

## REFERENCES

- [1] D. YLVIKAKER, *A note on the absence of tangencies in Gaussian sample paths*, Ann. Math. Statist., 39, 1 (1968), pp. 261–262.
- [2] A. V. SKOROKHOD, *A note on Gaussian measures in Banach space*, Theory Prob. Applications, 15, 3 (1970), p. 508.
- [3] X. FERNIQUE, *Intégrabilité des vecteurs gaussiens*, C.R. Acad. Sci. Paris, Ser. A, 270, 25 (1970), pp. 1698–1699.
- [4] H. J. LANDAU AND L. A. SHEPP, *On the supremum of a Gaussian process*, Sankhya, Ser. A, 32, 4 (1970), pp. 369–378.
- [5] V. N. SUDAKOV AND B. S. TSIREL'SON, *Extremal properties of half-spaces for spherically invariant measures*, Zapiski Nauchn. Seminarov LOMI, 41 (1974), pp. 14–24. (In Russian.)
- [6] V. I. SMIRNOV, *Course in Higher Mathematics*, Vol. 2, Addison-Wesley, Reading, Mass., 1964.

**ON CONTROLLED FINITE STATE MARKOV PROCESSES  
WITH COMPACT CONTROL SETS**

E. A. FAINBERG

(Translated by K. Durr)

**1.** In this paper we consider the maximization of the average gain per unit step in controlled finite state Markov chains with compact control sets.

In [1] and [2] a stationary optimal strategy was shown to exist under the assumption that the control sets are finite. In [3] it was proved that if the control sets are compact and coincide with the transition probability sets, the gain functions are continuous and any stationary strategy yields a Markov chain with one ergodic class and without transient states, then there exists a stationary optimal strategy. In the general case, conditions of compactness of the control sets and of continuity of the gain and transition functions are not sufficient for the existence of an optimal strategy (see [4], Example 3).

In this paper the existence is established of a stationary optimal strategy under the condition that the control sets are compact, the gain functions are upper semi-continuous, the transition functions depend continuously on the controls and that one of the following conditions holds: (i) any stationary strategy yields a Markov chain with one ergodic class, and possibly with transient states (Section 3); (ii) for each state the set of transition probabilities contains a finite set of extreme points (Section 4).

**2.** Let  $X$  be a state space consisting of a finite number of points ( $X = \{1, 2, \dots, s\}$ ). For each state there is given a control set  $A_x$  ( $x = 1, 2, \dots, s$ ). On the sets  $A_x$  there are defined functions  $q_x(a)$  (the gain from the control  $a \in A_x$  when the process is in the state  $x$ ), and probability measures  $p_x(\cdot | a)$  on  $X$  (the transition functions under the condition that the process is in the state  $x$  and the control  $a \in A_x$  is chosen). Set  $A = \bigcup_{x=1}^s A_x$ .

Let  $x_0, a_1, x_1, a_2, x_2, \dots$  be the succession of states and controls. At each moment  $t = 1, 2, \dots$  a choice of control is made which is given by a probability measure  $\pi_t(da_t | x_0, a_1, x_1, \dots, a_{t-1}, x_{t-1})$  on the set  $A_{x_{t-1}}$  measurably depending on the past. The collection of these measures for  $t = 1, 2, \dots$  defines the strategy  $\pi$ . The strategy  $\pi$  is called *stationary* if the measures  $\pi_t$  are concentrated at the points  $a_t = \varphi(x_{t-1})$ , where  $\varphi$  is a selector, i.e., a mapping of  $X$  into  $A$  such that  $\varphi(x) \in A_x$ . The corresponding strategy is also denoted  $\varphi$ . Clearly, each selector defines a homogeneous Markov chain. Let  $\Delta_A$  denote the collection of all strategies defined on the control sets  $A_x$ .

The gain yielded by the strategy  $\pi$  is estimated by the function

$$W_x(q, \pi) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbf{E}_x^\pi \sum_{t=1}^N q_{x_{t-1}}(a_t),$$

where  $E_x^\pi$  is the expectation for  $x_0 = x$  and the strategy  $\pi$ . Set

$$V_x(q, \Delta_A) = \sup_{\Delta_A} W_x(q, \pi).$$

The strategy  $\pi^*$  is called *optimal* if  $W_x(q, \pi^*) = V_x(q, \Delta_A)$ ,  $x = 1, 2, \dots, s$ . As long as we are dealing with a unique gain function  $q$ , we shall write  $W_x(\pi)$  and  $V_x(\Delta_A)$ .

