

ON SEQUENTIALLY CONTROLLED MARKOV PROCESSES

To cite this article: A K Zvonkin 1971 *Math. USSR Sb.* **15** 607

View the [article online](#) for updates and enhancements.

Related content

- [ON THE PAPER "MARKOV RANDOM FIELDS AND STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS"](#)
Ju A Rozanov
- [MARKOV RANDOM FIELDS AND STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS](#)
Ju A Rozanov
- [Switched Fault Diagnosis Approach for Industrial Processes based on Hidden Markov Model](#)
Lin Wang, Chunjie Yang, Youxian Sun et al.

Recent citations

- [A probabilistic verification theorem for the finite horizon two-player zero-sum optimal switching game in continuous time](#)
S. Hamadène *et al*
- [Controlled random sequences and Markov chains](#)
A A Yushkevich and R Ya Chitashvili
- [A TRANSFORMATION OF THE PHASE SPACE OF A DIFFUSION PROCESS THAT REMOVES THE DRIFT](#)
A K Zvonkin

ON SEQUENTIALLY CONTROLLED MARKOV PROCESSES

UDC 519.2

A. K. ZVONKIN

Abstract. We consider Markov processes with continuous time, where the switching of the controls takes place at random (independent of the future) moments of time. We derive Bellman's cost equation and the existence of (p, ϵ) optimal strategies, prove the measurability of cost and give an excessive characterization of cost.

Bibliography: 9 items.

§1. Introduction

On a probability space (Ω, \mathcal{F}, P) there is given a family of unstopped Markov processes $(1) X^d = \{x_t^d, \mathbb{M}_t^d, P_x^d\}$ with values in a common phase space (E, \mathcal{B}) and depending on some parameter $d \in D$ (the elements of the set D will be called controls). Being at the moment $t = 0$ in the state $x_0 \in E$, we choose a control $d_0 \in D$, and "insert" the process X^{d_0} . At some random moment τ_1 independent of the future we instantly "switch" to a control d_1 , and at this moment (and from the last state, i.e. x_{τ_1}) go to the process X^{d_1} up to the moment τ_2 , at which time we "switch" to the control d_2 , and so on. Our purpose is to maximize the value of $M_x \sum_{k=0}^{\infty} (y^{d_k})_{\tau_{k+1}}^{\tau_k}$, where y^d is an additive functional of the process X^d .

The problem, in such a formulation, on the one hand is a natural generalization of the problem of optimal stopping (see [2]); on the other hand it allows one to carry over a series of results from [3], [4] and [5] to the case of a process with continuous time.

In this article we derive Bellman's cost equation, prove the measurability of cost and the existence of (p, ϵ) -optimal strategies, and give an excessive characterization of cost.

§2. Assumptions

1.A. E is semicompact and \mathcal{B} is a σ -algebra of Borel sets of E .

1.B. D is a Borel subset of a complete separable metric space.

With respect to the processes X^d we shall assume the following:

AMS 1970 subject classifications. Primary 60G40, 62L15, 49C15, 60J25, 93E20; Secondary 60J60.

(1) We shall systematically make use of the definitions and notation of the monograph [1].

2.A. For each $d \in D$, the process X^d is standard.

2.B. For each $t \geq 0$, the σ -algebras \mathfrak{M}_t^d coincide for all d (in what follows, in place of \mathfrak{M}_t^d we shall write \mathfrak{M}_t), and for any $x \in E$, $d_1, d_2 \in D$, the measures $P_x^{d_1}$ and $P_x^{d_2}$ on (Ω, \mathfrak{M}_t) are mutually absolutely continuous.

2.C. For any $t \geq 0$ and $\Gamma \in \mathfrak{B}$, the function $P(t, d, x, \Gamma) = P_x^d\{x_t \in \Gamma\}$ is a $\mathfrak{B} \times \sigma(D)$ -measurable ⁽²⁾ function of (x, d) , and for each $d \in D$ the function $P(t, d, x, \Gamma)$ is a \hat{C} -function.

We shall define a *strategy* to be a sequence $\pi = (\delta_0, \tau_1, \delta_1, \tau_2, \delta_2, \dots)$ satisfying the following conditions.

3.A. $\tau_k = \tau_k(\omega; d_0, \dots, d_{k-1})$ is a function with values in $[0, \infty]$ such that for all $t \geq 0$

$$\{(\omega; d_0, \dots, d_{k-1}) : \tau_k(\omega; d_0, \dots, d_{k-1}) \leq t\} \in \mathcal{M}_t \times \sigma(D^k) \quad (3)$$

(we shall call such a function a *moment independent of the future*).

3.B. $\delta_k = \delta_k(\cdot | \cdot)$ is a (regular) conditional distribution on $(D, \sigma(D))$ for given $(\omega; d_0, \dots, d_{k-1})$ measurable with respect to the σ -algebra $\mathfrak{M}_{\tau_k} \times \sigma(D^k)$ in the space $\Omega \times D^k$.

