PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 13, Fasc. 1 (1992), pp. 97–121

A PROBABILISTIC APPROACH TO THE REDUITE IN OPTIMAL STOPPING

BY

N. EL KAROUI (PARIS), J.-P. LEPELTIER (LE MANS) AND A. MILLET (PARIS)

Abstract. In this work the approach to the reduite is made in a simple and unified way. More precisely, we use the same probabilistic technique to study the optimal stopping problem associated with the reduite, to prove the expression of the Snell envelope in terms of the reduite under very general assumptions, and to show continuity properties of the reduite. We finally describe the example of diffusion processes with jumps.

0. INTRODUCTION

During the sixties and seventies, the problem of optimal stopping of continuous time processes was the object of many papers. Without quoting all of them let us stress the importance of Mertens' works [21] and [22] which deal with the problem in its most general setting; in these articles, the Snell envelope is systematically studied and is characterized as the smallest supermartingale which is greater than or equal to the process. In this general setting, the works of Bismut and Skalli [8], Mainguenеau [19], and El Karoui [12] describe precisely an optimal stopping time (whenever it exists) as the "beginning" of the set where the process is equal to its Snell envelope and, more generally, give the exact default of optimality.

In the case of a Markov process (X_t), a new question arises: if the process one wants to stop depends only on the state of the process at time t, i.e., it can be written as $g(X_t)1_{\{t<\zeta\}}$, where ζ denotes the life-time of the process (X_t), does its Snell envelope have the same property? The function associated with this stopping problem is called the *reduite* of g, and denoted by Rg. The "general theory" of processes yields that the functions g and Rg are related by the following equation:

$$(0.1)\qquad Rg(x) = \sup\{E_x[g(X_S)1_{\{S<\zeta\}}];\ S \text{ stopping time}\},$$

and that Rg is the smallest strongly supermedian function, in the sense of Mertens, which dominates g. In order to prove that Rg exists and show its main properties, two different points of view have been used in the literature:

In potential theory, it is assumed that g is the difference of two excessive functions. Then Rg is set to be the smallest excessive function greater than or equal to g, and checking that Rg satisfies equation (0.1) is not too difficult. This point of view justifies the name of "reduite" given to the function $Rg(\cdot)$.

In [22], Mertens starts from the opposite point of view: He studies the right-hand side of (0.1) when g is the indicator function of some set A and he shows easily that

$$R(1_A)(x) = P_x(D_A < \zeta),$$

where $D_A = \inf\{t \geqslant 0;\ X_t \in A\}$ is the beginning of the set A. Then he extends his result to step functions and to arbitrary measurable functions g. However, the arguments required to show that (0.1) holds are more and more difficult.

In [23], Meyer solves the problem by using the dual point of view (in the sense of convex analysis) of that of potential theory: namely he introduces the family $M(x)$ of measures μ which satisfy

$$M(x) = \{\mu;\ \mu(f) \leqslant f(x) \text{ for excessive functions } f\},$$

and he uses the function

$$\hat{R}g(x) = \sup\{\mu(g);\ \mu \in M(x)\}.$$

Clearly, for any stopping time T, the measure μ^T defined by

$$\mu^T(g) = E_x[g(X_T);\ T < \zeta]$$

belongs to $M(x)$, and hence $Rg(x) \leqslant \hat{R}g(x)$.

The converse inequality is much more difficult to prove; it depends on a theorem of Rost [27] which gives a complete characterization of the set $M(x)$, at least in the transient case.

The interest of this formulation is to introduce a functional analysis setting which is well suited for solving the optimal stopping problem: under mild hypotheses, the set $M(x)$ as well as its graph are shown to be weakly compact.

Our point of view is quite similar to this last one: we introduce the set

$$\mathscr{A}(x) = \{\mu;\ \mu = \mu^T \text{ for some stopping time } T\}$$

which will be endowed with the topology induced by that of Baxter and Chacon [3] on processes; this topology is stronger than the topology of convergence in law. More precisely, let $\mathscr{R}(x)$ denote the set of stopping rules starting from x, which are measures on the set of processes $(Y_t(\omega);\ (\omega, t) \in \Omega \times \bar{R}_+)$, defined by

$$R \in \mathscr{R}(x) \Leftrightarrow \text{there exists a stopping time } T \text{ such that } R(Y) = E_x(Y_T),$$

and let $\mathscr{R}(x)$ be endowed with the Baxter–Chacon topology defined as follows: R^n converges to R if and only if $R^n(Y)$ converges to $R(Y)$ for every continuous process Y. The set $\mathscr{A}(x)$ is convex compact in the induced topology.

It is easy to see that the graphs of the multivalued mapping $x \rightsquigarrow \mathscr{R}(x)$ and $x \rightsquigarrow \mathscr{A}(x)$ are Borel subsets of products of compact spaces. Furthermore, if the map $x \mapsto P_x$ is weakly continuous, then these graphs are closed.

If g is measurable, then the function $\sup\{\mu(g);\ \mu \in \mathscr{A}(x)\}$ is also measurable, and it is identified with the reduite Rg of g (in the sense of Snell's envelope).

The same arguments show the continuity of Rg if g and X are continuous and if the map $x \mapsto P_x$ is weakly continuous.

Hence the novelty of the present method lies in the fact that the restriction of a topology, defined on a set of probabilities acting on a set of processes, to the set of stopping measures yields a "good" convex compact topology.

The paper is organized as follows. In the first section we define the optimal stopping problem and we show that the reduite is independent of the realization. In the second section we define and characterize the set of stopping rules; we also prove (0.1). Section 3 establishes the expression of the Snell envelope in terms of the reduite Rg. In the fourth section we study the continuity properties of the reduite, and the example of diffusion processes with jumps on $\mathbf{R}^d$ is described in the fifth section. Finally, in the Appendix, we prove the weak continuity of the map $x \mapsto P_x$ for Feller processes on a compact state space.

1. FIRST PROPERTIES OF THE OPTIMAL STOPPING PROBLEM

1.1. The Markov process. We consider a Markov process (X_t) taking on values in a metrizable Lusin space E; in some cases E is supposed to be LCCB. We denote its semigroup by (P_t) and its resolvent by (U^α). We suppose that the semigroup is conservative, i.e., $P_t 1 = 1$ for every t. Notice that this assumption is not restrictive. Indeed, if the semigroup (P_t) is sub-Markovian, we extend it to a Markovian semigroup over $E^\Delta = E \cup \{\Delta\}$, where Δ is a coffin-state (see, e.g., [9] or [28]). Let $\Pi(E)$ denote the set of probabilities on E.

In the paper we will make various assumptions on the semigroup (P_t); however, they all imply the existence of a strong Markov realization of (P_t) (see, e.g., [9], [28]). There is no uniqueness of strong Markov realizations of a semigroup (P_t). Thus on a given space there exists a realization which is the smallest one.

DEFINITION 1.1. Let $\mathscr{X} = (\Omega, \mathscr{G}_t, X_t, \theta_t, P_x;\ x \in E)$ be a strong Markov realization of (P_t). The *canonical realization associated with* $\mathscr{X}$ is defined on Ω with the filtration $(\mathscr{F}_t)$ deduced from $\mathscr{F}_t^0 = \sigma(X_s;\ s \leqslant t)$ by standard regularization procedures (completeness and right-continuity), say $\mathscr{C}(\mathscr{X}) = (\Omega, \mathscr{F}_t, X_t, \theta_t, P_x;\ x \in E)$.

1.2. Definition of the optimal stopping problem. Let $\mathscr{X} = (\Omega, \mathscr{G}_t, X_t, \theta_t, P_x;\ x \in E)$ be a strong Markov realization of the semigroup (P_t). Extend the process (X_t) to $\bar{\mathbf{R}}_+ = \mathbf{R}_+ \cup \{+\infty\}$ by setting $X_\infty = \Delta$, and let $(\mathscr{F}_t)$ denote the canonical filtration of $\mathscr{C}(\mathscr{X})$.

DEFINITION 1.2. Let $\mathscr{T}(\mathscr{G})$ denote the *set of stopping times* (finite or not) with respect to the filtration $(\mathscr{G}_t)$ (i.e., $T \in \mathscr{T}(\mathscr{G}) \Leftrightarrow \forall t, \{T \leqslant t\} \in \mathscr{G}_t$). For the canonical filtration, simply set $\mathscr{T} = \mathscr{T}(\mathscr{F})$.

Let μ be a probability on E. The aim is to stop the evolution of a reward process (Y_t) at a stopping time $T^* \in \mathscr{T}(\mathscr{G})$ which maximizes the expected reward, i.e., such that

$$E_\mu(Y_{T^*}) = \sup\{E_\mu(Y_T);\ T \in \mathscr{T}(\mathscr{G})\}.$$

The definition of the optimal stopping problem obviously requires integrability conditions on Y. The simplest condition, which will be often used, is to suppose that (Y_t) is bounded. The "good" hypothesis is to assume that (Y_t) is of class (D) with respect to the filtration $(\mathscr{G}_t)$, i.e., is uniformly integrable over the set of stopping times, uniformly with respect to the initial condition

$$\lim_{c \to +\infty} [\sup_{\mu \in \Pi(E)} \sup_{T \in \mathscr{T}(\mathscr{G})} E_\mu(|Y_T| 1_{\{|Y_T| > c\}})] = 0.$$

A process of class (D) with respect to the canonical filtration will simply be said to be of class (D).

We will study thoroughly the particular case of a process (Y_t) which depends only on the state of the process (X_t):

$$\forall t \in \boldsymbol{R}_+,\ Y_t = e^{-\alpha t} g(X_t); \qquad Y_\infty = 0,$$

where $\alpha > 0$, g: $E \to \boldsymbol{R}$ is borelian and either bounded or of class (D_α), i.e., such that (Y_t) is of class (D). We will also study the case where $Y_t = g(X_t)$ and $Y_\infty = \limsup g(X_t)$ as in [29]. We refer to these situations by the following notation:

$$\forall \alpha > 0,\ g_t^\alpha = e^{-\alpha t} g(X_t) \quad \text{or} \quad g_t = g(X_t) = g_t^0,$$

with the convention that $g_\infty^\alpha = \limsup g_t^\alpha$ for $\alpha \geqslant 0$, $e^{-\alpha\infty} = 0$ if $\alpha > 0$ ($e^{-\alpha\infty} = 1$ if $\alpha = 0$), and $g(\Delta) = g_\infty^0$.

DEFINITION 1.3. Let Y be a reward process of class (D) with respect to $(\mathscr{G}_t)$. The *reduite* of Y is the maximal payoff function

$$v(x, Y) = \sup\{E_x(Y_T);\ T \in \mathscr{T}(\mathscr{G})\}.$$

If $Y = g_t^\alpha$ ($\alpha \geqslant 0$), we will denote the reduite $v(x, Y)$ by $R^\alpha g(x)$.

