Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 121 (2011) 314–323
stochastic processes and their applications
www.elsevier.com/locate/spa

Hitting and returning to rare events for all alpha-mixing processes

Miguel Abadi[a], Benoit Saussol[b,*]

[a] *Universidade de São Paulo, São Paulo, Brazil*
[b] *Laboratoire de Mathématiques CNRS UMR 6205, Université de Bretagne Occidentale, Brest, France*

Received 10 March 2010; received in revised form 13 October 2010; accepted 3 November 2010
Available online 12 November 2010

Abstract

We prove that for any α-mixing stationary process the hitting time of any n-string A_n converges, when suitably normalized, to an exponential law. We identify the normalization constant $\lambda(A_n)$. A similar statement holds also for the return time.

To establish this result we prove two other results of independent interest. First, we show a relation between the rescaled hitting time and the rescaled return time, generalizing a theorem of Haydn, Lacroix and Vaienti. Second, we show that for positive entropy systems, the probability of observing any n-string in n consecutive observations goes to zero as n goes to infinity.

Keywords: Mixing processes; Hitting times; Repetition times; Return times; Rare event; Exponential approximation

1. Introduction

The study of the statistical properties of the time elapsed until the occurrence of an observable of positive measure in a stationary stochastic process and/or in a measure preserving dynamical system is a classical subject. The starting point of this study is the famous Poincaré recurrence theorem which states that in an ergodic system, any set of positive measure appears in the process infinitely many times. This is a qualitative result in the sense that no statistical properties of these returns are established. In the last twenty years many notions of return have been introduced and studied. These notions depend on the initial conditions, the observed set, and the measure

* Corresponding author.
E-mail addresses: leugim@ime.usp.br (M. Abadi), benoit.saussol@univ-brest.fr (B. Saussol).

0304-4149/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2010.11.001

of the system. There was intensive interest in studying their statistical properties to model physical phenomena like intermittency and metastability. Then, the applications were extended to other areas such biology, linguistics and computer science, to describe phenomena like gene occurrence in DNA and protein sequences, the rhythm of a language and data compression algorithms, to mention just some of them.

In the present paper we consider a fixed set A of positive measure $\mu(A)$ in an ergodic system. When the evolution starts outside A, the time elapsed until the first occurrence of the set is referred to as the hitting time of A. When the evolution starts inside A, the time is referred as the return time to A.

Our main result is that under the so called α or strongly mixing condition, the distribution of the hitting time of a set A can be well approximated by an exponential law. The approximation is in the supremum norm in the space of distribution functions. Although the exponential law is a classical subject our result is new and interesting:

(a) Our result holds for *any* cylinder set, namely, around any point, including periodic points, and not just around generic points.
(b) The result holds for any α-mixing systems, while the best previous works [1] assumed a polynomial rate of at least $(1+\sqrt{5})/2$. Moreover, this strong-mixing condition is the weakest among many kinds of mixing conditions, among them ψ, ϕ, ρ, β, or absolutely regular, or I or information regular. See Bradley [2].
(c) We also show that the exponential law holds when considering not just a cylinder set but even a set which is a union of cylinders. Moreover, the cardinal of this union can be exponentially large, with respect to the length of the cylinders.

Following the Galves and Schmitt [5] approach we get that the parameter of the exponential law is the product $\lambda(A)\mu(A)$, where $\lambda(A)$ is a positive number related to the short recurrence properties of the set A. For a description of these properties see Abadi [1]. In the aforementioned paper, the authors show that for ψ-mixing systems, there exist two positive constants K, K' such that $K \le \lambda(A) \le K'$. In our case, the constant K does not exist, and one can have $\lambda(A)$ arbitrarily small.

We prove our result by showing two other results which are interesting in themselves. In the first one, we establish an ergodic relationship between the rescaled hitting time $\lambda(A)\mu(A)\tau_A$ and the equally rescaled return time. The idea of this result comes from a paper of Haydn et al. [7], which established such a relationship for the rescaled $\mu(A)\tau_A$ hitting time and return time. This in general does not apply in our case since one can have $\lambda(A) \neq 1$, for instance, around periodic points. Even the proof follows a different approach.

The second result that we mentioned above reads as follows. The probability of observing a cylinder of rank n, or even a union of them, in n consecutive observations, goes to zero with n for α-mixing systems. Moreover, we show that the convergence is uniform on A. It only depends on the cardinality of the union, and not on the choice of the cylinders. This is natural when the measure of the set decays e.g. exponentially with n. But is far from obvious, and may even be anti-intuitive, when the measure decays just polynomially fast with power less than 1, as is covered by our case.