3.C. With probability 1, we have $0 \leq \tau_1 \leq \tau_2 \leq \dots \leq \infty$, and $\lim_{n \rightarrow \infty} \tau_n = \infty$.

As a price function we get the functional

$$r(\pi, \omega) = \sum_{k=0}^{\infty} (\gamma^d)_{\tau_{k+1}}^{\tau_k}, \quad (1)$$

where π is a usable strategy, and γ^d is a numerical functional satisfying the following conditions.

4.A. For each d , the functional γ^d is a homogeneous perfect W -functional.

4.B. For any $0 \leq s \leq t$ and any $\Gamma \in \sigma[0, \infty]$ ⁽⁴⁾ we have

$$\{(\omega, d) : (\gamma^d)_t^s \in \Gamma\} \in \mathcal{N}_t \times \sigma(D).$$

Let p be a probability measure on E , and let $\epsilon \geq 0$. We call a strategy $\pi^*(p, \epsilon)$ -optimal if

$$p\{x : v_{\pi^*}(x) \geq \sup_{\pi} v_{\pi}(x) - \epsilon\} = 1,$$

where $v_{\pi}(x) = M_x^{\pi}(r, \omega)$ is our mean ⁽⁵⁾ (expected) price using the strategy π . Such a definition of optimality makes sense since the function $v(x) = \sup_{\pi} v_{\pi}(x)$ is \mathfrak{B} -measurable (see the Corollary to Lemma 4). We shall call $v(x)$ a *cost* and shall assume, a priori, the condition

5.A. $v(x) < \infty$ for all $x \in E$.

⁽²⁾ Here $\sigma(D)$ is the σ -algebra of Borel sets of D .

⁽³⁾ Here $D^k = D \times D \times \dots \times D$ (k factors).

⁽⁴⁾ Here $\sigma[0, \infty]$ is the σ -algebra of Borel sets on the halfline.

⁽⁵⁾ If the process starts at the point x and we apply the strategy π , then by the same token there is generated in Ω some measure (the process corresponding to this measure of course will not in general be Markov). The mathematical expectation with respect to this measure will be denoted by M_x^{π} .

§ 3. Auxiliary results

The results set out in this section are basically of a technical nature.

I. Consider the process $\tilde{X}_n^d = \{\tilde{x}_t, \tilde{\mathbb{M}}_t, \tilde{P}_x^d\}$ obtained from the process $X^d = \{x_t, \mathbb{M}_t, P_x^d\}$ in the following way:

$$\begin{aligned} \tilde{\Omega} &= \Omega \times D^n, & \tilde{E} &= E, & \tilde{x}_t(\tilde{\omega}) &= \tilde{x}_t(\omega; d_1, \dots, d_n) = x_t(\omega), \\ \tilde{\mathbb{M}}_t &= \mathbb{M}_t \times \sigma(D^n) \text{ (the completion is carried out in the system of all measures } \tilde{P}_x^d, x \in E), \\ \tilde{P}_x^d &= P_x^d \times \nu \text{ (}\nu \text{ is an arbitrary but fixed probability measure on } (D^n, \sigma(D^n)) \text{).} \end{aligned}$$

Obviously all the processes $\tilde{X}_n^d, d \in D, n \geq 0$, are Markov.

- Lemma 1. a) For any $d \in D$ and $n \geq 0$, the process \tilde{X}_n^d is standard.
- b) The classes of functions $f(x)$ such that $\lim_{t \downarrow 0} f(x_t) \geq f(x)$ \tilde{P}_x^d -a.s. coincide for all processes $\tilde{X}_n^d, d \in D, n \geq 0$.
- c) For each $d \in D$, the classes of excessive functions for the processes $X_n^d, n \geq 0$, coincide.

Proof. Note that the transition function of the process $\tilde{X}_n^d, n \geq 0$, does not change with the process X^d . Hence assertion c) follows immediately; making use of Theorem 3.13 and Lemma 3.2 of [1], we obtain assertion a); using condition 2.B, we obtain assertion b).

The construction of our strategies is such that it is natural to think as though between the moments τ_k and τ_{k+1} we observe, not the process X^{d_k} , but the process $\tilde{X}_{k+1}^{d_k}, k = 0, 1, 2, \dots$ (then, in particular, all the moments independent of the future become simple Markov moments). In order not to encumber the construction excessively, we shall, as before, denote these processes by X^{d_k} , and omit almost all references to Lemma 1.

II. At some essential points in our reasoning, the technique of analytic sets (\mathcal{A} -sets) will be used. We cite some necessary facts. An \mathcal{A} -set is a continuous image (for example, a projection) of a Borel set.

