

*MULTIPLE RECURRENCE
FOR DISCRETE TIME MARKOV PROCESSES*

BY

A. IWANIK (WROCLAW)

In [1] Furstenberg proved a multiple recurrence theorem for measure automorphisms of standard probability spaces; this provided a measure theoretic proof of the celebrated Szemerédi theorem in combinatorial number theory. One version of Furstenberg's theorem asserts that if τ_1, \dots, τ_l are commuting measure preserving transformations of a standard probability space (X, Σ, μ) then

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(\tau_1^{-n} A \cap \dots \cap \tau_l^{-n} A) > 0$$

whenever $\mu(A) > 0$ (Thm. 7.14 in [2]).

Another interesting multiple recurrence result is the purely topological multiple Birkhoff recurrence theorem (Thm 2.6 in [2]). It states that for any finite set of commuting continuous transformations τ_1, \dots, τ_l on a compact metric space X there exist a point $x \in X$ and a sequence $n_k \rightarrow \infty$ such that $\tau_i^{n_k} x \rightarrow x$ simultaneously for $i = 1, \dots, l$. This theorem, although pertaining essentially to topological dynamics yields a new proof of the classical van der Waerden's theorem on arithmetic progressions and a number of other combinatorial results (see [2] and [3]).

The aim of the present paper is to apply the aforementioned recurrence theorems to the recurrence behavior of discrete time Markov processes.

In Section 2 a single Markov operator is considered and certain recurrence properties of the trajectories are proved, similar to those occurring for the iterations of a single continuous mapping. Our methods rely on a version of Furstenberg's multiple recurrence theorem represented by Lemma 1 and on applying an upper continuous function $F(x)$ which represents a measure of multiple recurrence. A similar but somewhat simpler function was exploited in [3] and [2] in the case of continuous mappings. We will also make use of the interplay between the Markov operator acting on the underlying phase space X and the shift transformation S on the space of

trajectories $\Omega = X \times X \times \dots$ of the canonical Markov process (see [5] for details).

In Section 3 we introduce a notion of multiple recurrence for families of Markov operators on a compact metric space and prove that, under additional assumptions, doubly recurrent points exist for two commuting Markov operators.

1. Preliminaries. Let (X, Σ) be a measurable space. A transition probability $p(x, \cdot)$ induces the Markov operator

$$Tf(x) = \int f(y) p(x, dy)$$

acting on the space of all bounded measurable functions. If X is a topological space endowed with the Borel σ -algebra then $p(x, \cdot)$ is called *Feller* if Tf is continuous for every continuous (bounded) f . (Only Feller transition probabilities will be considered on topological spaces.) A nonempty closed subset F of X is called *invariant* if $p(x, F) = 1$ for every $x \in F$. If X is compact Hausdorff then by Zorn's lemma minimal invariant subsets always exist. T is called *irreducible* if X is already minimal.

For any transition probability $p(x, \cdot)$ and any probability measure ν on (X, Σ) a canonical Markov process $(\xi_n)_{n \geq 0}$ is constructed on the product space $\Omega = X^{N_0}$ where $N_0 = \{0, 1, 2, \dots\}$. This means that $\xi_n(\omega) = \omega_n$ and there exists a unique probability measure P_ν on Ω (called *Markov measure*) such that

$$P_\nu \{ \xi_0 \in A_0, \dots, \xi_n \in A_n \} = \int \chi_0 T(\chi_1 T(\dots \chi_{n-1} T\chi_n)) d\nu$$

where $n \geq 0$, $A_i \in \Sigma$, and χ_i denote the characteristic functions of A_i ($i = 0, 1, \dots, n$). A probability measure μ on (X, Σ) is called *invariant* if

$$\int p(x, A) d\mu(x) = \mu(A)$$

for every $A \in \Sigma$. If μ is invariant then P_μ is *S-invariant* on Ω where S denotes the shift transformation; the canonical Markov process is then stationary with respect to P_μ . For Feller transition probabilities on compact spaces invariant probability measures always exist.

2. Multiple recurrence for a single process. The following lemma is a consequence of Furstenberg's multiple recurrence theorem ([2], Thm. 7.14).