Remark. This last definition refers to that in potential theory where the reduite of a function g is the smallest α-excessive function which is greater than g. One of the aims of this paper is to show that both notions coincide by means of probabilistic arguments, which are substantially simpler than the usual ones when g does not have any regularity property (cf., e.g., [12]).

1.3. Randomized stopping times. A classical approach of such problems consists in introducing a convex set of randomized stopping times containing

$\mathcal{T}(\mathcal{G})$ and extending the optimal stopping problem. This technique is indeed natural to study the relationship between the reduite and the realization of the semigroup.

DEFINITION 1.4. Let $\mathcal{X} = (\Omega, \mathcal{G}_t, X_t, \theta_t, \boldsymbol{P}_x; x \in E)$ be a strong Markov realization of (P_t). A *randomized stopping time* is an increasing, $(\mathcal{G}_t)$-adapted, right-continuous process (A_t) such that $A_\infty = 1$ $\boldsymbol{P}_\mu$ a.s. for every initial probability $\mu \in \Pi(E)$. Let $\mathcal{A}(\mathcal{G})$ denote the set of randomized stopping times. For the canonical filtration, simply set $\mathcal{A} = \mathcal{A}(\mathcal{F})$.

The set $\mathcal{T}(\mathcal{G})$ can be embedded in $\mathcal{A}(\mathcal{G})$; indeed, for $T \in \mathcal{T}(\mathcal{G})$ and $t \in \bar{\boldsymbol{R}}_+$, set $A_t = 1_{\{T \leqslant t\}}$. Changes of time describe the connection between stopping times and randomized stopping times (see also [3], [24]). Indeed, let $r_t = \inf\{s\colon A_s \geqslant t\}$ denote the pseudoinverse of (A_t), with the convention $\inf \emptyset = +\infty$; then each r_t is a stopping time. The following proposition shows that the maximal expected values are the same over the sets $\mathcal{T}$ and $\mathcal{A}$. It is essentially shown in [24], p. 419.

PROPOSITION 1.5. *Let Y be a process of class (D) with respect to the filtration $(\mathcal{G}_t)$. Then*

(i) *The family of random variables $\{Y_A = \int_{[0,+\infty]} Y_s dA_s;\ A \in \mathcal{A}(\mathcal{G})\}$ is uniformly integrable, uniformly with respect to the initial probability.*

(ii) *The reduite of Y is the same over the sets $\mathcal{T}(\mathcal{G})$ and $\mathcal{A}(\mathcal{G})$. More precisely,*

$$v(\mu, Y) = \sup\{\boldsymbol{E}_\mu(Y_T);\ T \in \mathcal{T}(\mathcal{G})\} = \sup\{\boldsymbol{E}_\mu(Y_A);\ A \in \mathcal{A}(\mathcal{G})\} \tag{1.1}$$

for every $\mu \in \Pi(E)$.

Finally, the following theorem shows that the reduite of an $(\mathcal{F}_t)$-adapted process is independent of the realization. It is a consequence of Property (K), which holds for the filtrations $(\mathcal{F}_t)$ and $(\mathcal{G}_t)$ (see [20]). We need first the following result:

PROPOSITION 1.6. *Let $\mathcal{X} = (\Omega, \mathcal{G}_t, X_t, \theta_t, \boldsymbol{P}_x; x \in E)$ be a strong Markov realization of (P_t) and let $(\mathcal{F}_t)$ denote its canonical filtration. Given any stopping time T with respect to the filtration $(\mathcal{G}_t)$, there exists a randomized stopping time $(A_t^T) \in \mathcal{A}$ such that $\boldsymbol{E}_\mu(Z_T) = \boldsymbol{E}_\mu(Z_{A^T})$ for every $\mu \in \Pi(E)$ and every $(\mathcal{F}_\infty \otimes \mathcal{B}(\boldsymbol{R}_+))$-measurable process Z which is positive (respectively, bounded).*

Proof. For every $t \in [0, +\infty]$, let $\hat{A}_t^T = \boldsymbol{E}_\mu(1_{\{T \leqslant t\}} \mid \mathcal{F}_\infty)$ be the version independent of μ. The Markov property yields that

$$\boldsymbol{E}_\mu(1_{\{T \leqslant t\}} \mid \mathcal{F}_\infty) = \boldsymbol{E}_\mu(1_{\{T \leqslant t\}} \mid \mathcal{F}_t)\ \boldsymbol{P}_\mu \text{ a.s.},$$

and therefore that $(\hat{A}^T)$ is $\boldsymbol{P}_\mu$ a.s. $(\mathcal{F}_t)$-adapted, increasing and such that $\hat{A}_\infty^T = 1$ $\boldsymbol{P}_\mu$ a.s. Let (A^T) denote its increasing, right-continuous, $(\mathcal{F}_t)$-adapted regularization such that $A_\infty^T = 1$ $\boldsymbol{P}_\mu$ a.s. The definition of A^T shows that, given any positive, $\mathcal{F}_\infty$-measurable random variable H and any interval $]s, t]$,

$$\boldsymbol{E}_\mu(H(A_t^T - A_s^T)) = \boldsymbol{E}_\mu(H 1_{\{s < T \leqslant t\}}).$$

The monotone class theorem shows that this equality extends to positive processes Z which are $(\mathscr{F}_\infty \otimes \mathscr{B}(\boldsymbol{R}_+))$-measurable and yields

$$\int_{[0,+\infty]} Z_s dA_s^T = E_\mu(Z_T \mid \mathscr{F}_\infty) \; P_\mu \text{ a.s. } \blacksquare$$

THEOREM 1.7. *Let $\mathscr{X} = (\Omega, \mathscr{G}_t, X_t, \theta_t, P_x; x \in E)$ be a strong Markov realization of (P_t) with canonical filtration $(\mathscr{F}_t)$. Let Y be an $(\mathscr{F}_\infty \otimes \mathscr{B}(\boldsymbol{R}_+))$-measurable reward process of class (D) (with respect to $(\mathscr{F}_t)$). Then Y is of class (D) with respect to $(\mathscr{G}_t)$ and the reduites of Y over both filtrations $(\mathscr{F}_t)$ and $(\mathscr{G}_t)$ coincide, i.e., for every $\mu \in \Pi(E)$,*

$$\sup\{E_\mu(Y_T);\ \mathscr{T} \in \mathscr{T}(\mathscr{G})\} = \sup\{E_\mu(Y_T);\ T \in \mathscr{T}(\mathscr{F})\}. \tag{1.2}$$

Proof. It is well known (see, e.g., [10], t. 1, p. 38) that the uniform integrability condition is equivalent to the existence of a convex function $\phi\colon \boldsymbol{R}_+ \to \boldsymbol{R}_+$ such that

$$\lim_{t\to\infty} \frac{\phi(t)}{t} = \infty \quad \text{and} \quad \sup_{\mu\in\Pi(E)} \sup_{T\in\mathscr{T}} E_\mu[\phi(|Y_T|)] < +\infty.$$

Let $T \in \mathscr{T}(\mathscr{G})$ and $\mu \in \Pi(E)$; then if $A^T \in \mathscr{A}$ is the increasing process constructed in Proposition 1.6, then

$$\begin{aligned} E_\mu[\phi(|Y_T|)] &= E_\mu\Big[\int_{[0,+\infty]} \phi(|Y_s|)dA_s^T\Big] \\ &\leqslant \sup\{E_\mu[\phi(|Y_T|)];\ T \in \mathscr{T}\} < +\infty \end{aligned}$$

by Proposition 1.5.

To show the equality of the reduites, we use a similar argument; indeed, by Proposition 1.5, we only need to compare the suprema over the sets $\mathscr{A}$ and $\mathscr{T}(\mathscr{G})$. Let $T \in \mathscr{T}(\mathscr{G})$, $\mu \in \Pi(E)$, and let $A^T \in \mathscr{A}$ be the process defined as above. Then

$$E_\mu(Y_T) = E_\mu(Y_{A^T}) \leqslant \sup\{E_\mu(Y_A);\ A \in \mathscr{A}\}.$$

Since $\mathscr{F} \subset \mathscr{G}$, the converse inequality is obvious. $\blacksquare$

This last result shows that we can work with the canonical filtration, whose topological properties we use.

2. STOPPING RULES AND REDUITE

2.1. Stopping rules. Following Baxter and Chacon [3], Bismut [6], and Meyer [24], we introduce the "good" set of parameters to solve the optimization problem, i.e., probabilities on $\Omega \times \bar{\boldsymbol{R}}_+$ which can be written as $P_\mu(d\omega)A(\omega, dt)$, where the process $A_t(\omega) = A(\omega, [0, t])$ is a randomized stopping time for the canonical filtration. These are Meyer's "temps d'arrêt flous".

DEFINITION 2.1. Let $\mu \in \Pi(E)$; a *stopping rule starting from* μ is a probability R on $\Omega \times \bar{\boldsymbol{R}}_+$ which can be desintegrated with respect to $\boldsymbol{P}_\mu$ as

$$R(d\omega, dt) = \boldsymbol{P}_\mu(d\omega) A(\omega, dt)$$

with a randomized stopping time $A \in \mathscr{A}$.

Let $\mathscr{R}(\mu)$ denote the set of stopping rules starting from μ. When μ is a Dirac measure δ_x, we write $\mathscr{R}(x)$ for simplicity. Baxter and Chacon [3] and Meyer [24] have given characterizations of stopping rules in the case of an abstract measurable space endowed with one probability. Our setting will be slightly different since we are working with the canonical realization, and thus with a topological space Ω of trajectories endowed with various probabilities $\boldsymbol{P}_\mu$.

Given an $\mathscr{F}_\infty$-measurable random variable h and a Borel function f on $\bar{\boldsymbol{R}}_+$, set $h \otimes f(\omega, t) = h(\omega) f(t)$.

THEOREM 2.2. *Let* $\mu \in \Pi(E)$; *a stopping rule starting from* μ *is a probability on* $\Omega \times \bar{\boldsymbol{R}}_+$ *satisfying the following conditions:*

(i) *For any bounded random variable* h, $R(h \otimes 1) = \boldsymbol{E}_\mu(h)$.

(ii) *For any bounded random variable* h *and any* t *in a countable dense subset* D *of* $\bar{\boldsymbol{R}}_+$,

$$R(h \otimes 1_{[0,t]}) = R\big(\boldsymbol{E}_\mu(h \mid \mathscr{F}_t) \otimes 1_{[0,t]}\big).$$

Remark. Condition (i) ensures that the projection of R on Ω is $\boldsymbol{P}_\mu$, and condition (ii) gives the suitable adaptation property of the desintegration of R with respect to $\boldsymbol{P}_\mu$.

