}} \colon L^1(\Omega, \mathcal{F}, P) \to L^1(\Omega, \mathcal{I}, P)$ characterized by

$$\forall B \in \mathcal{I}, \quad \int_B (\mathbf{E}^{\mathcal{I}} X)\, dP = \int_B X\, dP$$

is studied in most Probability textbooks. We wrote L^1, not $\mathcal{L}^1$, because the good spaces are the quotient ones whose elements are equivalence classes of random variables (r.v.) with respect to equality a.s. As starting value one can take $X \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ without any difficulty. On the contrary the value $\mathbf{E}^{\mathcal{I}} X$ is really a class. Usually one writes $(\mathbf{E}^{\mathcal{I}} X)(\omega)$, which supposes that a <u>version</u> of $\mathbf{E}^{\mathcal{I}} X$ has been chosen. As a precaution one could write "a.s." above = as in the following

$$Y(\omega) \overset{\text{a.s.}}{=} (\mathbf{E}^{\mathcal{I}} X)(\omega)\,. \tag{1}$$

When the r.v. X is ≥ 0 (the value $+\infty$ being possible) it does not need to be integrable: see Neveu [35] (even $X \in \mathcal{L}^0(\Omega, \mathcal{F}, P)$ is possible as soon as the negative part X^- is integrable). In all cases uniqueness holds up to equality a.s. (see Theorem 1 below for a more difficult result).

Now E is a "good" topological space, for example a Polish space or a metrizable Suslin space: we need a separable metrizable space in which the von Neumann-Aumann-Sainte-Beuve theorem applies. If f is a real ≥ 0 function (we will say an *integrand*) on $\Omega \times E$ which is $(\mathcal{F} \otimes \mathcal{B}(E))$-measurable, one wish defining $\mathbf{E}^{\mathcal{I}} f$ and showing its existence under general hypotheses. But one cannot only say: for each fixed x we take a version of $\mathbf{E}^{\mathcal{I}} f(.,x)$, which gives $\omega \mapsto \big(\mathbf{E}^{\mathcal{I}} f(.,x)\big)(\omega)$ and then $(\omega, x) \mapsto \big(\mathbf{E}^{\mathcal{I}} f(.,x)\big)(\omega)$ which would be $\mathbf{E}^{\mathcal{I}} f$. Indeed one has to get the $(\mathcal{I} \otimes \mathcal{B}(E))$-measurability of $\mathbf{E}^{\mathcal{I}} f$; and when for all (or almost all) ω, $f(\omega, .)$ is Lipschitz or lower semi-continuous, one expects the same property of $\mathbf{E}^{\mathcal{I}} f$. So the problem is to choose in a consistent way (or to prove the possibility of such a choice) versions of the classes $\mathbf{E}^{\mathcal{I}} f(.,x)$.

Look at the definition of $g := \mathbf{E}^{\mathcal{I}} f$. The minimum properties should be: g is $(\mathcal{I} \otimes \mathcal{B}(E))$-measurable and $\forall B \in \mathcal{I}$, $\forall x \in E$,

$$\int_B g(\omega, x)\, dP(\omega) = \int_B f(\omega, x)\, dP(\omega)\,.$$

But the good definition (due to Bismut [6]) is: for any $u \in \mathcal{L}^0(\Omega, \mathcal{I}, P; E)$, $\omega \mapsto g(\omega, u(\omega))$ is a version of $\mathbf{E}^{\mathcal{I}} f(., u(.))$ or, more precisely of $\mathbf{E}^{\mathcal{I}}(\widehat{f}(u))$ where $\widehat{f}$ denotes the Nemickii operator associated to f:

$$\widehat{f} \left| \begin{array}{r} \mathcal{L}^0(\Omega, \mathcal{I}, P; E) \to \mathcal{L}^0(\Omega, \mathcal{F}, P; [0, +\infty]) \\ u \mapsto [\omega \mapsto f(\omega, u(\omega))] \end{array} \right.$$

that is $\big[\widehat{f}(u)\big](\omega) = f(\omega, u(\omega))$. In other words the good characterization of "the" conditional expectation g of the integrand f is:

$$\forall u \in \mathcal{L}^0(\Omega, \mathcal{I}, P; E),\ \forall B \in \mathcal{I}, \quad \int_B g(\omega, u(\omega))\, dP(\omega) = \int_B f(\omega, u(\omega))\, dP(\omega) \tag{2}$$

or $\forall u \in \mathcal{L}^0(\Omega, \mathcal{I}, P; E)$, $\mathbf{E}^{\mathcal{I}}(\widehat{f}u) = \widehat{g}(u)$. So we are able to write

$$\forall u \in \mathcal{L}^0(\Omega, \mathcal{I}, P; E), \quad \mathbf{E}^{\mathcal{I}}(\widehat{f}u) = \widehat{\mathbf{E}^{\mathcal{I}} f}(u)\,. \tag{2'}$$

Now uniqueness of $\mathbf{E}^{\mathcal{I}} f$ is to be understood in the *indistinguishability* sense: two versions are indistinguishable if they coincide except on $N \times E$ where N is P-negligible. We prove uniqueness below. The analogous of formula (1) could be:

$$g(\omega, x) \overset{\text{indist.}}{=} (\mathbf{E}^{\mathcal{I}} f)(\omega, x)\,. \tag{3}$$