**3.** Let  $P$  be an  $s \times s$  stochastic matrix. Let  $P_i$  denote the matrix obtained from the matrix  $E - P$  by replacing the elements of the  $i$ -th column by ones.

**Lemma 1.** *If the Markov chain is given by the matrix  $P$  and contains a unique ergodic class, then the matrices  $P_i$  are non-singular.*

**PROOF.** Without loss of generality one may assume that  $i = s$ . Since the sum of all columns of  $E - P$  is equal to 0 the equations

$$(1) \quad m(E - P) = 0,$$

$$(2) \quad \sum_{x=1}^s m_x = 1,$$

which are satisfied by the unique stationary distribution  $m = (m_1, m_2, \dots, m_s)$  of the states of the chain, are equivalent to the equation

$$(3) \quad mP_s = (0, 0, \dots, 0, 1).$$

If the chain does not contain transient states, then equations (1) and (2) have a unique solution and therefore  $|P_s| \neq 0$ . If there are transient states, then we enumerate the states of the chain so that the states  $1, 2, \dots, s - k$  are transient while  $s - k + 1, s - k + 2, \dots, s$  are recurrent. Denote by  $P'$  the matrix of transition probabilities of this chain and the stationary distribution by  $m'$ :

$$(4) \quad m'P'_s = (0, 0, \dots, 0, 1).$$

The matrix  $P'_s$  has the form

$$P'_s = \begin{pmatrix} E - Q & T \\ 0 & R_k \end{pmatrix},$$

where  $R_k$  is the matrix of transition probabilities of an ergodic chain without transient states and  $|P'_s| = |E - Q| \cdot |R_k|$ ,  $|E - Q| \neq 0$  (see [5], p. 46),  $|R_k| \neq 0$ . Hence  $|P'_s| \neq 0$  and equation (4) has a unique solution. Equations (3) and (4) are equivalent since each of them is equivalent to equations (1) and (2), and so equation (3) has a unique solution and  $P_s$  is non-singular. The lemma is proved.

Let  $P(\varphi)$  denote the matrix of transition probabilities of the Markov chain given by the selector  $\varphi$ . Let  $W(\pi)$ ,  $V(\Delta_A)$  and  $q(\varphi)$  be the column vectors  $\{W_x(\pi)\}$ ,  $\{V_x(\Delta_A)\}$  and  $\{q_x(a_x)\}$ , where  $a_x = \varphi(x)$ . By the ergodic theorem for Markov chains,

$$(5) \quad W(\varphi) = M(\varphi)q(\varphi), \quad \text{where } M(\varphi) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N P^{t-1}(\varphi).$$

**Lemma 2.** *If the transition functions  $p_x(\cdot | a)$  are continuous and any selector  $\varphi$  yields a Markov chain with one ergodic class, then the elements of the matrix  $M(\varphi)$  are continuous functions on the set  $A_1 \times A_2 \times \dots \times A_s$ .*

**PROOF.** Since  $\varphi$  yields a Markov chain with one ergodic class, the matrix  $M(\varphi)$  consists of identical rows  $m(\varphi)$ . By Lemma 1 and (3),  $m(\varphi) = (0, 0, \dots, 0, 1)P_s^{-1}$ , and so the elements of the vector  $m(\varphi)$  are continuous functions on the set  $A_1 \times A_2 \times \dots \times A_s$ . The lemma is proved.

