Commun. Math. Phys. 206, 33 –55 (1999)

Communications in
Mathematical Physics

Statistics of Return Times: A General Framework and New Applications

Masaki Hirata[1], **Benoît Saussol**[2], **Sandro Vaienti**[2]

[1] Mathematical Department, Tokyo Metropolitan University, Japan. E-mail: mhirata@comp.metro-u.ac.jp
[2] Centre de Physique Théorique, Luminy, Marseille and PHYMAT, Mathematical Department, University of Toulon, France. E-mail: saussol@cpt.univ-mrs.fr; vaienti@cpt.univ-mrs.fr

Received: 4 August 1998 / Accepted: 9 March 1999

Abstract: In this paper we provide general estimates for the errors between the distribution of the first, and more generally, the K^{th} return time (suitably rescaled) and the Poisson law for measurable dynamical systems. In the case that the system exhibits strong mixing properties, these bounds are explicitly expressed in terms of the speed of mixing. Using these approximations, the Poisson law is finally proved to hold for a large class of non hyperbolic systems on the interval.

1. Introduction

The investigation of asymptotically rare events is growing up as a new direction in the understanding of statistical properties of dynamical systems. By "asymptotically rare" events we mean, in a wide sense and following the terminology in the review paper of [Coe97], those events which have asymptotically zero probability but which occur with a well determined asymptotic limit law. In the dynamical setting, where we have a probability space $(X, \mathcal{B}, \mu)$ with a measurable μ-preserving mapping T acting on it, the "events" will usually be the visits into a sequence of sets $\Omega_k \in \mathcal{B}$ of positive measure but with their measure going to zero in the limit of large k. We call the event "rare", when the expected entrance time in Ω_k diverges with k. A well-known result in ergodic theory shows how abundant are the "asymptotically rare" events. Let us consider in fact an ergodic measure μ for an endomorphism T and take a measurable subset Ω: then Kac's theorem [CFS82] says that the expectation of the return time to Ω, starting from Ω, is just $\mu(\Omega)^{-1}$.

Kac's theorem suggests the good normalization to keep in order to study the asymptotic distribution of the return time to Ω. The natural object will thus be the distribution:

$$F_\Omega(t) = \mu_\Omega\left(x \in \Omega \;\middle|\; \tau_\Omega(x)\mu(\Omega) > t\right), \tag{1}$$

where $\tau_\Omega(x)$ is the first return time to Ω provided that $x \in \Omega$ and μ_Ω is the normalized restriction of μ to Ω. The question will be whether the limit of $F_\Omega(t)$ exists when the measure goes to zero and what kind of distribution is recovered. The condition that the starting point x in (1) belongs to Ω could be relaxed by asking that x belongs to the whole space. In this case, $F_\Omega(t)$ will give the distribution of the "visiting time" into Ω, but in order to get its asymptotic distribution, a suitable normalization is needed [GS97]. The situations sketched above could be considerably refined, producing richer processes (see the quoted paper [Coe97] for an historical account of these questions and an exhaustive bibliography). We will however explore some of them in this paper under a more general perspective and successively by giving applications to class of systems never investigated before.

Let us first come back to formula (1) and replace Ω with a decreasing sequence of neighborhoods of a given point $z \in X$, $\Omega_\varepsilon(z)$, such that their measure goes to zero when $\varepsilon \to 0^+$. Then for some classes of hyperbolic dynamical systems, notably axiom A diffeomorphisms [Hir93], transitive Markov chains [Pit91], expanding maps of the interval with a spectral gap [Col96] and in the more general setting of systems verifying a strong mixing property ("self-mixing" condition and φ-mixing [Hir95]), and recently even in the case of rational maps with critical points in the Julia set [Hay98a], it is possible to prove that the distribution $F_{\Omega_\varepsilon(z)}(t)$ goes to the exponential-one law e^{-t} and this for μ-almost every $z \in X$. A strong improvement of this kind of result appears in the paper [GS97], where an upper bound for the difference

$$\left|\mu\left(\tau_A(x) > \frac{t}{\mu(A)\lambda(A)}\right) - \mathrm{e}^{-t}\right|$$

was explicitly computed in the case of φ-mixing systems and where A is a cylinder set, and $\lambda(A)$ a suitable normalizing factor. Recently [Hay98b] obtained an exponential error estimate for the quantity like (1) in the case of parabolic rational maps.

To enrich the process, and the statistics, one successively introduce the K^{th} return time, $\tau^K_{\Omega_\varepsilon}(x)$, from Ω_ε into itself (see the precise definition in the next section), where $\Omega_\varepsilon = \Omega_\varepsilon(z)$ is still a neighborhood of some point $z \in X$.

For the dynamical systems quoted above, a Poisson statistics can be proved, by showing that the distribution of successive return times into Ω_ε satisfies, for z μ-a.e.

$$\mu_{\Omega_\varepsilon}\left(x \in \Omega_\varepsilon \;\middle|\; \tau^K_{\Omega_\varepsilon}(x) \le t < \tau^{K+1}_{\Omega_\varepsilon}(x)\right) \underset{\varepsilon\to 0^+}{\longrightarrow} \frac{t^K}{K!}\mathrm{e}^{-t}. \tag{2}$$

The preceding results deserve further investigations at least in two directions:

1. extend them to non-hyperbolic dynamical systems and, more ambitiously, check their robustness when the system loses strong mixing properties.
2. prove an error estimate even for the distribution of successive return times (2) and relate this approximation rate, if possible, to the statistical properties of the system like correlations decay or spectral properties.

We try to give partial answer to these questions in this paper. The general setting we put in, is the return(s) times to the set Ω starting from itself, as expressed in formulas (1) and (2) (although in Theorem 2.1 we will also consider points starting everywhere). The first attempt was to give, for measure preserving dynamical systems, a general upper bound for the difference between the distribution of the (rescaled) first return time and the exponential-one law e^{-t} and then between the distribution of high-order (rescaled) return times and the Poisson law $\frac{t^K}{K!}\mathrm{e}^{-t}$. We do not make any hypothesis on the set Ω, nor on the ergodic properties of μ; nevertheless these bounds are expressed in terms of the self-interactions of the set Ω and can be explicitly computed when typical rates of mixing are known (uniform mixing, α-mixing or φ-mixing). In this context, our bounds greatly improve and simplify the hypothesis of self-mixing condition of [Hir95], which was a powerful tool to get sufficient condition for the Poisson statistics. This first part of the paper is essentially due to one of us (B.S.) and is part of his Ph.D. Thesis [Sau98b]. In the second part we apply the preceding bounds to new situations. The systems we treat are some non-uniformly hyperbolic maps of the interval; these maps are characterized by a structure parameter, say α, which measures the order of tangency at a neutral fixed point and governs the algebraic decay of correlations (in our example the order is $n^{1-1/\alpha}$). If μ denotes the absolutely continuous invariant measure, we prove Poisson statistics (in the sense precise above), by giving an explicit approximation of the asymptotic law in terms of the measure of the set Ω_n, where in this case Ω_n is a decreasing sequence of cylinder sets chosen around almost all points in the interval. To be precise the error is of the type: $\mu(\Omega_n)^\beta$, for any $\beta < 1 - \alpha$, and therefore β is explicitly related to α and optimized just by $1 - \alpha$. For the distributions of the K^{th} return times the bounds simply become $\mu(\Omega_n)^{\beta/K}$.

By inspecting these results, we could argue that the non-hyperbolic character of the maps reflects in the error term; to be more precise we think that as soon as the degree of non-uniform hyperbolicity of the map is monitored by a structure parameter α, this parameter will appear explicitly in the approximation to the Poisson law, which suggests, on the converse, that we could use Poissonian statistics to test lack of hyperbolicity. Our claim is motivated by two more observations: first, in getting these bounds we proved a sort of α-mixing for the map with a rate which was exactly the same as the algebraic rate for the correlations' decay. Second, in the forthcoming paper [Sau98a] the return times is analyzed for a class of piecewise expanding multidimensional maps. Although the mixing properties are much more difficult to handle with, especially for the presence of singularity lines and the geometry of their shape, the uniform dilatation will provide bounds on the form: $\mu(\Omega_n)^\beta$ and $\mu(\Omega_n)^{\beta/K}$ for all $\beta < 1$, which reflects the fact that all the quantities involved, and the correlations' decay too, admit exponential estimates. We will come back to these questions in Sect. 4. As a final remark, we address two questions:

1. Our analysis is local: the events are chosen around almost all points which we could call, following a widespread tradition, generic (for our statistics). What happens if we consider non-generic points (discarding of course some trivial situation like fixed points)? Could we see their (possibly different) statistics by involving some sort of large deviation argument ?
2. What is the place of Poissonian statistics regarding other ergodic characterizations of dynamical systems? For example: what is the largest class of ergodic dynamical systems enjoying a Poissonian statistics? Conversely, does an invariant measure satisfying that behavior imply strong ergodic properties too?