Theorem 1. *With the foregoing notations, if g_1 et g_2 are two ≥ 0 integrands which are conditional expectations of f* (i.e. $(\mathcal{I} \otimes \mathcal{B}(E))$-measurable and satisfying (2)), *they are indistinguishable.*

Proof. Let pr_Ω denote the projection of $\Omega \times E$ on Ω. Temporarily we will use the P-completion $\widehat{\mathcal{I}}$ of $\mathcal{I}$. We will show that $\mathrm{pr}_\Omega(\{g_1 < g_2\})$ is negligible, which implies that $\{g_1 < g_2\}$ is contained in a set as $N_0 \times E$ where N_0 is negligible. By symmetry the same holds for $\{g_1 > g_2\}$, hence $\{g_1 \neq g_2\}$ is contained is a set of the form $N \times E$ where N is negligible. Thanks to the von Neumann-Aumann-Sainte-Beuve theorem ([37], [11, Th.III-23]), $B_0 := \mathrm{pr}_\Omega(\{g_1 < g_2\})$ belongs to $\widehat{\mathcal{I}}$. Suppose $P(B_0) > 0$. Still by the von Neumann-Aumann-Sainte-Beuve theorem ([37], [11, Th.III-22]) there exists $\bar{u} : B_0 \to E$ measurable relatively to $\widehat{\mathcal{I}}$ (but modifying it on a negligible, one can get $\mathcal{I}$-measurability) such that on B_0, $g_1(\omega, \bar{u}(\omega)) \overset{\text{a.s.}}{<} g_2(\omega, \bar{u}(\omega))$. Changing if necessary B_0 by a subset of the form $B = \{\omega \in B_0 : g_1(\omega, \bar{u}(\omega)) \leq n\}$, one can suppose that, keeping $P(B) > 0$, $g_1(\omega, \bar{u}(\omega))$ is $\leq$ to a finite constant on B. It remain to set, x_0 being a fixed point in E,

$$u(\omega) = \begin{cases} \bar{u}(\omega), & \text{if } \omega \in B, \\ x_0\,, & \text{if } \omega \in \Omega \backslash B \end{cases}$$

to get the contradiction $\int_B g_1(\omega, u(\omega))\, dP(\omega) < \int_B g_2(\omega, u(\omega))\, dP(\omega)$. □

The existence for l.s.c. integrands has been proved using Lipschitz approximations (see Section 5 for the definition) by Castaing and Thibault, and using a lifting of L^∞ by Bismut. For general measurable integrands the existence has been proved by Evstigneev (see also [10]). For some references and a partial result (Theorem 2) see Section 3 and 4.

Remarks. 1) The difficulty of choosing a non countable family of versions of classes in a consistent way arise also in the disintegration problem (see Section 4).

2) In state of integrands one can also consider multifunctions (alias random sets) and look for their conditional expectation. This is the geometrical side of the problem. See next Section.

3. Geometrical versus analytical points of view and short historical review

When E is a locally convex topological vector space (in short l.c.t.v.s.), studying the conditional expectations of a convex l.s.c. integrand or of a closed convex random set is almost the same problem (see below).

The conditional expectation of a r.v. X can be defined as the Radon-Nikodým derivative, with respect to $P_{|\mathcal{I}}$, of the measure $B \mapsto \int_B X\,dP$ considered on $\mathcal{I}$. In this line, the multivalued Radon-Nikodým theorem has been stated and generalized several times: Valadier [40–41] in finite dimension, Castaing [7–8] in infinite dimension, Debreu-Schmeidler [17] in finite dimension but with non boundedness (their motivation coming from mathematical economics), then Costé-Pallu de La Barrière [13–14].

Bismut [6] is a basic text (after reading it I wrote [42]). The Lipschitz approximation technique introduced by Castaing [9] has been used also by Thibault [38]. Hiai-Umegaki [25] dared handling non-convex multifunctions (they used the concept of decomposable subsets of L^1). In similar questions Klei [26] used my result about the essential multivalued least upper bound. A. Truffert [39] wrote a very complete paper, probably optimal about minorations and growth conditions.

Surprisingly, when everybody assumed lower semi-continuity in x, Evstigneev [20] proved the existence of a good conditional expectation of any integrand f which is ≥ 0 and $\mathcal{F} \otimes \mathcal{B}(E)$-measurable.

More recently Castaing-Ezzaki in their paper [10] began with a synthetic account and then brought a lot of new results we will speak of later.

Assume that Γ is a multifunction (or random set) with closed convex values in a l.c.t.v.s. V. Let V^* denote the topological dual of V. To Γ is associated the integrand on $\Omega \times V^*$, $f(\omega, x') = \delta^*(x'|\Gamma(\omega))$ where $\delta^*(\,.\,|C)$ denotes the support function of C. Then one can recover the conditional expectation (in short c.e.) of Γ from the c.e. of f and, vice-versa one can obtain the c.e. of f from the c.e. of the multifunction $\omega \mapsto \operatorname{epi}(f^*(\omega,.))$ (recall $f(\omega,x) = \delta^*\big((x,-1)|\operatorname{epi}(f^*\omega,.)\big)$).