**Lemma 3.** *If all the sets  $A_x$  are compact, the transition functions  $p_x(\cdot | a)$  are continuous, the functions  $q_x(a)$  are upper semi-continuous and bounded and any selector  $\varphi$  defines a Markov chain with one ergodic class, then there exists a stationary optimal strategy.*

PROOF. Denote the total gain after  $n$  steps for the initial state  $x$ , strategy  $\pi$  and final reward  $r(x)$  by

$$F_x(\pi, n, r) = \mathbf{E}_x^\pi \left( \sum_{i=1}^n q_{x_{i-1}}(a_i) + r(x_n) \right).$$

Denote the maximum total gain after  $n$  steps for the initial state  $x$ , the control sets  $A_1, A_2, \dots, A_x$  and final reward  $r(x)$  by  $T_x(\Delta_A, n, r)$ . Let  $F(\pi, n, r)$ ,  $T(\Delta_A, n, r)$  and let  $r$  be the column-vectors  $\{F_x(\pi, n, r)\}$ ,  $\{T_x(\Delta_A, n, r)\}$  and  $\{r(x)\}$ . By Belman's equation (cf. [6]),

$$T_x(\Delta_A, n, 0) = \sup_{a \in A_x} (q_x(a) + \sum_{z=1}^s p_x(r|a) \cdot T_z(\Delta_A, n-1, 0)).$$

Since the  $p_x(\cdot | a)$  are continuous and the  $q_x(a)$  are upper semi-continuous, the supremum occurs at each step for some control  $a_x^n$ . Set  $A_x^N = \{a_x^n, 1 \leq n \leq N\}$ .

If the sets of controls  $A_x$  are restricted to  $A_x^N$  (while preserving the other characteristics of the process), then on the finite control sets  $A_x^N$  there exist, by [7] a selector  $\varphi_N$  and a final reward  $r_N$  such that

$$(6) \quad W(\varphi_N) = P(\varphi_N)W(\varphi_N) = \sup_{\varphi(x) \in A_x^N} P(\varphi)W(\varphi_N),$$

$$(7) \quad W(\varphi_N) + r_N = P(\varphi_N)r_N + q(\varphi_N) = \sup (P(\varphi)r_N + q(\varphi)).$$

In equation (7) the supremum is taken over all selectors satisfying equation (6). Relations (6) and (7) are preserved if the elements of the vector  $r_N$  are altered by a common amount, and thus one can choose  $r_N$  so that  $r_N(s) = 0$ . Since any selector  $\varphi$  defines a Markov chain with one ergodic class, the elements of the vector  $W(\varphi)$  are all equal for any  $\varphi$ , and hence equation (6) becomes an identity, while  $\varphi_N, r_N$  and  $W(\varphi_N)$  form a canonical triplet on the control sets  $A_x$  (cf. [8]):

$$\begin{aligned} F(\varphi_N, N, r_N) &= NW(\varphi_N) + r_N = T(\Delta_{A^N}, N, r_N) \\ &\cong T(\Delta_{A^N}, N, 0) - \|r_N\| = T(\Delta_A, N, 0) - \|r_N\|, \end{aligned}$$

where  $\|r_N\| = \max_{1 \leq x \leq s} |r_N(x)|$ . It follows that

$$(8) \quad T(\Delta_A, N, 0) \leq NW(\varphi_N) + 2\|r_N\|.$$

Let us show that the sequence  $\|r_N\|$  is bounded. From (5) and (7) it follows that

$$(9) \quad (E - P(\Phi_N))r_N = (E - M(\varphi_N))q(\varphi_N).$$

Since  $r_N(s) = 0$ , relation (9) is equivalent to the relation

$$P_s(\varphi_N)r_N = (E - M(\varphi_N))q(\varphi_N).$$

From Lemmas 1 and 2 it follows that the elements of the matrix  $P_s^{-1}(\varphi)(E - M(\varphi))$  are continuous functions on the compact set  $A_1 \times A_2 \times \dots \times A_s$  and hence are bounded, and  $q(\varphi)$  is bounded by hypothesis. Hence  $\|r_N\|$  is bounded.

Dividing (9) by  $N$ , choosing a convergent subsequence from  $\varphi_N$  and using Lemma 2 and the fact that the  $q_x(a)$  are upper semi-continuous, we obtain on passing to the limit the desired result.