LEMMA 1. *Let $p(x, \cdot)$ be a transition probability on a measurable space (X, Σ) and let μ be an invariant probability measure. If $\mu(A) > 0$ then for every $l \geq 1$ there exists $x \in A$ such that*

$$P_x \{ \xi_m \in A, \xi_{2m} \in A, \dots, \xi_{lm} \in A \text{ for infinitely many } m \}' > 0.$$

Proof. In order to apply Furstenberg's theorem we first reduce the problem to a separable space. By a standard argument, there exists a countably generated sub- σ -algebra Σ_0 of (the equivalence classes of) measurable sets such that (1) $A \in \Sigma_0$ and (2) the space $L^\infty(X, \Sigma_0, \mu)$ is closed under the action of the Markov operator T induced by $p(x, \cdot)$. Now let X_0 be a standard probability space such that the Boolean algebra of measurable sets (modulo null sets) in X_0 and in (X, Σ_0, μ) are isomorphic. The action of T can be carried over to X_0 and is induced by a transition probability on X_0 . We form the canonical Markov process $(\xi_n)_{n \geq 0}$ on

$$\Omega_0 = X_0 \times X_0 \times \dots$$

Since (Ω_0, P_μ) is a standard probability space and P_μ is shift invariant, Furstenberg's theorem is applicable to the iterations S, S^2, \dots, S^l and to the set

$$\{\xi_0 \in A\} \subset \Omega_0.$$

Therefore, for infinitely many m 's, we have

$$P_\mu(B_m) \geq \alpha > 0$$

where

$$B_m = \{\xi_0 \in A, \xi_m \in A, \dots, \xi_{lm} \in A\}.$$

In particular, $P_\mu(B) > 0$ where $B = \limsup B_m$. Since

$$P_\mu(B) = \int_A P_y(B) d\mu(y),$$

$P_y(B)$ must be positive on a subset of positive μ measure in A . Therefore, there exists a point $x \in A$ (in both X_0 and X) such that $P_x(B) > 0$. This proves the lemma.

Now we let (X, d) be a metric space and $p(x, \cdot)$ be a Feller transition probability on X . Consider the following nonnegative real valued function on the space X :

$$F(x) = \inf_{m \geq 1} \inf \{\alpha: P_x \{d(\xi_{im}, x) < \alpha, i = 1, \dots, l\} > 0\}.$$

LEMMA 2. F is upper semicontinuous.

Proof. Suppose $F(x) < \delta$. We want to show that the same inequality holds in a neighborhood of x . There exist $m \geq 1$ and $\alpha < \delta$ such that $P_x(B) > 0$ where

$$B = \{d(\xi_{im}, x) < \alpha, i = 1, \dots, l\}.$$

Since B is a finite dimensional open cylinder, there exists a continuous function $0 \leq f \leq 1$ such that

$$0 \leq f(\omega_m) f(\omega_{2m}) \dots f(\omega_{lm}) \leq \chi_B(\omega)$$

and

$$\int f(\omega_m) f(\omega_{2m}) \dots f(\omega_{lm}) dP_x(\omega) > 0.$$

The last integral is the continuous function

$$T^m(f T^m(\dots f T^m f))(x)$$

so the inequality $P_y(B) > 0$ holds in a neighborhood U of x . On the other hand, if z is in the open ball V of radius $\delta - \alpha$ and center x then $d(u, z) < \delta$ whenever $d(u, x) < \alpha$. This implies

$$P_y \{d(\xi_{im}, z) < \delta, i = 1, \dots, l\} > 0$$

for all $y \in U, z \in V$. In particular

$$P_y \{d(\xi_{im}, y) < \delta, i = 1, \dots, l\} > 0,$$

or

$$F(y) < \delta, \quad \text{in } U \cap V.$$

Now we are in a position to prove the existence of "multiple recurrent points" for stationary Feller processes. The following theorem is reminiscent of the multiple Birkhoff recurrence theorem for iterates of a single continuous map.