But generally studying the conditional expectations of Γ and of f are not equivalent problems. Indeed, in one direction, if Γ is not convex valued, the support function loses information. And in the other direction one can consider an integrand on $\Omega \times E$ even if E is not a l.c.t.v.s.

Finally note that when E is a l.c.t.v.s. the polar of $\mathbf{E}^{\mathcal{I}} f$ is the *conditional inf-convolution of* f^*. This notion has been studied by Bismut [6] and in Chapter 8 of [11].

4. Disintegration and ergodic theorem for positive r.v.

Under very general hypotheses there exists a disintegration, that is a family of probabilities $(Q_\omega)_{\omega \in \Omega}$ on $(\Omega, \mathcal{F})$ depending $\mathcal{I}$-measurably of ω and satisfying

$$\forall A \in \mathcal{F},\ \forall B \in \mathcal{I}, \quad P(A \cap B) = \int_B Q_\omega(A)\,dP(\omega)\,. \tag{4}$$

In other words the r.v. $\omega \mapsto Q_\omega(A)$ (A running through $\mathcal{F}$) are a consistent system of versions of the classes $\mathbf{E}^{\mathcal{I}}(1_A)$. (Due to the same difficulty, liftings of L^∞ has been used by Bismut [6].)

Classically for a real r.v. X (see for example Dudley [19, 10.2.5 p.272], Doob [18, Th.9.1 p.27], Kolmogorov [27, ch.V (12) and (14)]):

$$\int_\Omega X(\omega')\,dQ_\omega(\omega') \overset{\text{a.s.}}{=} (\mathbf{E}^{\mathcal{I}} X)(\omega)\,. \tag{5}$$

I extended this to non convex random sets in [45, 1980] (and wrote on the subject of disintegration: [43–44]; note that Bauer [5] clearly exposes the question of conditional laws but does not give (5)).

More easy than [45] is the case of an integrand treated in the next theorem. This theorem does not prove in general the existence de $\mathbf{E}^{\mathcal{I}} f$ since there exist cases where there is not any disintegration. But I enjoy it and it can help in suggesting conjectures and giving easy proofs under the hypothesis of existence of a disintegration. For general proofs of existence see Castaing-Ezzaki [10]. Writing this paper I rediscovered Theorem 2 forgetting that A. Truffert wrote it in 1991 [39, Th.1.7 p.133–134]. But Evstigneev already indicated it in 1986 [20, p.516].

Theorem 2. *If the family $(Q_\omega)_{\omega\in\Omega}$ satisfies* (4)*, the integrand g defined by:*

$$g(\omega, x) := \int_\Omega f(\omega', x)\, dQ_\omega(\omega')$$

is a conditional expectation of f. If the functions $f(\omega, .)$ are l.s.c., so are the $g(\omega, .)$; if the functions $f(\omega, .)$ are k-Lipschitz, so are the $g(\omega, .)$.

Proof. 1) Let us show that g is $(\mathcal{I} \otimes \mathcal{B}(E))$-measurable. The following arguments occur in usual proofs of Fubini's theorem. The set of all $C \in \mathcal{F} \otimes \mathcal{B}(E)$ such that $(\omega, x) \mapsto \int_\Omega 1_C(\omega', x)\, dQ_\omega(\omega')$ is $(\mathcal{I} \otimes \mathcal{B}(E))$-measurable constitute a Dynkin's system which contains $A \times U$ when $A \in \mathcal{F}$ and $U \in \mathcal{B}(E)$ (because then one gets $(\omega, x) \mapsto Q_\omega(A) 1_U(x)$); hence it coincides with $\mathcal{F} \otimes \mathcal{B}(E)$. By linearity, for any integrand f which is ≥ 0 and $(\mathcal{F} \otimes \mathcal{B}(E))$-simple, g is $(\mathcal{I} \otimes \mathcal{B}(E))$-measurable. By monotone convergence this holds for any f.

2) Now we prove that g is a conditional expectation of f. Firstly (4) implies some consequences: let R the image on $(\Omega^2, \mathcal{I} \otimes \mathcal{F})$ of P by the map $\omega \mapsto (\omega, \omega)$. For any $B \in \mathcal{I}$ and any $A \in \mathcal{F}$,

$$R(B \times A) = P(A \cap B) \overset{(4)}{=} \int_B Q_\omega(A)\, dP(\omega) = \int_\Omega \Big[\int_\Omega 1_B(\omega) 1_A(\omega')\, dQ_\omega(\omega') \Big]\, dP(\omega)$$

hence classically, for any ψ which is ≥ 0 on Ω^2 and $(\mathcal{I} \otimes \mathcal{F})$-measurable,

$$\int_{\Omega^2} \psi(\omega, \omega')\, dR(\omega, \omega') = \int_\Omega \Big[\int_\Omega \psi(\omega, \omega')\, dQ_\omega(\omega') \Big]\, dP(\omega)\,. \tag{6}$$

From the definition of R, the left-hand side of (6) equals $\int_\Omega \psi(\omega, \omega)\, dP(\omega)$. Let $u \in \mathcal{L}^0(\Omega, \mathcal{I}, P; E)$. Applying the foregoing observation and (6) to

$$\psi(\omega, \omega') := 1_B(\omega) f(\omega', u(\omega))\,,$$

one gets:

$$\begin{aligned} \int_B f(\omega, u(\omega))\, dP(\omega) &= \int_B \Big[\int_\Omega f(\omega', u(\omega))\, dQ_\omega(\omega') \Big]\, dP(\omega) \\ &= \int_B g(\omega, u(\omega))\, dP(\omega)\,. \end{aligned}$$

Thus we have proved that g is a conditional expectation of f.