2. Statement of the results

Let $\mathcal{A}$ be a finite or countable set and let $\Sigma = \mathcal{A}^{\mathbb{N}}$ be the set of sequences. We endow Σ with the shift map T. Given non-negative integers $m \le n$ and a point $x \in \Sigma$ we denote by $[x_m \dots x_n]$

the cylinder of rank (m, n) containing x, that is

$$[x_m \dots x_n] := \{y \in \Sigma : y_m = x_m, \dots, y_n = x_n\}.$$

A cylinder of rank $(0, n-1)$ will be simply called of rank n. We denote by $\mathcal{C}_m^n$ the collection of cylinders of rank (m, n) and by $\mathcal{F}_m^n$ the σ-algebra generated by the partition $\mathcal{C}_m^n$. Let $\mathcal{F}$ be the σ-algebra generated by the $\mathcal{F}_m^n$'s and μ be a T-invariant probability measure on $(\Sigma, \mathcal{F})$. Let

$$\alpha(g) = \sup_{m,n} \sup_{A \in \mathcal{F}_0^n, B \in \mathcal{F}_{n+g}^{m+g}} |\mu(A \cap B) - \mu(A)\mu(B)|$$

for any integer g. We assume that the system (Σ, T, μ) is α-mixing, in the sense that $\alpha(g) \to 0$ as $g \to \infty$. This is the weakest notion of mixing among ϕ-mixings and ψ-mixings. We emphasize that we do not assume any summability condition on the sequence $\alpha(g)$.

Let $A \in \Sigma$ be a measurable set. We define the hitting time of A by

$$\tau_A(x) = \inf\{k \geq 1 : T^k x \in A\}, \quad x \in \Sigma.$$

We are interested in the distribution of the hitting time τ_A on the probability space (Σ, μ), and the return time, defined with the same formula, but on the probability space $(A, \mu(\cdot|A))$ where $\mu(\cdot|A)$ denotes the conditional measure on A.

Theorem 1. *Suppose that the system (Σ, T, μ) is α-mixing. Then for any sequence $A_n \in \mathcal{F}_0^{n-1}$ such that $\mu(A_n) > 0$ and*

$$\mu(\tau_{A_n} \leq n) \to 0 \quad \text{as } n \to \infty, \tag{1}$$

there exists some normalizing constant $\lambda(A_n) > 0$ such that the following hold:

- *The hitting time of A_n, rescaled by $\lambda(A_n)\mu(A_n)$, converges in distribution to an exponential distribution. Namely,*

  $$\sup_{t \geq 0} \left| \mu(\lambda(A_n)\mu(A_n)\tau_{A_n} > t) - \exp(-t) \right| \to 0 \quad \text{as } n \to \infty.$$

  The convergence is uniform on families of sets A_n where the convergence in (1) is uniform.
- *The distribution of the return time is approximated by a convex combination of a Dirac mass at zero and an exponential distribution. More precisely,*

  $$\sup_{t \geq s} \left| \lambda(A_n)^{-1} \mu(\lambda(A_n)\mu(A_n)\tau_{A_n} > t | A_n) - \exp(-t) \right| \to 0 \quad \text{as } n \to \infty,$$

  for any $s > 0$.
- *we have* $\limsup \lambda(A_n) \leq 1$.

Remark 2. We emphasize that in the theorem the set A_n does not have to be a cylinder of rank n but only a union of cylinders of rank n. Moreover, even if A_n is a cylinder, n does not have to be equal to the rank of the cylinder; see Remark 9 for an instructive example.

The normalizing constant $\lambda(A_n)$ may not converge in general; thus we cannot simply say that the limiting distribution of the rescaled return time exists. Moreover, even if it converges, the limit may not be equal to 1. For example a case of interest is when $\lim \lambda(A_n) = 0$ where we still get a non-trivial exponential approximation, while without the extra factor $\lambda(A_n)$ one would just obtain the rough statement that the rescaled hitting time $\mu(A_n)\tau_{A_n} \to +\infty$ and the rescaled return time $\mu(A_n)\tau_{A_n} \to 0$ in distribution.

In the next section we show that the hypothesis in the theorem holds for a broad class of sequences of sets A_n.