Proof. A probability $R \in \mathscr{R}(\mu)$ clearly satisfies (i) and (ii). We prove only the converse implication. Condition (i) shows that the projection of R on Ω is $\boldsymbol{P}_\mu$, and hence that R can be desintegrated in the form $R(d\omega, dt) = \boldsymbol{P}_\mu(d\omega)\bar{A}(\omega, dt)$, where $\bar{A}$ is a transition probability from Ω to $\bar{\boldsymbol{R}}_+$, with repartition function $\bar{A}_t = \bar{A}(\cdot, [0, t])$. This fact has already been used in [15], p. 541, and in [24], p. 411; it comes from general results concerning the desintegration of measures ([10], t. 1, p. 125). Condition (ii) shows that, for each $a \in D$, $\bar{A}_a$ is $\boldsymbol{P}_\mu$ a.s. $\mathscr{F}_a$-measurable. Let $\tilde{A}_a$ be an $\mathscr{F}_a$-measurable random variable $\boldsymbol{P}_\mu$ a.s. equal to $\bar{A}_a$. Set $\tilde{A}_\infty = 1$ and let (A_t) be an increasing right-continuous extension of $(\tilde{A}_a, a \in D)$. Since both processes $A_.$ and $\bar{A}_.$ are right--continuous, they are $\boldsymbol{P}_\mu$ indistinguishable, and $R(d\omega, dt) = \boldsymbol{P}_\mu(d\omega) A(\omega, dt)$. Furthermore, the right-continuity of the filtration $(\mathscr{F}_t)$ shows that (A_t) is $(\mathscr{F}_t)$-adapted. ∎

Baxter and Chacon, and Meyer have endowed the set $\mathscr{R}(\mu)$ with a compact topology. Since we work with a topological space Ω and several probabilities $\boldsymbol{P}_\mu \in \Pi(\Omega)$, our definition of the topology on $\mathscr{R}(\mu)$ will look different from theirs. For a fixed probability $\boldsymbol{P}_\mu$, both topologies coincide.

THEOREM 2.3. *Let $\mathscr{R}(\mu)$ be endowed with the Baxter–Chacon topology, i.e., the coarsest topology such that the maps*

$$R \in \mathscr{R}(\mu) \mapsto R(X) = E_\mu\Big(\int_{[0,+\infty]} X_s dA_s\Big) \quad \text{with } R(d\omega, dt) = P_\mu(d\omega)A(\omega, dt)$$

are continuous for every bounded $(\mathscr{F}_\infty \otimes \mathscr{B}(\bar{R}_+))$-measurable process $(X_t(\omega))$ with continuous trajectories on $\bar{R}_+$ (i.e., $\forall \omega$, $t \mapsto X_t(\omega)$ is continuous). Then the set $\mathscr{R}(\mu)$ is compact in the Baxter–Chacon topology.

Proof. We briefly sketch the argument, and refer to [3] and [24] for details. Theorem 2.2 and approximations of $1_{[0,a]}$ by continuous functions clearly show that $\mathscr{R}(\mu)$ is closed in $\Pi(\Omega \times \bar{R}_+)$. To show that $\mathscr{R}(\mu)$ is relatively compact, we use a criterion of Jacod and Memin [15] and we prove that the set of projections of $\mathscr{R}(\mu)$ on Ω and $\bar{R}_+$ are both weakly relatively compact. This is obvious since $\{R^\Omega; R \in \mathscr{R}(\mu)\} = \{P_\mu\}$, and $\{R^{\bar{R}_+}; R \in \mathscr{R}(\mu)\}$ is a set of probabilities on a compact set. ■

2.2. Dependence on the initial condition. The characterization of stopping rules given in Theorem 2.2 allows us to precise the dependence of the reduite of a (not necessarily adapted) process Y on the initial condition. The proof depends on the properties of the graph of the multivalued mapping $x \rightsquigarrow \mathscr{R}(x)$. It is easy under an additional hypothesis. We consider at first the following particular case:

We suppose that the state space E is Polish. Let $\Omega^0 = D([0, +\infty], E)$ denote the set of right-continuous and left-limited (cad-lag) maps from $[0, +\infty]$ into E. The set Ω^0 is a Polish space (i.e., separable, complete metrizable when it is endowed with the Skorokhod topology). Let $\mathscr{F}^0$ denote its canonical σ-algebra and let $C(E)$ denote the set of continuous functions on E.

In the sequel, we will denote by $\mathscr{H}$ a countable family of bounded random variables which is stable by product, and generates the σ-algebra $\mathscr{F}^0_\infty = \sigma(X_s; s \in R_+)$. Depending on the special assumption made on the Markov process (Feller, right, ...) we will mainly consider several sets $\mathscr{H}$, which will be the best suited for the particular problem under study.

Let $\mathscr{H}_b(I)$ denote the set of random variables

$$h = \prod_{1 \leqslant i \leqslant k} f_i(X_{t_i}), \quad \text{where } k \geqslant 1,\ t_1 < t_2 < \ldots < t_k \in I, \tag{2.1}$$

and the functions f_i are bounded and measurable. We suppose that (t_i) belong to a countable dense subset $I \subset R_+$ and that (f_i) belong to a countable set generating $\mathscr{B}(E)$.

If E is compact, let $\mathscr{H}_c(I)$ denote the subset of $\mathscr{H}_b(I)$ corresponding to functions (f_i) that belong to a countable dense subset of $C(E)$. When no confusion is possible, simply set $\mathscr{H}_b$ and $\mathscr{H}_c$.

By the Markov property, for every $\mu \in \Pi(E)$ and for every $h \in \mathscr{H}_b$,

$$\hat{h}_t = E_{X_t(\omega)}[h(\omega/t/\cdot)] \tag{2.2}$$

is a version of $E_\mu(h \mid \mathscr{F}_t)$, where the trajectory $\omega/t/\omega'$ is defined by (cf., e.g., [28])

$$X_s(\omega/t/\omega') = \begin{cases} X_s(\omega) & \text{if } s \leqslant t, \\ X_{s-t}(\omega') & \text{if } s > t. \end{cases}$$

The assumptions we make in the following proposition ensure the measurability of the map $x \mapsto P_x$.

PROPOSITION 2.4. *Let (P_t) be a Borel semigroup on E. We suppose that there exists a family $(P_x;\ x \in E)$ of probabilities on $(\Omega^0, \mathscr{F}^0)$ such that $\mathscr{X} = (\Omega^0, \mathscr{F}^0_t, X_t, \theta_t, P_x;\ x \in E)$ is a strong Markov realization of the semigroup (P_t). Then, for each reward process Y of class (D), the reduite*

$$v(x, Y) = \sup\{E_x(Y_T);\ T \in \mathscr{T}\}$$

is an analytic function and, for each $\mu \in \Pi(E)$,

$$\int v(x, Y)\mu(dx) = \sup\{R(Y);\ R \in \mathscr{R}(\mu)\} = \sup\{E_\mu(Y_T);\ T \in \mathscr{T}\}. \tag{2.3}$$

Proof. We study the graph $G = \{(x, R);\ R \in \mathscr{R}(x)\}$. The map $x \mapsto P_x$ is Borel from E into $\Pi(\Omega^0)$, since the Borel σ-algebra on $\Pi(\Omega^0)$ is generated by the maps $\mu \mapsto \mu(Z)$, where Z is $\mathscr{F}^0$-measurable. The monotone class theorem shows that in the characterization of $\mathscr{R}(x)$ given in Theorem 2.2 it suffices to consider the maps $R(h\otimes 1)$, $E_x(h)$, $R(h\otimes 1_{[0,a]})$ and $R(E_x(h \mid \mathscr{F}_a)\otimes 1_{[0,a]}) = R(\hat{h}_a \otimes 1_{[0,a]})$ for $h \in \mathscr{H}_b$, and a in a countable dense subset of $\boldsymbol{R}_+$. Under our assumptions all these functions are Borel on $E\otimes\Pi(\Omega^0 \times \bar{\boldsymbol{R}}_+)$. Hence the graph G is a Borel subset of the product space $E\otimes\Pi(\Omega^0 \times \bar{\boldsymbol{R}}_+)$.

Suppose at first that Y is bounded. Since Y is $(\mathscr{F}^0\otimes\mathscr{B}(\bar{\boldsymbol{R}}_+))$-measurable, the map $R \mapsto R(Y)$ is Borel from $\Pi(\Omega^0 \times \bar{\boldsymbol{R}}_+)$ into $\boldsymbol{R}$, and the theory of analytic functions ([10], t. 1, p. 119, théorème 62) implies that the function

$$v(x, Y) = \sup\{R(Y);\ R \in \mathscr{R}(x)\}$$

is analytic. We follow the proof of lemme 17 in [10], chapitre X. The function v is universally measurable and, consequently, we can find a Borel function $w(x)$, majorized by $v(x, Y)$, and μ a.s. equal to $v(x, Y)$. Given $\varepsilon > 0$, the set

$$G^\varepsilon = \{(x, R);\ R(Y)+\varepsilon \geqslant w(x),\ R \in \mathscr{R}(x)\}$$

is Borel, and its section along a fixed x, say G^ε_x, is non-empty. From the section theorem ([10], t. 1, chapitre III. 44–45) there exists a Borel set A carrying μ, and a Borel section $R^\varepsilon(x, \cdot)$ of G^ε defined on A. We set $R^\varepsilon(x, \cdot) = \varepsilon_x$ on A^c. Since $\int R^\varepsilon(x, \cdot)\mu(dx)$ belongs to $\mathscr{R}(\mu)$ by Theorem 2.2, we deduce that

$$\int v(x, Y)\mu(dx) \leqslant \sup\{R(Y);\ R \in \mathscr{R}(\mu)\}.$$

Conversely, let $R = P_\mu(d\omega)\cdot dA_t(\omega) \in \mathscr{R}(\mu)$. Since (A_t) is a randomized stopping time, we have

$$\forall x \in E,\ P_x(d\omega)\cdot dA_t(\omega) \in \mathscr{R}(x).$$

Then

$$R(Y) = \int \mu(dx) \int P_x(d\omega)\, Y_t(\omega)\, dA_t(\omega) \leqslant \int \mu(dx)\, v(x, Y).$$

Let Y be of class (D); we compare $v(x, Y)$ and $v(x, Y^c)$, where Y^c is the (bounded) truncated process $(Y \wedge c) \vee (-c)$. Set

$$\varepsilon(c) = \sup\{E_x(|Y_T| 1_{\{|Y_T|>c\}});\ T \in \mathscr{T}\}$$

for fixed $x \in E$. Then, for every $T \in \mathscr{T}$,

$$|E_x(Y_T) - E_x(Y_T^c)| = |E_x[(Y_T - c) 1_{\{Y_T > c\}} + (Y_T + c) 1_{\{Y_T < -c\}}]|$$
$$\leqslant E_x[(|Y_T| - c) 1_{\{|Y_T| > c\}}] \leqslant \varepsilon(c),$$

and

$$|v(x, Y) - v(x, Y^c)| \leqslant \sup\{|E_x(Y_T) - E_x(Y_t^c)|;\ T \in \mathscr{T}\} \leqslant \varepsilon(c).$$

Hence $v(\cdot, Y)$ is the pointwise limit of $v(\cdot, Y^c)$ as $c \to \infty$, and the Lebesgue theorem implies that the first equality in (2.3) holds for Y. Finally, Proposition 1.5 concludes the proof. ∎