- a) Any \mathcal{A} -set is \mathcal{B} -measurable ([6], p. 601); there exist \mathcal{A} -sets which are not Borel ([7], §38, IV).
- b) *Theorem on uniformization:* Let G be an \mathcal{A} -set in a product $X \times Y, H$ the projection of G on X ; then there exists a $\sigma(X)$ -measurable function $f(x)$ such that $(x, f(x)) \in G$ for all $x \in H$ ([6], p. 603).
- c) \mathcal{A} -sets are invariant with respect to the Suslin operations ([7], Russian p. 37, and Theorem 4, §33, IX), in particular with respect to the operations of countable union and intersection (note that a set analytic together with its complement is Borel; see [6], p. 600), and also with respect to the operation of Cartesian product ([7], Theorem 1, §38, III).

III. Denote by \mathcal{L} the class of functions $f(x, d)$ such that

- (α) $f(x, d) \geq 0$,
- (β) $f(x, d)$ is an \mathcal{A} -function, i.e., for any $c \{ (x, d) : f(x, d) > c \}$ is an \mathcal{A} -set (then $\{ (x, d) : f(x, d) \geq c \} = \bigcap_{n=1}^{\infty} \{ (x, d) : f(x, d) > c - 1/n \}$ is an \mathcal{A} -set; see II.c)).

(γ) For all $d \in D$, $\lim_{t \downarrow 0} f(x_t, d) \geq f(x, d)$ P_x^d -a.s. (then, if τ is a Markov moment, $\lim_{t \downarrow 0} f(x_{\tau+t}, d) \geq f(x_\tau, d)$ P_x^d -a.s. for all d and x).

Lemma 2. a) If $f_1 \in \mathcal{L}$, $f_2 \in \mathcal{L}$, then $f_1 + f_2 \in \mathcal{L}$.

b) If $f_n \in \mathcal{L}$ and $f_n \uparrow f$, then $f \in \mathcal{L}$.

c) (⁶) If $u \in \mathcal{L}$ and $v(x) = \sup_{d \in D} u(x, d)$, then $v \in \mathcal{L}$ (more precisely, the function $w(x, d) \equiv v(x)$ belongs to the class \mathcal{L}).

Proof. a) That conditions (α) and (γ) are satisfied for the sum is obvious. Condition (β) follows from II.c) and the relation

$$\{(x, d) : f_1(x, d) + f_2(x, d) > c\} = \bigcup_r (\{(x, d) : f_1(x, d) > r\} \cap \{(x, d) : f_2(x, d) > c-r\}),$$

where r runs through the set of all rational numbers. Note that $f_1 - f_2 \in \mathcal{L}$ in the case where $f_1 \geq f_2$, the function f_2 is Borel, and $\lim_{t \downarrow 0} f_2(x_t, d) = f_2(x, d)$ P_x^d -a.s.

$$\text{b) } (\beta) : \{(x, d) : f(x, d) > c\} = \bigcup_{n=1}^{\infty} \{(x, d) : f_n(x, d) > c\},$$

$$(\gamma) : \lim_{t \downarrow 0} f(x_t, d) \geq \lim_{t \downarrow 0} f_n(x_t, d) \geq f_n(x, d)$$

P_x^d -a.s. for any $n \geq 1$; therefore $\lim_{t \downarrow 0} f(x_t, d) \geq f(x, d)$ P_x^d -a.s. also.

c) (β): the set $C = \{x : v(x) > c\}$ is the projection on E of the set $\{(x, d) : u(x, d) > c\}$; consequently C (and with it $C \times D = \{(x, d) : u(x, d) > c\}$) is an \mathcal{A} -set also; (γ) is verified in the same way as for assertion b), using Lemma 1.b).

IV. We shall call the function $f(x, d)$ *excessive* if it is $\mathcal{B} \times \sigma(D)$ -measurable and for each d it is excessive for the process X^d as a function of x . The operator which puts in correspondence to the function $g(x, d)$ its least excessive majorant will be denoted by Λ . The following theorem is a fundamental result in the theory of optimal stopping (see [2]).(⁷)

Theorem 1. Let $g(x, d) \in \mathcal{L}$ and

$$s(x, d) = \lim_{n \rightarrow \infty} \lim_{N \rightarrow \infty} Q_n^N g(x, d), \quad (2)$$

where Q_n^N is the N th power of the operator

$$Q_n g(x, d) = \max \{g(x, d), T_{2^{-n}}^d g(x, d)\}.$$

Then the following assertions are true.

a) $s(x, d) = \Lambda g(x, d)$,

b) $s(x, d) = \sup_{\tau} M_{x_\tau}^d g(x_\tau, d)$, where τ is a Markov moment.

Proof. We shall reason for each d separately.

(⁶) Namely, this requirement on the class \mathcal{L} forces us to include in it not only Borel functions but also \mathcal{A} -functions. If we did not make use of the technique of \mathcal{A} -sets, we would be able to prove all the results for countable sets D only.

