BULLETIN OF THE BRAZILIAN MATHEMATICAL SOCIETY

Bull Braz Math Soc, New Series 37(4), 593-609

Hitting, returning and the short correlation function

Miguel Abadi

Abstract. We consider a stochastic process with the weakest mixing condition: the so called α. For *any* fixed n-string we prove the following results. (1) The hitting time has approximately exponential law. (2) The return time has approximately a convex combination between a Dirac measure at the origin and an exponential law. In both cases the parameter of the exponential law is $\lambda(A)\mathbb{P}(A)$ where $\mathbb{P}(A)$ is the measure of the string and $\lambda(A)$ is a certain autocorrelation function of the string. We show also that the weight of the convex combination is approximately $\lambda(A)$. We describe the behavior of this autocorrelation function. Our results hold when the rate of mixing decays polinomially fast with power larger than the golden number.

Keywords: Mixing, recurrence, rare event, hitting time, return time, short correlation.

Mathematical subject classification: 60G10, 60F05, 60C05, 60G40, 37A50, 37A25.

1 Introduction

In the statistical analysis of Poincaré's recurrence it is well known that "occurrence times have exponential limit distribution law". This rough affirmation thus stated in some cases leaves many open questions and in others is miss-leading. For instance we would point out some questions:

(a) Under what hypothesis one has exponential times?

(b) What kind of occurrence time?

(c) Limit in which sense?

(d) Limit for what kind of sets and/or points?

(e) What is the parameter of the exponential law?

Received 31 March 2006.

With respect to (a) one would like to have the weakest hypothesis for which one has exponential times. This immediately impose a second question: we only have limit exponential law or have a rate of convergence of this limit which gives an error term for a finite approximation? And moreover, what kind of error term: total variation distance as was usually taken in the literature (see Abadi and Galves [3] for a review) or point-wise error as was introduced by the author in [2]?

With respect to (b) we could mention without exhausting the posibilities the hitting time (first time a process enters a fixed set), return time (first time the process comes back to a fixed set), repetition time (first time the process comes back to their first, non-fixed, state), waiting time (first time the process enters a set chosen independently from another copy of the process), among others.

The latters were proved to be of major importance due to their relationship with the entropy of the process (Wyner-Ziv [16], Orstein-Weiss [14], Shields [15]). Since them can be decomposed conditioning to the initial condition of the process, the formers are of major importance to describes the latters.

Of course (c) refers to limit in distribution, in probability or almost everywere.

With respect to (d), in general, results in the dynamical systems setting present nice results that hold for almost every point. However, it turns out that to present a full description of the repetition and waiting times it is necessary, as appears in Collet-Galves-Schmitt [6], Haydn-Vaienti [11], Abadi-Vaienti [4], to describe the statistics of a whole partition of the space without leaving any set or point.

With respect to (e), Kac's lemma ([13]) is commonly used to guess that the parameter of the exponential law should be the measure of the observable. However, since the seminal paper of Galves and Schmitt ([9]), it is known that for certain observables, the parameter it is not just the measure of the observable, but a certain correction factor must be introduced to get convergence to the exponential law, even when the observable could be very simple like a cylinder set. Later on, Abadi ([1]) shows that this correction factor is non-trivial and describes the short correlation of the process conditioned to the observable.

In the context (a)-(e), this paper is devoted to explore two questions:

- What is the largest class of systems which have exponential hitting and/or return times?
- What is the behavior of the short correlation function of a fixed observable?

α-mixing is the weakest hypothesis among several mixing conditions. We refer the reader to Doukhan ([8]) for a source of definitions and examples of the many mixing conditions. We prove that the hitting time of an n-string A converges in distribution, as n diverges, to an exponential law. The results holds for *every* string. The results holds for α-mixing processes with function α decreasing polynomially fast. It is quite surprising for us that the power of the polynomial must be larger than the golden number $(1+\sqrt{5})/2$. We recall that Chazottes ([5]) shows how to construct binary α-mixing processes with arbitrary polynomial rate of mixing. We concentrate our work in n-strings since any observable can be decomposed in n-strings, in particular the whole state space.

The convergence of the hitting time is obtained re-scaling it by a positive function $\lambda(A)\mathbb{P}(A)$ where $\mathbb{P}(A)$ is the measure of the string. $\lambda(A)$ is a certain function related to the short correlation of the process conditioned to start in A: physically it represent the mean probability that the process leaves the state A in a time not too big. We precise and describe this function.

For the return time to A, we prove that, under the same conditions, the return time law approaches to a law that is a convex combination of a Dirac measure concentrated at the origin and an exponential law. The re-scaling factor of the exponential law is the same as in the hitting time case. The weight of the convex combination is again a short correlation function related to $\lambda(A)$.

A remarkable point of our work is that our results hold for *every* string. Dynamically, this means that we prove exponential limit laws when the limit is taken along *any* point x, in contrast with previous works which find exponential law for *almost* every point. To get the exponential limit law we only have to consider the re-scaled function $\lambda(A)\mathbb{P}(A)\tau_A$ instead of the traditional re-scaled function $\mathbb{P}(A)\tau_A$. In a recent paper, Haydn, Lacroix and Vaienti ([10]) in a very general framework prove that the convergence of a sequence of hitting times is equivalent to the convergence of the return times, without introducing the auto-correlation function (case $\lambda(A) = 1$), and relate the limiting distribution of one with the other.

Another remarkable point of our results is the weakness of the hypothesis considered.