2. General Bounds on the Distribution of Return Times

We will consider in this section a probability space $(X, \mathcal{B}, \mu)$ together with a measure preserving transformation T acting on X. The basic object will be the return time into a positive measure set U starting from U defined by

$$\tau_U(x) = \inf \left\{ k \geq 1 |\, T^k x \in U \right\} \cup \{\infty\}.$$

We define as usual the conditional measure μ_U on U by $\mu_U(A) = \dfrac{\mu(A\cap U)}{\mu(U)}$. We then recall Kac's theorem which says that the conditional expectation of τ_U given U is finite, and equal to $1/\mu(U)$, when μ is ergodic. As indicated in the introduction, Kac's result suggests how to properly rescale the return time when we are interested in its distribution.

2.1. First return time. We begin to show that the distribution of the first return time into the set U starting from U is close to an exponential one law if and only if the two distributions of the first return time starting, respectively from U and everywhere, are close.

Theorem 2.1. *Let us define $c(k,U) = \mu_U(\tau_U > k) - \mu(\tau_U > k)$ and set $c(U) = \sup_k |c(k,U)|$. The distribution of the (rescaled) first return time into the set U differs from the exponential-one law by at most $d(U) := 4\mu(U) + c(U)(1 + \log c(U)^{-1})$, namely:*

$$\sup_{t\geq 0}\left|\mu\left(\tau_U > \frac{t}{\mu(U)}\right) - \mathrm{e}^{-t}\right| \leq d(U),$$

which is still true starting from U:

$$\sup_{t\geq 0}\left|\mu_U\left(\tau_U > \frac{t}{\mu(U)}\right) - \mathrm{e}^{-t}\right| \leq d(U).$$

Conversely, the difference between the two distributions (starting inside U and everywhere) can be bounded in terms of the distance $\widetilde{c}(U) := \sup_{t\geq 0} |\mu_U(\tau_U > t/\mu(U)) - \mathrm{e}^{-t}|$, precisely:

$$c(U) \leq 2\mu(U) + \widetilde{c}(U)(2 + \log \widetilde{c}(U)^{-1}).$$

Remark 2.2. Whenever $\mu(U) > 0$ the return time's law is discrete and this allow us to get a lower bound for the rate of convergence. More precisely, we have the following proposition:

Proposition 2.3. *For each $k \geq 0$,*

$$\begin{aligned}\varepsilon_{k,U} &:= \left|\mu\left(\tau_U > k\right) - \mathrm{e}^{-k\mu(U)}\right| + \left|\mu\left(\tau_U > k + 1/2\right) - \mathrm{e}^{-(k+1/2)\mu(U)}\right| \\ &\geq \frac{\mathrm{e}^{-k\mu(U)}}{4}\mu(U).\end{aligned}$$

In particular, $\varepsilon_{0,U} \geq \mu(U)/4$.

Proof of Proposition 2.3. Let $k \geq 0$ be an integer. Since τ_U takes only integer values, the distribution for $t = k\mu(U)$ and $t' = (k+1/2)\mu(U)$ is the same, then

$$\begin{aligned}\varepsilon_{k,U} &\geq |\exp(-k\mu(U)) - \exp(-(k+1/2)\mu(U))| \\ &\geq \exp(-k\mu(U))(1 - \mathrm{e}^{-\mu(U)/2}) \\ &\geq \frac{\mathrm{e}^{-k\mu(U)}}{4}\mu(U). \qquad \square\end{aligned}$$

Proof of Theorem 2.1. Let us remark that for any $k \geq 1$ we have

$$\mu(\tau_U = k) = \mu(U \cap \{\tau_U > k-1\}). \tag{3}$$

Since $\{\tau_U > k\} = T^{-1}(U^c \cap \{\tau_U > k-1\})$ by the invariance of μ we get that $\mu(\tau_U > k) = \mu(\tau_U > k-1) - \mu(U \cap \{\tau_U > k-1\})$, whence the result. Next, for all $k > 0$ we have

$$\begin{aligned}
\mu(\tau_U > k) &= \mu(\tau_U > k-1) - \mu(U)\mu_U(\tau_U > k-1) \\
&= \mu(\tau_U > k-1) - \mu(U)[\mu(\tau_U > k-1) + c(k,U)] \\
&= \mu(\tau_U > k-1)[1-\mu(U)] - \mu(U)c(k,U).
\end{aligned}$$

Then it follows by an immediate induction that

$$\mu(\tau_U > k) = (1-\mu(U))^k - \mu(U)\sum_{j=1}^{k} c(j,U)(1-\mu(U))^{k-j}.$$

Hence for all $t \geq 0$, putting $k_t = [t/\mu(U)]$, we have

$$\left|\mu(\tau_U > k_t) - (1-\mu(U))^{k_t}\right| \leq \mu(U)\sum_{j=1}^{k_t} |c(j,U)| \leq tc(U). \tag{4}$$

Setting $z = -\log c(U)$, and $k_z = [z/\mu(U)]$, we get

$$(1-\mu(U))^{k_z} \leq \mathrm{e}^{-k_z\mu(U)} \leq c(U)\mathrm{e}^{\mu(U)} \leq c(U) + 2\mu(U),$$

for any $t > z$,

$$\begin{aligned}
\mu(\tau_U > k_t) &\leq \mu(\tau_U > k_z) \\
&\leq (1-\mu(U))^{k_z} + zc(U) \\
&\leq 2\mu(U) + c(U)(1-\log c(U))
\end{aligned}$$

which gives $\left|\mu(\tau_U > k_t) - (1-\mu(U))^{k_t}\right| \leq 2\mu(U) + c(U)(1-\log c(U))$. Instead for any $t \leq z$ the same estimate holds by inequality (4). Since, by an easy computation

$$|(1-\mu(U))^{k_t} - \mathrm{e}^{-t}| \leq 2\mu(U),$$

we get for any $t \geq 0$,

$$\left|\mu(\tau_U > k_t) - \mathrm{e}^{-t}\right| \leq 4\mu(U) + c(U)(1-\log c(U)),$$

which proves the first part of the theorem. Moreover, since

$$\left|\mu_U(\tau_U > k_t) - \mu(\tau_U > k_t)\right| = |c(k_t,U)| \leq c(U),$$

we finally have for each $t \geq 0$,

$$\left|\mu_U(\tau_U > k_t) - \mathrm{e}^{-t}\right| \leq 4\mu(U) + c(U)(2-\log c(U)).$$

The converse part is proven in the same way. For $k \geq 1$,

$$\begin{aligned}\mu(\tau_U > k) &= 1 - \mu(\tau_U \leq k)\\ &= 1 - \sum_{j=1}^{k} \mu(\tau_U = j)\\ &= 1 - \mu(U) \sum_{j=1}^{k} \mu_U(\tau_U > j-1),\end{aligned}$$

where we used in the last equality the relation (3). Hence

$$\begin{aligned}|\mu(\tau_U > k) - \mathrm{e}^{-k\mu(U)}| &\leq \left|1 - \mu(U) \sum_{j=1}^{k} \mathrm{e}^{-(j-1)\mu(U)} - \mathrm{e}^{-k\mu(U)}\right| + k\mu(U)\widetilde{c}(U)\\ &\leq \left|1 - \mu(U) \frac{1 - \mathrm{e}^{-k\mu(U)}}{1 - \mathrm{e}^{-\mu(U)}} - \mathrm{e}^{-k\mu(U)}\right| + k\mu(U)\widetilde{c}(U)\\ &\leq (1 + \mathrm{e}^{-k\mu(U)}) \left|1 - \frac{\mu(U)}{1 - \mathrm{e}^{-\mu(U)}}\right| + k\mu(U)\widetilde{c}(U)\\ &\leq 2\mu(U) + k\mu(U)\widetilde{c}(U).\end{aligned}$$

This gives, whenever $k \leq k_0 := \log \widetilde{c}(U)^{-1}/\mu(U)$:

$$|c(k, U)| \leq 2\mu(U) + \widetilde{c}(U) \log \widetilde{c}(U)^{-1}.$$

For $k > k_0$ we simply have

$$\begin{aligned}|c(k, U)| &\leq \mu(\tau_U > k_0) + \mu_U(\tau_U > k_0)\\ &\leq 2\mu(U) + \widetilde{c}(U) \log \widetilde{c}(U)^{-1} + \mathrm{e}^{-k_0\mu(U)} + \widetilde{c}(U). \qquad \square\end{aligned}$$

The last theorem gives a *necessary and sufficient condition* to obtain the exponential law, that is $d(U) \to 0$. However, such a quantity is not very transparent for dynamical systems, that is why we give a criterion to estimate it. This kind of condition is a generalization of the so-called "self-mixing condition" introduced in [Hir95].

