The expected number of intersections of a four valued bounded martingale with any level may be infinite

Alexander Gordon and Isaac M. Sonin

According to the well-known Doob's lemma, the expected number of crossings of every fixed interval (a,b) by trajectories of a bounded martingale (X_n) is finite on the infinite time interval. For such a random sequence (r.s.) with an extra condition that X_n takes no more than N, $N < \infty$, values at each moment $n \geq 1$, this result was refined in Sonin (1987) by proving that inside any interval (a,b) there are non-random sequences (barriers) (d_n), such that the expected number of intersections of d_n by (X_n) is finite on the infinite time interval. This result left open the problem of whether for such r.s. any *constant* barriers $d_n \equiv d,\ n \geq 1,$ exist. The main result of this paper is an example of a bounded martingale X_n, $0 \leq X_n \leq 1$, with at most four values at each moment n, such that no constant d, $0 < d < 1$, is a barrier for (X_n). We also discuss the relationship of this problem with such problems as the behavior of a general finite nonhomogeneous Markov chain and the behavior of the simplest model of an irreversible process.

Key words: martingale, finite nonhomogeneous Markov chain, irreversible process

Mathematics Subject Classification (2000): 60G42, 60J10, 82B35

1 Introduction

In this note we present some results that shed light on particular properties of random sequences in discrete time (X_n) which satisfy two key assumptions. First, (X_n)

Alexander Gordon
Department of Mathematics and Statistics, UNC Charlotte
Charlotte, NC 28223, e-mail: aygordon@uncc.edu

Isaac M. Sonin
Department of Mathematics and Statistics, UNC Charlotte
Charlotte, NC 28223 e-mail: imsonin@uncc.edu

is a bounded (sub)(super)martingale in forward or reverse time. Second, (X_n), at each moment n, takes no more that N values, where $N < \infty$. In other words, there exists a sequence of finite sets (G_n) such that $P(X_n \in G_n) = 1$ and $|G_n| \leq N < \infty$, $n \geq 1$. The class of all random sequences that have the latter property is denoted by $\mathscr{G}^N$; the class of all random sequences that have both properties is denoted by $\mathscr{M}^N$.

The random sequences from $\mathscr{M}^N$ appear very naturally, for example, in the study of *finite nonhomogeneous Markov chains* (MC). Let S be a countable set, (P_n) be a sequence of stochastic matrices, $Z = (Z_n)$ be a Markov chain from a family of MCs defined by a Markov model $(S,\ (P_n))$. Let Φ be the corresponding "tail" σ-algebra for Z, i.e. $\Phi = \cap_n F_{n\infty}$, where $F_{n\infty}$ is a σ-algebra generated by $(Z_n, Z_{n+1}, \ldots)$. It is easy to check that if $A \in \Phi$ and $\beta_n(i) = P(A|Z_n = i)$, then the r.s. (X_n), where $X_n = \beta(Z_n)$, is a *martingale* (in forward time). Another, even more important family of (sub)martingales can be obtained as follows. Let D_1 be a subset of S. Let us set, for $n \geq 1$ and $i \in S$,

$$\alpha_n(i) = \begin{cases} P(Z_1 \in D_1 | Z_n = i) & \text{if } P(Z_n = i) > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

It is easy to verify that the r.s. (Y_n) specified by $Y_n = \alpha_n(Z_n)$, $n \geq 1$, is a *martingale in reverse time*. If a subset D_1 is replaced by a sequence of sets $D_n \subseteq S$, and $\alpha_n(i)$ is defined as $\alpha_n(i) = P(Z_s \in D_s, s = 1, 2, \ldots, n | Z_n = i)$, then (Y_n) is a *submartingale* in reverse time. Obviously, if $|S| = N < \infty$ then the martingales and submartingales described above belong to $\mathscr{M}^N$.

The random sequences from $\mathscr{M}^N$ have much stronger properties than implied by the well-known Doob's convergence theorem, i.e. a theorem about the existence of limits of trajectories of a bounded (sub)martingale when time tends to infinity. Theorem 1 below describing these properties played a key role in the proof of the final part of a general theorem describing the behavior of a family of *finite nonhomogeneous* Markov chains defined by a finite Markov model $(S,\ (P_n))$, where S is a finite state space and (P_n) is a sequence of stochastic matrices. The striking feature of this theorem called a Decomposition-Separation (DS) theorem in [9], is that *no assumptions* on the sequence of stochastic matrices (P_n) are made. The DS theorem was initiated by a small paper of A. Kolmogorov [5] and was proved in steps in a series of papers: D. Blackwell [1], H. Cohn ([3] and other papers) and I. Sonin ([6], [7], [9] and other papers). We refer the reader to [9], where the final version of the DS theorem was presented and a brief survey of related results was given, and to a current expository paper [10]. Theorem 1 left an open problem described below and the main goal of our paper is to give an answer to that problem.

Before formulating Theorem 1 and the main result of this paper, Theorem 2, let us recall the well-known Doob's Upcrossing Lemma (see [2]), which lies at the foundation of Doob's convergence theorem. If (X_n) is a r.s., then the number of upcrossings of an interval (a, b) by a trajectory $X_1, X_2, \ldots$ on the infinite time interval is the number of times when a transition, maybe in a few steps, occurs from values less than a to values larger than b.

