MARTINGALE THEORY
IN HARMONIC ANALYSIS
AND BANACH SPACES
Cleveland 1981
Springer's LNM 939

ON COMPACTNESS AND OPTIMALITY OF STOPPING TIMES

G. A. Edgar*, A. Millet and L. Sucheston**

Let B be a Banach space with norm $\| \ \|$. Suppose that we are allowed to view successively as many terms as we please of a sequence of B-valued random variables X_n. We stop viewing at a time n of our choice, and receive payoff X_n. Is there a non-anticipative stopping rule σ which would maximize a continuous convex function ϕ of the expected value of X_n? We allow stopping rules (= times) τ taking on the value ∞, and call σ optimal if the ϕ-<u>value</u>

$$V_\phi = \sup_\tau \phi[E(X_\tau)]$$

is achieved for σ. One interesting case is $X_n = \frac{1}{n}(Y_1+Y_2 \ldots +Y_n)$, where the B-valued process (Y_n) is stationary, and ϕ is the norm $\| \ \|$, or, more generally, the distance from a fixed convex set in B. We show that if $E(\|Y_1\| \log^+ \|Y_1\|) < \infty$, then an optimal σ exists. If the Y_n are independent (which implies that X_n is a descending martingale), ϕ is sublinear, and $E(\|Y_1\|^p) < \infty$ for some $p > 1$, then σ is finite a.s. If Y_n are real-valued, independent and identically distributed, and $E(|Y_1| \log^+ |Y_1|) = \infty$, then there exists a stopping time σ such that $E(|X_\sigma|) = \infty$ (B. Davis [7], and B. J. McCabe and L. A. Shepp [18]). This result is generalized here to Banach spaces, and the independence assumption is replaced by a weaker condition (I).

---

* The research of this author is supported by the National Science Foundation, Grant MCS-8003078.

** The research of this author is supported by the National Science Foundation, Grant MCS-8005395.

Except for the condition (I), our results are known in the real-valued case: see in particular D. Siegmund [24], Chow-Robbins-Siegmund [6], B. Davis [8], and M. Klass [14]. The article of Klass is a complete and self-contained presentation of the subject. It seems however that the real proofs do not extend; in particular, there are no analogues of admissible [6] (= regular [14]) stopping times, or of the Snell stopping time (see Snell [25], or Neveu [21], p.124). Instead, we apply a recent important theorem of Baxter-Chacon [1]: any sequence of stopping times τ_n (chosen here so that $\phi[E(X_{\tau_n})] \to V_\phi$) admits a subsequence, still denoted τ_n, which converges to a randomized stopping time γ in the Baxter-Chacon topology. We show that under proper boundedness assumptions this implies that $EX_{\tau_n} \to EX_\gamma$, hence γ is optimal. To "derandomize", we take a closer look at the set of randomized stopping times, noting that the non-random stopping times are exactly its extreme points. As an application, one proves the existence of a non-random optimal stopping time.

Section 1 discusses the Baxter-Chacon topology and extreme points. In Section 2 we prove a general theorem about the existence of optimal stopping times, and apply it. Section 3 considers the case when $E(\|Y_1\| \log^+ \|Y_1\|) = \infty$. A discussion of the continuous parameter case - the original setting of the Baxter-Chacon article - is given in Section 4.

1. Compactness and extreme points of stopping times. The following notation will be used throughout the paper. $\mathbb{R}$ is the set of real numbers; $\mathbb{N} = \{1,2,3,\ldots\}$ has its discrete topology; $\overline{\mathbb{N}} = \{1,2,3,\ldots,\infty\}$ is the one-point compactification of $\mathbb{N}$; $\mathcal{B}$ is the σ-algebra of Borel subsets of $[0,1]$; λ is Lebesgue measure on $\mathcal{B}$. If S is a topological space, then $\mathcal{C}(S)$ denotes the set of bounded continuous functions $f: S \to \mathbb{R}$.

Let $(\Omega,\mathcal{F},P)$ be a probability space, and let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be an increasing sequence of sub-σ-algebras of $\mathcal{F}$. By convention, we will write $\mathcal{F}_\infty$ for the σ-algebra generated by $\cup_{n=1}^{\infty}\mathcal{F}_n$. A sequence $(X_n)_{n\in\mathbb{N}}$ of random variables is said to be adapted to the sequence $(\mathcal{F}_n)$ iff X_n is $\mathcal{F}_n$-measurable for all n. This situation frequently occurs in the reverse order: a sequence $(X_n)_{n\in\mathbb{N}}$ of random variables is known, and $\mathcal{F}_n$ is defined to be the σ-algebra generated by $X_1,\ldots,X_n$. In this case, we call $(\mathcal{F}_n)$ the natural σ-algebras for

(X_n). Note that F_∞ is countably generated in this case.

A stopping time of (F_n) is a function $\sigma\colon \Omega \to \overline{\mathbb{N}}$ such that $\{\omega\colon \sigma(\omega) = n\} \in F_n$ for all $n \in \overline{\mathbb{N}}$. We will write Σ or $\Sigma((F_n)_{n \in \mathbb{N}})$ for the set of all stopping times.

We will often extend the probability space (Ω, F, P) to a larger one, namely $(\Omega \times [0,1], F \times B, P \times \lambda)$. A random variable $X\colon \Omega \to \mathbb{R}$ corresponds to a random variable $\tilde{X}\colon \Omega \times [0,1] \to \mathbb{R}$ defined by $\tilde{X}(\omega,v) = X(\omega)$; normally we will write X for both cases. The notation E for expectation will be used both for $\int \ldots dP$ and for $\iint \ldots dPd\lambda$. A randomized stopping time for (F_n) is simply a stopping time for the sequence $(F_n \times B)$. To every randomized stopping time $\gamma\colon \Omega \times [0,1] \to \overline{\mathbb{N}}$ there corresponds a unique increasing rearrangement $\tilde{\gamma}\colon \Omega \times [0,1] \to \overline{\mathbb{N}}$ such that

$$\lambda\{v\colon \gamma(\omega,v) = n\} = \lambda\{v\colon \tilde{\gamma}(\omega,v) = n\}$$

for all $\omega \in \Omega$, $n \in \overline{\mathbb{N}}$, and such that for each $\omega \in \Omega$, the function $\tilde{\gamma}(\omega,\cdot)$ is increasing and left-continuous. In most situations occurring in this paper, rearrangement with respect to the variable v will make no difference. For example, if $(X_n)_{n \in \overline{\mathbb{N}}}$ is adapted to $(F_n)_{n \in \overline{\mathbb{N}}}$ and γ, $\tilde{\gamma}$ are as above, then $E(X_\gamma) = E(X_{\tilde{\gamma}})$. We will write Γ or $\Gamma((F_n)_{n \in \mathbb{N}})$ for the set of all randomized stopping times, increasing and left-continuous in the second variable.

Baxter and Chacon [1] have defined a useful topology for the set Γ of randomized stopping times. For completeness, that definition is repeated here for discrete time. (See Section 4, below, for a brief discussion of the continuous time case.) For $\gamma \in \Gamma$, the ω-distribution of γ is defined by

$$M(\omega,K) = \lambda\{v\colon \gamma(\omega,v) \in K\}$$

for $\omega \in \Omega$, $K \subseteq \overline{\mathbb{N}}$. Then M has the following properties:

(a) For fixed $\omega \in \Omega$, the function $M(\omega,\cdot)$ is a probability measure on $\overline{\mathbb{N}}$:

(b) For fixed $n \in \mathbb{N}$, the function $M(\cdot,\{n\})$ is F_n-measurable.

We will write Γ' for the set of all functions M satisfying (a) and (b). (In order to define an element M of Γ', it suffices to define $M(\omega,\{n\})$ for $n \in \mathbb{N}$ and add for other sets $K \subseteq \overline{\mathbb{N}}$, or to define $M(\omega,\{1,\ldots,n\})$ for $n \in \mathbb{N}$ and subtract to obtain $M(\omega,\{n\})$.)