3) Suppose that $f(\omega,.)$ is l.s.c. and that x_n converges to $\bar{x}$. Then one has

$$\int_\Omega f(\omega',\bar{x})\,dQ_\omega(\omega') \le \int_\Omega \varliminf_{n\to\infty} f(\omega',x_n)\,dQ_\omega(\omega') \overset{\text{Fatou}}{\le} \varliminf_{n\to\infty} \int_\Omega f(\omega',x_n)\,dQ_\omega(\omega')\,,$$

hence the lower semi-continuity of $g(\omega,.)$.

4) Suppose that $f(\omega,.)$ is k-Lipschitz (with values $< +\infty$). Then

$$\begin{aligned} |g(\omega,x)-g(\omega,y)| &= \Big|\int_\Omega \big(f(\omega',x)-f(\omega',y)\big)\,dQ_\omega(\omega')\Big| \\ &\le \int_\Omega \big|f(\omega',x)-f(\omega',y)\big|\,dQ_\omega(\omega') \\ &\le k\,d(x,y)\,. \end{aligned}$$

□

Suppose T is a measurable map from Ω in itself preserving P (i.e. $P\circ T^{-1}=P$), and let $\mathcal{I}$ denote the σ-field of invariant sets (that is the set of all $A\in\mathcal{F}$ satisfying $T^{-1}A=A$).

Remark. The strong law of large numbers for ≥ 0 r.v. is easy: if the X_i are ≥ 0 and i.i.d., $\frac{1}{n}\sum_{i=1}^n X_i(\omega)$ converges almost surely to $\mathbf{E}(X_1)$ even if $\mathbf{E}(X_1)=+\infty$. This is easy to prove with Proposition 4 below. Note that Dudley [19, Th.8.3.5 p.206] proves, by a different method, better: if the X_i are real, i.i.d. and if $\mathbf{E}(|X_1|)=+\infty$, then almost surely $\frac{1}{n}\sum_{i=1}^n X_i(\omega)$ does not converge in $\mathbb{R}$. The following result is more difficult to find in the literature.

Theorem 3 (Birkhoff's ergodic theorem for a positive random variable)**.** *For any ≥ 0 random variable X (the value $+\infty$ being allowed),*

$$\frac{1}{n}\sum_{i=0}^{n-1} X(T^i\omega) \overset{a.s.}{\to} \mathbf{E}^{\mathcal{I}}\big(X\big)(\omega)\,.$$

Remark This is asserted without proof by L. Arnold [1, p.539]. This result is easy to prove when T is ergodic thanks to the fact that in this case the right-hand side is a not a function but a constant: see Proposition 4 and its application below. One could reduce the general case to the ergodic one when an ergodic decomposition exists (this goes back to J. von Neumann [33, 1932]; see also Kryloff and Bogoliouboff [30, 1937] and my survey [46]). A correct, short proof seems given by Choirat-Hess-Seri [12].

Proposition 4. *Let $r\wedge p$ denote* $\inf(r,p)$. *Let $(r_i)_{i\in\mathbb{N}^*}$ be a sequence in* $[0,+\infty]$. *Suppose that, for any $p\in\mathbb{N}$, $\displaystyle\frac{1}{n}\sum_{i=1}^n (r_i\wedge p)$ converges to $\ell_p\in[0,+\infty]$ as n tends to $+\infty$. Then $\displaystyle\varliminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^n r_i \ge \lim_{p\to\infty}\ell_p$.*

Proof. Firstly $(\ell_p)_p$ is increasing, hence $\ell := \lim_{p\to\infty} \ell_p$ exists in $[0, +\infty]$. And as, for any p, $\varliminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} r_i \geq \ell_p$, one has $\varliminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} r_i \geq \ell$. □

Consequence. Theorem 3 is easy to prove when T is ergodic (i.e. $\mathbf{E}^{\mathcal{I}} = \mathbf{E}$).

Proof. There is nothing new to prove when $\mathbf{E}(X)$ is finite. So assume $\mathbf{E}(X) = +\infty$. For any $p \in \mathbb{N}$, by the classical Birkhoff theorem

$$\frac{1}{n}\sum_{i=0}^{n-1} [X \wedge p](T^i\omega) \overset{\text{a.s.}}{\to} \mathbf{E}(X \wedge p)\,.$$

By the monotone convergence theorem $\mathbf{E}(X \wedge p) \nearrow \mathbf{E}(X) = +\infty$. So the result follows from Proposition 4. □

Example. This example shows that in general, $\frac{1}{n}\sum_{i=1}^{n} r_i$ does not converge to ℓ, excepted if $\ell = +\infty$ (in other words a monotone convergence theorem for Cesàro's limits is lacking). Let

$$r_i = \begin{cases} 2^j, & \text{if } i \text{ has the form } i = 2^j, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\forall p$, $\ell_p = 0$, $\varliminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} r_i = 1$ and $\varlimsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} r_i = 2$.