THEOREM 1. *Let $p(x, \cdot)$ be a Feller transition probability on a complete metric space (X, d) . Assume an invariant probability measure μ exists and $X = \text{supp } \mu$. Then there exists a point $x \in X$ such that for every $l \geq 1$ and every neighborhood U of x*

$$P_x \{\xi_m \in U, \xi_{2m} \in U, \dots, \xi_{lm} \in U\} > 0$$

for some $m \geq 1$.

Proof. We prove that the set of all such points x is residual. It suffices to prove the assertion for a single $l \geq 1$.

For every open ball V in X we have $\mu(V) > 0$, therefore by Lemma 1 there exist a point $x \in V$ and a number $m \geq 1$ such that

$$P_x \{\xi_m \in V, \xi_{2m} \in V, \dots, \xi_{lm} \in V\} > 0.$$

This implies that $F(x)$ assumes arbitrarily small values in every nonempty open subset of X . Now, by a standard argument, $F(x) = 0$ whenever x is a continuity point of F . As a consequence of semicontinuity, $F(x) = 0$ on a residual set in X . This ends the proof of the theorem.

Our next result is a recurrence property of irreducible Markov operators on compact spaces. All probability measures are assumed to be regular.

THEOREM 2. *Let T be an irreducible Markov operator on $C(X)$, X compact Hausdorff. Then for every initial probability measure ν and for every*

nonempty open set U the set

$$\{n: \xi_n \in U\}$$

contains arbitrarily long arithmetic progressions P_ν almost surely.

Proof. Fix $l \geq 1$. The function

$$h(x) = P_x \{ \exists n \geq 0 \exists m \geq 1 \xi_n \in U, \xi_{n+m} \in U, \dots, \xi_{n+lm} \in U \}$$

is lower semicontinuous since the event within the braces is an open subset of Ω . Therefore $h(x)$ attains its minimum on a nonempty closed subset F of X , say, $h(x) = \alpha$ on F . Next we prove $F = X$. This will follow from irreducibility as soon as F is shown to be invariant. But invariance of F is a consequence of the inequality:

$$Th(x) = P_x \{ \exists n \geq 1 \exists m \geq 1 \xi_n \in U, \xi_{n+m} \in U, \dots, \xi_{n+lm} \in U \} \leq h(x).$$

Now $h(x) = \text{const} = \alpha$ so we have $T^n h(x) = \alpha$ and

$$P_x \{ \forall k \geq 1 \exists n \geq k \exists m \geq 1$$

$$\xi_n \in U, \xi_{n+m} \in U, \dots, \xi_{n+lm} \in U \} = \lim T^n h(x) = \alpha.$$

Denote by B the invariant event defined in the last formula and let μ be any ergodic invariant probability measure (that such measures actually exist is a well-known fact and is a consequence of compactness and the Krein–Milman theorem). Since μ is ergodic, the Markov measure P_μ is also ergodic with respect to the shift transformation of Ω (see [5], Prop. V. 2.4). Therefore

$$\alpha = \int P_x(B) d\mu(x) = P_\mu(B)$$

is either 0 or 1. Now we prove that $\alpha \neq 0$. In fact, by Lemma 1 there exists $x \in U$ such that

$$P_x \{ \xi_m \in U, \xi_{2m} \in U, \dots, \xi_{lm} \in U \} > \delta > 0$$

for some $m \geq 1$. The event, say, C is an open subset of Ω so we have $P_y(C) > \delta$ on a neighborhood V of x . By irreducibility, the trajectories starting from any point y visit V infinitely often almost surely (see [4], Lemma 2). By conditioning on the first visit to V at a time $n \geq k$ we obtain by the Markov property that for every $k \geq 1$

$$P_y \{ \exists n \geq k \xi_{n+m} \in U, \xi_{n+2m} \in U, \dots, \xi_{n+lm} \in U \} > \delta.$$

Denote the last event by C_k . We have $C_1 \supset C_2 \supset \dots$, so

$$P_y(\cap C_k) \geq \delta > 0.$$

This means

$$P_y \{ \forall k \geq 1 \exists n \geq k \xi_{n+m} \in U, \xi_{n+2m} \in U, \dots, \xi_{n+lm} \in U \} > 0$$

so $P_\nu(B) > 0$, or $\alpha > 0$, and consequently $\alpha = 1$. Therefore $h(x) \equiv 1$ and

$$P_\nu \{ \exists n \geq 0 \exists m \geq 1 \xi_n \in U, \xi_{n+m} \in U, \dots, \xi_{n+lm} \in U \} = \int h d\nu = 1$$

for every initial distribution ν . Now it suffices to consider the intersection of the above events for $l = 1, 2, \dots$

It should be noted that the assertion of Theorem 2 can be strengthened if in addition to compactness X is metrizable (a countable basis of open sets exists): For every initial distribution almost every trajectory visits all nonempty open sets along arbitrarily long arithmetic progressions.