If $M \in \Gamma'$ is given, we may conversely define a randomized stopping time $\gamma \in \Gamma$ by

$$\gamma(\omega,v) = \inf\{n \in \overline{\mathbb{N}}: M(\omega,\{1,\ldots,n\}) \geq v\}.$$

Thus Γ and Γ' are in one-to-one correspondence. Notice that $M \in \Gamma'$ corresponds to a nonrandomized stopping time σ if and only if

$$M(\omega,\{n\}) = \begin{cases} 1, & \text{if } \sigma(\omega) = n \\ 0, & \text{if } \sigma(\omega) \neq n. \end{cases}$$

The Baxter-Chacon topology is the coarsest topology on Γ' such that, for all $n \in \mathbb{N}$ and all $Y \in L^1(\mathcal{F})$, the map $M \to \int Y(\omega)M(\omega,\{n\})P(d\omega)$ is continuous. Thus, for sequences, this means that M_k converges to M in the Baxter-Chacon topology iff

$$\lim_{k\to\infty} \int Y(\omega)M_k(\omega,\{n\})P(d\omega) = \int Y(\omega)M(\omega,\{n\})P(d\omega)$$

for all $n \in \mathbb{N}$ and all $Y \in L^1(\mathcal{F})$. We define the Baxter-Chacon topology on Γ via the bijection above. We write $\gamma_k \to \gamma(BC)$ iff $\lim_{k\to\infty} E(Y1_{\{n\}}(\gamma_k)) = E(Y1_{\{n\}}(\gamma))$ for all $n \in \mathbb{N}$ and all $Y \in L^1(\mathcal{F})$. (Of course, this is the topology induced on the set of randomized stopping times by a weak-star topology.) The usefulness of this topology is due largely to the following result of Baxter and Chacon [1]. For an early very general compactness argument see LeCam [16].

1.1. THEOREM. The set Γ of randomized stopping times is compact in the Baxter-Chacon topology. If $\mathcal{F}$ is countably generated, then Γ is metrizable, and therefore sequentially compact.

The set of all functions M such that

(a) For each $\omega \in \Omega$, $M(\omega,\cdot)$ is a signed measure on $\overline{\mathbb{N}}$;

(b) For each $n \in \mathbb{N}$, $M(\cdot,\{n\})$ is $\mathcal{F}_n$-measurable;

(c) There is a constant C such that $|M(\omega,k)| \leq C$ a.s. for all $K \subseteq \overline{\mathbb{N}}$;

is a topological vector space under the Baxter-Chacon topology. The set Γ' is a compact convex subset of it. The extreme points of Γ' are exactly the ω-distributions of the nonrandomized stopping times. Each element of Γ' can be represented as a continuous average of these extreme points. This can be proved using Choquet's theorem, but it can also be deduced from the equation

$$M = \int_0^1 M_{\gamma_0(\cdot,v)}dv \,, \tag{1.2}$$

where $M \in \Gamma'$ corresponds to $\gamma_0 \in \Gamma$, and for each $v \in [0,1]$, we write $M_{\gamma_0(\cdot,v)}$ for the ω-distribution of the nonrandomized stopping time $\omega \mapsto \gamma_0(\omega,v)$. Equation (1.2) can be interpreted to mean

$$(1.3) \qquad M(\omega,K) = \int_0^1 M_{\gamma_0(\cdot,v)}(\omega,K)\,dv$$

for all $\omega \in \Omega$, $K \subseteq \overline{\mathbb{N}}$. It follows from this that

$$(1.4) \qquad E(X_{\gamma_0}) = \int_0^1 E[X_{\gamma_0(\cdot,v)}]\,dv$$

for any adapted sequence $(X_n)_{n \in \overline{\mathbb{N}}}$ for which the right-hand side exists.

This equation can be used to "derandomize" optimal stopping times.

1.5. PROPOSITION. Let $(X_n)_{n \in \overline{\mathbb{N}}}$ be adapted to $(\mathcal{F}_n)_{n \in \mathbb{N}}$. Then

$$\sup_{\gamma \in \Gamma} E(X_\gamma) = \sup_{\sigma \in \Sigma} E(X_\sigma),$$

and if one of the suprema is achieved and finite, so is the other.

Proof. Write

$$V = \sup_{\sigma \in \Sigma} E(X_\sigma).$$

Assume that $V < \infty$. Suppose there exists $\gamma_0 \in \Gamma$ with $E(X_{\gamma_0}) \geq V$. Then from (1.4), we have

$$V \leq E(X_{\gamma_0}) = \int_0^1 E(X_{\gamma_0(\cdot,v)})\,dv$$

$$\leq \int_0^1 V\,dv = V.$$

Therefore $E(X_{\gamma_0(\cdot,v)}) = V$ for almost all $v \in [0,1]$, and hence for at least one v. But then we have $E(X_{\gamma_0}) \leq V$ for all $\gamma_0 \in \Gamma$, and if $\sup_{\gamma \in \Gamma} E(X_\gamma)$ is achieved, so is $\sup_{\alpha \in \Sigma} E(X_\sigma)$.

1.6. COROLLARY. If there exists $\gamma_0 \in \Gamma$, finite a.s., with

$$E(X_{\gamma_0}) = \sup_{\gamma \in \Gamma} E(X_\gamma) = V,$$

then there also exists $\sigma_0 \in \Sigma$, finite a.s., with

$$E(X_{\sigma_0}) = V.$$

Proof. Represent γ_0 as in (1.2). Then, for almost all ω, $0 = \lambda\{v: \gamma_0(\omega,v) = \infty\}$. So there exists v with both $P\{\gamma_0(\cdot,v) < \infty\} = 1$ and $E(X_{\gamma_0(\cdot,v)}) = V$.

For a derandomization in the vector-valued case, we use Jensen's inequality in a Banach space B.

1.7. THEOREM. Let $(X_n)_{n \in \overline{\mathbb{N}}}$ be an adapted sequence of Bochner integrable random variables in a Banach space B, and let $\phi: B \to \mathbb{R}$ be continuous and convex. Then

$$\sup_{\gamma \in \Gamma} \phi(E(X_\gamma)) = \sup_{\sigma \in \Sigma} \phi(E(X_\sigma)),$$

and if one of the suprema is achieved and finite, so is the other. If this supremum is achieved by $\gamma_0 \in \Gamma$ which is finite a.s., then it is also achieved by $\sigma_0 \in \Sigma$ which is finite a.s.

Proof. Write

$$V_\phi = \sup_{\sigma \in \Sigma} \phi(E(X_\sigma)).$$

Assume $V_\phi < \infty$. Suppose $\gamma_0 \in \Gamma$ and $\phi(E(X_{\gamma_0})) \geq V_\phi$. Represent γ_0 as in (1.4). Then

$$\begin{aligned} V_\phi \leq \phi(E(X_{\gamma_0})) &= \phi\left(\int_0^1 E(X_{\gamma_0(\cdot,v)})dv\right) \\ &\leq \int_0^1 \phi(E(X_{\gamma_0(\cdot,v)}))dv \\ &\leq \int_0^1 V_\phi dv = V_\phi, \end{aligned}$$

so $\phi(E(X_{\gamma_0(\cdot,v)})) = V_\phi$ for almost all v. The rest of the proof is as before.

2. Optimal stopping time: general case. In this section we study the optimization of $\phi(EX_\tau)$, where (X_n) is a Banach-valued process, and ϕ is a real-valued continuous convex function defined on the Banach space (e.g., the norm). Also conditions are given for the convergence of Banach-valued stopped processes X_{γ_n}, when γ_n

are randomized stopping times converging in the Baxter-Chacon topology. B will denote a Banach space with norm $\| \ \|$.

2.1. LEMMA. Let (γ_n) be a sequence of randomized stopping times that converges to a randomized stopping time γ in the Baxter-Chacon topology. Then for every Bochner integrable random variable Y, and for every function f continuous on $\overline{\mathbb{N}}$, $E[1_A Y f(\gamma_n)]$ converges strongly to $E[1_A Y f(\gamma)]$.

Proof. Fix $f \in C(\overline{\mathbb{N}})$, and first suppose that Y is a step-function, i.e., $Y = \sum_{1 \leq i \leq k} x_i 1_{A_i}$, where $x_i \in B$, and $A_i \in F$, $i = 1,\ldots,k$. Then for any $\eta \in \Gamma$, we have

$$E\, Yf(\eta) = \sum_{1 \leq i \leq k} x_i E[1_{A_i} f(\eta)].$$

By the definition of the Baxter-Chacon topology on Γ, the sequence $E[1_{A_i} f(\gamma_n)]$ converges to $E[1_{A_i} f(\gamma)]$ for every $i = 1,\ldots,k$, and hence the announced strong convergence holds for step-functions. Now let Y be a general Bochner integrable random variable. Fix $\varepsilon > 0$, and let Z be a step-function such that $E\|Y - Z\| \leq \varepsilon$. Then for every $\eta \in \Gamma$ we have

$$\|E[Yf(\eta)] - E[Zf(\eta)]\| \leq \varepsilon \|f\|_\infty .$$

Now apply this inequality with $\eta = \gamma_n$, and $\eta = \gamma$.