Doob's Lemma. *If $X = (X_n)$ is a (sub)martingale, then the expected number of upcrossings of a fixed interval (a,b) by the trajectories of X on the infinite time interval is bounded by* $\sup_n E(X_n - a)^+/(b-a)$.

Similar statements are true for the number of downcrossings and the number of crossings. The condition $\sup_n E(X_n - a)^+ < \infty$ obviously holds if (X_n) is bounded, so for simplicity we will consider only bounded random sequences with $0 \le X_n \le 1$ for all n.

Note that the width of the interval $(b-a)$ is in the denominator of the above estimate, so Doob's lemma does not imply that inside the interval there exists a *level* (constant barrier) d, such that the expected number of intersections of this level is finite, and in general such levels may not exist at all. But for the random sequences from $\mathscr{M}^N$ Doob's lemma can be substantially strengthened.

The following definition was introduced in [6]. A nonrandom sequence (d_n) is called a *barrier* for the r.s. $X = (X_n)$, if the *expected number of intersections* of (d_n) by trajectories of X on the infinite time interval is *finite*, i.e.

$$\sum_{n=1}^{\infty} [P(X_n \le d_n, X_{n+1} > d_{n+1}) + P(X_n > d_n, X_{n+1} \le d_{n+1})] < \infty. \tag{2}$$

Theorem 3 in [6] about the existence of barriers for processes with finite variation and a bounded number of values implies the following

Theorem 1. *Let (X_n) be a bounded r.s. from $\mathscr{M}^N$. Then inside each interval (a,b) there exists a barrier (d_n), $d_n \in (a,b)$, $n \ge 1$.*

An example in Sonin [9] shows that the barriers may not exist inside a given interval if a bounded martingale (X_n) takes a *countable* number of values, but for the random sequences from $\mathscr{M}^N$ Theorem 1 left an open problem.

Problem 1. Is it true that for any r.s. from $\mathscr{M}^N$ in any interval (a,b) there exists a constant barrier $d_n \equiv d$, $n \ge 1$?

For a r.s. $X = (X_n)$ defined on a finite or infinite time interval $\{1,2,\dots,T\}$, $T \le \infty$, with values in $[0,1]$, denote by $N_T(x,X)$ the expected number of intersections of level x by this sequence, i.e. the value of the sum in (2) when $d_n \equiv x$ for all n. We will omit the indication of X and T, if $T = \infty$ and X is clear from the context. Similarly, by $N^+(x)$ we denote the expected number of up-intersections, i.e. the first sum in (2) when $T = \infty$ and $d_n = x$ for all $n \ge 1$. Obviously, both $N(x)$ and $N^+(x)$ are finite or both are infinite.

In the sequel, the abbreviation MCM will mean a (nonhomogeneous) MC (X_i) defined on a finite or infinite time interval $[1,2,\dots,T]$, $T \le \infty$, which is also a *martingale*. We also assume that $0 \le X_i \le 1$ for all i. The main result of this paper is

Theorem 2. 1. *For any $X = (X_i) \in \mathscr{M}^2$, any value $x \in (0,1)$ is a constant barrier, i.e. $N(x) < \infty$ for all $x \in (0,1)$.*

2. For any $X = (X_i) \in \mathscr{M}^3$, $N(x) < \infty$ for Lebesgue almost every (a.e.) $x \in (0,1)$, and it is possible that $N(x) = \infty$ for all $x \in G$, where $G \subset (0,1)$ is a countable set.

3. *There is a MCM $X = (X_i) \in \mathscr{M}^4$, such that $N(x) = \infty$ for all $x \in (0,1)$.*

Remark 1. A Markov model $(S,(P_n))$ has a transparent deterministic interpretation (see [9]), and the DS theorem mentioned above has such an interpretation as well. According to this interpretation, the states of MC (Z_n) are represented by "cups" containing some solution (liquid), say tea. The entry $p_n(i,j)$ of the stochastic matrix P_n represents the proportion of the solution transferred from cup i to cup j at moment n. Correspondingly, $P(Z_n = i) = m_n(i)$ represents the volume of the solution in cup i at the moment n; $\alpha_n(i)$ introduced in (1) can be interpreted as the "concentration" of tea in the cup i at the moment n, and so on. Such a deterministic "colored" flow is the simplest example of an *irreversible process*. The DS theorem presented in the language of colored flows states that for any sequence (P_n) of $N \times N$ stochastic matrices the set of cups can be decomposed into a number of groups, with the decomposition possibly depending on time n, such that both the total volume and the concentration of tea in each group except possibly one tends to the limits. In the "exceptional group", the total volume tends to zero, but the concentration may oscillate. The total volume of solution exchanged between these groups is *finite* on the infinite time interval. The number of groups and the decomposition are unique (up to a certain equivalence) and depend only on the sequence (P_n). Problem 1 described above is equivalent to the question of whether such a decomposition can be provided by constant values of the concentration. Accordingly, our Theorem 2 can be reformulated as follows. If there are only two cups and the concentrations of tea in those cups do not tend to a common limit, then the total amount of liquid exchanged between the cups with the concentration higher than x (before the transfer) and lower than x (after the transfer), or vice versa, is finite for any x. For three cups – such values of x can be selected only from a subset of $(0,1)$ of full measure. For four or more cups, such x may not exist at all. We are going to present this and other possible interpretations, as well as some related results, in a separate paper.