**Theorem 1.** *If all the sets  $A_x$  are compact, the transition functions  $p_x(\cdot | a)$  are continuous, the functions  $q_x(a)$  are upper semi-continuous and any selector  $\varphi$  defines a Markov chain with one ergodic class, then there exists a stationary optimal strategy.*

PROOF. For each  $n = 1, 2, \dots$  we set

$$q_x^n(a) = \max \{q_x(a), -n\}.$$

Since an upper semi-continuous function on a compact set is bounded, the functions  $q_x^n(a)$  are upper semicontinuous and bounded. It follows from Lemma 3 that for each  $q_x^n(a)$  there

exists a stationary optimal strategy  $\varphi_n$ :

$$W_x(q^n, \varphi_n) = V_x(q^n, \Delta_A) \cong V_x(q, \Delta_A).$$

To complete the proof of the theorem it remains to consider a convergent subsequence of the sequence  $\varphi_n$ , and to pass, letting  $n$  tend to  $\infty$ , to the limit in the last inequality. We must use the upper semicontinuity of  $q_x(a)$  and Lemma 2.

4. Let us consider the vector-functions  $y = p_x(\cdot | a)$ . Set

$$Y_x = \{y: y = p_x(\cdot | a), a \in A_x\}.$$

Since  $\sum_{z=1}^s p_x(z|a) = 1$  for every  $a \in A_x$ , the set  $Y_x$  belongs to an  $(s-1)$ -dimensional simplex. Let the  $A_x$  be compact, the  $q_x(a)$  upper semi-continuous and the  $p_x(\cdot | a)$  continuous. For each point  $y \in Y_x$  consider the set  $A_{xy} = \{a \in A_x: p_x(\cdot | a) = y\}$ , which, as is not difficult to see, is compact. Set  $c_x(y) = \max_{a \in A_{xy}} q_x(a)$ . This definition is correct since an upper semi-continuous function on a compact set attains a maximum there.

**Lemma 4.** *The functions  $c_x(y)$  are upper semi-continuous.*

The proof of Lemma 4 is analogous to that of Lemma 2 in [9].

Let us consider, for a bounded function  $u = c(y)$ ,  $b \leq c(y) \leq b'$ , defined on the set  $Y \subset R^{s-1}$ , the set of points "lying under the graph of the function  $c(y)$  above  $b$ ":

$$D(Y, c(y)) = \{y, u: y \in Y, b \leq u \leq c(y)\}.$$

**Lemma 5.** *The set  $D(Y, c(y))$  is compact if and only if  $Y$  is compact and the function  $c(y)$  is upper semi-continuous.*

The proof is obvious.

Let  $[Y]$  denote the convex hull of  $Y$ . Suppose  $[Y]$  is a polyhedron.

**Lemma 6.** *If the bounded function  $c(y)$  is upper semi-continuous on the compact set  $Y \subset R^{s-1}$ , then there exists a concave continuous function  $\bar{c}(y)$  defined on  $[Y]$  such that  $\bar{c}(y) \geq c(y)$  for any  $y \in Y$  and that, for any  $y^0 \in [Y]$ , there exist  $s+1$  points  $y^k \in Y$  and  $s+1$  numbers  $\alpha_k$ ,  $k = 1, 2, \dots, s+1$ , satisfying the conditions*

$$\bar{c}(y^0) = \sum_{k=1}^{s+1} \alpha_k c(y^k), \quad \alpha_k \geq 0, \quad \sum_{k=1}^{s+1} \alpha_k = 1, \quad y^0 = \sum_{k=1}^{s+1} \alpha_k y^k.$$

Let us point out the main steps in the proof of the lemma. It follows from Lemma 5 that  $D(Y, c(y))$  is a compact set and hence  $[D(Y, c(y))]$  is also compact. Define the function

$$\bar{c}(y) = \max \{u: (y, u) \in [D(Y, c(y))]\}.$$

Since  $c(y) = \max \{u: (y, u) \in D(Y, c(y))\}$ , we have  $\bar{c}(y) \geq c(y)$ . One can see that, by the convexity of  $[D(Y, c(y))]$ , the function  $\bar{c}(y)$  is concave and  $[D(Y, c(y))] = D([Y], \bar{c}(y))$ . From the concavity of  $\bar{c}(y)$  it follows that it is lower semi-continuous, and from the compactness of  $D([Y], \bar{c}(y))$  it follows that  $\bar{c}(y)$  is upper semi-continuous. Hence  $\bar{c}(y)$  is continuous.