Let $(X_n, n \geq 1)$ and X be B-valued random variables, and let $A \in F$. We say that X_n <u>converges</u> to X <u>in distribution</u> on A, in symbols $X_n \Rightarrow X$ on A, if for every continuous and bounded real-valued function g defined on B, $E[1_A g(X_n)]$ converges to $E[1_A g(X)]$. The following proposition gives conditions for the convergence of the stopped process X_{γ_n} to X_γ if $\gamma_n \to \gamma(BC)$.

2.2. THEOREM. Let (γ_n) be a sequence of randomized stopping times that converges to $\gamma \in \Gamma$ in the Baxter-Chacon topology. Let $(B, \| \ \|)$ be a Banach space, and let $(X_n n \in \overline{\mathbb{N}})$ be a B-valued Bochner integrable adapted process, such that X_n converges strongly almost surely to X_∞ as $n \to \infty$. Then for every set $A \in F$, X_{γ_n} converges in distribution to X_γ on A. If furthermore $E(\sup \|X_n\|) < \infty$, then for every set $A \in F$, $E(1_A X_{\gamma_n})$ converges strongly to $E(1_A X_\gamma)$.

Proof. We prove the second part of the theorem first. Suppose that $E(\sup \|X_n\|) < \infty$. Fix $K > 1$; for every set $A \in F$ and for every n,

$$\|E(1_A X_{\gamma_n}) - E(1_A X_\gamma)\| \leq \|E(1_{A \cap \{\gamma_n \leq K\}} X_{\gamma_n}) - E(1_{A \cap \{\gamma \leq K\}} X_\gamma)\|$$

$$+ \|E(1_{A \cap \{\gamma_n > K\}} X_\infty) - E(1_{A \cap \{\gamma > K\}} X_\infty)\|$$

$$+ \|E[1_{A \cap \{K<\gamma_n<\infty\}}(X_{\gamma_n} - X_\infty)]\| + \|E[1_{A \cap \{K<\gamma<\infty\}}(X_\gamma - X_\infty)]\| .$$

Hence

$$\|E(1_A X_{\gamma_n}) - E(1_A X_\gamma)\| \leq \sum_{1 \leq i \leq K} \|E[1_A X_i 1_{\{i\}}(\gamma_n)] - E[1_A X_i 1_{\{i\}}(\gamma)]\|$$

$$+ \|E[1_A X_\infty 1_{[K+1,\infty]}(\gamma_n)] - E[1_A X_\infty 1_{[K+1,\infty]}(\gamma)]\|$$

$$+ 2E[1_A \sup_{K<i<\infty} \|X_i - X_\infty\|] .$$

Since X_n converges strongly to X_∞ a.s. on A, the sequence $1_A \sup_{K<i<\infty} \|X_i - X_\infty\|$, dominated by $\sup \|X_n\|$, converges to zero a.s. as $K \to \infty$. Fix $\varepsilon > 0$, and choose K such that $E[1_A \sup_{K<i<\infty} \|X_i - X_\infty\|] < \varepsilon$. Then applying Lemma 2.1 with $Y = 1_A X_i$, $f = 1_{\{i\}}$, and with $Y - 1_A X_\infty$, $f = 1_{[K+1,\infty]}$, one can choose n_0 such that for every $i = 1,\ldots,K$, one has

$$\sup_{n \geq n_0} \|E[1_A X_i 1_{\{i\}}(\gamma_n)] - E[1_A X_i 1_{\{i\}}(\gamma)]\| \leq \varepsilon/K,$$

and also

$$\sup_{n \geq n_0} \|E[1_A X_\infty 1_{[K+1,\infty]}(\gamma_n)] - E[1_A X_\infty 1_{[K+1,\infty]}(\gamma)]\| \leq \varepsilon.$$

Then $n \geq n_0$ implies $\|E(1_A X_{\gamma_n}) - E(1_A X_\gamma)\| \leq 4\varepsilon$, which proves the second statement in the proposition.

We now prove the first assertion of the theorem. Let g be a continuous bounded function from B to $\mathbb{R}$, and let $Z_n = g(X_n)$, $n \in \overline{\mathbb{N}}$. The real-valued process (Z_n) clearly satisfied the two assumptions $Z_n \to Z_\infty$ a.s., and $E(\sup|Z_n|) < \infty$. Hence by the first argument, for every $A \in F$, $E[1_A g(X_{\gamma_n})]$ converges to $E[1_A g(X_\gamma)]$. This completes the proof.

An example of a process (X_n) satisfying the hypotheses of Theorem 2.2 is an L_1-bounded martingale with values in a Banach space B with the Radon-Nikodym Property.

2.3. COROLLARY. Let $(B, \| \ \|)$ be a Banach space, let $\phi: B \to \mathbb{R}$ be a continuous function. Let $(X_n: n \in \overline{\mathbb{N}})$ be a B-valued adapted process such that $E(\sup \|X_n\|) < \infty$, and such that X_n converges strongly almost surely to X_∞. Then there exists a randomized stopping time γ such that

$$\phi(EX_\gamma) = V_\phi = \sup\{\phi(EX_\eta): \eta \in \Gamma\} < \infty.$$

Proof. Since only countably many Bochner integrable random variables are involved, we may assume that F is countably generated. Choose a sequence γ_n in Γ such that $\lim \phi(EX_{\gamma_n}) = V_\phi$. Since the set Γ is sequentially compact for the Baxter-Chacon topology, there exists a subsequence (γ_{n_k}) of (γ_n), and a $\gamma \in \Gamma$ such that $\gamma_{n_k} \to \gamma(BC)$, By Proposition 2.2, the sequence $EX_{\gamma_{n_k}}$ converges strongly to EX_γ; the continuity of ϕ implies that $\phi(EX_{\gamma_n}) \to \phi(EX_\gamma) = V_\phi < \infty$.

Using the results in Section 1, we obtain the existence of non-randomized optimal stopping times if ϕ is convex.

2.4. THEOREM. Let $(B, \| \ \|)$ be a Banach space and let $\phi: B \to \mathbb{R}$ be a convex continuous function. Let $(X_n: n \in \overline{\mathbb{N}})$ be a B-valued process such that $E(\sup \|X_n\|) < \infty$, and X_n converges strongly a.s. to X_∞ as $n \to \infty$. Then there exists a nonrandomized stopping time $\sigma \in \Sigma$ such that

$$\phi(EX_\sigma) = V_\phi = \sup\{\phi(EX_\tau): \tau \in \Sigma\} \tag{2.5}$$
$$= \sup\{\phi(EX_\gamma): \gamma \in \Gamma\} < \infty .$$

Proof. Since the function ϕ is continuous, Corollary 2.3 insures the existence of an optimal randomized stopping time $\gamma \in \Gamma$. By Theorem 1.7, the convexity of ϕ insures that $\sup\{\phi(EX_\tau): \tau \in \Sigma\} = \sup\{\phi(EX_\gamma): \gamma \in \Gamma\}$, and that there exists a nonrandomized stopping time $\sigma \in \Sigma$ such that $V_\phi = \phi(EX_\sigma)$.

We now give examples of processes $(X_n: n \in \overline{\mathbb{N}})$ and functions ϕ

that satisfy the assumptions of Theorem 2.4. Recall that if $(Y_n : n \in \mathbb{N})$ is a stationary B-valued process with $E\,\|Y_1\| < \infty$, then by E. Mourier's ergodic theorem [20], the Cesaro averages $X_n = \frac{1}{n}\sum_{i\leq n} Y_i$ converge strongly a.s. to a random variable X_∞ with $EX_\infty = EY_1$.