The authors would like to thank Robert Anderson and Joseph Quinn who read the first version of this paper and made valuable comments.

2 Proof of Theorem 2. Cases $N = 2$ and $N = 3$

To simplify the presentation, we will consider only martingales in forward time. We denote by EX and $V(X)$ the expected value and respectively the variance of a random variable (r.v.) X.

Case $N = 2$. The definition of $X = (X_n) \in \mathscr{M}^2$ implies that at each moment n, $X_n \in \{a_n, b_n\}$, where $0 \le a_n < b_n \le 1$, a_n is decreasing and b_n is increasing. Let $a_\infty = \lim a_n$, $b_\infty = \lim b_n$. For any $x \le a_\infty$ or $x \ge b_\infty$, obviously, $N(x) = 0$. If $x \in (a_\infty, b_\infty)$, select d_1 and d_2 so that $a_\infty < d_1 < x < d_2 < b_\infty$, and let n_0 be a number such that $a_n < d_1$, $b_n > d_2$ for all $n \ge n_0$. Then

$$N^+(x) = c + \sum_{n=n_0}^{\infty} P(X_n \le d_1, X_{n+1} > d_2). \tag{3}$$

Proposition 1. *For any two r.v's X_1, X_2 and any two numbers $d_1 < d_2$,*

$$P(X_1 \le d_1, X_2 > d_2) \le E(X_1 - X_2)^2/(d_2 - d_1)^2. \tag{4}$$

Proof. The assertion of Proposition 1 follows immediately from the implications $(X_1 \le d_1, X_2 \ge d_2) \subset (|X_1 - X_2| \ge h)$, $h = d_2 - d_1$, and the Chebyshev's inequality, $P(|Y| \ge h) \le EY^2/h^2$ for any r.v. Y.

To prove part 1) of Theorem 2, note that for a martingale (X_n) we have $E(X_{n+1}|X_n) = X_n$, and $V(X_{n+1} - X_n) = E(X_{n+1} - X_n)^2 = EX_{n+1}^2 - E{X_n}^2$ and hence for a bounded martingale (X_n), $0 \le X_n \le 1$,

$$\sum_{n=k}^{T-1} E(X_{n+1} - X_n)^2 = EX_T^2 - E{X_k}^2 \le 1. \tag{5}$$

Then formula (5) and Proposition 1 imply that the sum in (3) is finite. Part 1) of Theorem 2 is proved.

For any r.v. X with $EX = m$, let us denote $M^+(X) = E(X - m)^+$. Then $E(X - m)^+ = E(m - X)^+$, and $M(X)$, the mean absolute deviation of X, is

$$M(X) = E|X - m| = E(X - m)^+ + E(m - X)^+ = 2M^+(X). \tag{6}$$

Let us also put $M(X|c) = E|X - c|$ and $M^+(X|c) = E(X - c)^+$. If (X_n) is a martingale, then the equality $E(X_{n+1}|X_n) = X_n$ and (6) imply that

$$M(X_{n+1}|X_n) = 2M^+(X_{n+1}|X_n). \tag{7}$$

To study the case of $N = 3$, we need some simple properties of r.v's and martingales with two and three values. They are described in Propositions 2 - 4. Let $X \in G^2$, i.e. an r.v. with two values a and b, $a < b$, $b - a = d$ and $P(X = b) = p$, $P(X = a) = q = 1 - p$. Then it is easy to check that the following statement is true.

Proposition 2. *If $X \in G^2$, then $M(X) = 2pqd$, $V(X) = pqd^2$, and hence*

$$M^+(X) = V(X)/d. \tag{8}$$

Let X be an r.v. with three values $a \le e \le b$, and $P(X = b) = p$, $P(X = e) = r$, $P(X = a) = q$, where $p + r + q = 1$. Let us denote by X^U an r.v. obtained from X by averaging the two upper values, i.e. X^U takes two values a and $b' = (er + bp)/(r + p)$ with probabilities q and $p' = r + p$. Similarly, we define an r.v. X^L as an r.v. obtained from X by averaging the two lower values, i.e. X^L takes two values $a' = (aq + er)/(q + r)$ and b with probabilities $q' = q + r$ and p. It is easy to check that for such r.v's the following statement is true

Proposition 3. a) $a \le a' \le e \le b' \le b$, $EX = EX^A$, $V(X^A) \le V(X)$, $A = U$ *or* L,

b) $M(X^U) \le M(X)$, $M^+(X^U) \le M^+(X)$, *with equalities when* $EX \le e$,

c) $M(X^L) \le M(X)$, $M^+(X^L) \le M^+(X)$ *with equalities when* $EX \ge e$.