Since  $D(Y, c(y)) \subset R^s$  is bounded, there exists, for any  $(y^0, u^0) \in [D(Y, c(y))]$ ,  $s+1$  points  $(y^k, u^k) \in D(Y, c(y))$  and  $s+1$  numbers  $\alpha_k$  such that

$$(y^0, u^0) = \sum_{k=1}^{s+1} \alpha_k (y^k, u^k), \quad \alpha_k \geq 0, \quad \sum_{k=1}^{s+1} \alpha_k = 1$$

(see [10], p. 397). Let us take  $u^0 = \bar{c}(y^0)$  where  $y^0 \in [Y]$ . Then assuming that  $u^k \neq c(y^k)$  for at least one  $1 \leq k \leq s+1$ , i.e.,  $u^k < c(y^k)$ , we see that there exists a point

$$(y^0, u) = \sum_{k=1}^{s+1} \alpha_k (y^k, c(y^k)) \in D([Y], \bar{c}(y)),$$

such that  $u > \bar{c}(y^0)$ , which contradicts the definition of  $\bar{c}(y)$ .

**Corollary.** For each  $y^0 \in [Y]$  the following equality holds:

$$\bar{c}(y^0) = \max \sum_{k=1}^{\infty} \alpha_k c(y^k), \quad \text{where } \alpha_k \geq 0, \quad \sum_{k=1}^{\infty} \alpha_k = 1, \quad \sum_{k=1}^{\infty} \alpha_k y^k = y^0.$$

Indeed,

$$\sum_{k=1}^{\infty} \alpha_k c(y^k) \leq \sum_{k=1}^{\infty} \alpha_k \bar{c}(y^k) \leq \bar{c}(y^0).$$

The last inequality results from the concavity of  $\bar{c}(y)$ .

We call the function  $\bar{c}(y)$  the *concave majorant* of the function  $c(y)$ .

**Theorem 2.** If all the sets  $A_x$  are compact, the transition functions  $p_x(\cdot | a)$  are continuous, the  $q_x(a)$  are upper semi-continuous and all the sets  $Y_x$  contain a finite number of extreme points, then there exists a stationary optimal strategy.

**PROOF.** By the assumptions of the theorem, the sets  $[Y_x]$  are convex polyhedra. Let  $\bar{y}_x^1, \bar{y}_x^2, \dots, \bar{y}_x^{n_x}$  be the vertices of  $[Y_x]$ . Consider the functions

$$c_x^*(y) = \max \sum_{k=1}^{n_x} \alpha_k c_x(\bar{y}_x^k), \quad \text{where } \alpha_k \geq 0, \quad \sum_{k=1}^{n_x} \alpha_k = 1, \quad y = \sum_{k=1}^{n_x} \alpha_k \bar{y}_x^k.$$

By Lemma 6 and its corollary, the  $c_x^*(y)$  are continuous. Define the functions  $c'_x(y) = \max \{c_x(y), c_x^*(y)\}$ . For the upper semi-continuous and bounded functions  $c'_x(y)$  consider the concave majorants  $\bar{c}'_x(y)$ .

Consider a controlled process with state space  $X$ , control sets  $[Y_x]$  and gain functions  $\bar{c}'_x(y)$ . For this process the sets of transition probabilities coincide with the control sets. Since for each strategy  $\pi \in \Delta_A$  there exists a strategy  $\sigma \in \Delta_{[Y]}$  such that

$$\sigma_t(dy_t | x_0, y_1, x_1, \dots, y_{t-1}, x_{t-1}) = \pi_t(da_t | x_0, a_1, x_1, \dots, a_{t-1}, x_{t-1}),$$

where  $y_t = p_{x_{t-1}}(\cdot | a_t)$ , and  $\bar{c}'_{x_{t-1}}(y_t) \geq q_{x_{t-1}}(a_t)$ ,

$$(10) \quad W_x(\bar{c}', \sigma) \geq W_x(q, \pi) \quad \text{and} \quad V_x(\bar{c}', \Delta_{[Y]}) \geq V_x(q, \Delta_A).$$