3. Rare events do not appear too soon

We present some explicit examples of sequences A_n under which Theorem 1 applies, that is when the condition (1) of the theorem is satisfied. They are consequences of Proposition 7 presented below.

The first example was the motivation of our work:

Example 3. For any $a \in \mathcal{A}^{\mathbb{N}}$, the sequence of cylinders $A_n = [a_0, \dots, a_{n-1}]$ satisfies the hypothesis (1) of Theorem 1. Moreover, the convergence is uniform on a.

We emphasize that this approximation with an exponential distribution is valid for any point $a \in \Sigma$, including for example periodic points. This generalizes the result in [6] which concerns a.e. sequence a.

Returns to the cylinder $[a_0, \dots, a_{n-1}]$ in the example above means that there is a perfect matching of the first n symbols. It turns out that for some applications the approximate matching is more interesting:

Example 4. Approximate matching: Let $a \in \mathcal{A}^{\mathbb{N}}$ and $D \in (0, 1)$. Denote for $b \in \Sigma$ by $d_n(a, b) = \operatorname{card}\{i \le n - 1 : a_i \ne b_i\}$ the Hamming distance of the first n symbols. Let

$$A_n = \{b \in \Sigma : d_n(a, b) \le Dn\},$$

be the $D\%$ approximate matching of $[a_0, \dots, a_{n-1}]$. Then there exists $D_0 > 0$ such that for all $D \in (0, D_0)$, the sequence A_n satisfies the hypothesis (1) of Theorem 1.

In DNA sequence analysis the alphabet $\mathcal{A}$ is $\{A, C, G, T\}$. For some sequences the entropy is lower estimated by 1.7 bits per symbol (for example for the human gene HUMRETBLAS; see [8]), which means that $h_\mu = 1.7 \ln 2$. This gives a value of $D_0 \approx 41\%$.

Proof. We count the number κ_n of cylinders of rank n which compose the $D\%$ approximate matching A_n. We have

$$\begin{aligned} \kappa_n &\le \sum_{k=0}^{Dn} \binom{n}{k} (\operatorname{card} \mathcal{A} - 1)^k \\ &\le D^{-Dn} \sum_{k=0}^{n} \binom{n}{k} D^k (\operatorname{card} \mathcal{A} - 1)^k \\ &= \left(\frac{1 + D(\operatorname{card} \mathcal{A} - 1)}{D^D} \right)^n . \end{aligned}$$

We choose $D_0 > 0$ as the smallest solution of $(1 + D(\operatorname{card} \mathcal{A} - 1))/D^D = e^{h_\mu(T)}$ and then Proposition 7 applies for any $D < D_0$. □

Example 5. For a set $K \subset \Sigma$, define its topological entropy by

$$h_{\mathrm{top}}(K) = \limsup_{n \to \infty} \frac{1}{n} \log \#\{C : C \text{ cylinder of rank } n \text{ s.t. } C \cap K \ne \emptyset\}.$$

Denote by $\mathcal{F}_0^{n-1}(K)$ the union of those cylinders C of rank n such that $K \cap C \neq \emptyset$. The sequence $A_n = \mathcal{F}_0^{n-1}(K)$, under the assumption that $h_{\text{top}}(K) < h_\mu$, satisfies the hypothesis (1) of Theorem 1.

Example 6. Suppose that $A_n = A_n^0 \cup A_n^1$ where A_n^0 and A_n^1 are $\mathcal{F}_0^{n-1}$ measurable sets, and such that $\lim n\mu(A_n^0) = 0$ and A_n^1 satisfies the conditions of Example 5 above whenever it has positive probability. Then, we have

$$\mu(\tau_{A_n} \leq n) \leq \mu(\tau_{A_n^0} \leq n) + \mu(\tau_{A_n^1} \leq n) \leq n\mu(A_n^0) + \mu(\tau_{A_n^1} \leq n) \to 0,$$

and therefore the hypothesis (1) of Theorem 1 is satisfied.

We emphasize that, in this example, the exponential growth of the number of cylinders of rank n inside A_n is not a priori bounded by the entropy of the measure, contrary to the case for the preceding example.