2.3. Measurability of the reduite for right processes. We now consider a right semigroup (P_t) on a Lusin space E. The Ray compactification of a right semigroup is very long to be described completely, and we refer systematically to the notation and proofs of Getoor [14] (see also [28]). Let $\bar{E}$ denote the Ray–Knight compactification of E; the extension of (P_t) to $\bar{E}$ is denoted by $(\bar{P}_t)$. The relations between X and the Ray process associated with $(\bar{P}_t)$ are described in the following theorem:

THEOREM 2.5. *Let W be the set of applications w: $\mathbf{R}_+ \to E$ which are right-continuous both in the initial topology and in the Ray topology, and which have left limits in $\bar{E}$ in the Ray topology. Let (X_t) denote the coordinate process and set $\mathscr{F}_t^0 = \sigma(X_s, s \leqslant t)$. Then for each probability μ on E, P_μ is the measure constructed on $(W, \mathscr{F}^0)$ by using (P_t), and $\bar{P}_\mu$ is the corresponding one constructed by means of $(\bar{P}_t)$. These probabilities P_μ and $\bar{P}_\mu$ are equal, and $(X_t, \mathscr{F}_t^0, P_\mu)$ is a Markov process with semigroup (P_t) (resp. $(\bar{P}_t)$), if we consider E (resp. $\bar{E}$) as a state space.*

Following the remark of Getoor ([14], p. 80), by changing the topology on E into the Ray topology, the resolvent and the semigroup become Borel, and we can apply the argument above. The kind of measurability will be the following: a function v defined on E will be analytic if it is the restriction to E of a function $\bar{v}$ which is analytic on the Ray–Knight compactification $\bar{E}$ endowed with the metric ϱ.

THEOREM 2.6. *With the notation above, let* $\mathscr{C} = (\Omega, \mathscr{F}_t^{0,r}, X_t, \theta_t, P_x; x \in E)$ *denote the canonical realization of a right semigroup* (P_t). *For each* $(\mathscr{F}^{0,r} \otimes \mathscr{B}(R_+))$*-measurable process* Y *of class* (D), *the map*

$$v(x, Y) = \sup\{E_x(Y_T);\ T \in \mathscr{T}\}$$

is universally measurable on E, *and*

$$\int \mu(dx) v(x, Y) = \sup\{E_\mu(Y_T);\ T \in \mathscr{T}\}. \tag{2.4}$$

Proof. If $Y_t = g(X_t)$, let B denote the set of branching points, and $D = B^c$. If $\mu(B) = 0$, then (X_t) behaves like a right process taking on values in D. The reduite v can be defined on D and extended to R by setting $v = P_0 v$. Given any probability μ on E, the probability μP_0 does not charge B, and hence (see, e.g., [8])

$$\int \mu(dx) P_0 v(x, g(X)) = \sup\{E_{\mu P_0}(g(X_T));\ T \in \mathscr{T}\} = \sup\{E_\mu(g(X_t));\ T \in \mathscr{T}\}.$$

For a general process Y, let $\mu \in \Pi(E)$; Ω is not Lusin, but we can restrict ourselves to a Borel subset Ω' of Ω included in $W = D(R_+, \bar{E})$, where $\bar{E}$ is endowed with the Ray topology. Then replace E by a Borel subset E' such that $\mu(E') = 1$ and $P_x(\Omega') = 1$ for $x \in \bar{E}'$, and replace Y by a $(\mathscr{G}^0 \otimes \mathscr{B}(R_+))$-measurable process Y' which is P_μ indistinguishable of Y. The proof of Proposition 2.4 shows that $v(x, Y')$ is ϱ-analytic on E', and (2.4) holds by a selection theorem. ∎

3. REDUITE AND SNELL'S ENVELOPE

The most important application of Theorem 2.6 is the connection between the reduite and the Snell envelope for homogeneous processes and, more precisely, for processes of the form $g_t^\alpha = e^{-\alpha t} g(X_t)$, $t \in \bar{R}_+$ $(\alpha \geqslant 0)$, where g is a fixed nearly Borel function of class (D_α). We recall the conventions we made in Section 1:

$$X_\infty = \Delta, \quad g_\infty^\alpha = \limsup_{t \to \infty} g_t^\alpha, \quad g(\Delta) = g_\infty^0;$$

set

$$R^\alpha g(x) = v(x, g^\alpha) \text{ for } x \in E \quad \text{and} \quad R^\alpha g(\Delta) = g_\infty^\alpha.$$

We at first apply Theorem 2.6 to prove that $R^\alpha g$ is an α-strongly supermedian function.

PROPOSITION 3.1. *Let* $\mathscr{C} = (\Omega, \mathscr{F}_t, X_t, \theta_t, P_x; x \in E)$ *be the canonical realization of a right semigroup* (P_t). *For each stopping time* $S \in \mathscr{T}$ *and each probability* $\mu \in \Pi(E)$ *the following holds*:

$$E_\mu(e^{-\alpha S} R^\alpha g(X_S)) \leqslant \sup\{E_\mu(g_T^\alpha);\ T \geqslant S, T \in \mathscr{T}\} \leqslant \langle \mu, R^\alpha g \rangle. \tag{3.1}$$

Proof. Fix $S \in \mathscr{T}$, and let S_μ^α denote the measure on E defined by

$$S_\mu^\alpha(\phi) = E_\mu(e^{-\alpha S} \phi(X_S) 1_{\{S < +\infty\}}).$$

Theorem 2.6 applied to the probability $S_\mu^\alpha(1)^{-1}S_\mu^\alpha$ and the strong Markov property show that

$$S_\mu^\alpha(R^\alpha g) = \int v(x, g^\alpha) S_\mu^\alpha(dx) = \sup\{\int E_x(g_T^\alpha) S_\mu^\alpha(dx);\ T \in \mathscr{T}\}$$

$$= \sup\{E_\mu(e^{-\alpha S} E_{X_S}(g_T^\alpha) 1_{\{S < +\infty\}});\ T \in \mathscr{T}\}$$

$$= \sup\{E_\mu(e^{-\alpha(S + T\circ\theta_S)} g_{S+T\circ\theta_S};\ S < +\infty);\ T \in \mathscr{T}\}$$

$$\leqslant \sup\{E_\mu(e^{-\alpha T} g_T;\ S < +\infty);\ T \geqslant S,\ T \in \mathscr{T}\}.$$

On the set $\{S = +\infty\}$, for $T \geqslant S$ the random variables $e^{-\alpha S} R^\alpha g(X_S)$ and g_T^α are both null if $\alpha > 0$ and equal to g_∞^0 if $\alpha = 0$. Hence

$$E_\mu(e^{-\alpha S} R^\alpha g(X_S)) \leqslant \sup\{E_\mu(g_T^\alpha;\ S < \infty);\ T \geqslant S,\ T \in \mathscr{T}\}$$

$$+ E_\mu(e^{-\alpha S} R^\alpha g(\Delta);\ S = +\infty)$$

$$\leqslant \sup\{E_\mu(g_T^\alpha);\ T \geqslant S,\ T \in \mathscr{T}\} \leqslant \langle \mu, R^\alpha g \rangle. \blacksquare$$

Remark. Inequality (3.1) shows that the reduite $R^\alpha g$ is the smallest α-supermedian function greater than or equal to g. Its α-excessive regularization $\hat{R}^\alpha g$ is defined by

$$\hat{R}^\alpha g = \lim_{t \to 0} e^{-\alpha t} P_t R^\alpha g \leqslant R^\alpha g \text{ on } E \quad \text{and} \quad \hat{R}^\alpha g(\Delta) = g_\infty^\alpha.$$

The converse inequality of (3.1) is established in lemme 2.7.1 of [12]. It is clear if g is l.s.c. on trajectories, since $\hat{R}^\alpha g \geqslant g$ yields that $\hat{R}^\alpha g = R^\alpha g$. Let $(\mathscr{F}_t^0)$ denote the filtration generated by $(X_s;\ s \leqslant t)$.

LEMMA 3.2. *Let $\mathscr{C} = (\Omega, \mathscr{F}_t, X_t, \theta_t, P_x;\ x \in E)$ be the canonical realization of a right semigroup (P_t). Then, for all stopping times $T \geqslant S$ with $T \in \mathscr{F}_t^{0,+} = \bigcap_{\varepsilon > 0} \mathscr{F}_{t+\varepsilon}$ and $S \in \mathscr{T}(\mathscr{F}_t^0)$,*

$$E_\mu(g_T^\alpha) \leqslant E_\mu(e^{-\alpha S} R^\alpha g(X_S)). \tag{3.2}$$

Proof. We denote by $\mathscr{T}^0$ (resp. $\hat{\mathscr{T}}^0$) the set of stopping times with respect to $\mathscr{F}_t^0$ (resp. $\mathscr{F}_t^{0,+}$). Let S (resp. T) be a stopping time in $\mathscr{T}^0$ (resp. $\hat{\mathscr{T}}^0$), with $S \leqslant T$. By a slight modification of a theorem of Courrege–Priouret (cf. [10], t. 1, p. 237) there exists an $(\mathscr{F}_S^0 \otimes \mathscr{F}_\infty^0)$-measurable random variable $U\colon \Omega \times \Omega \to [0, +\infty]$ such that

(i) $U(\omega, w) = 0$ if $S(\omega) = +\infty$ or if $S(\omega) < +\infty$ and $X_0(w) \neq X_S(\omega)$.

(ii) $U(\omega, \cdot)$ belongs to $\hat{\mathscr{T}}^0$.

(iii) $T(\omega) = S(\omega) + U(\omega, \theta_S(\omega))$.

The proof is similar to that in [10], using the Galmarino test for $\mathscr{F}_t^{0,+}$ stopping times ([10], t. 1, p. 234, théorème 101) instead of the Galmarino test for $\mathscr{F}_t^0$ stopping times ([10, t. 1, p. 234, théorème 100). By the strong

Markov property,

$$E_\mu(g_T^\alpha;\ S < +\infty) = E_\mu\big(e^{-\alpha S} E_{X_S}(g^\alpha_{U(\omega,\cdot)});\ S(\omega) < +\infty\big) \qquad (3.3)$$
$$\leqslant E_\mu\big(e^{-\alpha S} R^\alpha g(X_S);\ S < +\infty\big).$$

Since g_T^α and $e^{-\alpha S} R^\alpha g(X_S)$ coincide on $\{S = +\infty\}$, this concludes the proof of (3.2). ∎

Hence, if g is l.s.c. on trajectories, approximating any stopping time S by a decreasing sequence in $\mathscr{T}_0$ yields that inequality (3.2) holds for any $S \in \mathscr{T}$. The following lemma gives a more precise inequality for any arbitrary nearly Borel function g.