2.6. THEOREM. Let $(B, \|\ \|)$ be a Banach space, and let (Y_n) be a B-valued stationary stochastic process with $E(\|Y_1\| \log^+ \|Y_1\|) < \infty$. For every $n \in \mathbb{N}$, set $X_n = \frac{1}{n}\sum_{i<n} Y_i$, and let X_∞ be the almost sure limit of X_n. Then given any continuous convex function $\phi: B \to \mathbb{R}$, there exists a nonrandomized stopping time $\sigma \in \Sigma$ such that

$$\phi(EX_\sigma) = V_\phi = \sup\{\phi(EX_\tau): \tau \in \Sigma\} < \infty. \tag{2.7}$$

Proof. By Wiener's dominated ergodic theorem applied to the real-valued stationary process $\|Y_n\|$, we have $\sup \|X_n\| \leq \sup \frac{1}{n} \sum_{1<i<n} \|Y_i\| \in L_1$ (see e.g. [10], p. 678). Now apply Mourier's theorem and Theorem 2.4.

2.8. COROLLARY. Let $(Y_n,\ n \in \mathbb{N})$, and $(X_n: n \in \overline{\mathbb{N}})$ be as in Theorem 2.6. Given any convex set $C \subset B$, and for every $x \in B$, let $\phi(x)$ denote the distance between x and C. Then there exists an optimal stopping time for ϕ, i.e., an element $\sigma \in \Sigma$ such that (2.7) is satisfied for ϕ.

Proof. It suffices to notice that the distance between x and a convex set is a continuous, convex, real-valued function.

The corollary shows in particular that there exists an optimal stopping time for the norm of the Cesaro averages X_n of a stationary, L Log L-bounded process taking values in $\mathbb{R}^2$. The following example implies that there does not exist a stopping time $\sigma \in \Sigma$ optimal for the X_n; i.e., such that $EX_\sigma = \sup\{EX_\tau: \tau \in \Sigma\}$, even in the case where the Y_i's are independent, identically distributed, positive and bounded. (The supremum is to be taken in the coordinate-wise ordering of $\mathbb{R}^2$.) This example also shows that the stopping times σ such that $\|EX_\sigma\| = V_{\|\ \|}$ depend on the choice of the norm $\|\ \|$, and that given a norm, the optimal stopping times for $\|EX_\sigma\|$ and for $E\,\|X_\sigma\|$ need not be the same.

2.9. Example. Let (A_n) be a sequence of independent events each of probability $1/2$. For every $n \in \mathbb{N}$, set

$$Y_n = (1,0)1_{A_n} + (0,1)1_{A_n^c}, \quad X_n = \frac{1}{n} \sum_{1 \le i \le n} Y_i,$$

and let $X_\infty = (\frac{1}{2},\frac{1}{2})$ be the a.s. limit of X_n. For every randomized stopping time $\gamma \in \Gamma$, and for every $(\omega,v) \in \Omega \times [0,1]$, $X_\gamma(\omega,v) = (a_\gamma(\omega,v), b_\gamma(\omega,v))$, with $a_\gamma(\omega,v) + b_\gamma(\omega,v) = 1$. Hence if $EX_\gamma = (x_\gamma, y_\gamma)$, then $x_\gamma + y_\gamma = 1$. However, let σ_1 [resp., σ_2] be the a.s. finite stopping time defined by $\sigma_1(\omega) = \inf\{n: \omega \in A_n\}$ [resp., $\sigma_2(\omega) = \inf\{n: \omega \in A_n^c\}$]. Then

$$[X_{\sigma_1}]_1 = 1_{A_1} + \frac{1}{2} 1_{A_1^c \cap A_2} + \ldots + \frac{1}{p} 1_{A_1^c \cap \cdots \cap A_{p-1}^c \cap A_p} + \ldots, \text{ and}$$

$EX_{\sigma_1} = (x_{\sigma_1}, y_{\sigma_1})$ with $x_{\sigma_1} = \sum_{i \ge 1} \frac{1}{i}(\frac{1}{2})^i > \frac{5}{8}$. A similar computation shows that $EX_{\sigma_2} = (x_{\sigma_2}, y_{\sigma_2})$ with $y_{\sigma_2} > \frac{5}{8}$. Hence

$\sup\{EX_\sigma: \sigma \in \Sigma, \ \sigma \text{ finite a.s.}\} > (\frac{5}{8},\frac{5}{8})$, and this supremum cannot be achieved by a randomized stopping time. Now set $\|(x,y)\|_1 = |x| + |y|$ and $\|(x,y)\|_\infty = \sup(|x|,|y|)$. Then for $i = 1, \infty$, and for every $\gamma \in \Gamma$, $\|EX_\gamma\|_i \le E(\|X_\tau\|_i) \le 1$. By the argument given above, for every stopping time $\tau \in \Sigma$, $\|EX_\tau\|_1 = E\|X_\tau\|_1 = 1$. Clearly $E\|X_1\|_\infty = 1$, $\|EX_1\|_\infty = \frac{1}{2}$, and if σ_1 is the stopping time defined above then $\|EX_{\sigma_1}\|_\infty > \frac{5}{8}$. Hence $\sigma = 1$ is optimal for $\| \ \|_1$, but not for $\| \ \|_\infty$.

The example shows that for a Banach lattice B there need not exist an optimal stopping time. However, for a large class of lattices there exists a maximal stopping time for B^+.

2.10. COROLLARY. Let B be a Banach lattice such that for any $x,y \in B^+$, $x < y$ implies $\|x\| < \|y\|$ (L_p-spaces have this property if $1 \le p < \infty$). Then under the conditions of Theorem 2.4, and assuming also the process positive, there exists a maximal stopping time σ; i.e., $\sigma \in \Sigma$ such that for every $\tau \in \Sigma$, the inequality $EX_\sigma < EX_\tau$ fails.

Proof. Set $\phi(x) = \|x\|$, and let σ be an optimal stopping time for ϕ, i.e., suppose that (2.5) holds. For any $\tau \in \Sigma$, the inequality $EX_\sigma < EX_\tau$ implies $\|EX_\sigma\| < \|EX_\tau\|$, which is a contradiction. Hence σ is maximal.

We show that if the process (X_n) takes values in $\mathbb{R}^p$, one can weaken the assumption that X_n converges a.s. and obtain a result similar to Proposition 2.2. The case of $\mathbb{R}^p$ can be also reduced to the case of $\mathbb{R}^1 = \mathbb{R}$ by consideration of linear functionals, but the proof given below is more in the spirit of the present paper.

2.11. PROPOSITION. Let (γ_n) be a sequence of randomized stopping times that converges to a randomized stopping time γ in the Baxter-Chacon topology. Let $(X_n : n \in \overline{\mathbb{N}})$ be a stochastic process taking values in $\mathbb{R}^p$, and let $A \in F$. Suppose that

(i) $X_\infty \geq \overline{\lim}\, X_n$ on A, and $1_A X_\infty$ is integrable,

(ii) $\sup(1_A X_n^+)$ is integrable,

(iii) $\sup E(1_A X_n^-) \in \mathbb{R}^p$,

(iv) $E(1_A X_\gamma^-) \in \mathbb{R}^p$.

Then $E(1_A X_\gamma) \geq \overline{\lim}\, E(1_A X_{\gamma_n})$.

Proof. The sequences $\sup\limits_{k<i<\infty} (1_A X_i)$ decreases to $1_A \overline{\lim}\, X_n$, as $k \to \infty$, and is bounded from above by $1_A \sup X_n^+ \in L_1$, and from below by $(1_A \overline{\lim}\, X_n)^-$, which is an integrable function by assumption (iii). Hence applying the dominated convergence theorem, we obtain $E(1_A \overline{\lim}\, X_n) = \lim\limits_k E[1_A \sup\limits_{k<i<\infty} X_i]$. For every $x \in \mathbb{R}^p$, set $\|x\| = \sup\{|x_j| : 1 \leq j \leq p\}$, and let $u = (1,\ldots,1)$ be the unit vector of $\mathbb{R}^p$. Fix $\varepsilon > 0$; choose $\alpha > 0$ such that $P \times \lambda(B) < \alpha$ implies $E[1_{A \cap B} \|X_\gamma^-\|] < \varepsilon$, and $E[1_{A \cap B} \|X_\infty\|] < \varepsilon$. Choose $K \in \mathbb{N}$ such that $P \times \lambda\{K < \gamma < \infty\} < \alpha$, and such that

$$E[1_A \sup_{K<i<\infty} X_i] \leq E[1_A \overline{\lim}\, X_n] + \varepsilon u.$$