Proposition 7. *Suppose that (Σ, T, μ) is an ergodic measure preserving system, not necessarily α-mixing. Let (κ_n) be a sequence of integers such that*

$$\limsup_n \frac{1}{n} \log \kappa_n < h_\mu(T).$$

Then there exists a sequence $\epsilon_n \to 0$ such that, for any $A_n \in \mathcal{F}_0^{n-1}$ which is the union of at most κ_n cylinders of rank n, we have

$$\mu(\tau_{A_n} \leq n) \leq \epsilon_n.$$

We emphasize that the bound ϵ_n does not depend on the particular set A_n but only on the number of cylinders which compose it. Note that the statement $\mu(\tau_{A_n} \leq n) \to 0$ is trivial whenever $\mu(A_n) \ll 1/n$. However, even for α-mixing systems, there can exist some cylinders A_n of rank n such that $\mu(A_n) \gg 1/n$ (see [3]).

When the system is α-mixing, the measure preserving transformation (T, μ) is an exact endomorphism and, in particular, its entropy $h_\mu(T)$ is positive (we refer the reader to [4] for details). In particular, Proposition 7 applies under the mixing hypotheses of Theorem 1.

Proof. Set $h_0 := \limsup_n \frac{1}{n} \log \kappa_n$ and let $h \in (h_0, h_\mu(T))$ and $k \in \mathbb{N}$ such that $h_0 < (1-1/k)h$. Let

$$\Gamma(N) = \{x : \forall n \geq N, \mu([x_0 \dots x_{n-1}]) \leq e^{-nh}\}.$$

By the Shannon–McMillan–Breiman theorem, $\mu(\Gamma(N)) \to 1$ as $N \to \infty$. Given an integer n, let $m = \lceil n/k \rceil$ be the smallest integer such that $km \geq n$. First, observe that by invariance we have

$$\mu(\tau_{A_n} \leq n) \leq k\mu(\tau_{A_n} \leq m). \tag{2}$$

By assumption we have $A_n = \bigcup_{i=1}^{\ell} C_i$ for some integer $\ell \leq \kappa_n$ and some cylinders $C_i \in \mathcal{F}_0^{n-1}$. For $j = 0$ to $m - 1$ let C_i^j be the element of $\mathcal{F}_0^{n-m-1}$ which contains $T^j C_i$. Let $U_n = \bigcup_{j=0}^{m-1} \bigcup_{i=1}^{\ell} C_i^j$. We have $\{\tau_{A_n} \leq m\} \subset T^{-m} U_n$; hence,

$$\mu(\tau_{A_n} \leq m) \leq \mu(U_n). \tag{3}$$

Moreover, the set U_n is contained by construction in at most $m\kappa_n$ cylinders of rank $n-m$; therefore

$$\mu(U_n \cap \Gamma(n-m)) \le m\kappa_n \mathrm{e}^{-(n-m)h}.$$

On the other hand,

$$\mu(U_n \setminus \Gamma(n-m)) \le 1-\mu(\Gamma(n-m)).$$

Setting ϵ_n equal to k times the sum of the last two upper bounds proves the proposition in view of (2) and (3). □

4. Proof of the main theorem

Our main theorem will be a direct application of this explicit estimation of the difference between the hitting time statistics and the exponential distribution.

Theorem 8. *Suppose that the system (Σ, T, μ) is α-mixing. Let n be an integer. For any $A \in \mathcal{F}_0^{n-1}$ there exists some constant $\lambda(A) \in (0,2]$ such that*

$$\sup_{k\in\mathbb{N}} \left|\mu(\tau_A > k) - \mathrm{e}^{-\lambda(A)\mu(A)k}\right| \le 12\sqrt{2\mu(\tau_A \le n)+\alpha(n)}.$$

The value of the upper bound is not intended to be optimal, but is just there to emphasize that it does not depend on the particular choice of the set $A \in \mathcal{F}_0^{n-1}$ but only on the probability of short hitting times $\mu(\tau_A \le n)$.

Remark 9. We emphasize that if $A \in \mathcal{F}_0^{m-1}$ for some integer m, one can apply the theorem with any integer $n \ge m$. It turns out that taking the minimal n is not always a good choice, even when A is itself a cylinder set.

Consider for example the full shift $(\mathbb{N}^{\mathbb{N}}, T)$ with an α-mixing measure μ such that $\alpha(1) \ne 0$. For any $k \in \mathbb{N}$ the set $[k]$ is a rank 1 cylinder, but applying Theorem 8 with $n=1$ has little interest.