(⁷) In [2] this theorem is proved with the additional assumption that the function $g(x, d)$ is almost Borel in x , so that in order for us to drop this assumption, we need to deal with some insignificant technical difficulties.

a) If $f(x) = T_s g(x)$, then according to Fatou's Lemma and property (y) we have for $g(x)$

$$\lim_{t \downarrow 0} T_t f(x) = \lim_{t \downarrow 0} M_x g(x_{s+t}) \geq M_x \lim_{t \downarrow 0} g(x_{s+t}) \geq M_x g(x_s) = f(x). \tag{3}$$

Inequality (3) remains true for the maximum of two functions and for monotone (increasing) passage to the limit, so that

$$\lim_{t \downarrow 0} T_t s(x) \geq s(x). \tag{4}$$

As was shown in the proof of Lemma III.1 of [2],

$$T_{m \cdot 2^{-n}} s(x) \leq s(x). \tag{5}$$

Take a sequence of dyadic rational numbers $r_i \downarrow t$, $i \rightarrow \infty$, and put $s_i = r_i - t$. Using sequence (5), Fatou's Lemma, and (4), we obtain

$$s(x) \geq \lim_{i \rightarrow \infty} M_x s(x_{r_i}) = \lim_{i \rightarrow \infty} M_x M_{x_i} s(x_{t+s_i}) \geq M_x \lim_{i \rightarrow \infty} M_{x_i} s(x_{t+s_i}) \geq M_x s(x_t), \tag{6}$$

i.e. $T_t s(x) \leq s(x)$. For $s(x)$ to be excessive, we must have $\lim_{t \downarrow 0} T_t s(x) = s(x)$, which follows from (4) and (6). Then the fact that $s(x)$ is the least of the excessive majorants is proved as in Lemma III.1 of [2].

b) The proof coincides with the proof of Theorem III.1 of [2]

V. Lemma 3. If $g(x, d) \in \mathcal{L}$, then $s(x, d) = \Lambda g(x, d) \in \mathcal{L}$.

Proof. Condition (y) is satisfied for $s(x, d)$ by virtue of its excessiveness (see [1], Theorem 12.4 and 4.9). In order to establish that condition (β) is satisfied for $s(x, d)$, it is sufficient, according to (2) and Lemma 2b), to show that the function $f(x, d) = T_t^d g(x, d)$ satisfies condition (β).

Let Z be a Borel subset of a complete separable metric space, and $P(Z)$ the set of all probability measures on Z . We determine in $P(Z)$ a σ -algebra of Borel sets, namely the minimal σ -algebra with respect to which the functions $f_B(\mu) = \mu(B)$ are measurable for all $B \in \sigma(Z)$ (see [9]). Let $\phi(z)$ be an \mathcal{U} -function (i.e. $\{z : \phi(z) > c\}$ is an \mathcal{U} -set for all c), let $\phi(z) \geq 0$ and let $\mu : Z \rightarrow P(Z)$ be a Borel function. We shall show that $f(z) = \int_Z \phi(y) \mu_z(dy)$ is an \mathcal{U} -function. It is sufficient to establish this fact for the function $\phi(z) = \chi_\Gamma(z)$, where Γ is an \mathcal{U} -set⁽⁸⁾, i.e. to show that the set $H = \{z : \mu_z(\Gamma) > c\}$ is an \mathcal{U} -set for any c . But H is the projection on Z of the set

$$\{(z, \mu_z)\} \cap (Z \times \{\mu : \mu(\Gamma) > c\})$$

in the space $Z \times P(Z)$, and the set $\{(z, \mu_z)\}$ is Borel since it is the graph of a Borel function (see [7], Russian p. 505); therefore, according to II.c), it is sufficient to prove that $\{\mu : \mu(\Gamma) > c\}$ is an \mathcal{U} -set in $P(Z)$.

As is seen from the proof of Theorem IV of Supplement III in [6] (pp. 601–602), there exists a family of compact sets $C_{n_1 \dots n_m} \subseteq Z$ (n_k is a natural number) such that for any measure μ , $\mu(\Gamma) > c$, there can be found a sequence p_1, \dots, p_m, \dots such that $\mu(C_{p_1 \dots p_k}) > c$ for all its initial segments p_1, \dots, p_k ($k \geq 1$). By the same token

(8) In fact, let $\phi_n(z) = n^{-1} \sum_{k=1}^{\infty} \chi_{\{\phi(z) > k/n\}}(z)$; then $\phi_n(z) \uparrow \phi(z)$ as $n \rightarrow \infty$.

$$\{\mu : \mu(\Gamma) > c\} = \bigcup_{p_1 \dots p_k \dots} \bigcap_{k=1}^{\infty} \{\mu : \mu(C_{p_1 \dots p_k}) > c\},$$

i.e. the set $\{\mu : \mu(\Gamma) > c\}$ is obtained from the Borel (by definition) set $\{\mu : \mu(C_{p_1 \dots p_k}) > c\}$ by means of an \mathcal{G} -operation; consequently it is an \mathcal{G} -set.

The assertion of the lemma follows from the fact that the function $\mu : E \times D \rightarrow P(E \times D)$ is such that if $B_1 \subseteq E$ and $B_2 \subseteq D$, then

$$\mu_{(x,d)}\{B_1 \times B_2\} = P_x^d\{x_i \in B_1\} \cdot \chi_{B_2}$$

is a Borel function (see condition 2.C).