Then

$$E(1_A X_\gamma) = E(1_{A \cap \{\gamma \leq K\}} X_\gamma) + E(1_{A \cap \{K<\gamma<\infty\}} X_\gamma) + E(1_{A \cap \{\gamma=\infty\}} X_\infty)$$

$$\geq E(1_{A \cap \{\gamma \leq K\}} X_\gamma) - \varepsilon u + E(1_{A \cap \{K<\gamma<\infty\}} X_\infty) - \varepsilon u + E(1_{A \cap \{\gamma=\infty\}} X_\infty).$$

Applying Lemma 2.1 with $Y = 1_A X_i$, $f = 1_{\{i\}}$, and with $Y = 1_A X_\infty$, $f = 1_{[K+1,\infty]}$, we obtain

$$E(1_A X_\gamma) \geq \lim_n E(1_{A \cap \{\gamma_n \leq K\}} X_{\gamma_n}) + \lim E(1_{A \cap \{\gamma_n > K\}} X_\infty) - 2\varepsilon u$$

$$\geq \overline{\lim}\{E(1_{A \cap \{\gamma_n \leq K\}} X_{\gamma_n}) + E(1_{A \cap \{\gamma_n = \infty\}} X_{\gamma_n})$$

$$+ E(1_{A \cap \{K < \gamma_n < \infty\}} \overline{\lim}\, X_k)\} - 2\varepsilon u$$

$$\geq \overline{\lim}\ \{E(1_A 1_{\{\gamma_n \leq K\} \cup \{\gamma_n = \infty\}} X_{\gamma_n})$$

$$+ E(1_{A \cap \{K < \gamma_n < \infty\}} \sup_{K < i < \infty} X_i)\} - 3\varepsilon u$$

$$\geq \overline{\lim}\ E(1_A X_{\gamma_n}) - 3\varepsilon u.$$

2.12. THEOREM. Let $(X_n : n \in \mathbb{N})$ be a stochastic process taking values in $\mathbb{R}^p$, and such that $E(\sup_n \sum_{i=1}^p |(X_n)_i|) < \infty$, and $X_\infty \geq \overline{\lim}\ X_n$. Let ϕ be an increasing continuous function defined on the closed convex hull of $\cup \{X_n(\Omega) : n \in \overline{\mathbb{N}}\}$. Then there exists an optimal randomized stopping time $\gamma \in \Gamma$ such that

$$\phi(EX_\gamma) = V_\phi = \sup\{\phi(EX_\eta) : \eta \in \Gamma\}.$$

If in addition ϕ is assumed to be convex, then the optimal time can be chosen non-random, i.e., there exists $\sigma \in \Sigma$ such that

$$\phi(EX_\sigma) = V_\phi = \sup\{\phi(EX_\tau) : \tau \in \Sigma\} = \sup\{\phi(EX_\gamma) : \gamma \in \Gamma\}.$$

Proof. Let (γ_n) be a sequence in Γ such that $V_\phi = \lim \phi(EX_{\gamma_n}) = \sup\{\phi(EX_\eta) : \eta \in \Gamma\}$. Let (γ_{n_k}) be a subsequence of (γ_n) that converges to an element $\gamma \in \Gamma$ in the Baxter-Chacon topology. Then by Proposition 2.11, $E(X_\gamma) \geq \overline{\lim}\ E(X_{\gamma_n})$. The monotonicity and continuity of ϕ implies that

$$\phi(EX_\gamma) \geq \phi(\overline{\lim}\ EX_{\gamma_n}) \geq \overline{\lim}\ \phi(EX_{\gamma_n}) = V_\phi.$$

If ϕ is convex, then Theorem 1.7 shows that there also exists an optimal <u>non-random</u> stopping time $\sigma \in \Sigma$.

Finally we show that there exists a <u>finite</u> optimal stopping time for the process $X_n = \|\frac{1}{n} \sum_{1\leq i\leq n} Y_i\|$, where Y_i is a Banach-valued independent identically distributed sequence and $\|Y_1\| \in L_p$, $p > 1$, thus generalizing the result of B. Davis [8] to Banach-valued processes. For this the condition $E(\|Y_1\| \log^+ \|Y_1\|) < \infty$ is not sufficient even in the real-valued case, as shown by M. Klass [14].

2.13. THEOREM. Let B be a Banach space, and let $\phi: B \to \mathbb{R}$ be a continuous function such that $\phi(x + y) \leq \phi(x) + \phi(y)$, and $\phi(\alpha x) = \alpha\phi(x)$ for every $\alpha \geq 0$, and for every $x,y \in B$. Let $(Y_n: n \in \mathbb{N})$ be an independent identically distributed B-valued sequence of random variables with mean 0, and with $E\|Y_1\|^p < \infty$ for some $p > 1$. For every $n \geq 1$ let $X_n = \frac{1}{n} \sum_{i\leq n} Y_i$, and let $X_\infty = 0$. Then every stopping time $\alpha \in \Sigma$ such that $\phi[EX_\sigma] = V_\phi = \sup\{\phi(EX_\tau): \tau \in \Sigma\}$ satisfies $P(\sigma < \infty) = 1$. Hence there exists an a.s. finite stopping time $\sigma \in \Sigma$ such that $\phi(EX_\sigma) = V_\phi$.

<u>Proof</u>. The assumptions made on ϕ clearly imply the convexity of ϕ. Hence by Theorem 2.6 there exists $\sigma \in \Sigma$ such that $\phi(EX_\sigma) = V_\phi$. It suffices to show that any such σ is finite a.s. Let $\sigma \in \Sigma$ satisfy $\phi(EX_\sigma) = V_\phi$, and let $x = EX_\sigma$. By the Hahn-Banach theorem, there exists a linear functional x^* on B such that $x^*(x) = \phi(x)$, and $-\phi(-y) \leq x^*(y) \leq \phi(y)$ for every $y \in B$. The continuity of ϕ clearly implies t-e continuity of x^*. For every $\tau \in \Sigma$, we have

$$E[x^*(X_\tau)] = x^*[EX_\tau] \leq \phi[EX_\tau] \leq V_\phi.$$

Since $E[x^*(X_\sigma)] = x^*(x) = \phi(x) = V_\phi$, the stopping time σ is optimal for the real-valued process $(x^*(X_n))$. Set $Z_n = x^*(Y_n)$; since x^* is continuous, and $E(\|Y_1\|^p) < \infty$, we have $E|Z_n|^p < \infty$. Now $x^*(X_n)$ are Cesaro averages of the real-valued process Z_n. Since B. Davis [8] has proved that an optimal stopping time σ for such a process $(x^*(X_n))$ is finite a.s., the proof is concluded.

3. Stopping rules under a weak independence assumption. In this section we introduce a condition (I) weaker than independence, and show that under (I) a stationary Banach-valued process (Y_n) is of class L Log L if and only if for every stopping time σ, $E(1_{\{\sigma<\infty\}}\frac{\|Y_\sigma\|}{\sigma}) < \infty$. A similar characterization is obtained in terms of the Cesaro averages.

Let $(A_n)_{n\geq 1}$ be a family of σ-algebras; typically A_n is generated by a single random variable Y_n. Let $F_n = \sigma(\bigcup_{i\leq n} A_i)$. We say that (A_n) [or (F_n)] satisfies the condition (I) if there exists a strictly increasing sequence of integers $(N_k)_{k\geq 1}$, and a constant α with $0 \leq \alpha < 1$, such that

(3.1) $$\overline{\lim}\ \frac{N_k}{k} < \infty,$$

(3.2) $$\forall k \geq 1, \quad \forall A \in F_{N_k}, \quad \forall C \in A_{N_{k+1}},$$
$$|P(A \cap C) - P(A)P(C)| \leq \alpha P(A)P(C).$$

The condition (I) is a weakening of the (*)-mixing condition introduced by Blum-Hanson-Koopmans [4]; see also [23], p. 140. We refer to [4] for examples. In the case when the A_n are atomic and $N_k = k$, the condition (I) coincides with the Vitali-Chow condition (see e.g. Neveu [21], p. 73).