**Lemma 7.** For any  $\varepsilon > 0$  there exists, under the assumptions of Theorem 2, a stationary strategy  $\varphi_\varepsilon \in \Delta_{[Y]}$  such that

$$V_x(\bar{c}', \Delta_{[Y]}) - W_x(\bar{c}', \varphi_\varepsilon) \leq \varepsilon, \quad x = 1, 2, \dots, s.$$

**PROOF.** A continuous function on a compact set is uniformly continuous and hence one can divide the convex polyhedra  $[Y_x]$  into a finite number of polyhedra in such a way that the oscillations of the  $\bar{c}'_x(y)$  on each of them do not exceed  $\varepsilon$ . Denote the vertices of these polyhedra by  $y'_x, y''_x, \dots, y_x^{N_x}$ . Consider the functions

$$c_x^\varepsilon(y) = \max \sum_{k=1}^{N_x} \alpha_k \bar{c}'_x(y_x^k), \quad \text{where } y = \sum_{k=1}^{N_x} \alpha_k y_x^k, \quad \alpha_k \geq 0, \quad \sum_{k=1}^{N_x} \alpha_k = 1.$$

Due to the choice of the  $y_x^k$ , we have  $\bar{c}'_x(y) - c_x^\varepsilon(y) \leq \varepsilon$ , and hence

$$(11) \quad V_x(\bar{c}', \Delta_{[Y]}) - V_x(c^\varepsilon, \Delta_{[Y]}) \leq \varepsilon, \quad x = 1, 2, \dots, s.$$

Since the  $\bar{c}'_x(y)$  are concave,  $c_x^\varepsilon(y_x^k) = \bar{c}'_x(y_x^k)$ . From the definition of the  $c_x^\varepsilon(y)$  it follows that each strategy  $\sigma$  on the control sets  $[Y_x]$  with gain functions  $c_x^\varepsilon(y)$  can be viewed as a strategy  $\pi$  on the control sets  $\bar{Y}_x = \{y'_x, y''_x, \dots, y_x^{N_x}\}$  with gain functions  $\bar{c}'_x(y)$  so that

$$(12) \quad W_x(c^\varepsilon, \sigma) = W_x(c^\varepsilon, \pi) = W_x(\bar{c}', \pi).$$

Since the  $\bar{Y}_x$  are finite, there exists by [1] a stationary strategy  $\varphi_\varepsilon \in \Delta_{\bar{Y}}$  such that

$$(13) \quad W_x(c^\varepsilon, \varphi_\varepsilon) = V_x(c^\varepsilon, \Delta_{\bar{Y}}) = V_x(c^\varepsilon, \Delta_{[Y]}).$$

From (11)–(13) it follows that

$$V_x(\bar{c}', \Delta_{[Y]}) - W_x(\bar{c}', \varphi_\varepsilon) \leq \varepsilon.$$

**Lemma 8.** *Under the assumptions of Theorem 2 there exists a stationary strategy  $\varphi^* \in \Delta_{[Y]}$  such that*

$$V_x(\bar{c}', \Delta_{[Y]}) = W_x(\bar{c}', \varphi^*), \quad x = 1, 2, \dots, s.$$

**PROOF.** Let  $\varphi_\varepsilon = \{\bar{y}_1^\varepsilon, \bar{y}_2^\varepsilon, \dots, \bar{y}_s^\varepsilon\}$ ,  $\bar{y}_x^\varepsilon \in [Y_x]$ . Consider the sequence of strategies  $\varphi_{\varepsilon_n}$  as  $\varepsilon_n \rightarrow 0$ . By the compactness of the  $[Y_x]$  and the finiteness of  $X$ , one can choose a subsequence  $\varphi_{\varepsilon_l}$  of the sequence  $\varphi_{\varepsilon_n}$  in such a way that, for any  $x = 1, 2, \dots, s$ ,  $\lim_{l \rightarrow \infty} \bar{y}_x^{\varepsilon_l} = \bar{y}_x^0$  exists.