The results presented in this paper are extensions of those in Hirata, Saussol and Vaienti ([12]) and Abadi ([2]) which basically proved that hitting and return time laws are exponentially distributed when the process is α-mixing with exponential mixing rate.

This paper is organized as follows. In section 2 we establish our framework. In section 3 we define several short correlation functions and establish some of their basic properties. In section 4 we establish the limiting hitting time distribution with its own rate of convergence. This is Theorem 6. In section 5.2 we establish the limiting return time distribution with its own rate of convergence. This is Theorem 7.I Since it depends on certain overlapping properties of the string, we first introduce them in section 5.1. The proof is in section 5.3.

2 Framework and notations

Let C be a (non-empty) finite or countable set. Put $\Omega = C^{\mathbb{Z}}$. For each $x = (x_m)_{m\in\mathbb{Z}} \in \Omega$ and $m \in \mathbb{Z}$, let $X_m : \Omega \to C$ be the m-th coordinate, that is $X_m(x) = x_m$. We denote by $T : \Omega \to \Omega$ the one-step-left shift operator, namely $(T(x))_m = x_{m+1}$.

We say that a subset $A \subseteq \Omega$ is a n-string if $A \in C^n$ and

$$A = \{X_0 = a_0, \dots, X_{n-1} = a_{n-1}\} ,$$

with $a_i \in C,\ i = 0, \dots, n-1$. We use the probabilistic notation: $\{X_n^m = x_n^m\} = \{X_n = x_n, \dots, X_m = x_m\}$. For $t \in \mathbb{Z}$ we write $\tau_A^{[t]}$ to mean $\tau_A \circ T^t$.

We consider an invariant probability measure $\mathbb{P}$ over the σ-algebra generated by the strings. We shall assume without loss of generality that there is no singleton of probability 0.

We say that the process $\{X_m\}_{m\in\mathbb{Z}}$ is α-mixing if the sequence

$$\alpha(l) = \sup |\mathbb{P}(B \cap C) - \mathbb{P}(B)\mathbb{P}(C)| \ ,$$

converges to zero. The supremum is taken over B and C such that $B \in \sigma(X_0^n)$, $C \in \sigma(X_{n+l+1}^{\infty})$.

For two measurables V and W, we denote as usual $\mathbb{P}(V|W) = \mathbb{P}_W(V) = \mathbb{P}(V; W)/\mathbb{P}(W)$ the conditional measure of V given W. We write $\mathbb{P}(V; W) = \mathbb{P}(V \cap W)$. We also write $V^c = \Omega \backslash V$, for the complement of V. Finally, logarithms will be taken in base e.

3 Short correlation functions

In this section we introduce several notions of short correlation functions. Despite the fact that they are interesting by itselves, in the next two sections it will appear their relationship with the hitting and return time distributions of an onservable.

Given $A \in C^n$, we define the *hitting time* τ_A as the first time the string A appears in the infinite sequence x. Namely, we define $\tau_A : \Omega \to \mathbb{N} \cup \{\infty\}$ as the following random variable: For any $x \in \Omega$

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

For each n-string A take any g and f such that $n \leq g \leq f \leq 1/\mathbb{P}(A)$, Let us define the short correlation function introduced by Galves and Schmitt in [9]: $\lambda_{g,f}(A) : \cup_n C^n \to (0,3)$ as follows:

$$\lambda_{f,g}(A) = \frac{-\log \mathbb{P}\left(\tau_A > f - g\right)}{f\mathbb{P}(A)} .$$

The next lemma presents an equivalent (up to $2f\mathbb{P}(A)$) expression for $\lambda_{f,g}(A)$.

Lemma 1. *For any process* $\{X_m\}_{m\in\mathbb{Z}}$

$$\frac{\mathbb{P}\left(\tau_A \leq f - g\right)}{f\mathbb{P}(A)} \leq \lambda_{f,g}(A) \leq \frac{\mathbb{P}\left(\tau_A \leq f - g\right)}{f\mathbb{P}(A)} + 2f\mathbb{P}(A) .$$

Proof. Taylor's expansion says that $1 - e^{-x} \leq x \leq 1 - e^{-x} + 2(1 - e^{-x})^2$ for $0 \leq x \leq \log 3$. Apply it with $x = -\log \mathbb{P}\left(\tau_A > f - g\right)$. The lemma follows. □

To understand the expression $\mathbb{P}(\tau_A \leq f - g)/f\mathbb{P}(A)$ we must compute the numerator $\mathbb{P}(\tau_A \leq f - g)$. In practice f has order much larger than g so we can replace $f - g$ by f with just introducing a small error.

Lemma 2. *For any process* $\{X_m\}_{m\in\mathbb{Z}}$

$$0 \leq \frac{\mathbb{P}(\tau_A \leq f)}{f\mathbb{P}(A)} - \frac{\mathbb{P}(\tau_A \leq f - g)}{f\mathbb{P}(A)} \leq \frac{g}{f} .$$

Proof. By stationarity we get

$$\begin{aligned}
\mathbb{P}(\tau_A \leq f) - \mathbb{P}(\tau_A \leq f - g) &= \mathbb{P}(\tau_A > f - g \,;\, \tau_A^{[\tau_A]} \leq g) \\
&\leq \mathbb{P}(\tau_A \leq g) \\
&\leq g\mathbb{P}(A) .
\end{aligned}$$

Of course, the other inequality is obvious. □

Lemma 1 and Lemma 2 say that asymptotically $\lambda_{f,g}(A)$ has the same behavior as $\mathbb{P}(\tau_A \leq f)/f\mathbb{P}(A)$. Our last lemma says that, physically speaking, $\lambda_{f,g}(A)$ can be regarded as the mean probability the process takes to leave the state A.