For a r.s. $(X_n) \in \mathscr{M}^3$, i.e. for a martingale with three values $a_n \le e_n \le b_n$ at moment n, similarly to Proposition 3, it is easy to obtain

Proposition 4. a) *If* $(X_n) \in \mathscr{M}^3$, then $(X_n^A) \in \mathscr{M}^2$, $V(X_{n+1}^A - X_n) \le V(X_{n+1} - X_n)$, $A = U$ or L,

b) $M^+(X_{n+1}^U|X_n) \le M^+(X_{n+1}|X_n)$, *with equality when* $X_n \le e_{n+1}$,

c) $M^+(X_{n+1}^L|X_n) \le M^+(X_{n+1}|X_n)$, *with equality when* $X_n \ge e_{n+1}$.

Now we can prove part 2) of Theorem 2 (case $N = 3$). The situation in this case is substantially different from $N = 2$ and $N > 3$. It is possible to have $N(x) = \infty$, for example, for all rational numbers; nevertheless the Lebesque measure of the set of all such x is always zero. We prove here only the latter statement.

Lemma 2. *For any r.s.* (X_n), $0 \le X_n \le 1$, *and* $T \le \infty$,

$$\int_0^1 N_T^+(x)dx = \sum_{n=1}^{T-1} EM^+(X_{n+1}|X_n). \tag{9}$$

Proof. The proof follows immediately from the definition of $N_T^+(x) = \sum_{n=1}^{T-1} P(X_n \le x, X_{n+1} > x)$ and the equalities: 1) $P(X_n \le x, X_{n+1} > x) = EI(X_n, X_{n+1}|x)$, where $I(A,B|x) = 1$ if $A \le x < B$, and 0 otherwise, 2) $\int_0^1 I(A,B|x)dx = (B-A)^+$ and 3) $E(X_{n+1} - X_n)^+ = EM^+(X_{n+1}|X_n)$.

Note that if, given a, b, $0 \le a < b \le 1$, in the left side of (9) we change the limits of integration from 0 and 1, to a and b, i.e. consider $\int_a^b$, then the equality (9) remains true with $M^+(X_{n+1}|X_n)$ replaced by $M^+(X_{n+1}|X_n, a, b) \equiv E(\min(b, X_{n+1}) - \max(a, X_n))^+$. For simplicity we will denote $M^+(X_{n+1}|X_n, 2\varepsilon, 1-2\varepsilon)$ as $M^+(X_{n+1}|X_n, \varepsilon)$.

To prove that the integral in (9) is finite and hence $N_T^+(x) < \infty$ almost surely, we will show that for any $(X_n) \in \mathscr{M}^3$ and any ε, $0 < \varepsilon < 1/4$,

$$M^+(X_{n+1}|X_n, \varepsilon) \le V((X_{n+1} - X_n)|X_n)/\varepsilon. \tag{10}$$

If (X_n) is a martingale, then $EV((X_{n+1} - X_n)|X_n) = E(E((X_{n+1} - X_n)^2|X_n)) = E(X_{n+1} - X_n)^2 = V(X_{n+1} - X_n)$. Since series in (5) is convergent, estimate (10) will prove that the integral in (9) is finite.

Let $(X_n) \in \mathscr{M}^3$ and $\{a_n, e_n, b_n\}$ be the ordered set of possible values of X_n at the moment n, $0 \le a_n \le e_n \le b_n \le 1$. The definition of a martingale again implies that the sequence (a_n) can only decrease, the sequence (b_n) can only increase but the sequence (e_n) may oscillate between a_n and b_n. Let $\varepsilon > 0$ and n_0 be a number such that $a_n < \varepsilon$, and $1 - \varepsilon < b_n$ for all $n \ge n_0$. In the sequel we will consider only $n \ge n_0$.

If $X_n > 1 - 2\varepsilon$, then, obviously, $M^+(X_{n+1}|X_n, \varepsilon) = 0$. Since $b_n > 1 - \varepsilon$, we need to consider further only the cases $X_n = a_n$ or $X_n = e_n$.

If $X_n = a_n$ or e_n, and $e_{n+1} \le 2\varepsilon$, then $M^+(X_{n+1}|X_n, \varepsilon) = M^+(X_{n+1}^L|X_n, \varepsilon)$. Using (8) applied to $X_{n+1}^L \in \mathscr{M}^2$, and part a) of Proposition 4, we have $EM^+(X_{n+1}^L|X_n, \varepsilon) \le V(X_{n+1}^L - X_n)/(b_{n+1} - a'_{n+1}) \le V(X_{n+1} - X_n)/(b_{n+1} - a'_{n+1})$. Since $a'_{n+1} \le e_{n+1} \le 2\varepsilon$ and $b_{n+1} > 1 - \varepsilon$, we have $b_{n+1} - a'_{n+1} \ge 1 - 3\varepsilon \ge \varepsilon$, and thus (10) holds.

If $e_{n+1} \geq 2\varepsilon$ and $X_n \leq 2\varepsilon$ then $M^+(X_{n+1}|X_n,\varepsilon) = M^+(X^U_{n+1}|X_n,\varepsilon)$. Using (8) applied to $X^U_{n+1} \in \mathscr{M}^2$, and point a) of Proposition 4, we have $EM^+(X^U_{n+1}|X_n,\varepsilon) \leq V(X^U_{n+1} - X_n)/(b'_{n+1} - a_{n+1}) \leq V(X_{n+1} - X_n)/(b'_{n+1} - a_{n+1})$. Since $b'_{n+1} \geq e_{n+1} \geq 2\varepsilon$ and $a_{n+1} < \varepsilon$, we have $b'_{n+1} - a_{n+1} \geq \varepsilon$, and thus (10) holds. Remaining is the case where $2\varepsilon \leq X_n = e_n \leq 1 - 2\varepsilon$ and $e_{n+1} \geq 2\varepsilon$.