However, choosing $n_k = \lfloor 1/\sqrt{\mu([k])}\rfloor$ we get that $\mu(\tau_{[k]} \le n_k) \le n_k\mu([k]) \to 0$ as $k\to\infty$. Therefore Theorem 8 applied with $n=n_k$ gives that the distribution of the hitting time $\tau_{[k]}$ is asymptotically equal to that of an exponential with parameter $\lambda([k])\mu([k])$.

In the proof of the theorem we make use of the following lemma.

Lemma 10. *Let n be an integer. For any $A \in \mathcal{F}_0^{n-1}$ such that*

$$\delta := 3\sqrt{2\mu(\tau_A \le n)+\alpha(n)} < 1/4,$$

there exists an integer $s > 2n$ such that

$$\mu(\tau_A \le s) \le \delta \quad \textit{and} \quad \frac{\mu(\tau_A \le 2n)+\alpha(n)}{\mu(\tau_A \le s-2n)} \le \delta. \tag{4}$$

Proof. Let us define $d = 2\mu(\tau_A \le n)+\alpha(n)$. By hypothesis, $d < 1/144$. By invariance we have

$$\mu(\tau_A \le 2n)+\alpha(n) \le d.$$

Let $s > 2n$ denote the smallest integer such that

$$\mu(\tau_A \le s - 2n) \ge \sqrt{d}.$$

With this choice we have

$$\frac{\mu(\tau_A \le 2n) + \alpha(n)}{\mu(\tau_A \le s - 2n)} \le \sqrt{d}.$$

Furthermore, since $\mu(\tau_A \le s - 2n - 1) < \sqrt{d}$, it follows from the invariance that

$$\mu(\tau_A \le s) \le \mu(\tau_A \le s - 2n - 1) + \mu(\tau_A \le 2n) + \mu(\tau_A \le 1) \le \sqrt{d} + 2d \le 3\sqrt{d}. \quad \square$$

Proof of Theorem 8. Let n be an integer and $A \in \mathcal{F}_0^{n-1}$. Let δ be as in Lemma 10. There is nothing to prove if $\delta \ge 1/4$ so we suppose that $\delta < 1/4$. Take $s > 2n$ given by Lemma 10 such that (4) holds.

To simplify the notation we drop the subscript A and write $\tau = \tau_A$. Set $H(k) = \mu(\tau > k)$, and denote by $\tau^{[t]} = \tau \circ T^t$ the first occurrence time starting at time t. For any integer $j \ge 1$ consider the modulus

$$|H(js) - H((j-1)s)H(s-2n)|. \tag{5}$$

The sets

$$\{\tau > js\} = \{\tau > (j-1)s\} \cap \{\tau^{[(j-1)s]} > s\}$$

and

$$\{\tau > (j-1)s\} \cap \{\tau^{[(j-1)s+2n]} > s - 2n\}$$

differ by a subset of $\{\tau^{[(j-1)s]} \le 2n\}$ whose measure is by invariance bounded by $\mu(\tau \le 2n)$. Furthermore, by mixing we get that

$$|\mu(\{\tau > (j-1)s; \tau^{[(j-1)s+2n]} > s - 2n\}) - H((j-1)s)H(s-2n)| \le \alpha(n).$$

Thus the above expression (5) is bounded by

$$\mu(\tau \le 2n) + \alpha(n).$$

Now, take q a positive integer. The absolute value

$$|H(qs) - H(s-2n)^q| \tag{6}$$

is bounded by

$$\sum_{j=1}^{q} |H(js) - H((j-1)s)H(s-2n)| H(s-2n)^{q-j}.$$

We have just proved that the modulus in the above sum is bounded by $\mu(\tau \le 2n) + \alpha(n)$. Summing over j we get that for all integers $k \ge 1$ the modulus in (6) is bounded by

$$\frac{\mu(\tau \le 2n) + \alpha(n)}{\mu(\tau \le s - 2n)} \le \delta.$$

Moreover, any non-negative integer k can be written as $qs + r$ with $q = [k/s]$ and $0 \le r < s$. Then

$$|H(k) - H(qs)| = \mu(\tau > qs; \tau^{[qs]} \le r), \tag{7}$$

which, by invariance, is bounded by $\mu(\tau \le s) \le \delta$.