LEMMA 3.3. *Let $\mathscr{C} = (\Omega, \mathscr{F}_t, X_t, \theta_t, P_x;\ x \in E)$ be the canonical realization of a right semigroup (P_t). Then, for all stopping times $T \geqslant S$ in $\mathscr{T}$,*

$$E_\mu(g_T^\alpha) \leqslant E_\mu\big(e^{-\alpha S}(\hat{R}^\alpha g \vee g)(X_S)\big) \qquad (3.4)$$

and $R^\alpha g = g \vee \hat{R}^\alpha g$.

Proof. At first we extend inequality (3.2) to stopping times $S \in \hat{\mathscr{T}}^0$. Apply (3.2) to $S = \varepsilon > 0$, and $T = \sup(V, \varepsilon)$, where V is a strictly positive stopping time in $\hat{\mathscr{T}}^0$. Then

$$E_x(g^\alpha_{\sup(V,\varepsilon)}) \leqslant E_x(e^{-\alpha\varepsilon} R^\alpha g(X_\varepsilon)).$$

Since g_t^α is of class (D^α), we can let $\varepsilon \to 0$, and obtain $E_x(g_V^\alpha) \leqslant \hat{R}^\alpha g(x)$. If we suppose that $T > S \in \mathscr{T}_0$ on $\{T < \infty\}$, then the stopping time $U(\omega, \cdot)$ is strictly positive, and we have

$$E_\mu(g_T^\alpha;\ S < +\infty) \leqslant E_\mu(e^{-\alpha S} \hat{R}^\alpha g(X_S);\ S < +\infty).$$

Fix $S \in \hat{\mathscr{T}}^0$, $T > S$ on $\{T < +\infty\}$, and approximate it by a strictly decreasing sequence (S_n) of $\mathscr{F}_t^0$ stopping times. Let (T_n) be a sequence in $\hat{\mathscr{T}}^0$ defined by

$$T_n = T \text{ on } \{T > S_n\} \quad\text{and}\quad T_n = +\infty \text{ otherwise.}$$

The sequence (T_n) decreases to T in a stationary way (i.e., for every ω there exists an integer $N(\omega)$ such that $T_n(\omega) = T(\omega)$ for all $n \geqslant N(\omega)$), and since $T_n > S_n$ on $\{T_n < \infty\}$, we have

$$E_\mu(g_{T_n}^\alpha;\ S_n < +\infty) \leqslant E_\mu(e^{-\alpha S_n} \hat{R}^\alpha g(X_{S_n});\ S_n < +\infty).$$

We use the right-continuity of the process $e^{-\alpha t} \hat{R}^\alpha g(X_t)$ to deduce that

$$E_\mu(g_T^\alpha;\ S < +\infty) \leqslant E_\mu(e^{-\alpha S} \hat{R}^\alpha g(X_S);\ S < +\infty)$$

by letting $n \to \infty$.

Suppose at last that $T \geqslant S$ belongs to $\mathscr{T}$, and set

$$\hat{T} = T \text{ on } \{S < T\} \quad\text{and}\quad \hat{T} = +\infty \text{ otherwise,}$$
$$\hat{S} = S \text{ on } \{S < T\} \quad\text{and}\quad \hat{S} = +\infty \text{ otherwise.}$$

Then

$$\begin{aligned}E_\mu(g_T^\alpha;\ \{S<+\infty\}) &= E_\mu(g_T^\alpha;\ \{S<T\}\cap\{S<+\infty\})\\ &\quad +E_\mu(e^{-\alpha S}g(X_S)1_{\{S=T\}};\ S<+\infty)\\ &= E_\mu(g_{\hat T}^\alpha;\ \hat S<+\infty)+E_\mu(e^{-\alpha S}g(X_S)1_{\{S=T\}};\ S<+\infty)\\ &\leqslant E_\mu(e^{-\alpha\hat S}\hat R^\alpha g(X_{\hat S});\ \hat S<+\infty)\\ &\quad +E_\mu(e^{-\alpha S}g(X_S)1_{\{S=T\}};\ S<+\infty)\\ &= E_\mu(e^{-\alpha S}[\hat R^\alpha g(X_S)1_{\{S<T\}}+g(X_S)1_{\{S=T\}}];\ S<+\infty)\\ &\leqslant E_\mu(e^{-\alpha S}(\hat R^\alpha g\vee g)(X_S);\ S<+\infty).\end{aligned}$$

Since $(\hat R^\alpha g \vee g)(\Delta) = g_\infty^\alpha$ on $\{S = +\infty\}$, the proof of (3.4) is complete. Finally, (3.4) applied with $S=0$ yields that $R^\alpha g \leqslant \hat R^\alpha g \vee g \leqslant R^\alpha g$. ■

The main consequence of those two results is the description of the Snell envelope in terms of the reduite. Let $\mathscr{E}^e$ denote the σ-algebra generated by excessive functions.

THEOREM 3.4. *Let $\mathscr{X} = (\Omega, \mathscr{G}_t, X_t, \theta_t, P_x;\ x\in E)$ be a strong Markov realization of a right semigroup (P_t) and let g be a universally measurable function of class (D_α). Then, if g is $\mathscr{E}^e$-measurable, the process*

$$\hat R_t^\alpha =: e^{-\alpha t}R^\alpha g(X_t) \ \text{if}\ t<\infty, \qquad \hat R_\infty^\alpha = g_\infty^\alpha$$

is a strong supermartingale which is the Snell envelope $J(g^\alpha)$ of $(g_t^\alpha,\ t\in[0,+\infty])$, i.e.,

$$\hat R_S^\alpha = \operatorname{ess\,sup}\{E_\mu(g_T^\alpha\,|\,\mathscr{G}_S);\ S\leqslant T\in\mathscr{T}(\mathscr{G})\}.$$

Proof. Theorem 1.7 shows that we can use the canonical realization of (P_t). By Proposition 3.1 and Lemma 3.3, we see that, given $S\in\mathscr{T}$ and $\mu\in\Pi(E)$,

$$\begin{aligned}\sup\{E_\mu(g_T^\alpha);\ S\leqslant T\in\mathscr{T}\} &\leqslant E_\mu(e^{-\alpha S}R^\alpha g(X_S))\\ &\leqslant \sup\{E_\mu(g_T^\alpha);\ S\leqslant T\in\mathscr{T}\}\\ &= E_\mu(J(g^\alpha)_S),\end{aligned}$$

where the last equality is deduced from the property of decreasing filtrations for the set $(E_\mu(g_T^\alpha),\ T\geqslant S,\ T\in\mathscr{T})$ (cf., e.g., [22]). Therefore,

$$(3.5)\quad e^{-\alpha S}R^\alpha g(X_S) = \operatorname{ess\,sup}\{E_\mu(g_T^\alpha\,|\,\mathscr{F}_S);\ S\leqslant T\in\mathscr{T}\} \quad \text{on } \{S<+\infty\}.$$

Since $\hat R^\alpha g$ is obviously $\mathscr{E}^e$-measurable, the equation $R^\alpha g = \hat R^\alpha g \vee g$ shows that $R^\alpha g$ is $\mathscr{E}^e$-measurable, and hence the process $\hat R^\alpha g$ is optional (cf., e.g., [10], t. 3, or [28]). Thus, equation (3.5) concludes the proof. ■

Remark 3.5. (1) The Snell envelope is a crucial tool in the theory of optimal stopping. Thus it is very important to relate the reduite and the Snell envelope. For example, optimal stopping times of a process (Y_t) can be characterized in terms of the Snell envelope $J(Y)$ of Y as follows (see, e.g., [12]): A stopping time T^* is *optimal* if and only if:

- $Y_{T^*} = J(Y)_{T^*}$;
- the process $(J(Y)_{t \wedge T^*}, t \geqslant 0)$ is a martingale.

The Snell envelope also gives an explicit construction of optimal stopping times under u.s.c. assumptions on the process (cf., e.g., [12]). More precisely, let (Y_t) be an optional process of class (D). Then:

- If Y is u.s.c. on trajectories, then $\inf\{t\colon Y_t = J(Y)_t\}$ is the smallest optimal stopping time.
- If Y is u.s.c. in expectation (i.e., $EY_T \geqslant \limsup EY_{T_n}$ for every sequence (T_n) of stopping times converging to T), let $J(Y)_t = M_t - A_t$ be the Doob decomposition of the supermartingale $J(Y)$. Then $\inf\{t\colon Y_t = J(Y)_t\}$ and $\inf\{t\colon J(Y)_t \neq M_t\}$ are optimal.

(2) Given an arbitrary function g, the reduite $R^\alpha g$ we have studied is the smallest strongly α-supermedian function greater than or equal to g. Then:

- For fixed $\alpha > 0$, the optimal stopping time for the process g^α and the probability P_μ is the entrance time in a subset of E which does not depend on the initial law μ.
- If g is l.s.c. or, more generally, l.s.c. on trajectories (e.g., the difference of two excessive functions), then $R^\alpha g$ is excessive, and hence it is the "excessive" reduite in the sense of potential theory (cf., e.g., [10], t. 3).

4. CONTINUITY PROPERTIES OF THE REDUITE

In this section we prove continuity results on the reduite when both the function f and the semigroup (P_t) have "continuity" properties, e.g., f is continuous and (P_t) is Feller. Thus we generalize some known results about the reduite, but the novelty lies mainly in the method which we develop. Instead of the penalization iterative method (cf. [26], [12]) or a discretization method (cf. [18], [7], [31]) we use the more probabilistic notion of stopping rule, which also yields functional results. Furthermore, our method depends only on the weak continuity of the map $x \mapsto P_x$, which is well known for diffusion Markov processes (cf., e.g., [25]). In the Appendix, we prove that this continuity property also holds for Feller processes on a compact state space E.

Since the initial condition $x \in E$ will not be fixed, we cannot use the Baxter–Chacon topology on $\mathscr{R} = \bigcup \mathscr{R}(x)$; indeed, this would imply that the laws of the processes are strongly continuous.

Throughout this section we suppose that the state space is Polish and that sometimes it is compact metrizable. Let $\Omega = D([0, +\infty], E)$ denote the canonical set of right-continuous and left-limited maps from $[0, +\infty]$ into E, and let ω_t denote the coordinate maps. The set Ω is endowed with the Skorokhod topology for which it is metrizable, complete and separable (cf., e.g., [4], [16]). We recall some well-known properties of this topology on Ω.

PROPOSITION 4.1. (a) *For every bounded continuous function f on E, every $\alpha > 0$ and $0 \leq t_1 < t_2 \leq +\infty$, the map $\omega \mapsto \int_{[t_1,t_2]} e^{-\alpha s} f(\omega_s) ds$ is continuous.*

(b) *For every given probability P on $(\Omega, \mathscr{F}^0_\infty)$, almost every t (with respect to Lebesgue measure), there exists a P-null set N_t such that the map $\omega \mapsto \omega_t$ is continuous for every $\omega \notin N_t$.*

In this section we will assume that the following assumptions (Ha) and (Hb) hold:

(Ha) One can define on Ω a family of probabilities $(\boldsymbol{P}_x;\ x \in E)$ such that $(\Omega, \mathscr{F}_t, \Omega_t, \theta_t, \boldsymbol{P}_x;\ x \in E)$ is a strong Markov realization of the semigroup (P_t).