Lemma 3. *For any stationary process* $\{X_m\}_{m\in\mathbb{Z}}$

$$\frac{\mathbb{P}(\tau_A \leq f)}{f\mathbb{P}(A)} = \frac{1}{f}\sum_{i=0}^{f-1} \mathbb{P}_A(\tau_A > i) .$$

Proof. Using stationarity it is very simply to see that $\mathbb{P}(\tau_A = t) = \mathbb{P}(A; \tau_A > t-1)$ which immediately implies the result. $\square$

The next two lemmas tend to bring some light on $\mathbb{P}(\tau_A \leq f)$. The first one is more a trivial observation that shows basics lower and upper bounds that hold for *any* process. The second one establishes for α-mixing processes that except of a factor belonging to the interval $[1/s, 1]$, with certain s, $\mathbb{P}(\tau_A \leq f)$ behaves like $f\mathbb{P}(A)$.

Lemma 4. *For any process* $\{X_m\}_{m\in\mathbb{Z}}$

$$\mathbb{P}(A) \leq \mathbb{P}(\tau_A \leq f) \leq f\mathbb{P}(A) . \tag{1}$$

Proof. Both inequalities are trivial.
– Lower bound:

$$\mathbb{P}(A) = \mathbb{P}(\tau_A = 1) \leq \mathbb{P}(\tau_A \leq f) .$$

– Upper bound:

$$\{\tau_A \leq f\} = \bigcup_{i=1}^{f} \{\tau_A = i\} \subseteq \bigcup_{i=1}^{f} T^{-i} A .$$

Thus the inequality follows by stationarity. $\square$

Lemma 5. *If* $\{X_m\}_{m\in\mathbb{Z}}$ *is* α*-mixing then for any* s *and* f *such that* $n \leq s < f < \mathbb{P}(A)$ *the following inequality holds*

$$\frac{f\mathbb{P}(A)}{s}\left(1 - \frac{\alpha(s-n)}{\mathbb{P}(A)}\right) \leq \mathbb{P}\left(\tau_A \leq f\right) . \tag{2}$$

Proof. Firstly we show a general inequality iterating the α-mixing property. Suppose that $S_i \in \sigma(X_{is+1}^{is+t})$ with some $t < s$ and for $i = 1, ..., m$. Then

$$\begin{aligned} \mathbb{P}\left(\bigcap_{i=1}^{m} S_i\right) &\leq \mathbb{P}\left(\bigcap_{i=1}^{m-1} S_i\right)\mathbb{P}(S_m) + \alpha(s-t) \\ &\leq \mathbb{P}\left(\bigcap_{i=1}^{m-2} S_i\right)\mathbb{P}(S_{m-1})\mathbb{P}(S_m) + \mathbb{P}(S_m)\alpha(s-t) + \alpha(s-t) . \end{aligned}$$

Iterating this procedure we get

$$\mathbb{P}\left(\bigcap_{i=1}^{m} S_i\right) \le \prod_{i=1}^{m} \mathbb{P}\,(S_i) + \alpha(s-t)\sum_{i=3}^{m+1}\prod_{j=i}^{m}\mathbb{P}(S_j)\ ,$$

where with some abuse of notation we $\prod_{j=m+1}^{m}\mathbb{P}(S_j)=1$.

Now consider the set $\{\tau_A > f\}$. Introducing gaps of length $s-n$ in between the sets $S_i = T^{-is+1}A^c$, with $i=0,\ldots,[f/s]$ we have

$$\{\tau_A > f\} \subseteq \bigcap_{i=1}^{[f/s]} T^{-is+1}A^c\ .$$

Applying the above inequality we get

$$\begin{aligned}\mathbb{P}\{\tau_A > f\} &\le (1-\mathbb{P}(A))^{f/s} + \alpha(s-n)\sum_{i=0}^{[f/s]-1}(1-\mathbb{P}(A))^i \\ &\le (1-\mathbb{P}(A))^{f/s} + \alpha(s-n)\frac{1-(1-\mathbb{P}(A))^{f/s}}{\mathbb{P}(A)}\ .\end{aligned}$$

Thus

$$\left(1-(1-\mathbb{P}(A))^{f/s}\right)\left(1-\frac{\alpha(s-n)}{\mathbb{P}(A)}\right) \le \mathbb{P}\{\tau_A \le f\}\ .$$

The conclusion follows by Taylor's expansion of e^{-x}. □

So far, we have proved that the short correlation function $\lambda_{f,g}(A)$ is bounded from above by a constant.

Under much stronger hypothesis it was firstly shown by Galves and Schmitt in [9] and further by the author in [1] and [2] that it is also bounded from below by a constant (which only depends on the properties of the measure $\mathbb{P}$). Under our current much weaker hypothesis, namely assuming without loose generality that α is decreasing, we only get the lower bound $1/s$ where $s = \alpha^{-1}(C\mathbb{P}(A)) + n$ for some constant $C \in (0,1)$.

The crucial point is that in general $\lambda_{f,g}(A)$ is difficult to compute explicitly. We would like to have a way to compute it. Under extra hypothesis, in the above referred papers [1] and [2], the author shows that it can be replaced by a much computable quantity (that we will introduce later on in section 5.1 for the convenience of the exposure) that depends on the overlapping properties of the considered string.