§4. Fundamental results

We consider the classes Π_N of strategies which can be switched not more than N times (in other words, $\tau_{N+1} = \infty$ with probability 1), $N = 0, 1, 2, \dots$. Let $v_N(x) = \sup_{\pi \in \Pi_N} v_{\pi}(x)$.

Theorem 2. *The functions $v_N(x)$ satisfy the following recurrence relations:*

$$v_N(x) = \sup_{d \in D} [\Lambda(v_{N-1}(x) - s_0(x, d)) + s_0(x, d)], \quad v_0(x) = \sup_{d \in D} s_0(x, d), \quad (7)$$

where $s_0(x, d) = M_x^d(\gamma^d)_{\infty}^0$.

Proof. Put $g_n(x, d) = v_n(x) - s_0(x, d)$ and $s_n(x, d) = \Lambda g_{n-1}(x, d)$. Fix an initial point x_0 . According to the usual scheme of dynamic programming, the proof consists of two parts: first show that if $v_N(x)$ is the solution of equation (7), then

$$v_{\pi}(x_0) \leq v_N(x_0) \quad \text{for any } \pi \in \Pi_N, \quad (8)$$

and then construct a strategy $\pi^* \in \Pi_N$ such that

$$v_{\pi^*}(x_0) \geq v_N(x_0) - \varepsilon. \quad (9)$$

I. If $\pi \in \Pi_N$, then

$$v_{\pi}(x_0) = M_{x_0}^{\pi}(\gamma^{d_0})_{\tau_1}^0 + (\gamma^{d_1})_{\tau_2}^{\tau_1} + \dots + (\gamma^{d_{N-1}})_{\tau_N}^{\tau_{N-1}} + (\gamma^{d_N})_{\infty}^{\tau_N}.$$

According to Theorem 6.12 of [1],

$$\begin{aligned} M_{x_0}^{\pi}(\gamma^{d_k})_{\tau_{k+1}}^{\tau_k} &= M_{x_0}^{\pi}[(\gamma^{d_k})_{\tau_{k+1}}^0 - (\gamma^{d_k})_{\tau_k}^0] \\ &= M_{x_0}^{\pi} s_0(x_0, d_k) - M_{x_0}^{\pi} s_0(x_{\tau_{k+1}}, d_k) - M_{x_0}^{\pi} s_0(x_0, d_k) + M_{x_0}^{\pi} s_0(x_{\tau_k}, d_k) \\ &= M_{x_0}^{\pi} [s_0(x_{\tau_k}, d_k) - s_0(x_{\tau_{k+1}}, d_k)], \end{aligned}$$

and analogously

$$\begin{aligned} M_{x_0}^{\pi}(\gamma^{d_0})_{\tau_1}^0 &= M_{x_0}^{\pi} [s_0(x_0, d_0) - s_0(x_{\tau_1}, d_0)], \\ M_{x_0}^{\pi}(\gamma^{d_N})_{\infty}^{\tau_N} &= M_{x_0}^{\pi} s_0(x_{\tau_N}, d_N). \end{aligned}$$

Therefore

$$\begin{aligned} v_{\pi}(x_0) &= M_{x_0}^{\pi} \{s_0(x_0, d_0) + [s_0(x_{\tau_1}, d_1) - s_0(x_{\tau_1}, d_0)] + \dots \\ &\quad \dots + [s_0(x_{\tau_N}, d_N) - s_0(x_{\tau_N}, d_{N-1})]\} \\ &= M_{x_0}^{\pi} \{s_0(x_0, d_0) + M_{x_0}^{\pi} \{[s_0(x_{\tau_1}, d_1) - s_0(x_{\tau_1}, d_0)] + \dots \\ &\quad + M_{x_0}^{\pi} \{[s_0(x_{\tau_N}, d_N) - s_0(x_{\tau_N}, d_{N-1})] | \tilde{\mathcal{M}}_{\tau_{N-1}}\} | \tilde{\mathcal{M}}_{\tau_{N-2}}\} | \dots | \tilde{\mathcal{M}}_0\}\}. \end{aligned}$$

The function $s_0(x, d)$ is Borel by virtue of conditions 2.C and 4.B, and $\lim_{t \downarrow 0} s_0(x, d) = s_0(x, d)$ P_x^d -a.s. by virtue of the excessiveness of $s_0(x, d)$ (see [1], §6.17 and Theorems 12.4 and 4.9). Therefore, according to Lemmas 2 and 3, all the functions $v_n(x)$, $g_n(x, d)$ and $s_n(x, d)$ belong to the class \mathcal{L} , and the inequality