5. Ergodic theorem for integrands

Besides the transformation T introduced in Section 4, we will consider E, a separable metric Suslin space (note that separability is a consequence of Suslin) and f, a l.s.c. ≥ 0 integrand on $\Omega \times E$ (this means: f is $(\mathcal{F} \otimes \mathcal{B}(E))$-measurable and l.s.c. in x). We will obtain, at least for almost all ω, the epi-graphical convergence of $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i\omega, .)$. Recall that, E being a metric space, the functions $\varphi_n : E \to \overline{\mathbb{R}}$ *epi-converge to* φ if their epigraphs converge in the *Painlevé-Kuratowski sense*. This allows some "horizontal deformations" and is a fundamental concept in optimization theory (see the books of Attouch [3] and Rockafellar-Wets [36]). This convergence expresses without the epigraphs: $\varphi_n \overset{\text{epi}}{\to} \varphi$ iff for any $x \in E$, both the following properties hold

(i) for any sequence $(x_n)_n$ converging to x, $\varphi(x) \leq \varliminf_{n\to\infty} \varphi_n(x_n)$,

(ii) there exists a sequence $(x_n)_n$ converging to x, such that $\varphi(x) \geq \varlimsup_{n\to\infty} \varphi_n(x_n)$.

Notations. For any real function h on $\Omega \times E$, $h_n(\omega, x) := \frac{1}{n}\sum_{i=0}^{n-1} h(T^i(\omega), x)$. And, assuming $h \geq 0$, $h^k(\omega, x) := \inf\{h(\omega, y) + k\, d(y, x) : y \in E\}$ which is the k-Lipschitz approximation (Baire-Hausdorff approximation) in x.

Note that $h^k(\omega, .)$ is finite valued as soon as $h(\omega, .)$ is not identically $+\infty$. One must differentiate $(h_n)^k$ and $(h^k)_n$. It is easy to check $(h_n)^k \geq (h^k)_n$ [10, (5.3.2)].

We will use several arguments of Castaing-Ezzaki [10]. This paper contains a lot of results and very ingenious proofs. The following Lemma is more simple than Lemma 5.1 and Proposition 5.2 of [10].

Lemma 5. *Let h be a ≥ 0 integrand (that is a $(\mathcal{F} \otimes \mathcal{B}(E))$-measurable function from $\Omega \times E$ to $[0, +\infty]$) and $u \in \mathcal{L}^0(\Omega, \mathcal{I}, P; E)$. Then*

$$\frac{1}{n} \sum_{i=0}^{n-1} h(T^i \omega, u(\omega)) \overset{a.s.}{\rightarrow} (\mathbf{E}^{\mathcal{I}} h)(\omega, u(\omega)) \,.$$

Proof. By Theorem 3 there exists a negligible set N_u such that, on $\Omega \backslash N_u$,

$$\frac{1}{n} \sum_{i=0}^{n-1} \big[\widehat{h}(u)\big](T^i \omega) \longrightarrow \Big[\mathbf{E}^{\mathcal{I}}\big(\widehat{h}(u)\big)\Big](\omega) \,.$$

Observing that, for any $i \in \mathbb{N}$, $u \circ T^i = u$ and $\mathbf{E}^{\mathcal{I}}\big(\widehat{h}(u)\big) = \big(\widehat{\mathbf{E}^{\mathcal{I}} h}\big)(u)$ (cf. (2')), one gets

$$\frac{1}{n} \sum_{i=0}^{n-1} h(T^i \omega, u(\omega)) \overset{\text{a.s.}}{\rightarrow} (\mathbf{E}^{\mathcal{I}} h)(\omega, u(\omega)) \,.$$

□

The following theorem corresponds to Theorem 5.3 and Lemma 5.4 of Castaing-Ezzaki [10]. For a spatial sub-additive process see (under the ergodic hypothesis) Theorem 4.1 in Licht-Michaille [31] or Theorem 5.1 of [32].

Theorem 6. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, T a measurable transformation of Ω preserving P, $\mathcal{I}$ the σ-algebra of invariant sets, E a Suslin metric space, f a l.s.c. ≥ 0 integrand on $\Omega \times E$. Then the following epi-convergence holds:*

$$\text{for } P\text{-almost all } \omega, \quad \frac{1}{n} \sum_{i=0}^{n-1} f(T^i \omega, .) \overset{epi}{\rightarrow} (\mathbf{E}^{\mathcal{I}} f)(\omega, .) \,. \tag{7}$$

Proof. 1) With the notation set above, (7) is equivalent to

$$f_n(\omega, .) \overset{\text{epi}}{\rightarrow} (\mathbf{E}^{\mathcal{I}} f)(\omega, .) \,.$$