(Hb) The map $x \mapsto \boldsymbol{P}_x$ is weakly continuous.

Note that the assumption (Hb) is satisfied when $E = \boldsymbol{R}^d$, and $(\boldsymbol{P}_x)$ is solution of a "good" martingale problem. The following theorem shows that (Hb) also holds for an arbitrary Feller process on a compact state space E (i.e., such that $P_t(C(E)) \subset C(E)$ for each $t > 0$, and $P_t f \to f$ pointwise as $t \to 0$ for each $f \in C(E)$), or Feller in the following weaker sense: $U^\alpha(C(E)) \subset C(E)$ for each $\alpha > 0$, and $\alpha U^\alpha f \to f$ pointwise as $\alpha \to +\infty$ for each $f \in C(E)$. Its proof is given in the Appendix.

THEOREM 4.2. *Let $(\Omega, \mathscr{F}_t, \omega_t, \theta_t, \boldsymbol{P}_x;\ x \in E)$ be the canonical realization of a Feller semigroup (P_t) on a compact metric space E. Then the map $x \mapsto \boldsymbol{P}_x$ is continuous when $\Pi(\Omega)$ is endowed with the weak topology.*

We now prove regularity properties of the sets $\mathscr{R}(\mu)$ of stopping rules when $\Omega = D([0, +\infty], E)$ under the assumptions (Ha) and (Hb) on the semigroup. The set $\mathscr{R} = \bigcup_{\mu \in \Pi(E)} \mathscr{R}(\mu)$ is included in $\Pi(\Omega \times [0, +\infty])$, and is endowed with the weak-star topology. Note that for fixed $\mu \in \Pi(E)$, the restriction to $\mathscr{R}(\mu)$ of the weak-star topology is the same as the Baxter–Chacon topology on $\mathscr{R}(\mu)$, since $\mu_n \to \mu$ in the Baxter–Chacon topology if $\mu_n(X) \to \mu(X)$ as $n \to +\infty$ for every bounded continuous process X (see [3] or [24], p. 419). Indeed, for fixed μ choose a countable dense subset $I \subset \boldsymbol{R}_+$ and a subset $N \subset \Omega$ such that $\boldsymbol{P}_\mu(N) = 0$, and $\omega \mapsto \omega_t$ is continuous for each $t \in I$ and $\omega \notin N$ (cf. Proposition 4.1). Then the random variables $h \in \mathscr{H}_c(I)$ and $\hat{h}_t = E_{X_t}(h(\omega/t/\cdot))$ are "continuous" on Ω if $t \in I$.

THEOREM 4.3. *Let $\Omega = D([0, +\infty], E)$, let K be a compact subset of E such that the map $x \mapsto \boldsymbol{P}_x$ is continuous on K. Then:*

(a) *The graph $G_K = \{(x, R);\ x \in K,\ R \in \mathscr{R}(x)\}$ is a compact subset of $K \times \Pi(\Omega \times [0, +\infty])$, where $\Pi(\Omega \times [0, +\infty])$ is endowed with the weak-star topology.*

(b) *The set $\mathscr{R}_K = \bigcup_{x \in K} \mathscr{R}(x)$ is a weakly compact subset of $\Pi(\Omega \times [0, +\infty])$.*

Proof. First we check that G_K is closed. Let (x_n, R_n) be such that $x_n \to x$ and $R_n \to R$. The sequence $(\boldsymbol{P}_{x_n})$ of projections of R_n on Ω converges weakly to $\boldsymbol{P}_x$, which is hence the projection of R on Ω. To show that $R \in \mathscr{R}(x)$, it suffices to prove that the second condition in Theorem 2.2 is satisfied. Let $I = (a_i)$ be a countable dense subset of $[0, +\infty]$ such that the maps $\omega \mapsto X_{a_i}(\omega)$ are continuous except on a P_x-null set N of Ω. Let $a \in I$, $h \in \mathscr{K}_c(I)$, and ϕ be a continuous function on $[0, +\infty]$ with support included in $[0, a]$. Then by Theorem 2.2 we have

$$R_n(h \otimes \phi) = R_n(\hat{h}_a \otimes \phi), \quad \forall n \geqslant 0.$$

Since the functions h and $\hat{h}_a$ are continuous except on $N \times [0, +\infty]$, which is an R-null set, letting $n \to +\infty$ yields that $R(h \otimes \phi) = R(\hat{h}_a \otimes \phi)$, and hence that G_K is closed.

Since G_K is closed, it suffices to prove that $\mathscr{R}_K$ is weakly compact, and hence that the sets $\mathscr{R}_K^{\Omega}$ and $\mathscr{R}_K^{\bar{\mathbb{R}}_+}$ of projections of $\mathscr{R}_K$ on $\Pi(\Omega)$ and $\Pi([0, +\infty])$, respectively, are tight. The continuity assumption made on the map $x \mapsto P_x$ shows that $\mathscr{R}_K^{\Omega} = \bigcup_{x \in K} P_x$ is a compact subset of $\Pi(\Omega)$. Since $[0, +\infty]$ is compact, the tightness of $\mathscr{R}_K^{\bar{\mathbb{R}}_+}$ is obvious, and $\mathscr{R}_K$ is compact. Since G_K is a closed subset of $K \times \mathscr{R}_K$, it is clearly compact. ∎

The following lemma gives sufficient conditions for the upper semicontinuity of a map defined in terms of suprema of continuous functions. The lower semicontinuity of such maps is intuitively expected. Similar results can be found in the more general setting of set-valued maps (cf. [2]).

LEMMA 4.4. *Let X be metrizable and let $\mathscr{R}$ be compact metrizable. Let $F\colon X \times \mathscr{R} \to \mathbb{R}$ be bounded u.s.c., and for every $x \in X$ let $\mathscr{R}(x) \subset \mathscr{R}$ be such that $G = \{(x, R);\ x \in X,\ R \in \mathscr{R}(x)\}$ is closed in $X \times \mathscr{R}$. Then*

$$v(x) = \sup\{F(x, R);\ R \in \mathscr{R}(x)\}$$

is u.s.c.

Thus we obtain the upper semicontinuity of the reduite of a u.s.c. process Y which is not necessarily a function of X.

THEOREM 4.5. *Let $\Omega = D([0, +\infty], E)$ and suppose that E is LCCB and that the map $x \mapsto P_x$ is continuous from E to $\Pi(\Omega)$. Let Y be an $(\mathscr{F}_\infty^0 \otimes \mathscr{B}(\bar{\mathbb{R}}_+))$--measurable process of class (D). If the map $(\omega, t) \mapsto Y_t(\omega)$ is u.s.c. on $\Omega \times [0, +\infty]$, then the reduite $v(x, Y) = \sup\{R(Y);\ R \in \mathscr{R}(x)\}$ is u.s.c.*

Proof. Fix $x_0 \in E$ and let K be a compact neighbourhood of x_0. First we suppose that Y is bounded. Set $F(x, R) = R(Y)$; then the map $R \mapsto R(Y)$ is u.s.c., and hence F is u.s.c. Therefore, Lemma 4.4 applied with $X = K$ and $\mathscr{R} = \mathscr{R}_K (= \bigcup_{x \in K} \mathscr{R}(x))$ shows that v is u.s.c. at x_0.

Suppose that Y is of class (D) and fix $c > 0$. Let Y^c be the truncated process $(-c) \vee (Y \wedge c)$. Then $F_c(R) = R(Y^c)$ converges uniformly to F as $c \to +\infty$; indeed,

$$\begin{aligned}\limsup_{c} \sup_{R \in \mathscr{R}} |R(Y) - R(Y^c)| &\leqslant \lim_{c} \sup_{R \in \mathscr{R}} R(|Y| 1_{\{|Y| > c\}}) \\ &= \lim_{c} \sup_{\mu \in \Pi(E)} \sup_{T \in \mathscr{T}} E_\mu(|Y_T| 1_{\{|Y_T| > c\}}) = 0.\end{aligned}$$

Hence the map $v(\cdot, Y)$ is also u.s.c. ∎

The extension of the optimal stopping problem to a larger convex compact set has yielded the upper semicontinuity of v. As might be expected, the lower semicontinuity of v is easier; it is obtained by restricting the study to the smaller class $\mathscr{T}_I$ of I-valued simple stopping times T (i.e., taking on fnitely many values in $I \subset \bar{R}_+$).

THEOREM 4.6. *Let* $\Omega = D([0, +\infty], E)$ *and let* Y *be a measurable (not necessarily adapted) process of class (D). Then:*

(a) *If* (Y_t) *has lower semicontinuous trajectories, then for every* $\mu \in \Pi(E)$ *and any dense subset* $I \subset \bar{R}_+$

$$v(\mu, Y) := \sup\{R(Y);\ R \in \mathscr{R}(\mu)\} = \sup\{E_\mu(Y_T);\ T \in \mathscr{T}_I\}.$$

(b) *Suppose that the map* $x \mapsto P_x$ *is continuous from* E *in* $\Pi(\Omega)$ *and let* (Y_t) *be l.s.c. on* $\Omega \times R_+$; *then* $v(\cdot, Y)$ *is l.s.c.*

Proof. We suppose first that Y is bounded.

(a) Let $T \in \mathscr{T}$ and let I be a countable dense subset of $\bar{R}_+$. Let T_n be a decreasing sequence of I-valued simple stopping times converging to T. Then since Y has l.c.s. trajectories, we obtain

$$E_\mu(Y_T) \leqslant \liminf E_\mu(Y_{T_n}) \leqslant \sup\{E_\mu(Y_S);\ S \in \mathscr{T}_I\},$$

so that $\sup\{E_\mu(Y_T);\ T \in \mathscr{T}\} \leqslant \sup\{E_\mu(Y_T);\ T \in \mathscr{T}_I\}$; the converse inequality is obvious, and Theorem 1.7 concludes the proof.

(b) Let (x_n) be a sequence in E converging to x. Let I be a countable dense subset of $\bar{R}_+$, and N be a P_x-null subset of Ω such that for every $t \in I$ the map $\omega \mapsto \omega_t$ is continuous on N^c. Let

$$T = \sum_{i \leqslant k} t_i 1_{\{T = t_i\}} \in \mathscr{T}_I$$

and fix $\varepsilon > 0$. For each $i \leqslant k$ choose an $\mathscr{F}_{t_i}$-measurable random variable H_i such that $\omega \mapsto H_i(\omega)$ is continuous except on N, and

$$|E_x(Y_T) - E_x(Y_A)| \leqslant \varepsilon \quad \text{for } A(\omega, dt) = \sum_{i \leqslant k} H_i(\omega)\delta_{t_i}(dt) \in \mathscr{A}.$$

Since the map Y is l.s.c. on $\Omega \times \bar{R}_+$, the map $\omega \mapsto Y_A(\omega)$ is l.s.c. except on N. Therefore, for $x_n \to x$

$$\liminf_n v(x_n, Y) \geqslant \liminf_n E_{x_n}(Y_A) \geqslant E_x(Y_A) \geqslant E_x(Y_T) - \varepsilon.$$

This clearly shows that $\liminf v(x_n, Y) \geqslant \sup\{E_x(Y_T);\ T \in \mathscr{T}_I\} = v(x, Y)$, and hence that the map $v(\cdot, Y)$ is l.s.c.