Lemma 2.4. *Let $U \subset X$ a measurable set. The following estimate holds:*

$$c(U) \leq \inf \{a_N(U) + b_N(U) + N\mu(U) | N \in \mathbb{N}\},$$

where the quantities are defined by

$$\begin{aligned}a_N(U) &= \mu_U(\bigcup_{j=1}^{N} T^{-j}U) = \mu_U(\tau_U \leq N),\\ b_N(U) &= \sup_{V \in \mathcal{U}_\infty} |\mu_U(T^{-N}V) - \mu(V)|\end{aligned}$$

with $\mathcal{U} = \{U, U^c\}$, $\mathcal{U}_n = \bigvee_{k=0}^{n-1} T^{-k}\mathcal{U}$ and $\mathcal{U}_\infty = \cup_n \sigma(\mathcal{U}_n)$.

Proof. Let $N \in \mathbb{N}$. If $k < N$, we just bound $c(k, U)$ by

$$\begin{aligned}|\mu_U(\tau_U > k) - \mu(\tau_U > k)| &= |\mu_U(\tau_U \le k) - \mu(\tau_U \le k)| \\ &\le |\mu_U(\tau_U \le k)| + |\mu(\tau_U \le k)| \\ &\le a_N(U) + k\mu(U) \le a_N(U) + N\mu(U).\end{aligned}$$

Otherwise, let us remark that $\{\tau_U > k\}$ and $\{\tau_U \circ T^N > k - N\}$ differ only on $\{\tau_U \le N\}$, and by hypothesis

$$|\mu_U(\tau_U > k) - \mu_U(\tau_U \circ T^N > k - N)| \le \mu_U(\tau_U \le N) = a_N(U).$$

Moreover

$$\begin{aligned}&|\mu_U(\tau_U \circ T^N > k - N) - \mu(\tau_U > k - N)| = \\ &|\mu_U(T^{-N}(\tau_U > k - N)) - \mu(\tau_U > k - N)| \le b_N(U).\end{aligned}$$

But $\{\tau_U > k - N\}$ and $\{\tau_U > k\}$ differs only on $\{\tau_U \circ T^{k-N} \le N\}$, hence

$$|\mu(\tau_U > k - N) - \mu(\tau_U > k)| \le \mu(\tau_U \circ T^{k-N} \le N) = \mu(\tau_U \le N) \le N\mu(U).$$

We finally get for each $k, N \in \mathbb{N}$,

$$|\mu_U(\tau_U > k) - \mu(\tau_U > k)| \le a_N(U) + b_N(U) + N\mu(U),$$

which concludes the proof, since N is arbitrary. □

We remark that $b_N(U)$ is bounded by $\alpha(N)$ if the partition $\mathcal{U} = \{U, U^c\}$ is α-mixing, and by $\gamma(N)$ if it is uniformly mixing (see Definition 2.1 below). To simplify, we could say that the exponential law holds when there exists some N so small that only few points of U come back in U before N steps, but large enough such that $T^N U$ is uniformly spread out.

Definition 2.1 (Speed of mixing). *Let $(X, \mathcal{B}, T, \mu)$ be a dynamical system and ξ a finite or countable measurable partition of X. We set $\xi_k = \bigvee_{j=0}^{k-1} T^{-j}\xi$ and $\sigma(\xi_k)$ the σ-algebra generated by ξ_k.*

1. **Uniform mixing.** *The partition ξ is uniformly mixing with speed $\gamma(n)$ going to zero for n going to infinity if for any n,*

$$\gamma(n) = \sup_{k,l} \sup_{\substack{R \in \sigma(\xi_k) \\ S \in T^{-(n+k)}\sigma(\xi_l)}} |\mu(R \cap S) - \mu(R)\mu(S)|.$$

2. **α-mixing.** *The partition ξ is α-mixing with speed $\alpha(n)$ going to zero for n going to infinity if for any n,*

$$\alpha(n) = \sup_{k,l} \sup_{\substack{R \in \xi_k \\ S \in T^{-(n+k)}\sigma(\xi_l)}} \left| \frac{\mu(R \cap S)}{\mu(R)} - \mu(S) \right|.$$

3. φ**-mixing.** *The partition ξ is φ-mixing with speed $\varphi(n)$ going to zero for n going to infinity if for any n,*

$$\varphi(n) = \sup_{k,l} \sup_{\substack{R\in\sigma(\xi_k)\\ S\in T^{-(n+k)}\xi_l}} \left| \frac{\mu(R\cap S)}{\mu(R)\mu(S)} - 1 \right|.$$

4. **Weak-Bernoulli.** *The partition ξ is weak-Bernoulli with speed $\beta(n)$ going to zero when n goes to infinity, if for any n,*

$$\beta(n) = \sup_{k,l} \sum_{\substack{R\in\xi_k\\ S\in T^{-(n+k)}\xi_l}} |\mu(R\cap S) - \mu(R)\mu(S)|.$$

Remark 2.5. We state some general implications and results verified by the preceding types of mixing.

1. φ-mixing implies α-mixing which implies uniform mixing. For any n, $\gamma(n) \le \alpha(n) \le \varphi(n)$.
2. φ-mixing implies weak-Bernoulli which implies uniform mixing. For any n, $\gamma(n) \le \beta(n) \le \varphi(n)$.
3. If ξ is a generating partition of an uniformly mixing dynamical system, then the system is mixing.
4. If ξ is a generating weak-Bernoulli partition then the system is metrically conjugated with a Bernoulli shift.

2.2. Successive return times. We will now investigate the properties of successive return times to the set U. For this purpose, let us define the k^{th} return time in U by

$$\tau_U^{(k)}(x) = \begin{cases} 0 & \text{if } k = 0, \\ \tau_U(x) + \tau_U^{(k-1)}(T^{\tau_U(x)}(x)) & \text{if } k > 1. \end{cases}$$

Observe that the difference between two consecutive return times follows the same law than the first, for the simple reason that

$$\tau_U^{(K+1)} - \tau_U^{(K)} = \tau_U \circ T^{\tau_U^{(K)}}$$

and the measure μ_U is invariant with respect to the induced application on U.

Theorem 2.6. *Let $U \subset X$ be a measurable set, and $\mathcal{U} = \{U, U^c\}$ the partition associated to it. Given an integer K and a rectangle Q_K in $\mathbb{R}^K$, the differences between successives normalized return times in U are independent and exponentially distributed up to $f(K,U)$ (see (5) below), where $f(K,U)$ is defined depending on the type of mixing by*

(α) *When (X, T, μ) is α-mixing for $\mathcal{U}$, with speed α* [1]*, then*

$$f(K,U) = K\left(3d(U) + \inf_{M\in\mathbb{N}} \{\alpha(M) + 3M\mu(U)\}\right).$$

[1] We just need that mixing property for some special sets, more precisely, we are interested by

$$\alpha'(N) = \sup\left\{ \left|\frac{\mu(R\cap S)}{\mu(R)} - \mu(S)\right| \;\middle|\; j, N\in\mathbb{N}, R\in\mathcal{U}_j, T^jR\subset U, V\in T^{-j-N}\mathcal{U}_\infty \right\}.$$

(γ) When the partition $\mathcal{U}$ is uniformly mixed by (X, T, μ) with speed γ, then

$$f(K, U) = K\left[4d(U) + \inf_{\substack{M\in\mathbb{N}\\ \gamma(M)<\mu(U)^2}} \left\{\frac{\gamma(M)}{\mu(U)^2}\left(2 - K\log\frac{\gamma(M)}{\mu(U)^2}\right) + 3M\mu(U)\right\}\right].$$

Indeed the following inequality holds:

$$\left|\mu_U\left((\tau_U^{(1)}, \tau_U^{(2)} - \tau_U^{(1)}, \dots, \tau_U^{(K)} - \tau_U^{(K-1)}) \in \frac{1}{\mu(U)} Q_K\right) - \int_{Q_K} \prod_{i=1}^{K} \mathrm{e}^{-s_i}\, ds^K\right| \leq f(K, U). \tag{5}$$

Remark 2.7. Note that the mixing assumption is made only for the special partition $\mathcal{U}$. If the system has a partition $\mathcal{Z}$ (not necessarily with two elements), uniformly mixing with speed $\gamma_{\mathcal{Z}}$, then for any cylinder $U \in \mathcal{Z}_n$ of order n, the partition $\mathcal{U} = \{U, U^c\}$ is still uniformly mixing with speed $\gamma_{\mathcal{U}}(M) \leq \gamma_{\mathcal{Z}}(M - n)$. The proof of the theorem is inspired by [CG93], with the following differences: 1) U is any measurable set; 2) we take care of the approximations to get an estimation of the error; 3) we still get an estimation even if the system is uniformly mixing; however, it is interesting whenever $\gamma(M) = o(1/M^2)$.