A class of Markov processes possessing strong recurrence properties are the Harris recurrent chains. The transition probability $p(x, \cdot)$ induces a positively recurrent Harris process if there exists an invariant probability measure μ such that every subset of positive μ measure is visited infinitely many times almost surely for every initial state (see [6] for the definition and properties of Harris processes). For aperiodic positively recurrent Harris processes we have the following sharper version of Theorem 1.

PROPOSITION. *Let $p(x, \cdot)$ be a positively recurrent aperiodic Harris transition probability with invariant probability measure μ . If $\mu(A) > 0$ then for every $l \geq 1$*

$$P_\nu \{ \xi_m \in A, \xi_{2m} \in A, \dots, \xi_{lm} \in A \text{ for infinitely many } m \}' = 1$$

for any distribution ν .

Proof. The event D within the braces is asymptotic, i.e., contained in the σ -algebra generated by ξ_n, ξ_{n+1}, \dots , for every $n \geq 1$. By Theorem 2.6 in [6], $P_x(D) \equiv 0$ or $P_x(D) \equiv 1$. Since, by Lemma 1, $P_x(D)$ is not identically zero, we obtain $P_\nu(D) = 1$.

COROLLARY. *Let $X, p(x, \cdot)$ be a positively recurrent discrete Markov chain (X is countable). Then for every $x \in X$ and every $l \geq 1$*

$$P_\nu \{ \xi_m = x, \xi_{2m} = x, \dots, \xi_{lm} = x \text{ for infinitely many } m \}' = 1.$$

Proof. Without loss of generality we may assume that X is irreducible (hence Harris recurrent). We may also assume that X is aperiodic – if necessary, consider the iterated transition probability $p^{(d)}(x, \cdot)$ on a cyclic subset of X containing x . Now the Proposition applies with $A = \{x\}$.

3. Multiple recurrence for families of operators. We introduce a general notion of multiple recurrence for a family $\Phi = \{T_i: i \in I\}$ of Markov operators on a compact metric space X . First recall that if τ is a continuous mapping on X then a point x is called *recurrent* if there exists a sequence $n_k \rightarrow \infty$ such that for every neighborhood U of x we have $\tau^{n_k} x \in U$ for all sufficiently large k . Similarly, if T is a Markov operator on $C(X)$, x will be

called *recurrent* if a sequence $n_k \rightarrow \infty$ exists such that

$$T^{n_k} \chi_U(x) > 0$$

or, equivalently, $p^{(n_k)}(x, U) > 0$ for k sufficiently large. If, still more generally, $\Phi = \{T_i: i \in I\}$ is a family of Markov operators on $C(X)$, we say that x is *multiply recurrent* with respect to Φ if there exists a sequence $n_k \rightarrow \infty$ such that for every neighborhood U of x and every $i \in I$

$$T_i^{n_k} \chi_U(x) > 0$$

for all sufficiently large k . Hence the notion of multiple recurrence for Markov operators generalizes the multiple recurrence of continuous mappings.

By Theorem 1, multiply recurrent points always exist if Φ consists of finitely many iterations of a single Markov operator. In fact, it follows from the proof of Theorem 1 that recurrent points with respect to $\Phi = \{T, T^2, \dots, T^l\}$ form a residual subset of $\text{supp } \mu$ for every invariant probability measure μ . This along with the following lemma and the Baire category theorem implies that multiply recurrent points always exist for cyclic semigroups of Markov operators.