The standard truncation argument used in the proof of Theorem 4.5 yields the lower semicontinuity of $v(\cdot, Y)$ for processes Y of class (D). ∎

Theorems 4.5 and 4.6 imply the continuity of the reduite in the particular case of continuous processes or of continuous functions of (X_t).

COROLLARY 4.7. *Let E be compact metrizable and let $\Omega = D([0, +\infty], E)$. Suppose that the map $x \mapsto P_x$ is weakly continuous and that Y is a process of class (D) which is continuous on $\Omega \times \bar{R}_+$. Then the reduite $v(x, Y) = \sup\{E_x(Y_T); T \in \mathcal{T}\}$ is continuous.*

Thus we obtain the continuity of the reduite q^α of a continuous function of a Feller process (see, e.g., [12], théorème 2.82).

COROLLARY 4.8. *Let $\mathcal{X} = (\Omega, \mathcal{G}_t, X_t, \theta_t, P_x;\ x \in E)$ be a strong Markov realization of a Feller semigroup (P_t) on a compact set E. Then, for every $g \in C(E)$ and $\alpha > 0$,*

$$R^\alpha g(x) = \sup\{E_x[e^{-\alpha T} g(X_T)];\ T \in \mathcal{T}(\mathcal{G})\}$$

is continuous.

Proof. We suppose at first that g is the α-potential of a continuous function f. Then for each $T \in \mathcal{T}$ and $x \in E$, by the strong Markov property,

$$E_x(e^{-\alpha T} U^\alpha f(X_T)) = E_x\Big(\int_{[T, +\infty]} e^{-\alpha s} f(X_s) ds\Big).$$

The map $(\omega, t) \mapsto Y_t(\omega) = \int_{[t, +\infty]} e^{-\alpha s} f(\omega_s) ds$ is continuous. Indeed, since f is bounded, the map $t \mapsto Y_t(\omega)$ is continuous uniformly in ω, and for fixed t Proposition 4.1 (a) implies that the map $\omega \mapsto Y_t(\omega)$ is continuous. Hence Corollary 4.7 shows that $R^\alpha g = v(\cdot, Y)$ is continuous. The uniform approximation of continuous functions by potentials concludes the proof. ∎

Remark. Suppose that the assumptions of Corollary 4.8 are satisfied and let g, h be continuous functions on E. Then the map

$$q^\alpha(x) = \sup\{E_x[e^{-\alpha T} g(X_T) + \int_{[0,T]} e^{-\alpha u} h(X_u) du];\ T \in \mathcal{T}(\mathcal{G})\}$$

is continuous. Indeed,

$$\begin{aligned} q^\alpha(x) &= E_x\Big[\int_{[0, +\infty]} e^{-\alpha u} h(X_u) du\Big] + \sup\{E_x[e^{-\alpha T}(g - U^\alpha h)(X_T)];\ T \in \mathcal{T}\} \\ &= U^\alpha h(x) - R^\alpha(g - U^\alpha h)(x). \end{aligned}$$

Since the maps h and g are continuous and (P_t) is Feller, $U^\alpha h \in C(E)$, and $R^\alpha(g - U^\alpha h) \in C(E)$. ∎

5. EXAMPLE: OPTIMAL STOPPING FOR DIFFUSION PROCESSES WITH JUMPS

In this section, we prove that in the case of a "good" diffusion with jumps the reduite of a continuous function in also continuous. We use the continuity results established in the previous section for potentials of continuous functions and an exponential estimation to conclude the proof.

5.1. Hypothesis and notation. Let L be an integro-differential operator of the form

$$Lf(t, x) = [\tfrac{1}{2} a_{ij} D_{ij}^2 f + b_j D_j f + D_t f](t, x) + \int_{R^d - \{0\}} [f(x+u) - f(x) - 1_{\{|u|<1\}} \langle u, \nabla f(t, x)\rangle] S(t, x, du),$$

where we assume that the following assumptions (H1) and (H2) hold (cf. [17] for details):

(H1) $\varrho(t, x) = (a_{ij}, b_j)(t, x)$ is Borel bounded by K, $\varrho(t, x)$ is continuous for each t, and (a_{ij}) is strictly positive.

(H2) $S(t, x, du)$ is a positive kernel on $R^d - \{0\}$ such that

$$\sup\{S(t, x, |u|^2 \wedge 1);\ t, x\} \leqslant K.$$

Furthermore, for each bounded continuous function f and each t, $S(t, \cdot, f)$ is continuous.

Then by [30], for each $x \in R^d$ the martingale problem corresponding to (x, a, b, S) is "well-set", i.e., there exists a unique probability P_x on the space $\Omega = D([0, +\infty], R^d)$ endowed with the canonical filtration $(\mathscr{F}_t)$ such that

(i) $P_x(X_0 = x) = 1$;

(ii) $\forall f \in C_b^{1,2}(R_+ \times R^d)$, $f(t, X_t) - f(0, X_0) - \int_{[0,t]} Lf(s, X_s)ds$ is a P_x-martingale.

The assumptions made on a, b, and S imply that $Lf(t, \cdot)$ is continuous for every function $f \in C_b^{1,2}(R_+ \times R^d)$.

5.2. Continuity properties. We use the following result which can be deduced from the exponential majorization in the setting of differential operators ([17], théorème 13).

LEMMA 5.1. *Given $\lambda \in R$ and positive numbers A, η, K, there exists a constant k which depends only on K such that*

$$P_x(\sup\{|X_s - X_0| > \eta;\ 0 \leqslant s \leqslant t\}) \leqslant 2d \exp\left[-\frac{\lambda}{d}(\eta - Kt - A) + \frac{\lambda^2}{2} kt(1 + e^{|\lambda|})\right] + \frac{Kt}{A}.$$

The following result generalizes a classical result for continuous diffusions (see, e.g., [25]):

LEMMA 5.2. *Under the assumptions* (H1) *and* (H2), *the map $x \mapsto P_x$ is weakly continuous from R^d to $\Pi(\Omega)$.*

Sketch of the proof. The technique used in [17], théorème 20, shows at first that the family $(P_x;\ x \in R^d)$ is weakly relatively compact. Let $x_x \to x$; the sequence (P_{x_n}) is weakly compact, and let P be one of its cluster points. Given $f \in C_b^{1,2}(R_+ \times R^d)$ set

$$H_t^f = f(t, X_t) - f(0, X_0) - \int_{[0,t]} Lf(s, X_s)ds.$$

The set $\{X_0 = x\}$ is closed in the Skorokhod topology, and $P(X_0 = x) = 1$. On the other hand, the right-continuity of (X_t) and the definition of P_x show that, in order to prove that $P = P_x$, it suffices to check that $E_{x_n}(\xi H_t^f) \to E_x(\xi H_t^f)$ for t in a countable dense subset of R_+ such that the maps $\omega \mapsto X_t(\omega)$ are continuous except on a P_x-null subset of Ω, for $f \in C_b^{1,2}(R_+ \times R^d)$, and for a bounded $\mathcal{F}_s$-measurable random variable ξ. This last convergence is a direct consequence of the continuity of H_t^f. ∎

Let $f \in C_b(R^d)$ and $\alpha > 0$; we want to prove that

$$R^\alpha f(x) = \sup\{E_x(e^{-\alpha T} f(X_T));\ T \in \mathcal{T}\}$$

is a bounded continuous function. We suppose at first that $f = U^\alpha g$ with $g \in C_b(R^d)$. Then Lemma 5.2 and Corollary 4.7 show that $R^\alpha f \in C_b(R^d)$. In order to deduce the general result, we need the following technical lemma:

Lemma 5.3. *Let $f \in C_b(R^d)$. Then for each $\varepsilon > 0$ and each compact set $K \subset R^d$ there exists a C^∞-function g_K^ε with compact support such that:*

(i) $\sup\{|f(x) - g_K^\varepsilon(x)|;\ x \in K\} < \varepsilon$;

(ii) $\|g_K^\varepsilon\| \leqslant \|f\|_\infty + \varepsilon$.

This yields the main result of this section.

Theorem 5.4. *Under the assumptions* (H1) *and* (H2), *given* $\alpha > 0$ *and a bounded continuous function g, the reduite $R^\alpha g$ is continuous.*

Proof. The continuity of the reduite is true for C^∞-functions with compact support, since they are the α-potentials of some continuous functions. For every n, let K_n denote the closure of the ball $B(0, n)$, and let $g_n = g_{K_n}^{1/n}$ be the function constructed in Lemma 5.3. Given $t > 0$ and $T \in \mathcal{T}$,

$$|E_x(e^{-\alpha T} f(X_T)) - E_x(e^{-\alpha T} g_n(X_T))| \leqslant I_1 + I_2 + I_3$$

with

$$I_1 = |E_x[e^{-\alpha T} f(X_T) - e^{-\alpha(T \wedge t)} f(X_{T \wedge t})]|,$$

$$I_2 = |E_x[e^{-\alpha(T \wedge t)}(f(X_{T \wedge t}) - g_n(X_{T \wedge t}))]|,$$

$$I_3 = |E_x[e^{-\alpha T} g_n(X_T) - e^{-\alpha(T \wedge t)} g_n(X_{T \wedge t})]|.$$

Then $I_1 \leqslant 2e^{-\alpha t}\|f\|_\infty$ and $I_3 \leqslant 2e^{-\alpha t}(\|f\|_\infty + n^{-1})$. Fix $\varepsilon > 0$ and choose t such that $I_1 + I_3 \leqslant \varepsilon$. Given any n, set $T_n = \inf\{t:\ X_t \notin K_n\}$. Then

$$I_2 \leqslant \sup\{|f - g_n|(x);\ x \in K_n\} + (2\|f\|_\infty + n^{-1}) P_x(T \wedge t \geqslant T_n)$$
$$\leqslant n^{-1} + (2\|f\|_\infty + n^{-1}) P_x(t \geqslant T_n).$$

Thus Lemma 5.1 shows that one can choose N such that $I_2 \leqslant \varepsilon$ for $n > N$. Hence

$$\sup\{|E_x(e^{-\alpha T}[f(X_T) - g_n(X_T)])|;\ T \in \mathcal{T},\ x \in R^d\} < 2\varepsilon$$

for each $n > N$, and the sequence of continuous functions $R^\alpha g_n$ converges uniformly to $R^\alpha g$. ∎

APPENDIX

In this Appendix we prove Theorem 4.2, i.e., the continuity of the map $x \mapsto P_x$ for a Feller semigroup on a compact space E.