4 Hitting time

Theorem 6. *Let* $\{X_m\}_{m\in\mathbb{Z}}$ *be an* α*-mixing process. Suppose that* $\alpha(x) \le x^{-\kappa}$ *with* $\kappa > (1+\sqrt{5})/2$*. Then, there exists a function* $\lambda(A) : \cup_n C^n \to (0,3]$ *such that for any* $A \in C^n$*,*

$$\lim_{n\to\infty} \sup_{t\ge 0} \left| \mathbb{P}\left(\tau_A > \frac{t}{\lambda(A)\mathbb{P}(A)}\right) - e^{-t} \right| = 0\,. \tag{3}$$

Moreover the rate of convergence of the above limit is bounded from above by

$$e_h(A) = C_0 \inf_{n\le g\le f\le 1/\mathbb{P}(A)} \left[f\mathbb{P}(A) + \frac{g\mathbb{P}(A)+\alpha(g)}{f\mathbb{P}(A)} s \right], \tag{4}$$

where $s = \alpha^{-1}(\mathbb{P}(A)) + n$ *and* C_0 *is a positive constant.*

Proof. First we prove the theorem for t of the form kf where k is a positive integer and f is a certain "scale", $n \le f \le 1/\mathbb{P}(A)$. Then we prove the theorem for a general t.

Step 1: First we prove that for all $M \ge 0$ and $M' \ge g \ge 0$

$$\left|\mathbb{P}\left(\tau_A > M+M'\right) - \mathbb{P}\left(\tau_A > M\right)\mathbb{P}\left(\tau_A > M'-g\right)\right| \le g\mathbb{P}(A) + \alpha(g). \tag{5}$$

To simplify notation denote $\tau_A \circ T^k$ by $\tau_A^{[k]}$. We introduce a gap of length g after coordinate M to construct the following triangular inequality

$$\begin{aligned} &\left|\mathbb{P}\left(\tau_A > M+M'\right) - \mathbb{P}\left(\tau_A > M\right)\mathbb{P}\left(\tau_A > M'-g\right)\right| \\ \le\ & \left|\mathbb{P}\left(\tau_A > M+M'\right) - \mathbb{P}\left(\tau_A > M;\, \tau_A^{[M+g]} > M'-g\right)\right| \qquad (6) \\ +\ & \left|\mathbb{P}\left(\tau_A > M;\, \tau_A^{[M+g]} > M'-g\right) - \mathbb{P}\left(\tau_A > M\right)\mathbb{P}\left(\tau_A > M'-g\right)\right|. \qquad (7) \end{aligned}$$

Term (6) is equal to

$$\mathbb{P}\left(\tau_A > M;\, \tau_A^{[M]} \le g;\, \tau_A^{[M+g]} > M'-g\right) \le \mathbb{P}\left(\tau_A \le g\right) \le g\mathbb{P}(A)\,.$$

First inequality follows by stationarity. Term (7) is bounded using the α-mixing property by $\alpha(g)$. Thus we conclude (5).

Now take any $n < g < f \leq 1/\mathbb{P}(A)$. The triangle inequality leads to

$$\left|\mathbb{P}\left(\tau_A > kf\right) - \mathbb{P}\left(\tau_A > f\right)\mathbb{P}\left(\tau_A > f-g\right)^{k-1}\right|$$
$$\leq \sum_{j=2}^{k} |\mathbb{P}(\tau_A > jf) - \mathbb{P}(\tau_A > (j-1)f)\mathbb{P}(\tau_A > f-g)|\, \mathbb{P}(\tau_A > f-g)^{k-j}.$$

By (5) the modulus in the above sum is bounded by

$$g\mathbb{P}(A) + \alpha(g)\ ,$$

for all j. Further

$$\sum_{j=2}^{k} \mathbb{P}\left(\tau_A > f-g\right)^{k-j} \leq \frac{1}{\mathbb{P}\left(\tau_A \leq f-g\right)}\ .$$

Step 1 follows.

Step 2: Remember that $\lambda_{f,g}(A) = -\log \mathbb{P}\left(\tau_A > f-g\right)/f\mathbb{P}(A)$. Write $t = kf + r$ with k positive integer. Consider the following triangle inequality

$$\begin{aligned}\left|\mathbb{P}\left(\tau_A > t\right) - e^{-\lambda_{f,g}(A)\mathbb{P}(A)t}\right| &\leq |\mathbb{P}\left(\tau_A > t\right) - \mathbb{P}\left(\tau_A > kf\right)| \\ &+ \left|\mathbb{P}\left(\tau_A > kf\right) - \mathbb{P}\left(\tau_A > f-g\right)^k\right| \\ &+ e^{-\lambda_{f,g}(A)\mathbb{P}(A)kf}\left|1 - e^{-\lambda_{f,g}(A)\mathbb{P}(A)r}\right|\ .\end{aligned}$$

The first term is $\mathbb{P}\left(\tau_A > kf\ ;\ \tau_A^{[kf]} \leq s\right)$ which is bounded by $\mathbb{P}\left(\tau_A \leq s\right) \leq r\mathbb{P}(A) \leq f\mathbb{P}(A)$. The second term was bounded in step 1. Finally, the modulus in the third term is bounded using the Mean Value Theorem by $\lambda_{f,g}(A)\mathbb{P}(A)f$. This ends step 2.