In Hess [23, Prop.3.4 p.1304], an analogous of the Attouch theorem concerning Moreau-Yosida approximation [3, Th.2.65 p.232] is given for the Lipschitz approximation. Thanks to this result it suffices to prove

$$\sup_{k \in \mathbb{N}} \Big[\varliminf_{n \to \infty} (f_n)^k(\omega, .)\Big] \geq (\mathbf{E}^{\mathcal{I}} f)(\omega, .) \tag{8}$$

and

$$\sup_{k \in \mathbb{N}} \Big[\varlimsup_{n \to \infty} (f_n)^k(\omega, .)\Big] \leq (\mathbf{E}^{\mathcal{I}} f)(\omega, .) \,. \tag{9}$$

2) Let us prove (8). Let $(x_p)_p$ denote a dense sequence in E. By Lemma 5 there exists a negligible set $N_{k,p}$ such that,

$$\forall \omega \in \Omega \backslash N_{k,p}, \quad \frac{1}{n} \sum_{i=0}^{n-1} f^k(T^i \omega, x_p) \longrightarrow (\mathbf{E}^{\mathcal{I}} f^k)(\omega, x_p) \,.$$

Since the functions $\mathbf{E}^{\mathcal{I}} f^k$, $(f_n)^k$ and $(f^k)_n$ are k-Lipschitz in x, one has for any $\omega \notin N_k := \bigcup_p N_{k,p}$ and any $x \in E$,

$$\underline{\lim}_{n\to\infty} (f_n)^k(\omega, x) \geq \underline{\lim}_{n\to\infty} (f^k)_n(\omega, x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f^k(T^i\omega, x) = (\mathbf{E}^{\mathcal{I}} f^k)(\omega, x) \quad (10)$$

Recall that, when $k \to \infty$, $(\mathbf{E}^{\mathcal{I}} f^k)(\omega, x) \nearrow (\mathbf{E}^{\mathcal{I}} f)(\omega, x)$: there is a monotone convergence theorem for conditional expectations [19, 10.1.7 p.266]. Taking sup over k in both extreme terms of (10), we get (8) with the negligible set $\bigcup_k N_k$.

3) Now we attack (9). Let $g = \mathbf{E}^{\mathcal{I}} f$ and let Φ be the multifunction $\omega \mapsto \operatorname{epi} g(\omega, .)$. We will restrict ourselves to $\Omega_0 := \{\omega \in \Omega : g(\omega, .) \not\equiv +\infty\} = \{\omega \in \Omega : \Phi(\omega) \neq \emptyset\}$. Let $\big(u_p(.), r_p(.)\big)_p$ be a sequence of $\mathcal{I}$-measurable functions which constitute a Castaing representation of Φ (i.e. for all ω, $\big(u_p(\omega), r_p(\omega)\big)_p$ is dense in $\Phi(\omega)$). Let $m \in \mathbb{N}^*$ be fixed for a time. For all k and p, the multifunction

$$\Gamma_{k,p,m}(\omega) = \{y \in E : g(\omega, y) + k\, d(y, u_p(\omega)) \leq g^k(\omega, u_p(\omega)) + m^{-1}\}$$

is non empty valued and its graph belongs to $\widehat{\mathcal{I}} \otimes \mathcal{B}(E)$ hence admits a $\widehat{\mathcal{I}}$-measurable selection $v_{k,p,m}$. Modifying it on a negligible set one can suppose it is $\mathcal{I}$-measurable. By its definition $v_{k,p,m}$ satisfies

$$g(\omega, v_{k,p,m}(\omega)) + k\, d(v_{k,p,m}(\omega), u_p(\omega)) \leq g^k(\omega, u_p(\omega)) + m^{-1}\,.$$

Thanks to Lemma 5 there exists a negligible set $N_{k,p,m}$ outside of which the following convergence holds:

$$\begin{aligned} f_n(\omega, v_{k,p,m}(\omega)) = \frac{1}{n} \sum_{i=0}^{n-1} f(T^i\omega, v_{k,p,m}(\omega)) &\longrightarrow (\mathbf{E}^{\mathcal{I}} f)(\omega, v_{k,p,m}(\omega)) \\ &= g(\omega, v_{k,p,m}(\omega))\,. \end{aligned}$$

Then, for $\omega \in \Omega_0 \backslash N_{k,p,m}$

$$\begin{aligned} \overline{\lim}_{n\to\infty} (f_n)^k(\omega, u_p(\omega)) &\leq \overline{\lim}_{n\to\infty} \big[f_n(\omega, v_{k,p,m}(\omega)) + k\, d(v_{k,p,m}(\omega), u_p(\omega))\big] \\ &= g(\omega, v_{k,p,m}(\omega)) + k\, d(v_{k,p,m}(\omega), u_p(\omega)) \\ &\leq g^k(\omega, u_p(\omega)) + m^{-1}\,. \end{aligned}$$

Taking sup over k in extreme terms one gets a negligible set N_m $(= \bigcup_{k,p} N_{k,p,m})$ such that for all k, p et $\omega \in \Omega_0 \backslash N_m$,

$$\sup_{k\in\mathbb{N}} \Big[\overline{\lim}_{n\to\infty} (f_n)^k(\omega, u_p(\omega))\Big] \leq g(\omega, u_p(\omega)) + m^{-1}\,.$$