$$\begin{aligned} s_0(x_{\tau_N}, d_N) &\leq v_0(x_{\tau_N}), \\ v_0(x_{\tau_N}) - s_0(x_{\tau_N}, d_{N-1}) &= g_0(x_{\tau_N}, d_{N-1}) \leq s_1(x_{\tau_N}, d_{N-1}), \\ M_{x_0}^\pi \{s_1(x_{\tau_N}, d_{N-1}) \mid \tilde{\mathcal{M}}_{\tau_{N-1}}\} &\leq s_1(x_{\tau_{N-1}}, d_{N-1}) \quad P_{x_0}^\pi\text{-a.s.} \end{aligned}$$

holds (this inequality follows from Theorem 1.5 and property V on p. 119 of [2]);

$$\begin{aligned} s_0(x_{\tau_{N-1}}, d_{N-1}) + s_1(x_{\tau_{N-1}}, d_{N-1}) &\leq v_1(x_{\tau_{N-1}}), \\ v_1(x_{\tau_{N-1}}) - s_0(x_{\tau_{N-1}}, d_{N-2}) &= g_1(x_{\tau_{N-1}}, d_{N-2}) \leq s_2(x_{\tau_{N-1}}, d_{N-2}), \\ M_{x_0}^\pi \{s_2(x_{\tau_{N-1}}, d_{N-2}) \mid \tilde{\mathcal{M}}_{\tau_{N-2}}\} &\leq s_2(x_{\tau_{N-2}}, d_{N-2}) \quad P_{x_0}^\pi\text{-a.s.}, \end{aligned}$$

and so on. In sum we get inequality (8).

II. We construct the strategy π^* . Let $\epsilon > 0$ be given; then proceed as follows.

d_0) Choose d_0 so that

$$s_N(x_0, d_0) + s_0(x_0, d_0) \geq v_N(x_0) - \frac{\epsilon}{2N+1};$$

such a d_0 exists since $v_N(x) = \sup_{d \in D} [s_N(x, d) + s_0(x, d)]$.

τ_1) Choose $\tau_1 \geq 0$ so that

$$M_{x_0}^{d_0} g_{N-1}(x_{\tau_1}, d_0) \geq s_N(x_0, d_0) - \frac{\epsilon}{2N+1};$$

such a τ_1 exists according to Theorem 1.b).

d_1) Let p_1 be a measure on E corresponding to the distribution of the random variable x_{τ_1} . We change the function $v_{n-1}(x)$ on a set of p_1 -measure 0 so that it becomes Borel; then the set

$$\left\{ (x, d) : s_{N-1}(x, d) + s_0(x, d) \geq v_{N-1}(x) - \frac{\epsilon}{2N+1} \right\}$$

becomes an \mathcal{L} -set (see Lemma 2.a)). This means we can uniformize it (see §3, II.b)), i.e. there exists a \mathcal{B} -measurable function $d_1 = d_1(x)$ (it is even possible to make it Borel, changing it on a set of p_1 -measure 0) such that

$$p_1 \left\{ x : s_{N-1}(x, d_1(x)) + s_0(x, d_1(x)) \geq v_{N-1}(x) - \frac{\epsilon}{2N+1} \right\} = 1$$

.....

τ_N) Analogously to τ_1 , we choose $\tau_N \geq \tau_{N-1}$ such that

$$M_{x_{\tau_{N-1}}}^{d_{N-1}} g_0(x_{\tau_N}, d_{N-1}) \geq s_1(x_{\tau_{N-1}}, d_{N-1}) - \frac{\epsilon}{2N+1}.$$

d_N) Analogously to d_1 , we choose $d_N = d_N(x_{\tau_N})$ so that

$$p_N \left\{ x : s_0(x, d_N(x)) \geq v_0(x) - \frac{\epsilon}{2N+1} \right\} = 1,$$

where p_N is a measure on E corresponding to the distribution of the random variable x_{τ_N} .

The strategy constructed obviously satisfies inequality (9). The theorem is proved.

Lemma 4. *If $v(x)$ is a cost function, then $v_n(x) \uparrow v(x)$ as $n \rightarrow \infty$.*

The proof follows from condition 5.A and (1) and from the nonnegativity of the functional γ^d .

Corollary. *The function $v(x)$ is $\bar{\mathcal{B}}$ -measurable (and, moreover, $v(x) \in \mathcal{L}$; see Lemma 2.b)).*

Theorem 3. *The cost $v(x)$ is the least solution of the equation (9)*

$$v(x) = \sup_{d \in D} [\Lambda(v(x) - s_0(x, d)) + s_0(x, d)]. \quad (10)$$

Proof. 1) The left-hand side of equation (7) converges to the function $\sup_{d \in D} [s(x, d) + s_0(x, d)]$, where

$$s(x, d) = \lim_{n \rightarrow \infty} s_n(x, d) = \lim_{n \rightarrow \infty} \Lambda g_n(x, d).$$

We must show that $s(x, d) = \Lambda(\lim_{n \rightarrow \infty} g_n(x, d))$. The equality

$$\lim_{n \rightarrow \infty} \Lambda g_n(x, d) = \Lambda(\lim_{n \rightarrow \infty} g_n(x, d))$$

(the sequence $g_n(x, d)$ is nondecreasing) is proved exactly as in [2] (Remark 2 on p. 51).