Putting together steps 1 and 2 we get that for $t \geq 0$

$$\left|\mathbb{P}\left(\tau_A > t\right) - e^{-\lambda_{f,g}(A)\mathbb{P}(A)t}\right| \leq f\mathbb{P}(A) + \frac{g\mathbb{P}(A) + \alpha(g)}{\mathbb{P}(\tau_A \leq f-g)}\ .$$

Now we recall Lemma 5 to bound the above expression by

$$f\mathbb{P}(A) + \frac{g\mathbb{P}(A) + \alpha(g)}{(f-g)\mathbb{P}(A)}\, C_1\, s\ , \tag{8}$$

with a constant

$$C_1 = \frac{1}{1 - \frac{\alpha(s-n)}{\mathbb{P}(A)}} ,$$

provided that $\alpha(s-n) < \mathbb{P}(A)$ for all n. Thus, in order to prove (3), we have to chose f, g, s for each A such that they satisfy the following four constraints:

(a) $f\mathbb{P}(A) \to 0$, as $n \to \infty$,

(b) $gs/(f-g) \to 0$, as $n \to \infty$,

(c) $\alpha(g)s/(f-g)\mathbb{P}(A) \to 0$, as $n \to \infty$,

(d) there exists $C \in (0,1)$ such that $\alpha(s-n) = C\mathbb{P}(A)$.

Since $\alpha(x) \le x^{-\kappa}$ we first chose $s = (C\mathbb{P}(A))^{-1/\kappa} + n$ with any constant $C \in (0,1)$. This implies (d).

Choose $f = \mathbb{P}(A)^{-1+\varepsilon}$, thus we have (a). Then choose $g = \mathbb{P}(A)^{-1+\delta}$ with $0 < \varepsilon < \delta < 1$. With this choice of f and g we have that for large enough n, there exists a positive constant C_2 such that $f - g \ge C_2 f$. Constraints (b) and (c) become

$$\delta - \varepsilon - 1/\kappa > 0 \quad \mathit{and} \quad (1-\delta)\kappa - 1/\kappa - \varepsilon > 0 ,$$

respectively, where κ is given. Solving these inequalities we find that there exist such ϵ and δ if and only if $\kappa > (1+\sqrt{5})/2$. For a given $\kappa > (1+\sqrt{5})/2$, among the possible solutions of ε and δ, that is of f and g, chose those that minimize (8). Then define $\lambda(A) = \lambda_{f,g}(A)$ for these f and g. Now make the change of variables $t' = \lambda(A)\mathbb{P}(A)t$ Since this holds for all $C \in (0,1)$ of constraint (d), we can take $s = \alpha^{-1}(\mathbb{P}(A)) + n$. This ends the proof. □

5 Return times

5.1 Overlapping

For $A \in C^n$ define the first overlapping position of A as

$$\tau(A) = \min\left\{k \in \{1, \ldots, n\} \mid A \cap T^{-k}(A) \neq \emptyset\right\} .$$

Write $n = q\,\tau(A) + r$, with $q = [n/\tau(A)]$ and $0 \le r < \tau(A)$. Thus

$$A = \left\{X_0^{\tau(A)-1} = X_{\tau(A)}^{2\tau(A)-1} = \cdots = X_{(q-1)\tau(A)}^{q\tau(A)-1} = a_0^{\tau(A)-1} \; ; \; X_{q\tau(A)}^{n-1} = a_0^{r-1}\right\} .$$

For instance, in the following 15-string one has $\tau(A) = 6$

$$A = (\overbrace{\mathsf{aaaabb}}^{\tau(\mathrm{A})}\overbrace{\mathsf{aaaabb}}^{\tau(\mathrm{A})}\overbrace{\mathsf{aaa}}^{\mathrm{r}}) \ . \tag{9}$$

Consider the set of overlapping positions of A:

$$\begin{aligned} &\left\{k \in \{1, \dots, n-1\} \mid A \cap \ T^{-k}(A) \neq \emptyset\right\} \\ &= \{\tau(A), \dots, [n/\tau(A)]\tau(A)\} \cup \mathcal{R}(A) \ , \end{aligned}$$

where

$$\mathcal{R}(A) = \left\{k \in \{[n/\tau(A)]\tau(A) + 1, \dots, n-1\} \mid A \cap \ T^{-k}(A) \neq \emptyset\right\} \ .$$

Observe that $\#\mathcal{R}(A) \leq r < n/2$. For instance, in the string given in (9), one has $\mathcal{R}(A) = \{13, 14\}$. Further, consider an infinite sequence that begins with A. In such a sequence A can not reappear before $\tau(A)$. Thus, $\mathbb{P}_A\left(\tau_A < \tau(A)\right) = 0$. Still, if A does not reappear at time $\tau(A)$, then it can not reappear at times $k\tau(A)$, with $1 \leq k \leq [n/\tau(A)]$, so one has

$$\mathbb{P}_A\left(\tau(A) < \tau_A \leq [n/\tau(A)]\tau(A)\right) = 0.$$

One concludes that the *first* possible return after $\tau(A)$ is

$$n_A = \begin{cases} \min \mathcal{R}(A) & \mathcal{R}(A) \neq \emptyset \\ n_A = n & \mathcal{R}(A) = \emptyset \end{cases} \ .$$

Observe that by construction $n_A > n/2$.

5.2 Results

The *return time* is the hitting time restricted to the set A, namely $\tau_A|_A$. Formally, given $A \in C^n$, we define the *return time* $\tau_A : A \to \mathbb{N} \cup \{\infty\}$ as the following random variable: For any $x \in A$

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

We remark the difference between τ_A and $\tau(A)$ defined in the previous section: while $\tau_A(x)$ is the first time A appears in the infinite sequence x, $\tau(A)$ is the first overlapping position of A.