To finish the proof, set

$$\lambda_n(A) = -\frac{\ln H(s-2n)}{s\mu(A)},$$

and note that the mean value theorem gives

$$|H(s-2n)^{[k/s]} - H(s-2n)^{k/s}| \le -\ln H(s-2n). \tag{8}$$

Note that $H(s-2n)^{k/s} = \mathrm{e}^{-\lambda_n(A)\mu(A)k}$. By convexity we have $-\ln(1-u) \le u/(1-\delta)$ whenever $0 \le u \le \delta$, and therefore

$$-\ln H(s-2n) \le \frac{1}{1-\delta}\mu(\tau \le s-2n) \le \frac{1}{1-\delta}\mu(\tau \le s) \le 2\delta.$$

Putting together the three estimates for (6)–(8) gives the conclusion. Observe in addition that $\lambda_n(A) \le 1/(1-\delta) \le 2$ since $\mu(\tau \le s) \le s\mu(A)$.

To drop the dependence of $\lambda_n(A)$ on n and finish the proof, we define $\lambda(A)$ as the minimal λ which realizes the infimum

$$\inf_{\lambda\in(0,2]} \sup_{k\in\mathbb{N}} \left|\mu(\tau_A > k) - \mathrm{e}^{-\lambda\mu(A)k}\right|. \quad \square$$

Remark 11. The upper bound $\lambda(A) \le 2$ can be sharpened when δ is small. In particular if $\delta_n \to 0$ as $n \to \infty$ we get $\limsup \lambda(A_n) \le 1$.

We conclude this section with the proof of the main theorem. In view of Theorem 12, the statement for hitting times in the main theorem (Theorem 1) and that for return times are equivalent; hence it is sufficient to prove the first statement with $F(t) = 1 - \mathrm{e}^{-t}$, which will imply the second statement with $G(s) = \mathrm{e}^{-s}$.

Proof of Theorem 1. For any real $t > 0$, taking $k = \lfloor t/\mu(A_n)\rfloor$ in Theorem 8 gives

$$\left|\mu(\lambda(A_n)\mu(A_n)\tau_{A_n} > t) - \mathrm{e}^{-t}\right| \le 12\sqrt{2\mu(\tau_{A_n} \le n) + \alpha(n)} + 2\mu(A_n),$$

which proves the first statement. The uniform convergence in (1) implies that of this upper bound, since

$$\mu(A_n) = \mu(\tau_{A_n} = 1) \le \mu(\tau_{A_n} \le n).$$

The second statement follows from Theorem 12. The third statement follows from Remark 11. $\square$

5. Hitting and returning: an adaptation of the Haydn–Lacroix–Vaienti theorem

Haydn et al. [7] have proved that the asymptotic distributions of hitting and return times τ_{A_n}, rescaled by the measure $\mu(A_n)$, are related by an integral equation. Their result does not apply to our setting because the asymptotic distribution does not exist in general, because the normalizing constant does not converge in general.

We now give the generalization of their result adapted to our case, which deserves a new proof since the technique needs to be somewhat different. Let

$$F_A(t) = \mu(\lambda(A)\mu(A)\tau_A \le t),$$

$$G_A(s) = \frac{1}{\lambda(A)}\mu(\lambda(A)\mu(A)\tau_A > s|A).$$

F_A is the usual non-decreasing cumulative distribution function of the rescaled hitting time $\lambda(A)\mu(A)\tau_A$ while G_A is a normalized non-increasing distribution function of the rescaled return time $\lambda(A)\mu(A)\tau_A$. We recall that since F_A and G_A are monotonic, their convergences when $\mu(A) \to 0$ on a dense set or on all but countably many points are equivalent and we will simply say that they converge.

Theorem 12. *Suppose that the measure preserving system* (Σ, T, μ) *is ergodic.*

Let A_n *be a sequence of measurable sets such that* $\mu(A_n) \to 0$. *If* F_{A_n} *converges to* F *as* $n \to \infty$ *then* G_{A_n} *converges to some function* G, *and the limits are related by the integral equation*

$$F(t) = F(0+) + \int_0^t G(s)\mathrm{d}s \quad (\forall t > 0).$$

In particular, if the solution G *is continuous then the convergence is uniform on* $[s, +\infty)$ *for any* $s > 0$.