Proof of Theorem 2.6. Let us remark first that if we denote by $F = T^{\tau_U}$ the induced application on U, then for each $k \in \mathbb{N}$,

$$\tau_U^{(k+1)} - \tau_U^{(k)} = \tau_U \circ F^k.$$

We set $\tau_k = (\tau_U, \tau_U \circ F, \cdots, \tau_U \circ F^{k-1})$. We will show that the inequality (5) holds by induction on K.

For $K = 1$, we apply Theorem 2.1 which gives, setting $Q_1 = [u, v]$,

$$\begin{aligned}&|\mu_U(\tau_U \in [u, v]) - \int_u^v \mathrm{e}^{-s} ds| \\ &= |\mu_U(\tau_U > v) - \mu_U(\tau_U > u) - (\mathrm{e}^{-u} - \mathrm{e}^{-v})| \leq 2d(U).\end{aligned}$$

Let's suppose that the inequality (5) is true for K; we want to prove that it is also true for $K + 1$. Let $[r, s]$ be the projection of Q_{K+1} onto the last coordinate, and for $k = K, K + 1$ denote:

$$D_k = U \cap \tau_k^{-1}\left(\frac{1}{\mu(U)} Q_k\right).$$

For any $M \in \mathbb{N}$, the set defined by

$$E_{K+1}(M) = D_K \cap \left\{x \in U \mid \tau_U \circ T^M \circ F^K(x) \in [r, s]/\mu(U) - M\right\}$$

verifies the inclusions

$$E_{K+1}(M) \cap \{\tau_U \circ F^K > M\} \subset D_{K+1} \subset E_{K+1}(M) \cup \{\tau_U \circ F^K \leq M\}.$$

Theorem 2.1 shows that the two sets which bound D_{K+1} do not differ too much, namely,

$$\mu_U(\tau_U \circ F^K \le M) = \mu_U(\tau_U \le M) \le 1 - \mathrm{e}^{-M\mu(U)} + d(U) \le M\mu(U) + d(U).$$

Therefore we get the first bound

$$|\mu_U(D_{K+1}) - \mu_U(E_{K+1}(M))| \le M\mu(U) + d(U). \tag{6}$$

So the problem reduces to prove that $\mu_U(E_{K+1}(M))$ follows the expected law. We decompose the sets $E_{K+1}(M)$ over $A_K^j = U \cap \{\tau_U^{(K)} = j\}$. We have

$$E_{K+1}(M) \cap A_K^j = D_K \cap A_K^j \cap T^{-(M+j)}\{\tau_U \in \frac{[r,s]}{\mu(U)} - M\}.$$

We can now use the mixing with $R = D_K \cap A_K^j \in \sigma(\mathcal{U}_j)$ and $S = T^{-(M+j)}\{\tau_U \in [r,s]/\mu(U) - M\}$. According to the type of mixing, we get two approximations:

(α) When the partition $\mathcal{U}$ is α-mixing:

$$\begin{aligned}&|\mu_U(E_{K+1}(M) \cap A_K^j) - \mu_U(D_K \cap A_K^j)\mu(\tau_U \in \frac{[r,s]}{\mu(U)} - M)|\\ &\le \alpha(M)\mu_U(D_K \cap A_K^j).\end{aligned}$$

Summing over the possible values of j we get:

$$|\mu_U(E_{K+1}(M)) - \mu_U(D_K)\mu(\tau_U \in \frac{[r,s]}{\mu(U)} - M)| \le \alpha(M)\mu_U(D_K) \le \alpha(M). \tag{7}$$

Now Theorem 2.1 gives

$$\begin{aligned}|\mu(\tau_U \in \frac{[r,s]}{\mu(U)} - M) - (\mathrm{e}^{-r} - \mathrm{e}^{-s})| &\le |\mu(\tau_U \in \frac{[r,s]}{\mu(U)}) - (\mathrm{e}^{-r} - \mathrm{e}^{-s})| + 2M\mu(U)\\ &\le 2(M\mu(U) + d(U)).\end{aligned}$$

We briefly recall the approximations done with their respective errors

$$\begin{array}{ccccccc}\mu_U(D_{K+1}) & \rightarrow & \mu_U(E_{K+1}(M)) & \rightarrow & \mu_U(D_K)\mu\{\tau_U \in \frac{[r,s]}{\mu(U)}\} & \rightarrow & \mu_U(D_K)(\mathrm{e}^{-r} - \mathrm{e}^{-s})\\ & \downarrow & & \downarrow & & \downarrow & \\ & M\mu(U) + d(U) & & \alpha(M) & & 2(M\mu(U) + d(U)) & \end{array}$$

This allows us to show that the difference

$$\left|\mu_U(D_{K+1}) - \int_{Q_{K+1}} \prod_{i=1}^{K+1} \mathrm{e}^{-s_i} ds^{K+1}\right| \tag{8}$$

is bounded by the quantity $f(K,U) + 3M\mu(U) + \alpha(M) + 3d(U) \le f(K+1,U)$, which proves the induction and concludes the proof of this first case.

(γ) We now consider the case when $\mathcal{U}$ is uniformly mixing:

Let M be such that $\gamma(M) < \mu(U)^2$. As a first step, we can restrict ourselves to the case when $Q_K \subset [0,z]^K$, with $z = -\log \dfrac{\gamma(M)}{\mu(U)^2} > 0$. In fact,

$$Q_K \setminus [0,z]^K \subset \bigcup_{k=1}^{K} \mathbb{R}_+^{k-1} \times]z,\infty] \times \mathbb{R}_+^{K-k}$$

which implies using Theorem 2.1

$$\begin{aligned}\mu_U(\mu(U)\tau_K \in Q_K \setminus [0,z]^K) &\le \sum_{k=1}^{K} \mu_U(\tau_U^{(k+1)} - \tau_U^{(k)} > z/\mu(U)) \\ &= K\mu_U(\tau_U > z/\mu(U)) \\ &\le K(\mathrm{e}^{-z} + d(U)).\end{aligned}$$

Moreover

$$\int_{Q_K \setminus [0,z]^K} \prod_{i=1}^{K} \mathrm{e}^{-s_i} ds^K \le \sum_{k=1}^{K} \int_{\mathbb{R}_+^{k-1} \times]z,\infty] \times \mathbb{R}_+^{K-k}} \prod_{i=1}^{K} \mathrm{e}^{-s_i} ds^K \le K\mathrm{e}^{-z}.$$

Next, by decomposing according to

$$\begin{aligned}\mu_U(\mu(U)\tau_K \in Q_K) = \mu_U(\mu(U)\tau_K \in Q_K \cap [0,z]^K) \\ + \mu(\mu(U)\tau_K \in Q_K \setminus [0,z]^K),\end{aligned}$$

we get $f(K,U) \le K(2\mathrm{e}^{-z} + d(U)) + f'(K,U)$, where $f'(K,U)$ is the maximum of the difference (5) for the boxes $Q_K \subset [0,z]^K$. We then estimate $f'(K,U)$. First by uniform mixing we get

$$|\mu_U(E_{K+1}(M) \cap A_K^j) - \mu_U(D_K \cap A_K^j)\mu(\tau_U \in [r,s]/\mu(U) - M)| \le \frac{\gamma(M)}{\mu(U)}$$

and then we sum over all possible[2] values j of $\tau^{(K)}$,

$$|\mu_U(E_{K+1}(M)) - \mu_U(D_K)\mu(\tau_U \in [r,s]/\mu(U) - M)| \le \frac{Kz\gamma(M)}{\mu(U)^2}.$$

The same computation performed after estimation (7) (where now $\alpha(M)$ is replaced by $Kz\gamma(M)/\mu(U)^2$ in inequality (7)), gives the bound $f'(K+1,U) \le K\dfrac{z\gamma(M)}{\mu(U)^2} + 3(d(U) + M\mu(U))$. Then for each M,

$$f'(K,U) \le K^2 \frac{z\gamma(M)}{\mu(U)^2} + 3K(d(U) + M\mu(U)).$$

Since M is arbitrary, our choice of z implies that the inequality (5) is verified with

$$f(K,U) = K\left[4d(U) + \inf_{\substack{M\in\mathbb{N}\\ \gamma(M)<\mu(U)^2}} \left\{\frac{\gamma(M)}{\mu(U)^2}\left(2 - K\log\frac{\gamma(M)}{\mu(U)^2}\right) + 3M\mu(U)\right\}\right]. \quad \square$$

[2] Since $Q_K \subset [0,z]^K$, the K^{th} return time is less or equal to Kz, hence it takes at most $[Kz]$ different values.