If (A_n) satisfied the condition (I), then for every Banach space $(B, \|\ \|)$ we have

(3.3) $$\begin{cases} \forall k \geq 2, \quad \forall X \in L_1^B(F_{N_{k-1}}), \quad \forall D \in A_{N_k} \\ \quad \| E(1_D X) - P(D)E(X)\| \leq \alpha P(D)E\|X\| \end{cases}$$

and

(3.4) $$\begin{cases} \forall k \geq 2, \quad \forall A \in F_{N_{k-1}}, \quad \forall X \in L_1^B(A_{N_k}), \\ \quad \|E(1_A X) - P(A)E(X)\| \leq \alpha P(A)E\|X\| . \end{cases}$$

We only show the assertion 3.3. Let X be an F_{N_k}-measurable step function, say $X = \sum_{i \leq n} x_i 1_{A_i}$, with $x_i \in B$, $A_i \in F_{N_{k-1}}$, $i = 1,\ldots,n$, and let $D \in A_{N_k}$. Then

$$\|E(1_D X) - P(D)E(X)\| \leq \sum_{i \leq n} \|x_i\| \; |P(A_i \cap D) - P(A_i)P(D)|$$

$$\leq \alpha \sum_{i \leq n} \|x_i\| P(A_i)P(D) = \alpha P(D)E\,\|X\| .$$

Fix $\varepsilon > 0$, let $X \in L_1^B(F_{N_{k-1}})$ and let Z be an $F_{N_{k-1}}$-measurable step function with $E\,\|X - Z\| < \varepsilon$. Then

$$\|E(1_B X) - P(D)EX\| \leq 2\varepsilon + \|E(1_D Z) - P(D)EZ\|$$

$$\leq 2\varepsilon + \alpha P(D)E\,\|Z\| \leq 4\varepsilon + \alpha P(D)E\,\|X\| .$$

A process $(Y_n)_{n>1}$ is said to satisfy condition (I) if the family $A_n = \sigma(Y_n)$ does.

We prove next our main result on condition (I). If the Y_n are real-valued, independent, and identically distributed, the implications (ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (i) are due to Burkholder [5]; the implications (iv) $\Rightarrow$ (i) and (v) $\Rightarrow$ (i) to B. Davis [7] and independently to McCabe-Shepp [18].

3.5. THEOREM. Let B be a Banach space, and let $(Y_n)_{n>1}$ be a stationary B-valued Bochner integrable process. Suppose that (Y_n) satisfies the condition (I) for an increasing sequence of integers (N_k). Then the following conditions are equivalent:

(i) $E(\|Y_1\| \operatorname{Log}^+ \|Y_1\|) < \infty$;

(ii) $E(\sup \|\frac{Y_n}{n}\|) < \infty$;

(iii) $E \sup(\frac{1}{n} \sum_{i \leq n} \|Y_i\|) < \infty$;

(iv) $\forall \sigma \in \Sigma$, $E(1_{\{\sigma<\infty\}} \frac{\|Y_\tau\|}{\tau}) < \infty$;

(v) $\forall \sigma \in \Sigma$, $E(1_{\{\sigma<\infty\}} \|\frac{1}{\sigma} \sum_{N_k \leq \sigma} Y_{N_k}\|) < \infty$.

Proof. By the pointwise ergodic theorem, the sequence $\frac{1}{n}\sum_{i<n} \|Y_i\|$ converges a.s., and hence $\|\frac{Y_n}{n}\|$ converges a.s. to zero. By Wiener's dominated ergodic theorem applied to the process $\|Y_n\|$, (i) $\Rightarrow$ (iii)

(see e.g. [10], p. 678). Since

$$\left\|\frac{Y_n}{n}\right\| = \left\|\frac{1}{n}\sum_{i\leq n} Y_i - \frac{1}{n}\sum_{i\leq n-1} Y_i\right\|$$

$$\leq \left\|\frac{1}{n}\sum_{i\leq n} Y_i\right\| + \left\|\frac{1}{n-1}\sum_{i\leq n-1} Y_i\right\|,$$

(iii) implies (ii). Obviously (ii) $\Rightarrow$ (iv), and (iii) $\Rightarrow$ (v). We only prove (iv) $\Rightarrow$ (i), and (v) $\Rightarrow$ (i). Given α with $0 \leq \alpha < 1$, let (N_k) be the strictly increasing sequence such that $A_n = \sigma(Y_n)$ satisfied the conditions (3.1) and (3.2).

<u>Proof of (iv) $\Rightarrow$ (i)</u>. Fix $\varepsilon > 0$; applying the pointwise ergodic theorem, choose $N_i \geq 1$ and a set $A \in F_\infty$ such that for every $n \geq N_i$, we have $\|Y_n\| \leq n/2$ on A. By assumption (3.1), we also may assume that there exists a constant c such that $N_k \leq ck$ for every k. Define a stopping time $\sigma \in \Sigma$ by

$$\sigma = \inf\{N_k : N_k \geq N_i,\ \|Y_{N_k}\| \geq N_k\},$$

with the convention $\inf \emptyset = +\infty$. Clearly $\sigma = +\infty$ on A. Given a process (X_n), and a stopping time $\tau \in \Sigma$, we write EX_τ for $E(1_{\{\tau<\infty\}}X_\tau)$. By assumption (iv),

$$\infty > E\left(\frac{\|Y_\sigma\|}{\sigma}\right) = \sum_{k\geq i} \frac{1}{N_k} \int_{\{\sigma=N_k\}} \|Y_{N_k}\|\, dP.$$

The set $\{\sigma \geq N_k\} = \bigcap_{i\leq j\leq k-1} \{\|Y_{N_j}\| < N_j\}$ is $F_{N_{k-1}}$ measurable.

Applying the relation (3.4) with the set $A = \{\sigma \geq N_k\}$ and $X = 1_{\{\|Y_{N_k}\| \geq N_k\}} \|Y_{N_k}\|$, we obtain

$$\infty > (1-\alpha) \sum_{k\geq i} \frac{1}{N_k} P(\sigma \geq N_k) E(1_{\{\|Y_{N_k}\|\geq N_k\}} \|Y_{N_k}\|)$$

$$\geq (1-\alpha) \sum_{k\geq i} \frac{1}{N_k} P(A) E(1_{\{\|Y_1\|\geq N_k\}} \|Y_1\|).$$

Since $\alpha < 1$ and $P(A) > 0$, we have for fixed i

$$\infty > \sum_{k\geq i} \frac{1}{N_k} \int_{\{\|Y_1\|\geq N_k\}} \|Y_1\| \, dP = \int_0^\infty x\Big(\sum_{N_i\leq N_k\leq x} \frac{1}{N_k}\Big) P_{\|Y_1\|}(dx),$$

where $P_{\|Y_1\|}$ denotes the distribution of $\|Y_1\|$. We have that $ck \leq x$ implies $N_k \leq x$. Hence if $x \geq N_i$, we have

$$\sum_{N_i\leq N_k\leq x} \frac{1}{N_k} \geq \sum_{i\leq k\leq x/c} \frac{1}{ck} \geq \frac{1}{c} \int_i^{[\frac{x}{c}]+1} \frac{dx}{x}$$

$$\geq \frac{1}{c}[\mathrm{Log}\ x - \mathrm{Log}\ ci].$$

This clearly implies that $E[\ \|Y_1\| \,\mathrm{Log}^+ \|Y_1\|] < \infty$.

Proof of (v) $\Rightarrow$ (i). Let σ be the stopping time defined in the proof above. Then

$$E(\frac{1}{\sigma}\|\sum_{N_k\leq\sigma} Y_{N_k}\|) \geq E(\frac{\|Y_\sigma\|}{\sigma}) - E(\frac{1}{\sigma}\|\sum_{N_{k+1}\leq\sigma} Y_{N_k}\|)$$

$$\geq E(\frac{\|Y_\sigma\|}{\sigma}) - E(\frac{1}{\sigma}\sum_{N_{k+1}\leq\sigma} \|Y_{N_k}\|).$$

By the proof of (iv) $\Rightarrow$ (i), it suffices to show that $E(\frac{\|Y_\sigma\|}{\sigma}) < \infty$, and hence that $E(\frac{1}{\sigma}\sum_{N_{j+1}\leq\sigma} \|Y_{N_j}\|) < \infty$. We have

$$E(\frac{1}{\sigma}\sum_{N_{j+1}\leq\sigma} \|Y_{N_j}\|) = \sum_{k\geq i} \frac{1}{N_k} E(1_{\{\sigma=N_k\}} \sum_{j\leq k-1} \|Y_{N_j}\|)$$

$$\leq \sum_{k\geq i} \frac{1}{N_k} \sum_{1\leq j\leq k-1} E(\ \|Y_{N_j}\|\, 1_{\{\|Y_{N_k}\|\geq N_k\}}).$$

Applying the relation (3.3) with $X = \|Y_{N_j}\|$, which is measurable with respect to $F_{N_{k-1}}$, and with $D = \{\|Y_{N_k}\|\geq N_k\} \in A_{N_k}$, we have

$$E(\frac{1}{\sigma}\sum_{N_{k-1}\leq\sigma} \|Y_{N_k}\|) \leq (1+\alpha) \sum_{k\geq i} \frac{1}{N_k} \sum_{1\leq j\leq k-1} E\|Y_{N_j}\|\, P(\|Y_{N_k}\|\geq N_k)$$

$$\leq (1+\alpha) E\|Y_1\| \sum_{k\geq i} P(\|Y_1\|\geq N_k)$$

$$\leq (1+\alpha)(E\,||Y_1||\,)^2 < \infty.$$

This completes the proof of the theorem.