It would be useful for the reader to note now that according to the comments of the previous section, one has

$$\tau_A|_A \ \in \ \{\tau(A)\} \cup \mathcal{R}(A) \cup \{k \in \mathbb{N} \mid k \geq n\} \ .$$

To simplify notation, for any $n \leq f \leq 1/\mathbb{P}(A)$ put

$$\zeta_{A,f} \stackrel{def}{=} \mathbb{P}_A(\tau_A > \tau(A) + f) \ .$$

Theorem 7. *Let* $\{X_m\}_{m\in\mathbb{Z}}$ *be a* α*-mixing process. Then for any* $A \in C^n$*, the following holds:*

$$\lim_{n\to\infty} \mathbb{P}_A(\tau_A \geq 0) = 1 \ ,$$

and

$$\sup_{t>0} \left| \mathbb{P}_A \left(\tau_A > \frac{t}{\lambda(A)\mathbb{P}(A)} \right) - \zeta_{A,f} e^{-t} \right| \leq e_r(A) \ , \tag{10}$$

where

$$e_r(A) = \frac{\alpha(f)}{\mathbb{P}(A)} + 6e_h(A) \ ,$$

and f *defines* $e_h(A)$ *and* $\lambda(A)$*. Further, if* $\alpha(x) \leq x^{-\kappa}$ *with* $\kappa > (1+\sqrt{5})/2$*, then* $e_r(A)$ *goes to zero as n goes to infinity.*

Remark 8. *Theorem 7says that in contrast with the (re-scaled) hitting time that has exponential limit law for any string, the (re-scaled) return time can present different limiting behaviors.*

- *When* $\zeta_{A,f}$ *remains bounded away from zero and one,* $\lambda(A)\mathbb{P}(A)\tau_A$ *approaches to* $(1-\zeta_{A,f})\delta_0 + \zeta_{A,f}X$ *where* δ_0 *is the Dirac measure at the origin and* $X \sim \exp(1)$.
- *When* $\zeta_{A,f}$ *goes to one (and therefore* $\lambda(A)$ *does it too by Lemma 1 and Lemma 3), then* $\lambda(A)\mathbb{P}(A)\tau_A$ *(and therefore* $\mathbb{P}(A)\tau_A$*) converges to a purely* $\exp(1)$ *law.*
- *When* $\zeta_{A,f}$ *goes to zero, then* $\lambda(A)\mathbb{P}(A)\tau_A$ *converges to a degenerated law at the origin.*

We say something more about this in the next two lemmas.

As explained at the end of section 3, $\lambda(A)$ and also $\zeta_{A,f}$ are in practice, difficult to handle. Under extra hypothesis on the mixing rate of the process a much easier quantity can replace them.

Lemma 9. *Suppose that* $\{X_m\}_{m\in\mathbb{Z}}$ *is* α*-mixing. Then*

$$|\mathbb{P}_A(\tau_A > \tau(A)) - \zeta_{A,f}| \leq f\mathbb{P}(A) + 2 \inf_{0\leq w\leq n_A} \left\{ n\mathbb{P}(A^{(w)}) + \frac{\alpha(n_A - w)}{\mathbb{P}(A)} \right\} \ .$$

Lemma 10. *Suppose that* $\{X_m\}_{m\in\mathbb{Z}}$ *is* α*-mixing. Then*

$$|\mathbb{P}_A(\tau_A > \tau(A)) - \lambda(A)| \le 3f\mathbb{P}(A) + \frac{g}{f} + 2\inf_{0\le w\le n_A}\left\{n\mathbb{P}(A^{(w)}) + \frac{\alpha(n_A - w)}{\mathbb{P}(A)}\right\}.$$

Remark 11. *According to Shanon-Mac-Millan-Breiman Theorem (see e.g. [7]), almost every string has exponential measure (with rate close to the entropy). Basically, the above two lemmas say that if* α *decays exponentially fast, then it is ok to approximate* $\zeta_{A,f}$ *and* $\lambda(A)$ *by* $\mathbb{P}(\tau_A > \tau(A))$ *(as observed at the end of section 5.1, one has* $n_A > n/2$*).*

Under extra conditions on the rate of mixing of the process *and on the overlapping properties of* A we have a purely exponential limit law for both hitting and return times.

Lemma 12. *Suppose that* $\{X_m\}_{m\in\mathbb{Z}}$ *is* α*-mixing. Then*

$$|\mathbb{P}_A(\tau_A > \tau(A)) - 1| \le \inf_{0\le w\le \tau(A)}\left\{\mathbb{P}(A^{(w)}) + \frac{\alpha(\tau(A) - w)}{\mathbb{P}(A)}\right\}.$$

Remark 13. *We remark strongly the above three lemmas hold just under the* α*-mixing hypothesis. However, lemmas (9) and (9) are only useful whenever* $\alpha(n_A - w)/\mathbb{P}(A)$ *is small for some* w*. In Lemma (12) we need a stronger condition:* $\tau(A)$ *must be large enough to make* $\mathbb{P}(A^{(w)})$ *and* $\alpha(\tau(A) - w)/\mathbb{P}(A)$ *small for some* w*. This means basically* $\tau(A) \ge Cn$ *for some positive constant* C *and* α *decaying exponentially fast.*

5.3 Proofs

Proof of Theorem 7. We observe that the distribution of $\lambda(A)\mathbb{P}(A)\tau_A$ is a discrete one over the set $\lambda(A)\mathbb{P}(A)\mathbb{N}$ and its limit is a distribution over $\Re \ge 0$. Thus $\mathbb{P}_A(\lambda(A)\mathbb{P}(A)\tau_A \ge 0) = 1$. Now we proside to prove the theorem for $t > 0$.