For ending let $\omega \in \Omega_0 \backslash \bigcup_m N_m$ fixed. For any p,

$$\sup_{k\in\mathbb{N}} \Big[\overline{\lim}_{n\to\infty} (f_n)^k(\omega, u_p(\omega))\Big] \leq g(\omega, u_p(\omega))\,. \quad (11)$$

Let $x \in E$. If $g(\omega, x) = +\infty$, (9) is trivially satisfied. If $g(\omega, x) < +\infty$, there exists a sequence $(p_j)_j$ such that $(x, g(\omega, x)) = \lim_{j\to\infty} (u_{p_j}(\omega), r_{p_j}(\omega))$. One knows that

$$\sup_{k\in\mathbb{N}}\Big[\overline{\lim_{n\to\infty}}(f_n)^k(\omega,.)\Big]$$

is l.s.c. (for example, by [23, Prop.3.4 p.1304], it is the epigraphical limsup of the sequence $(f_n(\omega,.))_n$), hence

$$\sup_{k\in\mathbb{N}}\Big[\overline{\lim_{n\to\infty}}(f_n)^k(\omega,x)\Big] \le \underline{\lim}_{j\to\infty}\Big[\sup_{k\in\mathbb{N}}\Big[\overline{\lim_{n\to\infty}}(f_n)^k(\omega,u_{p_j}(\omega))\Big]\Big]$$
$$\overset{(11)}{\le} \underline{\lim}_{j\to\infty}\, g(\omega,u_{p_j}(\omega)) \le \underline{\lim}_{j\to\infty}\, r_{p_j}(\omega)) = g(\omega,x)\,.$$

This proves (9). □

References

[1] L. Arnold, *Random dynamical systems*, Springer, Berlin, 1998.

[2] Z. Artstein & S. Hart, *Law of large numbers for random sets and allocation processes*, Math. Oper. Res. **6** (1981), 485–492.

[3] H. Attouch, *Variational convergence for Functions and Operators*, Applicable Mathematics Series, Pitman, Boston-London-Melbourne, 1984.

[4] H. Attouch & R.J.-B. Wets, *Epigraphical processes: laws of large numbers for random LSC functions*, Sém. Anal. Convexe **20** (1990), 13.1–13.29.

[5] H. Bauer, *Probability theory and elements of measure theory*, Holt, Rinehart & Winston, New York, 1972 (translation of *Wahrscheinlichkeitstheorie und Grundzge der Masstheorie*, De Gruyter, Berlin, 1968).

[6] J.-M. Bismut, *Intégrales convexes et probabilités*, J. Math. Anal. Appl. **42** (1973), 639–673.

[7] C. Castaing, *Le théorème de Dunford-Pettis généralisé*, C. R. Acad. Sci. Paris Sér. A-B **268** (1969), A327–A329.

[8] C. Castaing, *Le théorème de Dunford-Pettis généralisé*, Séminaire Pierre Lelong (Analyse) (année 1969), Lecture Notes in Math., vol. 116, Springer-Verlag, Berlin, 1970, pp. 133–151.

[9] C. Castaing, *À propos de l'existence des sections séparément mesurables et séparément continues d'une multiapplication séparément mesurable et séparément semi-continue inférieurement*, Sém. Anal. Convexe **6** (1976), 6.1–6.6.

[10] C. Castaing & F. Ezzaki, *Variational inequalities for integrand martingales and additive random sequences*, Acta Math. Vietnam. **18** (1993), 137–171 (see also, a bit shorter: Sém. Anal. Convexe **22** (1992) 1.1–1.32 and the Doctorat Thesis of F. Ezzaki, where some misprints are corrected: *Contributions aux problèmes de convergence des suites adaptées et des ensembles aléatoires*, Montpellier, 1993, chapitre II)

[11] C. Castaing & M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Math. vol. 580, Springer-Verlag, Berlin, 1977.

[12] C. Choirat, C. Hess & R. Seri, *A Birkhoff's ergodic theorem for a normal integrand: a variational approach*, Manuscrit, Université Paris-Dauphine, Février 2000.

[13] A. Costé & R. Pallu de La Barrière, *Un théorème de Radon-Nikodým pour les multimesures à valeurs convexes fermées localement compactes sans droite*, C. R. Acad. Sci. Paris Sér. A-B **280** (1975), A255–A258.

[14] A. Costé & R. Pallu de La Barrière, *Radon-Nikodým theorems for set-valued measures whose values are convex and closed*, Comment. Math. Prace Mat. **20** (1977/78), 283–309.

[15] G. Dal Maso & L. Modica, *Nonlinear stochastic homogenization*, Annali Mat. Pura ed Appl. **144** (1986), 347–389.

[16] G. Dal Maso & L. Modica, *Nonlinear stochastic homogenization and ergodic theory*, J. Reine Angew. Math. **368** (1986), 28–42.

[17] G. Debreu & D. Schmeidler, *The Radon-Nikodým derivative of a correspondence*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California,

Berkeley, Calif., 1970/1971), Vol. II: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 41–56.

[18] J.L. Doob, *Stochastic processes*, Wiley, New York, 1965 (sixth printing. First printing, 1953).

[19] R.M. Dudley, *Real analysis and probability*, Wadsworth & Brooks/Cole Mathematics Series, Pacific Grove, California, 1989.