If $X_n = e_n \leq e_{n+1}$, then $M^+(X_{n+1}|X_n,\varepsilon) = M^+(X^U_{n+1}|X_n,\varepsilon)$. Using (8) applied to $X^U_{n+1} \in \mathscr{M}^2$, and point a) of Proposition 4, we have $EM^+(X^U_{n+1}|X_n,\varepsilon) \leq V(X^U_{n+1} - X_n)/(b'_{n+1} - a_{n+1}) \leq V(X_{n+1} - X_n)/(b'_{n+1} - a_{n+1})$. Since $b'_{n+1} \geq e_{n+1} \geq 2\varepsilon$ and $a_{n+1} \leq \varepsilon$, we have $b'_{n+1} - a_{n+1} \geq \varepsilon$ and thus (10) holds. If $X_n = e_n \geq e_{n+1}$ then $M^+(X_{n+1}|X_n,\varepsilon) = M^+(X^L_{n+1}|X_n,\varepsilon)$. Using (8) applied to $X^L_{n+1} \in \mathscr{M}^2$, and point a) of Proposition 4, we have $EM^+(X^L_{n+1}|X_n,\varepsilon) \leq V(X^L_{n+1} - X_n)/(b_{n+1} - a'_{n+1}) \leq V(X_{n+1} - X_n)/(b_{n+1} - a'_{n+1})$. Since $a'_{n+1} \leq e_{n+1} \leq e_n \leq 1 - 2\varepsilon$ and $b_{n+1} \geq 1 - \varepsilon$, we have $b_{n+1} - a'_{n+1} \geq \varepsilon$ and thus (10) holds. Part 2 of Theorem 2 is also proved.

3 Proof of Theorem 2. Case $N > 3$. An example

We prove part 3) of Theorem 2 by a direct construction of the MCM $X = (X_i)$ for $N = 4$. First, we construct an auxiliary MCM $U = (U_i)$. Let $(a_k),(b_k), k = 1,2...$ be two deterministic sequences such that:

$$1 > a_1 > a_2 > ... > 0,\ a_1 < b_1 < b_2 < ... < 1,\ \lim a_k = 0,\ \lim b_k = 1. \tag{11}$$

Given such sequences (a_k) and (b_k), we can define a MC $U = (U_i)$, $i = 1,2,...,$ such that $P(U_1 = b_1) = 1$, $U_{2k-1} \in \{a_k, b_k, 1\}$, $U_{2k} \in \{a_k, 1\}$, and the transition probabilities $u_i(x,y)$ are: $u_k(1,1) = 1$, $u_{2k-1}(a_k,a_k) = 1$, $u_{2k-1}(b_k,1) + u_{2k-1}(b_k,a_k) = 1$, $u_{2k}(a_k,b_{k+1}) + u_{2k}(a_k,a_{k+1}) = 1$, $k \geq 1$. To obtain not just a MC but a (unique) MCM, it is sufficient to define

$$u_{2k-1}(b_k,1) = \frac{b_k - a_k}{1 - a_k},\ u_{2k}(a_k,b_{k+1}) = \frac{a_k - a_{k+1}}{b_{k+1} - a_{k+1}},\ k \geq 1.$$

Assumptions (11) imply that this is possible and that $m_k = P(U_{2k-1} = b_k) > 0$ for all $k \geq 1$.

The MCM $U = (U_i)$ will serve as a "frame sequence" for MCM $X = (X_i)$, i.e. (X_i) will consist of "blocks" (X^k_i), $k \geq 1$, where each (X^k_i) is a MCM defined on a time interval $[t_k, e_k]$, $t_1 = 1$, $t_{k+1} = e_k + 1$, $k \geq 1$, and each block is "inserted" into a constructed above "frame sequence" $U = (U_i)$ so that the time interval $[2k-1, 2k]$ "stretches" into the time interval $[t_k, e_k]$, $k \geq 1$. More precisely, the values and the transition probabilities $p_i(x,y)$ for MCM (X_i) are defined as follows: $P(X_1 = X^1_1 = b_1) = 1$, $P(X_{e_1} = X^k_{e_1} \in \{a_1, 1\}) = 1$. Any other k-th block, $k \geq 2$, has three entrance points $\{a_k, b_k, 1\}$ and two exit points $\{a_k, 1\}$, i.e. $P(X_{t_k} = X^k_{t_k} \in \{a_k, b_k, 1\}) = 1$, and $P(X_{e_k} = X^k_{e_k} \in \{a_k, 1\}) = 1$. The state 1 is an absorbing state, $p_i(1,1) = 1$ for all $i \geq 1$. The transition probabilities *between* blocks, i.e. at moments e_k, $k \geq 1$, are defined using the transition probabilities from MCM U : if $i = e_k$ then $p_i(a_k, y) = u_{2k}(a_k, y)$, where $y = a_{k+1}$ or b_{k+1}. The transition probabilities of k-th block $p_i(x,y)$, $t_k \le i < e_k$, are as follows: $p_i(a_k, a_k) = 1$ for all i, the other transition probabilities are the "shifted" probabilities from MCM $Y^k = (Y_i^k), k \ge 1$, where (Y_i^k) is defined on the time interval $[1, T_k]$, $T_k = e_k - t_k + 1$, i.e. $p_{t_k+i-1}(x,y) = q_i^k(x,y)$, where $i = 1, 2, ..., T_k$ and $q_i^k(x,y)$ are transition probabilities for (Y_i^k). We say that block X^k is obtained from a block Y^k by a *shift* from interval $[1, T_k]$ to interval $[t_k, e_k]$, $e_k = t_k + T_k - 1$.