First we prove that for all $M \ge \tau(A) + f$ and $M' \ge f \ge 0$ the following inequality holds

$$\begin{aligned}&\left|\mathbb{P}_A\left(\tau_A > M + M'\right) - \mathbb{P}_A\left(\tau_A > M\right)\mathbb{P}\left(\tau_A > M' - f\right)\right| \\ &\qquad \le f\mathbb{P}(A) + 2\frac{\alpha(f)}{\mathbb{P}(A)}.\end{aligned} \tag{11}$$

We use again $\tau_A^{[t]}$ to mean $\tau_A \circ T^t$. The proof follows the steps of (5). We introduce a gap of length f after M.

$$\begin{aligned} &\left|\mathbb{P}_A\left(\tau_A > M + M'\right) - \mathbb{P}_A\left(\tau_A > M\right)\mathbb{P}\left(\tau_A > M' - f\right)\right| \\ \leq &\left|\mathbb{P}_A\left(\tau_A > M + M'\right) - \mathbb{P}_A\left(\tau_A > M; \tau_A^{[M+f]} > M' - f\right)\right| \qquad (12) \\ + &\left|\mathbb{P}_A\left(\tau_A > M; \tau_A^{[M+f]} > M' - f\right) - \mathbb{P}_A\left(\tau_A > M\right)\mathbb{P}\left(\tau_A > M' - f\right)\right|. \qquad (13) \end{aligned}$$

Term (12) is equal to

$$\mathbb{P}_A\left(\tau_A > M; \tau_A^{[M]} \leq f; \tau_A^{[M+f]} > M' - f\right) \leq \mathbb{P}_A\left(\tau_A > \tau(A); \tau_A^{[M]} \leq f\right).$$

The α-mixing property applied over the last term bounds it by $\mathbb{P}(\tau_A \leq f) + \alpha(f)/\mathbb{P}(A)$. Term (13) is bounded using the α-mixing property by $\alpha(f)/\mathbb{P}(A)$. Thus we conclude (11).

Now we prove the theorem for $t \geq \tau(A)+2f$. Consider the triangle inequality

$$\begin{aligned} &\left|\mathbb{P}_A\left(\tau_A > t\right) - \mathbb{P}_A\left(\tau_A > \tau(A) + f\right) e^{-\lambda(A)\mathbb{P}(A)t}\right| \\ \leq\ &\left|\mathbb{P}_A\left(\tau_A > t\right) - \mathbb{P}_A\left(\tau_A > \tau(A) + f\right)\mathbb{P}\left(\tau_A > t - (\tau(A) + 2f)\right)\right| \\ +\ &\mathbb{P}_A\left(\tau_A > \tau(A) + f\right)\left|\mathbb{P}\left(\tau_A > t - (\tau(A) + 2f)\right) - e^{-\lambda(A)\mathbb{P}(A)t}\right| . \end{aligned}$$

The first term is bounded applying (11) by $f\mathbb{P}(A)+2\alpha(f)/\mathbb{P}(A)$. The second one is bounded applying Theorem 6 and then the Mean Value Theorem by $e_h(A) + \lambda(A)\mathbb{P}(A)(\tau(A) + 2f)$. The change of variables $t' = \lambda(A)\mathbb{P}(A)t$ shows that for $t' > \lambda(A)\mathbb{P}(A)(\tau(A) + 2f)$ one has

$$\left|\mathbb{P}\left(\tau_A > \frac{t'}{\lambda(A)\mathbb{P}(A)}\right) - e^{-t'}\right| \leq 4\left(f\mathbb{P}(A) + \frac{\alpha(f)}{\mathbb{P}(A)}\right) + e_h(A) .$$

Since $\lambda(A)\mathbb{P}(A)(\tau(A)+2f) \leq 6f\mathbb{P}(A)$ which goes to zero as n goes to infinity, (10) follows. We note that with respect to the proof of Theorem 6, we have the extra constrains

(e) $f\mathbb{P}(A) \to 0$ and (f) $\alpha(f)/\mathbb{P}(A)$.

Of course (e) is the same that (a) in Theorem 6. A straightforward computation shows that under a polynomial mixing rate (f) is weaker than (b) of Theorem 6. This concludes the proof. □

Proof of Lemma 9. A direct computation gives

$$\mathbb{P}_A(\tau_A > \tau(A)) - \mathbb{P}_A(\tau_A > \tau(A) + f) = \mathbb{P}_A(\tau_A > \tau(A);\, \tau_A^{[\tau(A)]} \leq f)\ .$$

For any $0 \leq w \leq n$, consider the w-string $A^{(w)} = \{X_{n-w}^{n-1} = a_{n-w}^{n-1}\}$. Namely, the string constructed with the *last* w-letters of A belonging to $\sigma(X_{n-w}^{n-1})$. Thus, according to the description of section 5.1

$$\begin{aligned} & A \cap \{\tau_A > \tau(A)\} \cap \{\tau_A^{[\tau(A)]} \leq f\} \\ \subseteq\ & A \cap \left(\bigcup_{i \in \mathcal{R}(A), i=n}^{2n-1} T^{-i} A^{(w)} \bigcup_{i=2n}^{\tau(A)+f} T^{-i} A \right) \\ =\ & \left(A \cap \bigcup_{i \in \mathcal{R}(A), i=n}^{2n-1} T^{-i} A^{(w)} \right) \cup \left(A \cap \bigcup_{i=2n}^{\tau(A)+f} T^{-i} A \right) . \end{aligned}$$