For each  $[Y_x]$  consider the set  $Y'_x$  consisting of the vertices of  $[Y_x]$  and the points  $\bar{y}_x^0$ .  $Y'_x = \{\bar{y}_x^0, \bar{y}_x^1, \dots, \bar{y}_x^{n_x}\}$ . The control sets  $[Y'_x]$  are finite. Hence among the strategies defined on these sets there exists a stationary optimal strategy  $\varphi^*$ . Let us show that  $\varphi^*$  is optimal among the  $\sigma \in \Delta_{[Y]}$ .

On each of the  $[Y_x]$  consider the function

$$\tilde{c}_x(y) = \max_{k=0}^{n_x} \sum \alpha_k \bar{c}'_x(\bar{y}_x^k), \quad \alpha_k \geq 0, \quad \sum_{k=0}^{n_x} \alpha_k = 1, \quad y = \sum_{k=0}^{n_x} \alpha_k \bar{y}_x^k.$$

By Lemma 6 and its corollary the  $\tilde{c}_x(y)$  are concave and continuous. In the same way as in the proof of Lemma 7 one can show that

$$W_x(\bar{c}', \varphi^*) = W_x(\tilde{c}, \varphi^*) = V_x(\tilde{c}, \Delta_{[Y]}).$$

For any  $\varepsilon > 0$  we select in each  $[Y_x]$  a neighborhood  $U(\bar{y}_x^0, \delta_x)$  of points  $\bar{y}_x^0$  such that  $\bar{c}'(y) - \tilde{c}(y) \leq \varepsilon/2$  if  $y \in U(\bar{y}_x^0, \delta_x)$ . Consider the strategy  $\varphi_{\varepsilon_N}$  where

$$\varepsilon_N \leq \varepsilon/2 \quad \text{and} \quad \bar{y}_x^{\varepsilon_N} \in U(\bar{y}_x^0, \delta_x), \quad x = 1, 2, \dots, s.$$

Then

$$W_x(\bar{c}', \varphi_{\varepsilon_N}) - W_x(\tilde{c}, \varphi_{\varepsilon_N}) \leq \varepsilon/2.$$

Finally we have

$$V_x(\bar{c}', \Delta_{[Y]}) \leq W_x(\bar{c}', \varphi_{\varepsilon_N}) + \frac{\varepsilon}{2} \leq W_x(\tilde{c}, \varphi_{\varepsilon_N}) + \varepsilon \leq V_x(\tilde{c}, \Delta_{[Y]}) + \varepsilon = W_x(\bar{c}', \varphi^*) + \varepsilon.$$

Since this inequality is valid for all  $\varepsilon > 0$  and  $V_x(\bar{c}', \Delta_{[Y]}) \geq W_x(\bar{c}', \varphi^*)$ ,

$$W_x(\bar{c}', \varphi^*) = V_x(\bar{c}', \Delta_{[Y]}), \quad x = 1, 2, \dots, s.$$

The lemma is proved.

Let us continue the proof of Theorem 2. Suppose  $\varphi^*$  is defined by the controls  $\{y'_1, y'_2, \dots, y'_s\}$  where  $y'_x \in [Y_x]$ . Then from the definition of the  $\bar{c}'_x(y)$  and Lemma 6 it follows that for each  $x$  there exists a finite number of points  $\tilde{y}_x^k \in Y_x$  and a finite collection of numbers  $\alpha_{x,k}$  satisfying the conditions

$$\bar{c}'(y'_x) = \sum_k \alpha_{x,k} c_x(\tilde{y}_x^k), \quad \alpha_{x,k} \geq 0, \quad \sum_k \alpha_{x,k} = 1, \quad y'_x = \sum_k \alpha_{x,k} \tilde{y}_x^k.$$

Hence on the control sets  $[Y_x]$  there exists a stationary optimal strategy  $\varphi' = \{y_1, y_2, \dots, y_s\}$  where  $y_x \in Y_x$  and  $\bar{c}'_x(y_x) = c_x(y_x)$ . By considering on the control sets  $A_x$  the stationary strategy  $\varphi = \{a_1, a_2, \dots, a_s\}$  where  $a_x \in A_x$ ,  $p_x(\cdot | a_x) = y_x$  and  $q_x(a_x) = c_x(y_x)$ , we see that  $W_x(q, \varphi) = V_x(\bar{c}', \Delta_{[Y]})$ . From (10) it follows that  $W_x(q, \varphi) = V_x(q, \Delta_A)$ ,  $x = 1, 2, \dots, s$ . The theorem is proved.