We are now ready to give the most important result of this section, namely, to prove the Poisson statistics for successive return times. Let $N(t)$ be the number of visits into U up to the normalized time $t/\mu(U)$,

$$N(t) = \sup\left\{ K > 0 |\ \tau_U^{(K)} \leq t/\mu(U)\right\}.$$

It turns out that $N(t)$ is a discrete random variable whose law is close to a Poissonian one, more precisely we have

Theorem 2.8. *The distribution of the number of visits $N(t)$ differs from the Poissonian law by*

$$\left|\mu_U(N(t) = K) - \frac{t^K}{K!}\mathrm{e}^{-t}\right| \leq g(t, K, U) + g(t, K+1, U),$$

where for each $k \geq 0$ $g(t, k, U) = \left(12t^k/k + k^{k-1}\right)\sqrt[k]{f(k, U)}$.

Proof. It is a consequence of the weak dependence of the differences of successives return times established by Theorem 2.6. We first remark that

$$\begin{aligned}\mu_U(N(t) = K) &= \mu_U\left(\{\tau_U^{(K)} \leq \frac{t}{\mu(U)}\} \cap \{\tau_U^{(K+1)} > \frac{t}{\mu(U)}\}\right)\\ &= \mu_U\left(\tau^{(K)} \leq t/\mu(U)\right) - \mu_U\left(\tau^{(K+1)} \leq t/\mu(U)\right).\end{aligned}$$

It is then sufficient to compute the measure of points whose k^{th} rescaled return time is smaller than t, for $k = K, K+1$. If we put $\widetilde{P}_k(t)$ the distribution of the sum of the differences of successive return times, we know that when the latter are i.i.d. random variables with the same exponential law, then setting

$$L_k(t) = \left\{(s_1, \ldots, s_k) \in \mathbb{R}_+^k \,\middle|\, s_1 + \cdots + s_k \leq t\right\}$$

we get

$$\widetilde{P}_k(t) = P_k(t) := \int_{L_k(t)} \prod_{i=1}^{k} \mathrm{e}^{-s_i} ds_i$$

which gives the classical result $P_K(t) - P_{K+1}(t) = \dfrac{t^K}{K!}\mathrm{e}^{-t}$.

The difficulty comes now from the fact that we have to translate Theorem 2.6 given for boxes on the simplex $L_k(t)$.

Let's suppose that $f(k, U) < 1$, otherwise there is nothing to prove. Hence the integer defined by $N = [k/f(k, U)^{k+1}]$ is bigger than k. We consider the uniform partition of $[0, t]^k$ by cubes of size t/N. Let Δ_k be the union of those cubes Q_k included in the interior of $L_k(t)$, for which for any $(s_1, \ldots, s_k) \in Q_k$, $\sum_{i=1}^k s_i < t$ and Σ_k those which intersect the boundary, i.e. the union of those cubes such that there exists

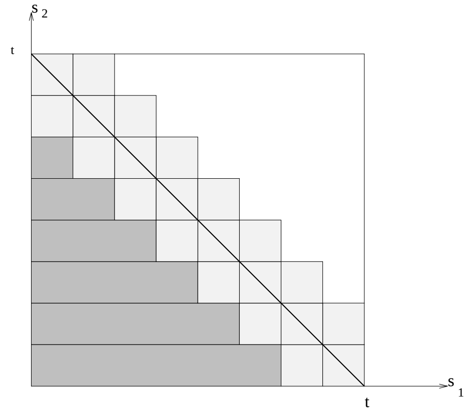

Fig. 1. Partition of the cube $[0, t]^k$ for $k = 2$. Σ_k is the union of dotted squares and Δ_k the union of shaded rectangles $R_k(Q_k)$.

$(s_1, \dots, s_k) \in Q_k$ with $\sum_{i=1}^k s_i = t$. By using the notation τ_k introduced in the proof of Theorem 2.6 we have,

$$\delta := \left| \mu_U(\tau_U^{(k)} \le t/\mu(U)) - \int_{L_k(t)} \prod_{i=1}^k \mathrm{e}^{-s_i} ds^k \right|$$
$$\le \left| \mu_U(\tau_k \in \frac{\Delta_k}{\mu(U)}) - \int_{\Delta_k} \prod_{i=1}^k \mathrm{e}^{-s_i} ds^k \right| + \mu_U(\tau_k \in \frac{\Sigma_k}{\mu(U)}) + \int_{\Sigma_k} \prod_{i=1}^k \mathrm{e}^{-s_i} ds^k$$
$$\le \delta_1 + \delta_2 + \delta_3.$$

To estimate δ_1, we put Π for the projection over the $k-1$ last coordinates; then the sets $R_k(Q_k) = \{Q'_k \in \Delta_k | \Pi(Q'_k) = \Pi(Q_k)\}$ are boxes, and their number is bounded by N^{k-1} (see Fig. 1). For each of these boxes Theorem 2.6 gives an error smaller than $f(k, U)$, and then we get $\delta_1 \le N^{k-1} f(k, U)$.

To compute δ_2 and δ_3, we first remark that a straightforward combinatorial calculus gives, for the number C_N^k of cubes inside Σ_k, $C_N^k \le 6N^{k-1}$ (see [Sau98b]). But for each cube $Q_k \subset \Sigma_k$ Theorem 2.6 gives

$$\mu_U(\tau_k \in Q_k) \le \int_{Q_k} \prod_{i=1}^k \mathrm{e}^{-s_i} ds^k + f(k, U).$$

Summing over all the cubes contained in Σ_k one has $\delta_2 \le 6N^{k-1} f(k, U) + \delta_3$. Moreover the integral $\int_{Q_k} \prod_{i=1}^k \mathrm{e}^{-s_i} ds^k$ is bounded by the volume of Q_k equal to $(t/N)^k$, which gives $\delta_3 \le 6N^{k-1} t^k / N^k$. We then deduce that

$$\delta \le \delta_1 + \delta_2 + \delta_3 \le N^{k-1} f(k, U) + 12 t^k / N$$

which implies $\delta \le \left(12 t^k / k + k^k\right) f(k, U)$ by the previous choice of N. □

3. Applications

In the preceding chapter we gave general estimates for the error between the distribution of the number of visits into a set U and the Poissonian law. We could wonder whether this law is attained in the limit of $\mu(U) \to 0$. Put in this way the question is not very clear. What we need is instead to localize a sequence of neighborhoods $U_\varepsilon(z)$ shrinking to zero and ask whether the Poisson law holds in the limit $\varepsilon \to 0$. This approach was successfully carried out by several authors as reminded in the introduction. Although their results were applied to dynamical systems, the inspiration and some of the techniques of the proofs were of probabilistic nature (theory of moments, Laplace transform). Here we follow a purely dynamical direction, trying to extract all the statistical information by the ergodic properties of the system. In this way we are able, for example, to exhibit the Poissonian statistics for a large class of non uniformly hyperbolic maps of the interval, widely studied in the last years especially to determine the rate of decay of correlations and the central limit theorem.

Some statistical properties of these maps have been studied in the paper [LSV97] (this paper contains a quite complete bibliography on the subject), where an absolutely continuous invariant probability measure (ACIM) is first constructed, and then it is shown that it enjoys a polynomial decay of correlations.

One feature of these maps is that they are characterized by a structure parameter (the order of tangency at an indifferent fixed point), which governs the statistical properties, and that can be viewed as an indicator of the "weak" hyperbolicity of the map. Actually, it turns out that this parameter appears even in the approximation to the Poissonian law.