Now we bound the probability of the last expression using the α-mixing property with a gap of size $n_A - w$ over the first set and with a gap of size n over the second one in between A and the remaining set. Thus

$$\begin{aligned} & \mathbb{P}_A \left(\tau_A > \tau(A);\, \tau_A^{[\tau(A)]} \leq f \right) \\ \leq\ & 2n\mathbb{P}(A^{(w)}) + \frac{\alpha(n_A - w)}{\mathbb{P}(A)} + (f - n)\mathbb{P}(A) + \frac{\alpha(n)}{\mathbb{P}(A)}\ . \end{aligned}$$

This ends the proof. □

Proof of Lemma 10. This follows directly by Lemma 1, Lemma 2, Lemma 3, Lemma 9 and the fact that $\mathbb{P}_A(\tau_A > \tau(A)) \geq \mathbb{P}_A(\tau_A > j) \leq \mathbb{P}_A(\tau_A > \tau(A) + f)$ for all j such that $\tau(A) \leq j \leq \tau(A) + f$. □

Proof of Lemma 12. By definition of $\tau(A)$

$$1 - \mathbb{P}_A(\tau_A > \tau(A)) = \mathbb{P}_A(\tau_A = \tau(A)) = \mathbb{P}_A(T^{-n} A^{(\tau(A))})\ .$$

The last equality follows since

$$A \bigcap_{i=1}^{\tau(A)-1} T^{-i} A^c \cap T^{-\tau(A)} A = A \cap T^{-\tau(A)} A = A \cap T^{-n} A^{(\tau(A))}\ .$$

Now, for any $0 \le w \le \tau(A)$ one has $A^{(\tau(A))} \subseteq A^{(w)}$. Therefore, by the α-mixing property

$$\mathbb{P}_A(T^{-n}A^{(\tau(A))}) \le \mathbb{P}(A^{(w)}) + \frac{\alpha(\tau(A)-w)}{\mathbb{P}(A)}\,.$$

The proof follows. □

Acknowledgements. Work done as part of the activities of the project Projeto Temático-PRONEX *"Stochastic behavior, critical phenomena and rhythmic patterns identification in natural languages"*. Pros. FAPESP 03/09930-9.

References

[1] M. Abadi. Exponential approximation for hitting times in mixing processes. *Math. Phys. Elec. J.* **7**(2): (2001).

[2] M. Abadi. Sharp error terms and necessary conditions for exponential hitting times in mixing processes. *Ann. Probab.* **32**(1A): (2004), 243–264.

[3] M. Abadi and A. Galves. Inequalities for the occurrence times of rare events in mixing processes. The state of the art . *Markov Proc. Relat. Fields.* **7**(1): (2001), 97–112.

[4] M. Abadi and S. Vaienti. Full large deviation properties for return times. *Preprint.* (2006).

[5] J. R. Chazottes. Hitting and returning to non-rare events in mixing dynamical systems. *Nonlinearity*, **16** (2003), 1017–1034.

[6] P. Collet, A. Galves and B. Schmitt. Repetition times for gibbsian sources. *Nonlinearity*, **12** (1999), 1225–1237.

[7] I. Cornfeld, S. Fomin and Y. Sinai. Ergodic theory. *Grundlähren der Mathematischen Wissenschaften*, **245**. Springer-Verlag, New York. (1982).

[8] P. Doukhan. Mixing. Properties and examples. *Lecture Notes in Statistics*, **85** (1995), Springer-Verlag.

[9] A. Galves and B. Schmitt. Inequalities for hitting times in mixing dynamical systems. *Random Comput. Dyn.* **5** (1997), 337–348.

[10] N. Haydn, Y. Lacroix and S. Vaienti. Hitting and return times in ergodic dynamical systems. *Ann. Prob.* (2004).

[11] N. Haydn and S. Vaienti. The limiting distribution and error terms for return times of dynamical systems. *Discrete Contin. Dyn. Syst.* **33**(5): (2005), 2043–2050.

[12] M. Hirata, B. Saussol and S. Vaienti. Statistics of return times: a general framework and new applications. *Comm. Math. Phys.* **206** (1999), 33–55.

[13] M. Kac. On the notion of recurrence in discrete stochastic processes. *Bull. Amer. Math. Soc.* **53** (1947), 1002–1010.

[14] D. Ornstein and B. Weiss. Entropy and data compression schemes. *IEEE Trans. Inform. Theory*, **39**(1): (1993), 78–83.

[15] P. Shields. Waiting times: positive and negative results on the Wyner-Zyv problem. *J. Theo. Prob.* **6**(3): (1993), 499–519.

[16] A. Wyner and J. Ziv. Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression. *IEEE Trans. Inform. Theory*, **35**(6): (1989), 1250–1258.

Miguel Abadi
IMECC, Universidade Estadual de Campinas
Pça Sérgio Buarque de Holanda 651
Cidade Universitária, Caixa Postal 6065
13083-859 Campinas, SP
BRAZIL

E-mail: miguel@ime.unicamp.br