**Corollary.** *If the space  $X$  consists of two points, the sets  $A_x$  are compact, the transition functions  $p_x(\cdot | a)$  are continuous, while the functions  $q_x(a)$  are upper semi-continuous, then there exists a stationary optimal strategy.*

In fact,  $Y_x \subset R^1$ ,  $x = 1, 2$ , and every set on the line contains at most two extreme points.

**REMARK.** If in estimating the gain one considers the test

$$(14) \quad W_x(q, \pi) = \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \mathbf{E}_x^\pi \sum_{t=1}^N q_{x_{t-1}}(a_t),$$

then, according to [11], p. 26, every stationary optimal strategy for testing the lower limit is also optimal for testing the upper limit if the set of states and the control sets are finite. Noting this fact it is not hard to see that the proofs adduced in this paper remain valid also for the test (14).

The author expresses sincere thanks to A. A. Yushkevich for the statement of the problem and useful discussions.

*Received by the editors  
June 26, 1974*

#### REFERENCES

- [1] O. V. VISKOV AND A. N. SHIRYAEV, *On controls leading to optimal stationary regimes*, Trudy MIAN, 71 (1964), pp. 35–45. (In Russian.)
- [2] C. DERMAN, *On sequential decisions and Markov chains*, Manag. Sci., 9 (1962), pp. 16–24.
- [3] A. MARTIN-LÖF, *Existence of a stationary control for a Markov chain maximizing the average reward*, Oper. Res. 15 (1967), pp. 886–871.
- [4] J. BATHER, *Optimal decision procedures for finite Markov chains, Part I: examples*, Adv. Appl. Prob., 5 (1973), pp. 328–339.
- [5] J. G. KEMENY AND J. L. SNELL, *Finite Markov Chains*, Van Nostrand, Princeton, N.Y., 1965.
- [6] R. A. HOWARD, *Dynamic Programming and Markov Processes*, Wiley, N.Y., 1960.
- [7] E. V. DENARDO AND B. L. FOX, *Multichain Markov renewal programs*, SIAM J. Appl. Math., 16 (1968), pp. 468–487.
- [8] A. A. YUSHKEVICH, *On a class of strategies in general controlled Markov models*, Theory Prob. Applications, 18, 4 (1973), pp. 777–778.
- [9] L. G. GUBENKO AND E. S. SHTATLAND, *On controlled Markov processes with discrete time*, Teoriya Veroyatn. i Matem. Statistika, Kiev, 7 (1972), pp. 51–64. (In Russian.)
- [10] S. KARLIN, *Mathematical Methods in the Theory of Games, Programming, and Economics*, Addison-Wesley, Reading, Mass., 1959.
- [11] C. DERMAN, *Finite State Markovian Decision Processes*, Acad. Press, New York, 1970.

### A CENTRAL LIMIT THEOREM FOR THE NUMBER OF PARTIAL LONG REPETITIONS

V. G. MIKHAILOV

(Translated by W. U. Sirk)

Let  $X_1, X_2, \dots$  be a sequence of random variables assuming values  $1, 2, \dots$ . We introduce the concept of an  $m$ -fold repetition of length  $s$  and rank  $d$  ( $(m, s, d)$ -repetition) as an event consisting of the fulfilment of exactly  $s - d$  sequences of equation in the system

$$X_{i_1+k} = \dots = X_{i_m+k}, \quad k = 0, \dots, s-1.$$

Let  $\eta_d(i_1, \dots, i_m)$  be the indicator of an  $(m, s, d)$ -repetition (the numbers  $i_1, \dots, i_m$  indicate the position of the portions being repeated in the sequence  $X_1, X_2, \dots$ ) and let  $\xi_n(m, s, d)$  be the number of all  $(m, s, d)$ -repetitions in the set  $X_1, \dots, X_{n+s-1}$ . It is obvious that

$$\xi_n(m, s, d) = \sum_{1 \leq i_1 < \dots < i_m \leq n} \eta_d(i_1, \dots, i_m), \quad d = 0, \dots, s.$$