Let's then consider for $0 < \alpha < 1$ the following map of the unit interval:

$$T(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \forall x \in [0,\ 1/2) \\ 2x - 1 & \forall x \in [1/2,\ 1] \end{cases}.$$

We recall some properties and results which we will need in the following, and we refer the reader to the quoted paper for more informations and proofs. This application has a finite Markov partition (with two elements), but for our purposes it is more convenient to work with the countable one ξ generated by the left preimages a_n of 1, $\xi = \{A_m | m \in \mathbb{N}\}$ with $A_n =]a_{n+1}, a_n]$. We will often use in the following the easy bound $\frac{a_n}{a_{n+1}} \le 2$.

We can associate to each point $z \in X =]0, 1]$ an unique infinite sequence $\underline{\omega} = \omega_1\omega_2...$ with the property that $T^{m-1}z \in A_{\omega_m}$ for all integer $m \ge 1$. We denote by ξ_m the dynamical partition $\xi \vee T^{-1}\xi \cdots T^{-m+1}\xi$ and call its elements m-cylinders. We denote with $\xi_m(z) \in \xi_m$ the m-cylinder which contains z. The sequence $\underline{\omega}$ satisfies the admissibility condition: $\omega_m\omega_{m+1}$ appears in $\underline{\omega}$ if and only if $\omega_m = 0$ or $\omega_{m+1} = \omega_m - 1$. We say that a non empty cylinder $C = [\omega_1 \dots \omega_k] \in \xi_k$ is *maximal* if it maps onto X after exactly k iterations, which is easily seen to be equivalent to $\omega_k = 0$.

3.1. Some mixing properties. We begin with a brief survey of some results proved by two of us (B.S., S.V) in the joint paper [LSV97] with Carlangelo Liverani. We showed that the density h of the ACIM belongs to a certain cone of functions $C_*(a)$, which will be characterized later (see Lemma 3.2), provided a is big enough, and satisfies[3]:

[3] We recall the formal definition of the Perron Frobenius operator P acting on function $f : [0, 1] \to \mathbb{R}$: $Pf(x) = \sum_{Ty=x} \frac{1}{D_yT} f(y)$. One easily check that μ is an ACIM iff $h = \frac{d\mu}{dx}$ is a fixed point of P on $L^1(dx)$.

Lemma A (Lemma 2.2 in [LSV97]**).** *The cone $\mathcal{C}_*(a)$ is left invariant by the Perron-Frobenius operator P, i.e. $P(\mathcal{C}_*(a)) \subset \mathcal{C}_*(a)$.*

Lemma B (Lemma 2.3 in [LSV97]**).** *The density h belongs to the cone $\mathcal{C}_*(a)$, and verifies in particular whenever $x \le y$,*

$$\frac{h(x)}{h(y)} \le (y/x)^{\alpha+1}, \tag{9}$$

$$h(x) \le a x^{-\alpha}. \tag{10}$$

Proposition C (Distortion inequality, proof of Proposition 3.3 in [LSV97]**).** *There exists some constant Δ such that for all k and $x, y \in C \in \xi_k$,*

$$\frac{D_x T^k}{D_y T^k} \le \Delta < \infty. \tag{11}$$

We will suppose without loss of generality that $a \ge 4\Delta$.

Theorem D (Theorem 4.1 in [LSV97]**).** *In the proof of this theorem we in particular got that for $f \in \mathcal{C}_*(a)$,*

$$\left\| P^n\Big(f - \lambda(f)\Big)\right\|_{L^1(\lambda)} \le \Phi(n)\|f\|_{L^1(\lambda)} \tag{12}$$

with $\Phi(n) = Cn^{-\frac{1}{\alpha}+1}(\log n)^{\frac{1}{\alpha}} = \mathcal{O}_L(n^{-\frac{1}{\alpha}+1})$, where we define by

$$\mathcal{O}_L(\varepsilon) = \mathcal{O}(\varepsilon(\log \varepsilon^{-1})^r)$$

in the limit $\varepsilon \to 0$, for any constant r.

We then need a few more results on the speed of mixing which turn out to be useful for the statistics of return times and also to establish the weak-bernoullicity of the map.

Lemma 3.1. *For any $z \in X$, and for any m such that $\xi_m(z)$ is maximal, the partition $\mathcal{U} = \{\xi_m(z), \xi_m(z)^c\}$ satisfies a property close to the α-mixing, namely*

$$\alpha'(N) = \sup_{j\in\mathbb{N}} \sup_{\substack{R\in\mathcal{U}_j \\ T^jR\subset U}} \sup_{S\in\mathcal{U}_\infty} \left| \frac{\mu(R\cap T^{-N-j}S)}{\mu(R)} - \mu(S)\right| = \mathcal{O}_L((N-m)^{1-\frac{1}{\alpha}}).$$

Proof. Let z be a point of X and m be an integer such that $\xi_m(z)$ is maximal. Let $\mathcal{U}$ be the partition given by $\xi_m(z)$ and its complement, and $\mathcal{U}_j$ the refinement of $\mathcal{U}$. For $R \in \mathcal{U}_j$ such that $T^jR \subset U$, we have $R \in \sigma(\xi_{m+j})$ and R is a union of maximal cylinders $V^k_{m+j} \in \xi_{m+j}$; choose $V \in \xi_{m+j}$ one of these maximal cylinders. For any $S \in T^{-(N+j)}\mathcal{B}$ there exists a set $W \in \mathcal{B}$ such that $R = T^{-(N+j)}W$. We then have

$$\begin{aligned}
(*) &:= \mu(V\cap S) - \mu(V)\mu(S) \\
&= \int \mathbf{1}_V \mathbf{1}_W \circ T^{N+j} h d\lambda - \int \mu(V) h \mathbf{1}_W d\lambda \\
&= \int P^{N+j}[h(\mathbf{1}_V - \mu(V))]\mathbf{1}_W d\lambda \\
&\le \|P^{N+j}[h(\mathbf{1}_V - \mu(V))]\|_{L^1(\lambda)}.
\end{aligned}$$

By exploiting the fact that V is maximal we continue the preceding bound as

$$\begin{aligned}(*) &\leq \left\| P^{N-m}[P^{j+m}(h\mathbf{1}_V) - \mu(V)] \right\|_{L^1(\lambda)} + \left\| P^{N-m}[\mu(V)h - \mu(V)] \right\|_{L^1(\lambda)} \\ &\leq 4a\Phi(N-m)\mu(V),\end{aligned}$$

with Φ given by inequality (12), provided $P^{m+j}(h\mathbf{1}_V) \in \mathcal{C}_*(a)$, which is the case by Lemma 3.2 below. We conclude the proof by summing over all the maximal cylinders of R. $\square$

Lemma 3.2. *For any maximal cylinder* $V \in \xi_p$,

$$P^p(h\mathbf{1}_V) \in \mathcal{C}_*(a).$$

Proof. We first set $f := P^p(h\mathbf{1}_V)$ and $T_V^p : V \to X$ the restriction of T^p to V. Since T^p is injective over V we can rewrite f as

$$f(x) = h \circ T_V^{-p}(x) D_x T_V^{-p}$$

which in particular shows that f is continuous. To prove that f belongs to the cone of smooth functions $\mathcal{C}_*(a)$ we must verify the following four properties which just define the cone:

1. f is continuous and positive, that is clear in our case.
2. f is decreasing. Since $h \in \mathcal{C}_*(a)$, h decreases. In addition, T_V^{-p} is decreasing and concave, therefore $h \circ T_V^{-p}$ and DT_V^{-p} decrease.
3. $x \mapsto x^{\alpha+1} f(x)$ increases. Since $T_V^{-p} : X \to V$ is increasing, an equivalent statement is that
$$(T^p u)^{\alpha+1} h(u) \frac{1}{D_u T^p}$$
is increasing with $u \in V$. Observing that
$$\left(\frac{T^p u}{u}\right)^{\alpha+1} \frac{1}{D_u T^p}$$
increases over $V \in \xi_p$ (which is true for $p = 1$ and the general case is proved by recurrence), and $u \mapsto u^{\alpha+1} h(u)$ increases, we obtain the result.
4. $f(x) \leq ax^{-\alpha} \int f$. Since f is continuous, there exists $v \in V$ such that
$$\int f = f(T^p v) = h(v) \frac{1}{D_v T^p}.$$
The distortion estimate (11) for $u \in V \in \xi_p$ gives
$$\frac{D_v T^p}{D_u T^p} \leq \Delta.$$
Moreover since h decreases, inequality (9) yields
$$\frac{h(u)}{h(v)} \leq \frac{h(a_{\omega_1+1})}{h(a_{\omega_1})} \leq \left(\frac{a_{\omega_1}}{a_{\omega_1+1}}\right)^{\alpha+1} \leq 4.$$

As a consequence, we get for $u = T_V^{-p} x$,

$$f(x) = h(u)\frac{1}{D_u T^p} \leq 4h(v)\frac{\Delta}{D_v T^p} \leq a x^{-\alpha} \int f,$$

because $x \leq 1$ and $4\Delta \leq a$. □

We finally prove that the countable partition ξ, and therefore the two-elements one, is weakly Bernoulli.