The theorem of Davis [7] and McCabe-Shepp [18], and the theorem of Dvoretzky (see e.g.[6], p. 86) were extended to tactics on directed sets, in particular to stopping times in $\mathbb{N} \times \mathbb{N}$, by Krengel-Sucheston [15], by application of their linear embedding theorem. Since, as noted in [16], the linear embedding preserves also vector-valued integrals, the results of the present paper concerning averages of vector-valued independent identically distributed random variables extend similarly to directed sets.

4. Continuous parameter. In this section we extend some results of Sections 1 and 2 to continuous-parameter processes. Let (Ω, F, P) be a probability space, and let $(F_t: t \in [0,\infty])$ be an increasing right-continuous family of sub-σ-algebras of F such that F_0 contains the null sets. A stopping time of (F_t) is a map $\sigma: \Omega \to [0,\infty]$ such that $\{\omega: \sigma(\omega) \leq t\} \in F_t$ for every $t \in [0,\infty]$; we again write Σ for the set of all stopping times. A randomized stopping time for (F_t) is a map $\gamma: \Omega \times [0,1] \to [0,\infty]$ which is a stopping time for $(F_t \times B)$; we assume that γ is non-decreasing and left-continuous in the second variable. We will write Γ for the set of randomized stopping times. For every $\gamma \in \Gamma$, the ω-distribution of γ is defined by

$$M(\omega,[0,t]) = \sup\{v: \gamma(\omega,v) \leq t\}$$

for $\omega \in \Omega$, $t \in [0,\infty]$. Then M has the following properties:

(a) For fixed $\omega \in \Omega$, the function $M(\omega,\cdot)$ defines a probability measure on $\overline{\mathbb{R}}$.

(b) For fixed $t \in [0,\infty]$, the function $M(\cdot,[0,t])$ is F_t-measurable.

We will write Γ' for the set of all functions M satisfying (a) and (b). There is a one-to-one correspondence between Γ and Γ' (see [1] for details). As in the discrete case, $M \in \Gamma'$ corresponds to a nonrandomized stopping time σ if and only if $M(\omega,[0,t]) = 1_{\{\sigma<t\}}(\omega)$. The Baxter-Chacon topology on Γ is the coarsest topology such that for every $Y \in L^1(F)$, for every $f \in C([0,\infty])$, the map $\gamma \to \int Y(\omega) f[\gamma(\omega,v)] P(d\omega)\lambda(dv)$ is continuous. Baxter and

Chacon have shown that the set Γ is compact for this topology ([1] Theorem 1.5).

The following theorem is analogous to a result used in Section 1.

4.1. THEOREM. The extreme points of Γ' are exactly the ω-distributions of the nonrandomized stopping times.

In the sequel we will consider an adapted Banach-valued process $(X_t, F_t, t \in [0,\infty])$ such that

(4.2) $$E(\sup \|X_y\|) < \infty,$$

(4.3) $$X_t \text{ is right-continuous},$$

(4.4) $$\forall \sigma_n \in \Sigma,\quad \sigma_n \nearrow \sigma \text{ implies } X_{\sigma_n} \to X_\sigma \text{ a.s.},$$

For real-valued processes (X_t), condition (4.4) is equivalent with quasi-left-continuity, and the property $X_\infty = \lim_{t\to\infty} X_t$ a.s.

The following theorem is an analog of (1.4), above.

4.5. THEOREM. Let $\gamma \in \Gamma$ be a randomized stopping time. If (X_t) is an adapted process satisfying the conditions (4.2) and (4.3), then

$$E(X_\gamma) = \int_0^1 E\, X_{\gamma(\cdot,v)}\, dv.$$

Proof. As in (1.2), we have

$$M_\gamma = \int_0^1 M_{\gamma(\cdot,v)}\, dv.$$

Then for every $A \in F$, for every $t \in [0,\infty]$.

$$\int 1_A M_\gamma(\omega,[0,t])P(d\omega) = \int_0^1 \int 1_A(\omega) M_{\gamma(\cdot,v)}(\omega,[0,t])P(d\omega)\,dv$$

$$= \int 1_A(\omega)\lambda\{v: \gamma(\omega,v) \le t\}P(d\omega).$$

Hence $M_\gamma(\omega,[0,t]) = \lambda\{v: \gamma(\omega,v) \le t\}$ a.s. For every $j \ge 1$, let $\gamma[j]$ be the smallest element $k/2^j$ such that $\gamma \le k/2^j$. Then for every $j \ge 1$,

$$E(X_{\gamma[j]}) = \int X_{2^{-j}}(\omega)M(\omega, [0,2^{-j}])P(d\omega)$$

$$+ \int \sum_{k\geq 2} X_{k2^{-j}}(\omega)M(\omega,](k-1)2^{-j}, k2^{-j}])P(d\omega)$$

$$= \int X_{2^{-j}}(\omega)\lambda\{v: \gamma(\omega,v) \leq 2^{-j}\}P(d\omega)$$

$$+ \int \sum_{k>2} X_{k2^{-j}}(\omega)\lambda\{v: (k-1)2^{-j} < \gamma(\omega,v) \leq k2^{-j}\}P(d\omega)$$

$$= \int_0^1 E(X_{\gamma(\cdot,v)[j]})dv.$$

Letting $j \to \infty$, and using the properties (4.2) and (4.3), we obtain

$$E(X_\gamma) = \lim_j EX_{\gamma[j]} = \lim \int_0^1 E(X_{\gamma(\cdot,v)[j]})dv$$

$$= \int_0^1 E(X_{\gamma(\cdot,v)})dv.$$

The following theorem allows us to "derandomize" continuous-parameter optimal stopping times. The proof, similar to the proof of Theorem 1.7, is omitted.

4.6. THEOREM. Let B be a Banach space, let $\phi: B \to \mathbb{R}$ be a continuous and convex function. Let (X_t) be a B-valued adapted process with the properties (4.2) and (4.3). Then

$$V_\phi = \sup_{\gamma \in \Gamma} \phi[E(X_\gamma)] = \sup_{\sigma \in \Sigma} \phi[E(X_\sigma)],$$

and if one of these suprema is achieved and finite, so is the other one. If the supremum is achieved by $\gamma_0 \in \Gamma$ which is finite a.s., then it is also achieved by $\sigma_0 \in \Sigma$ which is finite a.s.

We now study the convergence of the stopped sequence X_{γ_n}, when $\gamma_n \to \gamma$ (BC) and (X_t) is an adapted process with the Properties (4.2) - (4.4).

The following theorem is a generalization of Corollary 1.15 in [1], and the proof is similar.

4.7. THEOREM. Let B be a Banach space, let $(X_t: t \in [0,\infty])$ be a B-valued right-continuous, quasi-left continuous process such that $E(\sup \|X_t\|) < \infty$. Let γ_n be a sequence of randomized stopping times such that γ_n converges to a randomized stopping time γ in the Baxter-Chacon topology. Suppose that $\lim_{a\to\infty} \sup_n P \times \lambda(\gamma_n > a) = 0$. Then for every $A \in F$, $E(1_A X_{\gamma_n})$ converges to $E(1_A X_\gamma)$.

Proof. The proof of this theorem depends on several lemmas stated below.

4.8. LEMMA. Let (γ_n) be a sequence of randomized stopping times taking values in a finite set $K \subset [0,\infty]$, and converging (BC) to a randomized stopping time γ. Suppose that the random variables $(X_t,\ t \in K)$ are Bochner integrable. Then

$$E(1_A X_{\gamma_n}) \to E(1_A X_\gamma), \quad A \in F.$$

Proof. The lemma is first proved for step functions $(X_t\colon t \in K)$. The proof is similar to the one in [1], Lemma 4.1.