LEMMA 3. *Let Φ be the union of an ascending sequence of finite families Φ_l ($l = 1, 2, \dots$) of Markov operators on $C(X)$. If x is multiply recurrent with respect to all the Φ_l ($l = 1, 2, \dots$) then x is multiply recurrent with respect to Φ .*

Proof. Let U_1, U_2, \dots be a basic neighborhood sequence for x and let χ_i denote the characteristic function of U_i . For every $i \geq 1$ there exists $n_i \geq 1$ such that

$$T^{n_i} \chi_i(x) > 0$$

for all $T \in \Phi_i$ and we may assume that $n_1 < n_2 < \dots$. Now for every $T \in \Phi$ and every $j \geq 1$ the sequence n_i satisfies the inequality $T^{n_i} \chi_j(x) > 0$ for all i sufficiently large.

It should be noted that monothetic semigroups need not possess multiply recurrent points.

Example. Let X be the circle group and consider the translation homeomorphisms $\tau_z x = zx$ ($z \in X$). Let $\Phi = \{\tau_z: z \in X\}$. Suppose there exist $y \in X$ and a sequence $n_k \rightarrow \infty$ such that $\tau_z^{n_k} y \rightarrow y$ for every z . This would imply $z^{n_k} \rightarrow 1$ for all $z \in X$. To prove that this is never possible, choose an increasing subsequence m_k of n_k such that $\sum m_k/m_{k+1} < 1/4$. Define by induction a sequence $0 \leq a_k \leq 1$ such that the fractional part of the number

$$m_k \left(\frac{a_1}{m_1} + \dots + \frac{a_{k-1}}{m_{k-1}} \right) + a_k$$

differs from $1/2$ by less than $1/8$. Now let $t = \sum a_k/m_k$ and $z = e^{2\pi i t}$. It is easy to see that the fractional part of $m_k t$ differs from $1/2$ by less than $3/8$ so the sequence

$$z^{m_k} = e^{2\pi i m_k t}$$

does not converge to 1.

By the Birkhoff multiple recurrence theorem any finite set of commuting continuous transformations possesses a multiply recurrent point. (In fact the same is true of all countable commuting sets due to Lemma 3.) In the case of two commuting Markov operators we are able to prove that under additional conditions – always satisfied by continuous transformations – doubly recurrent points exist.

To specify these conditions, for every Markov operator T on $C(X)$ we define the compact valued mapping by letting

$$T(x) = \text{supp } T^* \delta_x.$$

The mapping $x \rightarrow T(x)$ is easily seen to be lower semicontinuous, i.e., for every open set U in X if the intersection $T(x) \cap U$ is nonempty then it remains nonempty in some neighborhood of x . It is also not hard to see that the graph of the relation $y \in T(x)$ is closed in $X \times X$ if and only if $x \rightarrow T(x)$ is in addition upper semicontinuous (which is also equivalent to saying that $x \rightarrow T(x)$ is continuous for the Hausdorff topology on the space of nonempty closed subsets of X).

THEOREM 3. *Let T_1, T_2 be commuting Markov operators on $C(X)$, X compact metric. Assume in addition that the relations $y \in T_i(x)$ are closed ($i = 1, 2$). Then there exists a doubly recurrent point x_0 for T_1, T_2 .*

Proof. Let $\Omega = X^{N_0 \times N_0}$. For every $n \geq 1$ we define Ω_n to be the set of all $\omega \in \Omega$ satisfying

$$\begin{aligned} \omega(k+1, l) \in T_1(\omega(k, l)); & \quad 0 \leq k < n, \quad 0 \leq l \leq n, \\ \omega(k, l+1) \in T_2(\omega(k, l)); & \quad 0 \leq k \leq n, \quad 0 \leq l < n. \end{aligned}$$

Since the relations are closed, the sets Ω_n are also closed. Next we prove $\Omega_n \neq \emptyset$ for $n \geq 1$. Choose any $x \in X$ and let $\omega(0, 0) = x$, $\omega(n, n) \in (T_1 T_2)^n(x)$. By assumption, the images of closed sets are closed so by compactness

$$(T_1 T_2)^n(x) = (T_1 T_2)[(T_1 T_2)^{n-1}(x)].$$

Therefore there exists $y \in (T_1 T_2)^{n-1}(x)$ such that $\omega(n, n) \in (T_1 T_2)(y)$. We let $\omega(n-1, n-1) = y$.