[20] I.V. Evstigneev, *Regular conditional expectations of random variables depending on parameters*, Theory Probab. Appl. **31** (1986), 515–518 (in Russian in Teor. Veroyatnost. i Primenen. **31** (1986) 586–589).

[21] C. Hess, *Théorème ergodique et loi forte des grands nombres pour des ensembles aléatoires*, C. R. Acad. Sci. Paris Sér. A-B **288** (1979), A519–A522.

[22] C. Hess, *Loi forte des grands nombres pour des ensembles aléatoires non bornés à valeurs dans un espace de Banach séparable*, C. R. Acad. Sci. Paris Sér. I **300** (1985), 177–180.

[23] C. Hess, *Epi-convergence of sequences of normal integrands and strong consistency of the maximum likelihood estimator*, Ann. Statist. **24** (1996), 1298–1315.

[24] F. Hiai, *Convergence of conditional expectations and strong laws of large numbers for multivalued random variables*, Trans. Amer. Math. Soc. **291** (1985), 613–627.

[25] F. Hiai & H. Umegaki, *Integrals, conditional expectations, and martingales of multivalued functions*, J. Multivariate Anal. **7** (1977), 149–182.

[26] H.-A. Klei, *Le théorème de Radon Nikodým pour des multi-mesures à valeurs faiblement compactes*, C. R. Acad. Sci. Paris Sér. I Math. **297** (1983), 643–646.

[27] A.N. Kolmogorov, *Foundations of the theory of probability*, Chelsea, New York, 1950 (translation of *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933).

[28] L.A. Korf & R.J.-B. Wets, *Random lsc functions: scalarization*, Manuscript, University of California, Davis, 1999.

[29] L.A. Korf & R.J.-B. Wets, *Random lsc functions: an ergodic theorem*, Manuscript, University of California, Davis, 1999.

[30] N. Kryloff & N. Bogoliouboff, *La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire*, Ann. of Math. **38** (1937), 65–113.

[31] C. Licht & G. Michaille, *Convergence results for subadditive processes*, Prépublication n° 28 (15 pages) (1994), Départ. de Math., Université de Montpellier, France.

[32] C. Licht & G. Michaille, *Global-local subadditive ergodic theorems and application to homogenization in elasticity*, submitted (1999).

[33] J. von Neumann, *Zur Operatorenmethode in der klassischen Mechanik*, Ann. of Math. **33** (1932), 587–642, 789–791 (pages 307–365 in [34]).

[34] J. von Neumann, *Collected works*, vol. II, Operators, Ergodic theory and almost periodic functions in a group (A.H. Taub, ed.), Pergamon Press, Oxford, 1961.

[35] J. Neveu, *Martingales à temps discret*, Masson, Paris, 1972 (in English: *Discrete-parameter martingales*, North-Holland, Amsterdam; Elsevier, New York, 1975).

[36] R.T. Rockafellar & R.J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[37] M.-F. Sainte-Beuve, *On the extension of von Neumann-Aumann's theorem*, J. Functional Analysis **17** (1974), 112–129.

[38] L. Thibault, *Espérances conditionnelles d'intégrandes semi-continus*, Ann. Inst. H. Poincaré Sect. B (N.S.) **17** (1981), 337–350.

[39] A. Truffert, *Conditional expectation of integrands and random sets*, Stochastic Programming Part I (Ann Arbor, MI, 1989), Ann. Op. Res. **30** (1991), 117–156.

[40] M. Valadier, *Contribution à l'analyse convexe*, Thèse de Doctorat d'État, Paris, 1970.

[41] M. Valadier, *Multi-applications mesurables à valeurs convexes compactes*, J. Math. Pures Appl. **50** (1971), 265–297.

[42] M. Valadier, *Espérance conditionnelle d'un convexe fermé aléatoire*, C. R. Acad. Sci. Paris Sér. A-B **273** (1971), A1265–A1267 (detailed version of *Espérance conditionnelle d'un convexe fermé aléatoire*, Sém. Anal. Convexe **2**, (1972) 1.1–1.15).

[43] M. Valadier, *Désintégration d'une mesure sur un produit*, C. R. Acad. Sci. Paris Sér. A **276** (1973), 33–35 (detailed version of *Comparaison de trois théorèmes de désintégration*, Sém. Anal. Convexe **2** (1972) 10.1–10.21).

[44] M. Valadier, *Structure des atomes conditionnels d'un espace de probabilité*, Sém. Anal. Convexe **10** (1980), 1.1–1.17.
[45] M. Valadier, *Sur l'espérance conditionnelle multivoque non convexe*, Ann. Inst. H. Poincaré Sect. B (N.S.) **16** (1980), 109–116 (published version of *Sur l'espérance conditionnelle multivoque non convexe*, Sém. Anal. Convexe **10** (1980) 3.1–3.10).
[46] M. Valadier, *What differentiates stationary stochastic processes from ergodic ones: a survey*, submitted.

Manuscript received July 17, 2000

M. VALADIER
Département de Mathématiques — case 051, Université Montpellier II, place Eugène Bataillon, 34 095 MONTPELLIER CEDEX 5, France
E-mail address: valadier@math.univ-montp2.fr