For every $\gamma \in \Gamma$, and every $j \geq 1$, let $\gamma[j]$ be the j-th dyadic approximation of γ from above, i.e., $\gamma[j](\omega,v) = \inf\{k/2^j;\ k/2^j \geq \gamma(\omega,v)\}$ (with the convention $\inf \emptyset = +\infty$).

The following lemma is similar to Theorem 1.8 [1], and the proof is omitted.

4.9. LEMMA. Let $\gamma_n \in \Gamma$ converge (BC) to $\gamma \in \Gamma$, and satisfy $\lim_{a\to\infty} \sup_n P \times \lambda\{\gamma_n \geq a\} = 0$. Let B be a Banach space, and let (X_t) be a B-valued right-continuous process such that $E(\sup \|X_t\|) < \infty$. Suppose that for every $A \in F$

$$\lim_j E(1_A X_{\gamma_n[j]}) = E(1_A X_{\gamma_n}) \quad \text{uniformly in } n.$$

Then for every $A \in F$, $E(1_A X_\gamma) = \lim E(1_A X_{\gamma_n})$.

The following lemma is similar to Lemma 5.1 [1]; again the proof is omitted.

4.10. LEMMA. Let (X_t) be a Banach valued process such that $E(\sup \|X_t\|) < \infty$. Suppose that for every $\varepsilon > 0$, and for every sequence (σ_n) in Σ such that $\lim_{a\to\infty} \sup_n P(\sigma_n > a) = 0$, one has $\lim_j \sup_n P\{ \|X_{\sigma_n[j]} - X_{\sigma_n}\| > \varepsilon\} = 0$. Then for every $A \in F$, and for every sequence $\gamma_n \in \Gamma$ such that $\gamma_n \to \gamma$ (BC), and $\lim_{a\to\infty} \sup_n P \times \lambda(\gamma_n > a) = 0$, one has

$$\lim_j E(1_A X_{\gamma_n[j]}) = E(1_A X_{\gamma_n}) \text{ uniformly in } n.$$

Sketch of Proof of Theorem 4.7. Because of Lemmas 4.8-4.10 and the condition $E(\sup \|X_t\|) < \infty$, the proof of the theorem reduces to the proof of the following: for every $\varepsilon > 0$, and for every sequence (σ_n) in Σ such that $\lim \sup P(\sigma_n > a) = 0$, one has $\lim_j \sup_n P(\|X_{\sigma_n[j]} - X_{\sigma_n}\| > \varepsilon) = 0$. This is proved via the arguments given in Lemma 5.9 and 5.11 in [1], setting $f(x,y) = \|x-y\| / (1+ \|x-y\|)$, and $Y(s,t) = \sup\{f(X_s,X_r): s \leq r \leq t\}$.

4.11. COROLLARY. Let (γ_n) be a sequence of randomized stopping times such that γ_n converges to a randomized stopping time γ in the Baxter-Chacon topology. Let (X_t) be a Banach-valued process with the properties (4.2)-(4.4). Then for every $A \in F$, $E(1_A X_\gamma) = \lim E(1_A X_{\gamma_n})$, and X_{γ_n} converges in distribution to X_γ on A.

Proof. Let $T: [0,\infty] \to [0,1]$ be a continuous bijective and increasing function. For every $s \in [0,1]$, set $Y_s = X_{T^{-1}(s)}$ and $G_s = F_{T^{-1}(s)}$. For every $s \geq 1$, set $Y_s = X_\infty$ and $G_s = F_\infty$. The proof of the theorem reduces to the proof of the convergence of $E(1_A Y_{T\circ\gamma_n})$ to $E(1_A Y_{T\circ\gamma})$. The process (Y_s) clearly satisfied the assumptions (4.2)-(4.4), and the sequence $T\circ\gamma_n$ is bounded by 1, so that Theorem 4.7 applies.

We now deduce the following theorem about the existence of an optimal stopping time for $\phi(EX_\sigma)$. It is a continuous parameter analog of Theorem 2.4.

4.12. THEOREM. Let $(B, \|\ \|)$ be a Banach space, and let $\phi: B \to \mathbb{R}$ be a convex continuous function. Let $(X_t, t \in [0,\infty])$ be an adapted process with the properties (4.2)-(4.4). Then there exists a non-randomized stopping time $\sigma \in \Sigma$ such that

$$\phi(EX_\sigma) = V_\phi = \sup\{\phi(EX_\tau): \tau \in \Sigma\}$$

$$= \sup\{\phi(EX_\gamma): \gamma \in \Gamma\} < \infty.$$

Proof. The theorem is a consequence of Theorem 4.6, and Corollary 4.11. We refer to the proof of Theorem 2.4 for details.

REFERENCES

[1] J. R. Baxter and R. V. Chacon, Compactness of stopping times. Z. Wahrscheinlichkeitstheorie verw. Gebiete 40(1977), 169-181.

[2] J. R. Baxter and R. V. Chacon, Englargement of σ-algebras and compactness of time changes, Canadian J. Math. 29(1977), 1055-1065.

[3] A. Beck, On the strong law of large numbers, Ergodic Theory (Proc. International Sympos., Tulane Univ., New Orleans, La., 1961), Academic Press, New York (1961), 21-53.

[4] J. R. Blum, D. L. Hanson, and L. H. Koopmans, On the strong law of large numbers for a class of stochastic processes, Z. Wahrscheinlichkeitstheorie verw. Geb. 2(1963), 1-11.

[5] D. L. Burkholder, Successive conditional expectations of an integrable function, Ann. Math. Stat. 33(1962), 887-893.

[6] Y. S. Chow, H. Robbins, and D. Siegmund, Great expectations: The theory of optimal stopping, Boston: Houghton Mifflin Co. 1971.

[7] B. Davis, Stopping rules for S_n/n and the class L log L, Z. Wahrscheinlichkeitstheorie verw. Geb. 17(1971), 147-150.

[8] B. Davis, Moments of random walk having infinite variance and the existence of certain optimal stopping rules for S_n/n, Illinois J. Math. 17(1973), 75-81.

[9] C. Dellacherie, Convergence en probabilité et topologie de Baxter-Chacon, Université de Strasbourg, Séminaire de Probabilites, année 1976/7, Springer Verlag, Lecture Notes in Math., Vol 649, 1978.

[10] N. Dunford and J. T. Schwartz, Linear Operators, Part I, Interscience, New York 1957.

[11] W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, Wiley, New York 1971.

[12] R. F. Gundy, On the class L log L martingales and singular integrals, Studia Math. 33(1969), 109-118.

[13] J. L. Kelley, I. Namioka, et al, Linear Topological Spaces, Van Nostrand, Princeton 1963.

[14] M. J. Klass, Properties of optimal extended-valued stopping rules for S_n/n, Ann. Probability 5(1973), 719-757.

[15] U. Krengel and L. Sucheston, Stopping rules and tactics for processes indexed by directed sets, J. Multivariate Analysis 11(1981), 199-229.

[16] L. LeCam, An extension of Wald's theory of statistical decision functions, Ann. Math. Stat. 26(1955), 69-81.

[17] L. H. Loomis, Dilations and extremal measures, Advances in Math. 17(1975), 1-13.

[18] B. J. McCabe and L. A. Shepp, On supremum of S_n/n, Ann. Math. Stat. 41(1970), 2166-2168.

[19] P. A. Meyer, Convergence faible et compacité des temps d'arrêt d'apres Baxter et Chacon, Université de Strasbourg, Séminaire de Probabilités, anneé 1976/7, Springer Verlag, Lecture Notes in Math., Vol 649, 1978.

[20] E. Mourier, Les éléments aléatoires dans un espace de Banach, Ann. Ins. Henri Poincaré 13(1953), 161.

[21] J. Neveu, Discrete Parameter Martingales, North Holland, Amsterdam 1975.

[22] R. R. Phelps, Lectures on Choquet's theorem, Van Nostrand, New York 1966.

[23] P. Revesz, The laws of large numbers, Academic Press, New York and London 1968.

[24] D. O. Siegmund, Some problems in the theory of optimal stopping, Ann. Math. Stat. 38(1967), 1627-1640.

[25] J. L. Snell, Application of martingale system theorems, Trans. Amer. Math. Soc. 13(1952), 293-312.

Department of Mathematics
Ohio State University
Columbus, Ohio 43210

Laboratoire du Calcul des Probabilités
Université Paris-VI
Tour 56, 4, Place Jussieu
75230 Paris Cedex 05