Theorem 3.3. *The partition ξ is weakly Bernoulli for (X, T, μ) with speed $\beta(n) = \mathcal{O}_L(n^{1-1/\alpha})$.*

Proof. We begin to recall the following result by Hofbauer and Keller [HK82] which permits to bound $\beta(n)$ as

$$\beta(n) \leq \sup_{m\in\mathbb{N}} \sum_{R\in\xi_m} \| P^{n+m}((\mathbf{1}_R - \mu(R))h)\|_{L^1_\lambda}. \tag{13}$$

Then it will be enough to bound

$$\| P^{m+n}((\mathbf{1}_R - \mu(R))h)\|$$

with $R \in \xi_m$. Let $p_R \geq m$ be the integer for which $R \in \xi_{p_R}$ is maximal. We decompose the sum over all the cylinders $R \in \xi_m$ into two blocks. Let $M(m, n)$ be the set of maximal cylinders for $p_R < m + n/2$. When $R \in M(m, n)$, the same computation performed in Lemma 3.1 gives

$$\| P^{m+n}((\mathbf{1}_R - \mu(R))h)\|_{L^1_\lambda} \leq \mu(R)\mathcal{O}_L((m+n-p_R)^{1-1/\alpha}) = \mu(R)\mathcal{O}_L(n^{1-1/\alpha}).$$

Then the set of cylinders which do not belong to $M(m, n)$ is exactly $T^{-m+1}[0, a_{n/2}]$, whose measure is equal to

$$\mu(T^{-m+1}[0, a_{n/2}]) = \mu([0, a_{n/2}]) = \int_0^{a_{n/2}} h(x)dx = \mathcal{O}(n^{1-1/\alpha}).$$

This proves the theorem. □

3.2. Statistics of return times. We now come back to the study of return times and the first step will be the estimation of the quantities involved in the error term given by Lemma 2.4.

Lemma 3.4. *There exists a constant B such that for any k and $C \in \xi_k$ with $T^{-k}C \cap C \neq \emptyset$,*

$$\sup P^k \mathbf{1}_C \leq B k^{-1-1/\alpha}. \tag{14}$$

Proof. Let k_0 be such that $D_{a_{k_0}}T \leq 2$, and put $r = D_{a_{k_0}}T > 1$. Let $C = [\omega_1 ... \omega_k]$ be a k-cylinder such that $T^{-k}C \cap C \neq \emptyset$. This implies that $\omega_k\omega_1$ is admissible. We want to estimate $\sup P^k \mathbf{1}_C = 1/\inf_C DT^k$. If $\omega_j \leq k_0$ for all $j = 1..k$, then $DT^k \geq r^k$. Else, take j such that $\omega_j = \max_{1\leq i\leq k}\omega_i$. Either $j = 1$, and consequently $\omega_k = 0$ or $\omega_{j-1} = 0$. In the last case we have

$$\inf_C DT^k \geq \inf_{[\omega_1...\omega_{j-1}]} DT^{j-1} \inf_{[\omega_j...\omega_k]} DT^{k+1-j} \geq \Delta^{-1} \inf_{[\omega_j...\omega_k\omega_1...\omega_{j-1}]} DT^k.$$

By this argument we are led to consider the worst case which is given by a cylinder of type $C = [(k-1)(k-2)...0]$. For $T^kC = [0,1]$, the distortion formula (11) and the estimation $a_k \leq ck^{-1/\alpha}$ given by Lemma 3.2 in [LSV97] we get $D_{a_k}T^k = c'k^{1+1/\alpha}$ for some constant c', from which the lemma follows by taking $B \geq \Delta/c'$ such that $Bk^{1-1/\alpha} \geq r^k$ for all $k > 0$. □

We now introduce the first return time of a cylinder U which plays a crucial role in [Hir95]. We define it as $\tau(U) = \inf\{\tau_U(x) | x \in U\}$.

Lemma 3.5. *The quantity $a_N(U)$ defined in Lemma 2.4 for $U = \xi_m(z)$ is bounded by,*

$$a_N(U) = \frac{4\Delta}{\inf h}\ \frac{N\mu(U)}{\lambda(T^{\tau(U)}U)}.$$

Proof. We suppose $N > \tau(U)$ otherwise $a_N(U) = 0$. Set $\tau = \tau(U)$; for each z in X we have

$$\begin{aligned} a_N(U) &\leq \sum_{j=1}^N \frac{1}{\mu(U)}\mu(T^{-j}U \cap U) \\ &= \sum_{j=\tau}^N \frac{1}{\mu(U)} \int P^j(\mathbf{1}_U h)\mathbf{1}_U d\lambda \\ &\leq N \sup_{j=\tau..N} \sup_U \frac{P^j(\mathbf{1}_U h)}{h}. \end{aligned}$$

Now the distortion (11) and the regularity of the density (9) give

$$\begin{aligned} P^\tau(\mathbf{1}_U h) &= h \circ T_U^{-\tau} DT_U^{-\tau} \mathbf{1}_{T^\tau U} \\ &\leq 4\Delta \frac{1}{\lambda(T^\tau U)} \int_{T^\tau U} h \circ T_U^{-\tau} DT_U^{-\tau} \mathbf{1}_{T^\tau U} d\lambda \\ &\leq 4\Delta \frac{\mu(U)}{\lambda(T^\tau U)}. \end{aligned}$$

Finally, $Ph = h$ and since P is a positive operator one has

$$\frac{P^j(\mathbf{1}_U h)}{h} \leq \frac{P^{j-\tau}\mathbf{1}}{h} \sup P^\tau(\mathbf{1}_U h) \leq \frac{P^{j-\tau}\frac{h}{\inf h}}{h} \sup P^\tau(\mathbf{1}_U h) \leq \frac{4\Delta}{\inf h}\ \frac{\mu(U)}{\lambda(T^\tau U)}.$$

□

The next step will be to show that $\tau(U)$ is almost everywhere big enough to give a good upper bound in the previous lemma for $a_N(U)$.

We first define in full generality the *local rate of return for cylinders*. As a matter of fact, we would like to point out that the first return time of a set into itself allows to define and compute an interesting dimension-like characteristic which we called the Afraimovich-Pesin dimension in [PSV98].

Definition 3.1. *Let ζ a partition of X. Denote with $\zeta_n(x)$ the element of $\zeta \vee T^{-1}\zeta \vee \cdots \vee T^{-n+1}\zeta$ which contains $x \in X$. We then define the local (lower and upper) rate of return for cylinders as*

$$\overline{\underline{R}}_\zeta(x) = \overline{\underline{\lim}}_{n\to\infty} \frac{\tau(\zeta_n(x))}{n}.$$

Proposition 3.6. (i) *Both $\underline{R}_\zeta$ and $\overline{R}_\zeta$ are sub-invariant, namely $\underline{R}_\zeta \circ T \le \underline{R}_\zeta$ and $\overline{R}_\zeta \circ T \le \overline{R}_\zeta$.*

(ii) *Assume that ζ is a measurable partition of the measurable space X, and μ is an invariant probability, then $\underline{R}_\zeta$ and $\overline{R}_\zeta$ are μ-a.e. invariant.*

(iii) *Moreover, whenever μ is ergodic $\underline{R}_\zeta$ and $\overline{R}_\zeta$ are μ-a.e. constant*

Proof. (i) Let $x \in X$. For each integer $n > 0$, we have:

$$\zeta_n(x) \cap T^k\zeta_n(x) \neq \emptyset \Longrightarrow \zeta_{n-1}(Tx) \cap T^k\zeta_{n-1}(Tx) \neq \emptyset,$$

which implies that $\tau(\zeta_{n-1}(Tx)) \le \tau(\zeta_n(x))$.
(ii) is a standard property of sub-invariant functions on finite measure spaces and then (iii) follows immediately. □

We state the following result which can be improved for some subshifts[4].