The structure of each MCM $Y^k = (Y_i^k), k = 1, 2, ...$ is similar and its properties are described in Lemma 1 which is the key element of our construction.

Lemma 1. *For every tuple* $\beta = (a, b, \varepsilon, C)$, $0 \le a < b < 1$, $0 < \varepsilon < b - a$, $C > 0$, *there is a MCM* $Y = (Y_i)$ *defined on a finite time interval* $[1, 2, ..., T]$, $T = T(\beta)$, *and such that*

1) $P(Y_1 = b) = 1$, $P(Y_T \in \{a, 1\}) = 1$, *and for all other* i, $1 < i < T$, Y_i *takes no more than three values* $r_i, s_i, 1, \quad a \le r_i < s_i \le 1$.

2) $N_T^Y(x) \ge C$ *for each* $x \in (a + \varepsilon, b)$.

We will prove Lemma 1 later. Assuming that Lemma 1 holds, we next construct a MCM (X_i) satisfying part 3) of Theorem 2.

Let (ε_k) be a sequence, $\varepsilon_k > 0$, $\lim \varepsilon_k = 0$, and let $(a_k), (b_k)$ be two sequences satisfying conditions (11). Let (U_i) be a corresponding "frame" MCM, $u_i(x,y)$ its transitional probabilities, $i = 1, 2, ...$ and $m_k = P(U_{2k-1} = b_k) > 0, k \ge 1$. Given $k \ge 1$, let $Y^k = (Y_i^k)$, $i = 1, 2, ..., T_k$, be a MCM satisfying the conditions of Lemma 1 with parameters $(a_k, b_k, \varepsilon_k, C_k)$, where $C_k = 1/m_k$. We define sequences (t_k) and (e_k) by: $t_1 = 1$, $e_k = t_k + T_k - 1$, $t_{k+1} = e_k + 1$, $k \ge 1$. Let us denote by (X_i) the combined MC consisting of blocks X^k obtained by the corresponding shift from Y^k and connected by the frame sequence (U_i) as described above.

Let us denote by $N^k(x) \equiv N_{T_k}^{Y_k}(x)$ the expected number of intersections of level x by a r.s. $(Y_i^k), k \ \ge 1$, and $N(x) \equiv N^X(x)$ the expected number of intersections of level x by a r.s. $X = (X_i)$. By our choice of C_k we have $N^k(x) \ge 1/m_k$ for each $x \in (a_k + \varepsilon_k, b_k)$, where $m_k = P(X_{t_k} = b_k), k \ge 1$. Let $x \in (0,1)$ and let $k(x)$ be a number such that $x \in (a_k + \varepsilon_k, b_k)$ for all $k \ge k(x)$. Then by our construction

$$N(x) \ge \sum_{k=1}^{\infty} P(X_{t_k} = b_k) N^k(x) \ge \sum_{k \ge k(x)}^{\infty} m_k / m_k = \infty.$$

Thus, to prove part 3 of Theorem 2 we only need to prove Lemma 1. From now on, the numbers (indices) i, k and n and such notation as $p_i(x,y)$ have a new meaning.

We prove Lemma 1 for a special case when $a = 0$, $b = \frac{1}{2}$. The general case requires only minor changes in notation.

We will construct (Y_i) combining the finite number of MCM, having similar structure. To avoid confusion with the "blocks" used above, we call these MCM *modules*. Each module $(L_i^{k,r})$ is a MCM characterized by two parameters (k, r), $k \ge 1$, $0 \le r < 1$, and defined on the time interval $[1, 2, ..., k]$.

First we describe the *standard module* with parameters $(k, 0)$. This is a MC $(L_i^{k,0}) \equiv (S_i)$ defined on $[1, 2, ..., k]$, and taking at each moment i two values 0 and

s_i, where (s_i) is a deterministic sequence given by formula

$$s_i = \frac{1}{k+1-i}, \qquad i = 1,2,...,k. \tag{12}$$

Obviously $0 < \frac{1}{k} = s_1 < ... < s_{k-1} = \frac{1}{2} < s_k = 1$, and s_i satisfy $s_{i+1}(1-s_i) = s_i$. Point s_1 is an initial point for a r.s. (S_i), i.e. $P(S_1 = s_1) = 1$. The transition probabilities $p_i(x,y)$ are defined as follows. State 0 is absorbing for all i, i.e. $p_i(0,0) = 1$, $i = 2,3,...,k-1$. The other transition probabilities are given by

$$p_i(s_i,0) = s_i, p_i(s_i,s_{i+1}) = 1-s_i = \frac{k-i}{k+1-i}, \quad i = 1,...,k-1. \tag{13}$$

It is easy to see that $E(S_{i+1}|S_i = s_i) = s_{i+1}(1-s_i) = s_i$. Therefore, the r.s. (S_i) is also a martingale, i.e (S_i) is a MCM.