Reciprocally, if G_{A_n} *converges to* G *as* $n \to \infty$ *and* $\int_0^\infty G(s)\mathrm{d}s = 1$ *then* F_{A_n} *converges to some function* F, *and the limits are related by the same integral relation with* $F(0+) = 0$. *In particular,* F *is continuous on* $[0, \infty)$ *and the convergence is uniform.*

Proof. Let A be any measurable set with $\mu(A) > 0$. Note that $0 \le F_A(t) \le 1$ and $0 \le G_A(s) \le 1/s$ for any $s > 0$, where this last upper bound follows from Markov inequality and Kac's lemma:

$$G_A(s) = \frac{1}{\lambda(A)}\mu(\lambda(A)\mu(A)\tau_A > s|A) \le \frac{\mu(A)}{s}\int \tau_A \mathrm{d}\mu(\cdot|A) \le \frac{1}{s}.$$

First observe that by invariance one has for every integer n

$$\mu(\tau_A = n) = \mu(A \cap \{\tau_A \ge n\}).$$

Therefore

$$F_A(t) = \sum_{n=1}^{t/\lambda(A)\mu(A)} \mu(A \cap \{\tau_A \ge n\}) = \int_0^{\lfloor t/\lambda(A)\mu(A)\rfloor} \mu(A \cap \{\tau_A > r\})\mathrm{d}r.$$

Since $\mu(A \cap \{\tau_A > r\}) \le \mu(A)$ we get by a change of variable

$$F_A(t) \le \int_0^t G_A(s)\mathrm{d}s \le F_A(t) + \mu(A).$$

For any $0 < t < t'$ we get the relation

$$\int_t^{t'} G_A(s)\mathrm{d}s - \mu(A) \le F_A(t') - F_A(t) \le \int_t^{t'} G_A(s)\mathrm{d}s + \mu(A). \tag{9}$$

• Assume that F_{A_n} converges to some function F and suppose for a contradiction that G_{A_n} does not converge. By Helly's selection principle, each subsequence of function must have an accumulation point.[1] Therefore G_{A_n} must have at least two different accumulation points G_1

[1] Indeed, the space of decreasing functions g from $(0, \infty)$ to itself such that $g(s) \le s$, under the equivalence relation of equality outside countable sets, is metrizable (e.g. a slight modification of the Levy metric) and compact (Helly's selection principle) and an accumulation point refers to this notion of convergence.

and G_2. By dominated convergence, (9) gives that for all $0 < t < t'$

$$F(t') - F(t) = \int_t^{t'} G_i(s)\mathrm{d}s \quad (i = 1, 2). \tag{10}$$

Hence $G_1 = G_2$ a.e., a contradiction; thus G_{A_n} converges. Lastly, the integral relation follows from (10) by monotone convergence.

• Assume that G_{A_n} converges to some function G. By Fatou's lemma, the leftmost inequality in (9) gives that for all $t > 0$,

$$\int_0^t G(s)\mathrm{d}s \le \liminf_{n\to\infty} F_{A_n}(t); \qquad \int_t^\infty G(s)\mathrm{d}s \le \liminf_{n\to\infty} \left(1 - F_{A_n}(t)\right)$$

and therefore under our assumption on the limit G, F_{A_n} converges to F and

$$F(t) = \int_0^t G(s)\mathrm{d}s. \quad \square$$

Acknowledgements

The authors received a Capes-Cofecub grant, no. 545/07. MA is partially supported by CNPq grant 312904/2009-6.

References

[1] M. Abadi, Hitting, returning and the short correlation function, Bull. Braz. Math. Soc. (NS) 37 (2006) 593–609.
[2] R. Bradley, Basic properties of strong mixing conditions. A survey and some open questions. Update of, and a supplement to, the 1986 original, Probab. Surv. 2 (2005) 107–144.
[3] L. Cardeño, Comportamentos globais et locais de tempos de entrada curtos, Thesis, Univ. São Paulo, Brazil, 2007.
[4] I.F. Cornfeld, S.V. Fomin, Ya.G. Sinai, Ergodic Theory, Springer, Berlin, Heidelberg, New York, 1982.
[5] A. Galves, B. Schmitt, Occurrence times of rare events for mixing dynamical systems, Ann. Inst. Henri Poincare Sect. A Phys. Theor. 52 (1990) 267–281.
[6] P. Grzegorek, M. Kupsa, Return times in a process generated by a typical partition, Nonlinearity 22 (2) (2009) 371–379.
[7] N. Haydn, Y. Lacroix, S. Vaienti, Hitting and returning in ergodic dynamical systems, Ann. Probab. 33 (2005) 2043–2050.
[8] D. Loewenstern, P. Yianilos, Significantly lower entropy estimates for natural DNA sequences, J. Comput. Biol. 6 (1) (1999) 125–142.