(a) The proof reduces to showing that $(P_x;\ x \in E)$ is tight. Indeed, let (P_{x_n}) converge to a probability P on Ω. Since E is compact, extract a subsequence (still denoted by (x_n)) such that x_n converges to x. By Proposition 4.1 (b) choose a countable dense $I \subset \mathbf{R}_+$ and a P-null N such that $\omega \mapsto \omega_t$ is continuous for $\omega \notin N$ and $t \in I$. Let

$$h = \prod_{1 \leqslant i \leqslant k} f_i(X_{t_i}) \in \mathscr{H}_c(I).$$

The choice of I ensures that $E_{x_n}(h) \to \int h\,dP$. On the other hand, for every $y \in E$,

$$E_y(h) = P_{t_1}\big(f_1 P_{t_2 - t_1}(f_2 \ldots (P_{t_k - t_{k-1}} f_k)) \ldots\big)(y).$$

Since (P_t) is Feller, the function $E_\cdot(h)$ is continuous, and $E_{x_n}(h) \to E_x(h)$. Hence both probabilities P and P_x coincide on $\mathscr{H}_c(I)$, and the monotone class theorem shows that they are equal on $\sigma(\mathscr{H}_c(I)) = \mathscr{F}_\infty$.

(b) To establish the tightness of a sequence $(P_{x_n};\ n \in N)$ of probabilities on Ω, we use the following criteria due to Aldous [1] (see also [16]). The following conditions are satisfied:

(i) The laws of ω_0 are tight on E.

(ii) For every $\varepsilon > 0$,

$$\lim_{h \to 0} \limsup_{n} \sup_{T \in \mathscr{T}} \sup_{0 \leqslant s \leqslant h} P_{x_n}\big(d(\omega_{T+s}, \omega_T) > \varepsilon\big) = 0,$$

where d denotes the Skorokhod metric on Ω.

Since E is compact, condition (i) is trivially satisfied. To prove that (ii) holds, we apply the strong Markov property; for every $y \in E$,

$$P_y\big(d(\omega_T, \omega_{T+s}) > \varepsilon\big) = E_y\big(P_{\omega_T}(d(\omega_s, \omega_0) > \varepsilon)\big) \leqslant \sup_{y \in E} P_y\big(d(\omega_s, \omega_0) > \varepsilon\big).$$

Hence the proof of (ii) reduces to showing that

$$\lim_{h \to 0} \sup_{x \in E} \sup_{0 \leqslant s \leqslant h} P_x\big(d(\omega_s, \omega_0) > \varepsilon\big) = 0.$$

This is Dynkin's stochastic continuity property [11], which is always satisfied by a Feller process on a compact set. We briefly sketch the proof. Set

$$q(u) = 1 - u/\varepsilon \text{ if } u \leqslant \varepsilon \quad \text{and} \quad q(u) = 0 \text{ if } u > \varepsilon.$$

The function $\phi_x(\cdot) = q(d(x, \cdot))$ is continuous, and $\|\phi_x - \phi_y\| \leqslant d(x, y)/\varepsilon$. Fix $0 < \alpha < \varepsilon^2$, and let $(x_i;\ 1 \leqslant i \leqslant k)$ be the centers of balls of radius α covering the compact set E. Given $x \in E$, let x_i be such that $d(x, x_i) < \alpha$. Then, for every s,

$$\begin{aligned} P_x(d(\omega_0, \omega_s) > \varepsilon) = P_x(d(x, \omega_s) > \varepsilon) &\leqslant \phi_x(x) - E_x(\phi_x(\omega_s)) \\ &\leqslant |\phi_x(x) - \phi_{x_i}(x)| + |\phi_{x_i}(x) - E_x(\phi_{x_i}(\omega_s))| + E_x(|\phi_{x_i} - \phi_x|(\omega_s)) \\ &\leqslant 2\alpha/\varepsilon + |\phi_{x_i}(x) - E_x(\phi_{x_i}(\omega_s))|. \end{aligned}$$

Since (P_t) is a Feller semigroup and since ϕ_{x_i} is continuous, we have

$$\lim_{h \to 0} \sup_{0 \leqslant s \leqslant h} \sup_x |\phi_{x_i}(x) - P_s \phi_{x_i}(x)| = 0. \blacksquare$$

Remark. Theorem 4.2 is also true for Markov processes which are Feller in the following weaker sense: $U^\alpha(C(E)) \subset C(E)$ for each $\alpha > 0$, and $\alpha U^\alpha f \to f$ pointwise as $\alpha \to \infty$ for each $f \in C(E)$. Indeed, in part (a) of the proof, replace the class of r.v.'s $\mathscr{H}_c(I)$ by the following class $\mathscr{H}_u$ made of random variables h:

$$h = I_k(\lambda_1, f_1; \ldots; \lambda_k, f_k) = \prod_{1 \leqslant i \leqslant k} \int_{[0, +\infty]} e^{-\lambda_i t} f_i(X_t) dt$$

for $k \geqslant 1$, $\lambda_i \in Q$ and f_i in a dense subset of $C(E)$.

The characterization of potentials given in [14], p. 38, shows that $E_x(h)$ is the $\lambda = \sum \lambda_i$ potential of the function $P_0 g$, with

$$g(x) = \sum_{1 \leqslant i \leqslant k} f_i E_x(I_{k-1}(\lambda_1, f_1; \ldots; \hat{\lambda}_i, \hat{f}_i; \ldots; \lambda_k, f_k)),$$

in wich ^ indicates the quantity which has been omitted. Since $U^\lambda P_0 = U^\lambda$, an easy induction argument together with the weaker Feller property shows that $E_x(h)$ is the potential of a continuous function g on E. The continuity of potentials shows that

$$E_{x_n}(h) = U^\lambda g(x_n) \to U^\lambda g(x) = E_x(h).$$

The proof of (b) carries over without change. Indeed, the last convergence property can be checked on potentials which are dense in $C(E)$ for the uniform topology.

REFERENCES

[1] D. Aldous, *Stopping times and tightness*, Ann. Probab. 6 (1978), pp. 335–340.
[2] J. P. Aubin and A. Celina, *Differential Inclusions*, Springer-Verlag, 1984.
[3] J. R. Baxter and R. V. Chacon, *Compactness of stopping times*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 40 (1981), pp. 169–181.
[4] P. Billingsley, *Convergence of Probability Measures*, Wiley, 1969.

[5] J. M. Bismut, *Dualité convexe, temps d'arrêt optimal et contrôle stochastique*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 38 (1977), pp. 169–198.
[6] — *Temps d'arrêt optimal, quasi-temps d'arrêt et retournement du temps*, Ann. Probab. 7 (1979), pp. 933–964.
[7] — *Contrôle de processus alternants et applications*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 47 (1979), pp. 241–288.
[8] — et B. Skalli, *Temps d'arrêt optimal, théorie générale des processus et processus de Markov*, ibidem 39 (1977), pp. 301–313.
[9] R. M. Blumental and R. K. Getoor, *Markov Processes and Potential Theory*, Academic Press, 1968.
[10] C. Dellacherie et P. A. Meyer, *Probabilité et potentiel*, Hermann, t. 1, 1975; t. 2, 1980; t. 3, 1984.
[11] E. B. Dynkin, *Markov Processes*, Spriner-Verlag, 1965.
[12] N. El Karoui, *Les aspects probabilistes du contrôle stochastique*, Ecole d'Eté de Probabilités de Saint Flour IX, Lecture Notes in Math. 876 (1981), pp. 74–239.
[13] — D. Huu Nguyen and M. Jeanblanc-Picque, *Compactification methods in the control of degenerate diffusions: existence of an optimal control*, Stochastics 20 (1987), pp. 169–219.
[14] R. Getoor, *Markov processes: Ray processes and right processes*, Lecture Notes in Math. 440 (1975).
[15] J. Jacod et J. Memin, *Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité*, Séminaire de Probabilités XV, ibidem 851 (1981), pp. 529–546.
[16] A. Joffe and M. Metivier, *Weak convergence of semimartingales with applications to multitype branching processes*, Adv. in Appl. Probab. 18 (1986), pp. 20–65.
[17] J. P. Lepeltier et B. Marchal, *Problème de martingale et équations différentielles stochastiques associés à un opérateur intégrodifférentiel*, Ann. Inst. H. Poincaré 12 (1976), pp. 43–103.
[18] V. Mackevicius, *Passing to the limit in the optimal stopping of Markov processes*, Lietuvos Mathamatikos Rinkinys 13 (1973), pp. 115–118.
[19] M. A. Mainguèneau, *Temps d'arrêt optimaux et théorie générale*, Séminaire de Probabilités XII, Lecture Notes in Math. 649 (1978), pp. 457–468.
[20] G. Mazziotto et J. Spzirglas, *Théorème de séparation dans le problème d'arrêt optimal*, Séminaire de Probabilités XIII, ibidem 721 (1979), pp. 378–384.
[21] J. F. Mertens, *Théorie des processus stochastiques généraux, application aux surmartingales*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 22 (1972), pp. 45–68.
[22] — *Strongly supermedian functions and optimal stopping*, ibidem 26 (1973), pp. 119–139.
[23] P. A. Meyer, *Réduite et jeux de hasard*, Séminaire de Probabilités VII, Lecture Notes in Math. 321 (1973), pp. 155–172.
[24] — *Convergence faible et compacité des temps d'arrêt d'après Baxter et Chacon*, Séminaire de Probabilités XII, ibidem 649 (1978), pp. 411–423.
[25] P. Priouret, *Processus de diffusion et équations différentielles stochastiques*, Ecole d'Eté de Probabilités de Saint-Flour III, ibidem 390 (1974), pp. 38–113.
[26] M. Robin, Thèse d'Etat Paris Dauphine, 1978.
[27] H. Rost, *Markov Ketten bei sich füllenden Löchern in Zustandraum*, Ann. Inst. Fourier 21 (1971).
[28] M. Sharpe, *General Theory of Markov Processes*, Academic Press, 1988.
[29] A. N. Shiryayev, *Optimal Stopping Rules*, Springer-Verlag, 1976.

[30] D. Stroock, *Diffusion processes associated with Lévy generators*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete 32 (1975), pp. 209–244.
[31] J. Zabczyk, *Stopping Problems in Stochastic Control*, Proceedings ICM-88, Vol. II, PWN–North-Holland 1984, pp. 1425–1437.

Nicole El Karoui
Laboratoire de Probabilités
Tour 56, 3ème etage
Université Paris VI
4, place Jussieu
75230 Paris Cedex O5

Jean-Pierre Lepeltier
Université du Maine
Faculté des Sciences
Route de Laval
72017 Le Mans Cedex

Annie Millet
Université Paris X
200, avenue de la République
92001 Nanterre Cedex
and
Laboratoire de Probabilités
URA 224

Received on 12.10.1989;
revised version on 20.8.1990