2) Let $u(x)$ be any solution of equation (10). Then

$$u(x) \geq \sup_{d \in D} [0 + s_0(x, d)] = v_0(x).$$

If $v(x) \geq v_k(x)$, then

$$u(x) = \sup_{d \in D} [\Lambda(u(x) - s_0(x, d)) + s_0(x, d)] \geq \sup_{d \in D} [\Lambda(v_k(x) - s_0(x, d)) + s_0(x, d)] = v_{k+1}(x).$$

By the same token, it is shown by induction that $u(x) \geq v_n(x)$ for all n : consequently $u(x) \geq v(x)$.

Theorem 4. *$v(x)$ is the least function such that the function $g(x, d) = v(x) - s_0(x, d)$ is excessive (i.e. $\Lambda g(x, d) = g(x, d)$).*

Proof. 1) If $\Lambda(v(x_0) - s_0(x_0, d_0)) > v(x_0) - s_0(x_0, d_0)$ for some pair (x_0, d_0) , then

$$\begin{aligned} v(x_0) &= \sup_{d \in D} [\Lambda(v(x_0) - s_0(x_0, d)) + s_0(x_0, d)] \\ &\geq \Lambda(v(x_0) - s_0(x_0, d_0)) + s_0(x_0, d_0) > v(x_0) - s_0(x_0, d_0) + s_0(x_0, d_0) = v(x_0). \end{aligned}$$

2) If the function $u(x) - s_0(x, d)$ is excessive, then

$$\sup [\Lambda(u(x) - s_0(x, d)) + s_0(x, d)] = \sup [u(x) - s_0(x, d) + s_0(x, d)] = u(x),$$

and, according to Theorem 3, $u(x) \geq v(x)$.

Theorem 5. *For any probability measure p on E and for any $\epsilon > 0$, there exists a (p, ϵ) -optimal strategy.*

⁽⁹⁾ The question of $\bar{\mathcal{B}}$ -measurability of a cost function and the validity of equation (10) for random discrete variables was posed by David Blackwell [3] and solved by R. E. Strauch [4].

Proof. We change all the functions $v_n(x)$, $n = 0, 1, 2, \dots$, and $v(x)$ on a set of p -measure 0 so that they become Borel, and separate E into sets $E_n = \{x : v_n(x) > v(x) - \epsilon/2\}$. Let $G_0 = E_0$ and $G_n = E_n \setminus E_{n-1}$. Choose $d_0 = d_0(x)$ to be a Borel function of x such that

$$s_n(x, d_0(x)) + s_0(x, d_0(x)) > v_n(x) - \frac{\epsilon}{2(2n + 1)} \quad p\text{-a.s. on } G_n,$$

and proceed as is described in the second part of the proof of Theorem 2, choosing an $\epsilon/2$ -optimal strategy in class Π_n if we go out of the set G_n .

Remark. According to our formulation of the problem, in controlling the process we have the right to use the information on all the preceding history of the process. However, as is seen from Theorems 2 and 5, it is automatically sufficient for (p, ϵ) -optimal control to have the information on where we got out of (x_0) from, what process we observe at the present moment (d_k), the number of transitions (k) already made, and also the information necessary for optimal stopping (as a rule, this is settled by knowledge of the current state (x_t) ; for details see [2]).

§ 5. Example

The process X^d is a diffusion process on the segment $E = [r_1, r_2]$, with absorption at the boundary, and given in the interior of the segment by the equation

$$dx_t = a(x_t, d)dt + dw_t,$$

where w_t is a standard Wiener process. The set D is finite, and the function $a(x, d)$ is continuous in x for each d . Our purpose is to maximize the value of

$$v_\pi(x) = M_x^\pi \sum_k \int_{\tau_k}^{\tau_{k+1}} f(x_t, d_k) dt,$$

where the function $f(x, d)$ is nonnegative and continuous in x , and $f(r_i, d) = 0$ (by the same token $v_\pi(x) = M_x^\pi \int_0^\zeta f(x_t, d_t) dt$, where ζ is the moment of exit to the boundary of the segment $[r_1, r_2]$, and $d_t = d_k$ for $\tau_k \leq t < \tau_{k+1}$). According to Theorem 13.16 of [1], for each d

$$\frac{1}{2} s_0''(x, d) + a(x, d) s_0'(x, d) + f(x, d) = 0, \quad s_0(r_i, d) = 0$$

(the differentiation is carried out with respect to x). Let $u(x)$ be a solution of the equation

$$\frac{1}{2} u''(x) + \max_{d \in D} [a(x, d)u'(x) + f(x, d)] = 0, \quad u(r_i) = 0. \tag{11}$$

Then for the function $b(x, d) = u(x) - s_0(x, d)$ we have, for each d ,

$$\frac{1}{2} h''(x, d) + a(x, d)h'(x, d) \leq 0, \quad h(r_i, d) = 0,$$

i.e. the function (x, d) is excessive (this follows trivially from Theorems 15.9 and 13.16 in [1]); therefore, according to Theorem 4, $u(x) \geq v(x)$.