It is easy to check that

$$P(S_i = s_i) = \prod_{j=1}^{i-1} p_j(s_j,s_{j+1}) = \frac{k+1-i}{k} = \frac{1}{ks_i}, \;\; i = 1,2,...,k. \tag{14}$$

Let us denote by $N^k(x)$ the expected number of intersections of level x by the r.s. (S_i). If $x \in (s_i,s_{i+1})$, $i = 1,2,...,k-2$, then every trajectory can intersect x on the way up and after that on the way down, so $N^k(x) = P(S_{i+1} = s_{i+1}) + P(S_{i+1} = s_{i+1}, S_k = 0) = 2P(S_{i+1} = s_{i+1}) - P(S_k = 1) = \frac{2(k-i)-1}{k} \geq 1/ks_i$. These relations imply that

$$N^k(x) \geq f^k(x), \text{ where } f^k(x) = \frac{1}{kx}, \text{ if } \frac{1}{k} < x \leq 1/2. \tag{15}$$

The *module* $(L_i^{k,r})$ with parameters (k,r), $0 \leq r < 1$, $k \geq 1$ is a r.s. defined on the finite time interval $i = 1,2,..,k$ by equalities

$$L_i^{k,r} = r + (1-r)S_i, \qquad i = 1,2,...,k, \tag{16}$$

where $(S_i) = (L_i^{k,0})$ is a standard module with parameters $(k,0)$.

Formula (16) implies that the initial point for $(L_i^{k,r})$ is

$$r + \frac{1-r}{k} \tag{17}$$

and that $(L_i^{k,r})$ is also a MCM with the same transitional probabilities as in (13) but with possible values r and $r+(1-r)s_i$ instead of 0 and s_i. The value r is the smallest of possible values for this module, so later we will refer to r as to the "floor" of this module. The intersection function $N^{k,r}(x)$ for $(L_i^{k,r})$, instead of (15), satisfies the inequality

$$N^{k,r}(x) \geq f^k\left(\frac{x-r}{1-r}\right), \quad r + \frac{1-r}{k} \leq x < r + \frac{1-r}{2} = \frac{1+r}{2}. \tag{18}$$

Formula (14) for $i = k$ and formula (16) imply that

$$P(L_k^{k,r} = 1) = \frac{1}{k}, \quad P(L_k^{k,r} = r) = \frac{k-1}{k}. \tag{19}$$

Now we will construct a sequence of MCM (Y_i^n), $n = 1,2,...$, $i = 1,2,...,T_n$, and we will show that for any $\varepsilon > 0$ and any C for large n each of these MCM will satisfy the condition of Lemma 1. Each (Y_i^n) consists of n modules connected subsequently, each with parameters (k_j, r_j), $j = 1,2,...,n$. The parameters (k_j, r_j), $j = 1,2,...,n$, given $n = 1,2,...$, are selected as follows

$$k_j = n + j, \quad r_j = \frac{n-j}{2n}, \quad j = 1,...,n. \tag{20}$$

It is easy to check that

$$r_{j-1} = r_j + \frac{1-r_j}{k_j} = r_j + \frac{1}{2n}, \quad j = 2,...,n. \tag{21}$$

Thus, for each n, points r_j divide the interval $(0, \frac{1}{2})$ into n equal parts of size $1/2n$ and the interval $(1-r_j, 1)$ contains k_j subintervals of this size. Let us denote, for the sake of brevity, $(L_i^{k_j,r_j})$ by (L_i^j). Formulas (17) and (21) imply that the floor r_j of the module (L_i^{j-1}) serves as the initial point for the next module (L_i^j).

Let $T_n = \sum_{j=1}^n k_j - n + 1$ be the total length of the time interval where these modules are sequentially defined. We define $(Y_i) \equiv (Y_i^n)$, $1 \le i \le T_n$, as follows. State 1 is absorbing for all i. At moment 1 r.s. (Y_i) starts at $r_0 = \frac{1}{2}$ and on the time interval $[1, k_1]$ coincides with the module $(L_i^{k_1,r_1}) = (L_i^1)$. Then, at moment k_1 according to (19), we have

$$P(Y_{k_1} = 1) = P(L_{k_1}^1 = 1) = \frac{1}{k_1}, \; P(Y_{k_1} = r_1) = P(L_{k_1}^1 = r_1) = \frac{k_1 - 1}{k_1}. \tag{22}$$

On the time interval $[k_1, k_1 + k_2 - 1]$ r.s. (Y_i) stays at 1 with probability $\frac{1}{k_1}$ and with probability $m_2 = \frac{k_1-1}{k_1}$ coincides with the module (L_i^2). As mentioned above, the floor r_1 of the module (L_i^1) serves as the initial point for the next module (L_i^2). And so on. Obviously, MCM (Y_i) satisfies the condition 1) of Lemma 1 with $b = \frac{1}{2}$ and $a = 0$.