Similarly we obtain $\omega(k, k)$, $k = 0, 1, \dots, n-1$, such that

$$\omega(k+1, k+1) \in (T_1 T_2)(\omega(k, k)).$$

In like manner (using commutativity) there exist elements $\omega(k, k+1)$ and $\omega(k+1, k)$ such that

$$\begin{aligned} \omega(k+1, k) \in T_1(\omega(k, k)), & \quad \omega(k+1, k+1) \in T_2(\omega(k+1, k)), \\ \omega(k, k+1) \in T_2(\omega(k, k)), & \quad \omega(k+1, k+1) \in T_1(\omega(k, k+1)) \end{aligned}$$

for $k = 0, 1, \dots, n-1$. Since, in general, $w \in T_j(v)$, $v \in T_i(u)$ implies $w \in (T_i T_j)(u)$, we have

$$\begin{aligned} \omega(k+1, k+2) &\in (T_1 T_2)(\omega(k, k+1)), \\ \omega(k+2, k+1) &\in (T_1 T_2)(\omega(k+1, k)) \end{aligned}$$

for $k = 0, 1, \dots, n-2$. In the same way we can construct a whole array $\omega(k, l)$, $0 \leq k \leq n$, $0 \leq l \leq n$, such that

$$\begin{aligned} \omega(k+1, l) &\in T_1(\omega(k, l)); & k < n, \\ \omega(k, l+1) &\in T_2(\omega(k, l)); & l < n. \end{aligned}$$

This proves $\Omega_n \neq \emptyset$ and, by compactness, $\Omega_\infty = \bigcap \Omega_n \neq \emptyset$. It is now obvious that Ω_∞ is invariant under the shift transformations $S_1 \omega(k, l) = \omega(k+1, l)$ and $S_2 \omega(k, l) = \omega(k, l+1)$. By the multiple Birkhoff recurrence theorem (Thm. 2.6 in [2]) there exists $\omega_0 \in \Omega_\infty$ such that $S_i^{n_k} \omega_0 \rightarrow \omega_0$ in Ω for a sequence $n_k \rightarrow \infty$. If, in particular, $x_0 = \omega_0(0, 0)$ then for every neighborhood U of x_0 there exists an n_k such that $\omega_0(n_k, 0) \in U$ and $\omega_0(0, n_k) \in U$. Since $\omega_0 \in \Omega_\infty$, this implies

$$T_i^{n_k}(x_0) \cap U \neq \emptyset$$

so x_0 is doubly recurrent with respect to $\{T_1, T_2\}$.

Added in proof.

1. Remark. If X is finite, then any finite family of Markov operators (i.e. stochastic matrices) having a common invariant probability measure μ is multiply recurrent. In fact, we may assume $\text{supp } \mu = X$. Now there are no transient states, so each T_i is completely reducible and for every x in X we have $T_i^{m_i} \chi_{\{x\}}(x) > 0$ for some $m_i > 0$. Consequently, $T_i^m \chi_{\{x\}}(x) > 0$ for $m = m_1 \dots m_l$ and $i = 1, \dots, l$. In particular, any commuting stochastic matrices T_1, \dots, T_l are multiply recurrent.

2. The study of multiple recurrence will be continued in a forthcoming paper.

REFERENCES

[1] H. Furstenberg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, Journal d'Analyse Mathématique 31 (1977), p. 204–256.

- [2] —, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, N. J., 1981.
- [3] — and B. Weiss, *Topological dynamics and combinatorial number theory*, *Journal d'Analyse Mathématique* 34 (1978), p. 61–85.
- [4] A. Iwanik, *On pointwise convergence of Cesàro means and separation properties for Markov operators on $C(X)$* , *Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques* 29 (1981), p. 515–520.
- [5] J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco 1965.
- [6] D. Revuz, *Markov Chains*, North-Holland, Amsterdam 1975.

TECHNICAL UNIVERSITY
WROCLAW, POLAND

AND

CALIFORNIA STATE UNIVERSITY
FULLERTON, CALIFORNIA, U.S.A.

Reçu par la Rédaction le 20. 01. 1984