We shall separate $[r_1, r_2]$ into sets A^d such that for $x \in A^d$

$$\max_d [a(x, d)u'(x) + f(x, d)] = a(x, d_0)u'(x) + f(x, d_0)$$

(each A^d is the union of not more than a countable number of segments). We cover all the boundary points of all the A^d by intervals of common length l ; the union of these intervals will be denoted by J_l . Let the strategy π consist of the following: at the moment $t = 0$ we choose d_0 so that $x_0 \in A^{d_0}$, and subsequently, having observed the process X^{d_k} , we stop at the moment τ_{k+1} of the first (after τ_k) exit from the set $A^{d_k} \cup J_l$ and switch to a d_{k+1} such that $x_{\tau_{k+1}} \in A^{d_{k+1}}$ ($k = 0, 1, 2, \dots$). After making some simple but rather tedious estimates, we can show that for any $\epsilon > 0$ we have $v_\pi(x) \geq u(x) - \epsilon$ for sufficiently small l . By the same token we show that $u(x) = v(x)$, i.e. the solution of equation (11) is a cost function.

§6. Concluding remarks

1) The requirement that the transition function of the process X^d be a \hat{C} -function (see 2.C) can be discarded; this is not reflected in the results but follows nontrivially from the proof of Lemma 1.a).

2) All the results of this article can be placed without difficulty in a more general situation:

a) all the processes X^d have a (common) cut-off moment $\zeta = \zeta(\omega)$ satisfying the condition $P_x\{\zeta > 0\} = 1, x \in E$ (i.e. the process X^d is normal);

b) as a price function we obtain the functional

$$r(\pi, \omega) = \sum_{k=0}^{\infty} \int_{\tau_k}^{\tau_{k+1}} e^{-\phi_t} \gamma^{d_k}(dt),$$

where ϕ_t is an additive, homogeneous, finite, continuous, nonnegative, perfect functional, and $P_x\{\phi_0 = 0\} = 1, x \in E$. Here we must assume that $s_0(x, d) = M_x^d \int_0^\infty e^{\phi_t} \gamma^d(dt)$, and in all the formulas and reasoning we must replace the term "excessive function" by the term " (α, ζ) -excessive function", i.e. an excessive function for a standard α -subprocess (terminating) of the Markov process $\{x_t, \zeta, \mathbb{M}_t, P_x^d\}$, where $\alpha_t = e^{-\phi_t}$ is a multiplicative functional.

This work was completed under the guidance of A. N. Širjaev, to whom the author is grateful.

Received 29/DEC/70

BIBLIOGRAPHY

- [1] E. B. Dynkin, *Markov processes*, Fizmatgiz, Moscow, 1963; English transl. Vols. I, II, Die Grundlehren der math. Wissenschaften, Bände 121, 122, Academic Press, New York; Springer-Verlag, Berlin, 1965. MR 33 #1886; #1887.
- [2] A. N. Širjaev, *Statistical sequential analysis*, "Nauka", Moscow, 1969; English transl., Transl. Math. Monographs, vol. 38, Amer. Math. Soc., Providence, R. I. (to appear).

- [3] D. Blackwell, *Discounted dynamic programming*, Ann. Math. Statist. 36 (1965) 226–235; Russian transl., Matematika 11 (1967), no. 4, 151–160. MR 30 #3749.
- [4] R. E. Strauch, *Negative dynamic programming*, Ann. Math. Statist. 37 (1966), 871–890; Russian transl., Matematika 13 (1969), no. 5, 107–127. MR 33 #2456.
- [5] G. W. Haggstrom, *Optimal sequential procedures when more than one stop is required*, Ann. Math. Statist. 38 (1967), 1618–1626. MR 36 #1035.
- [6]* M. A. Naïmark, *Normed rings*, 2nd rev. ed., “Nauka”, Moscow, 1968; English transl. of 1st ed., Noordhoff, Groningen, 1959. MR 22 #1824.
- [7] K. Kuratowski, *Topologie*. Vol. I, PWN, Warsaw, 1958; English transl., Academic Press, New York; PWN, Warsaw, 1966; Russian transl., “Mir”, Moscow, 1966. MR 19, 873; 36 #839; #840.
- [8] E. B. Dynkin, *Foundations of the theory of Markov processes*, Fizmatgiz, Moscow, 1959; English transl., Prentice-Hall, Englewood Cliffs, N.J.; Pergamon Press, Oxford, 1961. MR 24 #A1745; A#A1747.
- [9] L. Dubins and D. Freeman, *Measurable sets of measures*, Pacific J. Math. 14 (1964), 1211–1222. MR 30 #4887.

Translated by:
E. Nishiura

* *Editor's note.* All page references to [6] are to the Russian 1968 edition.