From the above construction, using the last equality in formula (14) for $i = k = k_j$ and denoting $m_0 = 1$, we also obtain that for $j = 1,2,...,n$,

$$m_j = P(Y_{k_1+...+k_j-j+1} = r_j) = m_{j-1} P(L_{k_j}^j = r_j) = m_{j-1} \frac{k_j - 1}{k_j} = \frac{n}{n+j}. \tag{23}$$

Our last step is to estimate $N^{(n)}(x)$, the expected number of intersections of level x by the constructed MCM (Y_i^n), $i = 1,2,...,T_n$, and to show that for any ε, $0 < \varepsilon <$

$1/2, \lim_n N^{(n)}(x) = \infty$ uniformly for all x, $\varepsilon \le x \le 1/2$. Therefore, given any number C, for sufficiently large n, MCM (Y_i^n) will satisfy the condition 2) of Lemma 1.

By our construction, we have obviously $N^{(n)}(x) = \sum_{j=1}^n m_{j-1} N^j(x)$, where $N^j(x)$ is the expected number of intersections of level x by module (L_i^j). Using for each j the estimate (18) with $k = k_j$ and $r = r_j$ taken from (20), and taking into account that by (23), $\frac{1}{2} \le m_j \le 1$ for all j, we obtain that

$$N^{(n)}(x) \ge \frac{1}{2} \sum_{j=1}^{n} f^{k_j}\left(\frac{x - r_j}{1 - r_j}\right), \tag{24}$$

where $f^{k_j}\left(\frac{x-r_j}{1-r_j}\right)$ is defined by (15) (see also (18)) for $r_{j-1} \le x < (1+r_j)/2$ and for other x's can be defined to be equal to zero. Hence $f^{k_j}\left(\frac{x-r_j}{1-r_j}\right) \ge \frac{1-r_j}{k_j(x-r_j)} = \frac{1-r_j}{2n(x-r_j)}$ for $r_{j-1} \le x \le (1+r_j)/2$. Using formulas (20) and (21), we obtain that for any $x, 0 < x \le 1/2$,

$$N^{(n)}(x) \ge \frac{1}{2} \sum_{j: r_{j-1} \le x}^{n} \frac{1}{2n(x - r_j)} = \frac{1}{2} \sum_{j: n-j+1 \le 2nx}^{n} \frac{1}{2n(x - \frac{n-j}{2n})} = \frac{1}{2} \sum_{k=0}^{[2nx]-1} \frac{1}{2n(x - \frac{k}{2n})}, \tag{25}$$

where $[a]$ is an integer part of a.

The last sum is just a Riemann sum of the integral $\int_0^x \frac{dy}{x-y} = \infty$. Thus, for large n the sum $N^{(n)}(x)$ in (25) can be made arbitrarily large uniformly for all x, $0 < \varepsilon \le x \le 1/2$. This proves Lemma 1 and therefore part 3) of Theorem 2.

Remark. A slightly different but similar construction proves the analog of Theorem 2 for the case where (X_i) is a martingale in reverse time.

References

1. Blackwell D. (1945). Finite nonhomogeneous Markov chains. *Ann. Math*. **46** 594-599.
2. Strook D. W. (1999). Probability Theory, an analytic view. Cambridge Univ. Press.
3. Cohn H. (1989). Products of stochastic matrices and applications, *Int. J. Math. Sci*. **12** 209-333.
4. Cohn H. (1976). Finite nonhomogeneous Markov chains: asymptotic behavior. *Adv. Appl. Prob*. **8** 502-516.
5. Kolmogoroff A. N. (1936). Zur Theorie der Markoffschen Ketten. *Math. Ann. Bd*. **112**. Selected Works of A.N. Kolmogorov, v. 2, Prob. Theory and Math. Stat., ed. A.N. Shiryaev, Kluwer Acad. Publ.
6. Sonin I. (1987). Theorem on separation of jets and some properties of random sequences. Stochastics, **21**, pp. 231-250.
7. Sonin I. (1996). The Asymptotic Behaviour of a General Finite Nonhomogeneous Markov Chain (The Decomposition-Separation Theorem), IMS, Lecture Notes-Monograph Series, v. 30, Statistics, Probability and Game Theory, papers in Honor of David Blackwell, eds. T. S. Fergusson, L. S. Shapley and J. B. MacQueen, 337-346.

8. Sonin I. M. (1991b). An arbitrary nonhomogeneous Markov chain with bounded number of states may be decomposed into asymptotically noncommunicating components having the mixing property. *Theory Probab. Appl*. **36** 74-85.
9. Sonin I. (1997). On Some Asymptotic Properties of Nonhomogeneous Markov Chains and Random Sequences with Countable Number of Values, in Statistics and Control of Stochastic Processes, The Liptser Festschrift, Proc. of Steklov Math. Inst. Seminar, eds. Y. Kabanov, B. Rozovskii, A. Shiryaev, World Scientific, 297–313.
10. Sonin, I. (2007). The Decomposition-Separation Theorem for Finite Nonhomogeneous Markov Chains and Related Problems, to appear in Markov Processes and Related Fields: A Festschrift in Honor of Thomas G. Kurtz, eds. S. Ethier, J. Feng and R.H. Stockbridge, IMS.