Proposition 3.7. *For μ-almost every $z \in X$, the lower rate of return for cylinders is equal to* 1.

$$\underline{R}_\xi(z) = 1.$$

Proof. Let $1/2 < \delta < 1$. Consider the set (we denote $N_m(z) = \tau(\xi_m(z))$),

$$L_m := \{z \in A_0 |\, N_m(z) \le \delta m\}.$$

If

$$\sum_{m=1}^{\infty} \mu(L_m) < \infty, \tag{15}$$

then the Borel-Cantelli Lemma ensures that for almost every $z \in A_0$, we have $N_m > \delta m$, up to finitely many m. By sending δ to 1 we show that $\underline{R}_\xi(z) \ge 1$ almost everywhere on A_0. Then for the preceding proposition (iii) and the ergodicity of the measure μ, we

[4] We have in fact the following:
Theorem. *Suppose that μ is a Gibbs state for the Hölder potential φ on some irreducible and aperiodic subshift of finite type with finite alphabet ζ, then μ-almost everywhere, $\overline{R}_\zeta = \underline{R}_\zeta = 1$.*

Proof. An easiest version of the Proposition 3.7 gives the lower bound, while the uniform upper bound $\tau(C_n) \le n + n_0$ holds, where C_n is a cylinder of order n, and n_0 is the lowest power for which the transition matrix becomes strictly positive.

get the same bound almost everywhere. The equality finally follows since each time that $T^{m-1}z \in A_0$, we have $T^m \xi_m(z) = X$ hence $N_m(z) \le m$.

In order to prove (15) it is sufficient to consider the Lebesgue measure instead of μ (since the density h is bounded from below). We have

$$\lambda(L_m) = \underbrace{\sum_{k=1}^{[m/2]} \lambda(N_m = k)}_{(1)} + \underbrace{\sum_{k=[m/2]+1}^{\delta m} \lambda(N_m = k)}_{(2)}.$$

We now perform a detailed analysis of the sets appearing in the preceding formula.
(1): In this case, the cylinder $\xi_m(z)$ with $N_m = k$ must be of the form

$$\xi_m(z) = [(\omega_1 .. \underbrace{.\omega_k)(\omega_1 ... \omega_k) ... (\omega_1 .}_{[m/k]} .. \omega_k) ...].$$

Therefore when $k \le [m/2]$, the cylinder is completely determined by its first k symbols. Put $C = [\omega_1 ... \omega_k]$; we say that a cylinder of length k is admissible (admis) when it is the beginning of a cylinder of L_m with $N_m = k$. Then we can bound (1) by

$$\begin{aligned}(1) &\le \sum_{k=1}^{[m/2]} \sum_{C \text{ admis}} \lambda(C \cap T^{-k}C \cap \cdots \cap T^{-[m/k-1]k}C) \\ &\le \sum_{k=1}^{[m/2]} \sum_{C \text{ admis}} \left(\sup_C P^k \mathbf{1}_C\right)^{[m/k]-1} \lambda(C) \\ &\le \sum_{k=1}^{[m/2]} \sup_{C \text{ admis}} \left(\sup_C P^k \mathbf{1}_C\right)^{[m/k]-1}.\end{aligned}$$

We first remark that T^k being injective over $C \in \xi_k$, we have

$$P^k \mathbf{1}_C \le 1/\inf_{A_0} DT^k \le 1/2.$$

We split the last sum in three pieces by fixing k_0 as the biggest integer for which $k_0^{1+\frac{1}{\alpha}} \ge e^B$, where B is the constant in Lemma 3.4. We then have by using Lemma 3.4,

$$(1) \le \sum_{k=1}^{k_0} (1/2)^{[m/k]-1} + \sum_{k=k_0}^{m/3} (Bk^{-1-1/\alpha})^{m/k-2} + \sum_{m/3}^{[m/2]} Bk^{-1-1/\alpha}.$$

The first and the last sum are easily shown to be summable with respect to m. For the second term, we observe that the terms $(Bk^{-1-1/\alpha})^{m/k-2}$ are increasing in k when k is bigger than k_0. A direct estimation of the sum is $B3^{1/\alpha}m^{-1/\alpha}$ which is summable with respect to m.
(2): In this case, the cylinder $\xi_m(z)$ has the form

$$\xi_m(z) = [\underbrace{\omega_1 ... \omega_{m-k}}_{m-k} \underbrace{\omega_{m-k+1} ... \omega_k}_{2k-m} \underbrace{\omega_1 ... \omega_{m-k}}_{m-k}].$$

As before, we set $C = [\omega_1 ... \omega_{m-k}]$, and we say that C is admissible (admis) when it is the beginning of a cylinder of L_m with $N_m = k$,

$$(2) \leq \sum_{k=[m/2]+1}^{\delta m} \sum_{C \text{ admis}} \lambda(C \cap T^{-k} C)$$
$$\leq \sum_{k=[m/2]+1}^{\delta m} \sup_{C \text{ admis}} \sup_{C} P^k \boldsymbol{1}_C.$$

Let first $p = p(C) \geq m - k$ be such that $C \in \xi_p$ is maximal (i.e. $p(C)$ is the smallest p for which $C \in \xi_p$). When $p < k$, since $1 \in \mathcal{C}_*(a)$ the inequality (12) and Lemma 3.4 give

$$\sup_C P^k \boldsymbol{1}_C \leq \sup P^p \boldsymbol{1}_C \sup_C P^{k-p} 1 \leq a 2^\alpha B p^{-1-1/\alpha} \leq a 2^\alpha B (m-k)^{-1-1/\alpha}.$$

When $p \geq k$, $C \in \xi_k$ and $T^{-k} C \cap C \neq \emptyset$ we have

$$P^k \boldsymbol{1}_C \leq B k^{-1-1/\alpha}.$$

But $k \geq m - k \geq (1-\delta)m$, and then the sum (15) is summable for any $\delta < 1$. □

We are now ready to state and prove the main theorems of this section

Theorem 3.8. *For μ-almost every $z \in X$ and $\beta < \bar{\beta}(\alpha)$,*

$$\sup_{t \geq 0} \left| \mu_{\xi_m(z)} \left(\tau_{\xi_m(z)} > \frac{t}{\mu(\xi_m(z))} \right) - \exp(-t) \right| = \mathcal{O}(\mu(\xi_m(z))^\beta),$$

where the critical exponent $\bar{\beta}(\alpha) = 1 - \alpha$.

Proof. Let ε be a positive number. Let z be a typical point for Proposition 3.7 and for the Shannon–McMillan–Breiman theorem. We want to apply Lemma 2.4; Let $m(\varepsilon)$ such that for any $m > m(\varepsilon)$ we have $(1-\varepsilon)m \leq \tau(\xi_m(z))$, $\mu(\xi_m(z)) \leq \exp(-m 2 h_\mu / 3)$ and also $\mu(\xi_{\varepsilon m}(T^{[(1-\varepsilon)m]} z)) \geq \exp(-(2[\varepsilon m]) h_\mu)$.

For the sake of simplicity, we put for any m, $U_m = \xi_m(z)$. For any $m > m(\varepsilon)$ such that U_m is maximal, we have $(1-\varepsilon)m \leq \tau(U_m) \leq m$, and all the iterates $T^j U_m$ for $1 \leq j < m$ are at a distance bigger than a_m from the neutral fixed point (because U_m is maximal). If $\tau(U_m) < m$ then the density stays bounded on the orbit $T^j U_m$ by $b a_m^{-\alpha}$ so we have

$$\lambda(T^{\tau(U_m)} U_m) \geq \frac{a_m^\alpha}{b} \mu(T^{\tau(U_m)} U_m) \geq \frac{a_m^\alpha}{b} \exp(-2\varepsilon m h_\mu).$$

On the other hand, when $\tau(U_m) = m$ we still get

$$\lambda(T^{\tau(U_m)} U_m) = 1 \geq \frac{a_m^\alpha}{b} \exp(-2\varepsilon m h_\mu).$$

Lemma 3.5 gives us the following estimation with $N = \mu(U_m)^{-\alpha+\varepsilon}$,

$$a_N(U_m) = \mathcal{O}(\mu(U_m)^{1-\alpha-3\varepsilon}).$$

Lemma 3.1 with $R = U_m$ gives us

$$b_N(U_m) = \mathcal{O}_L((\mu(U_m)^{-\alpha+\varepsilon} - m)^{1-\frac{1}{\alpha}}) = \mathcal{O}_L(\mu(U_m)^{(-\alpha+\varepsilon)(1-\frac{1}{\alpha})}).$$

We can then apply Lemma 2.4, which gives

$$c(U_m) =\leq a_N(U_m) + b_N(U_m) = \mathcal{O}(\mu(U_m)^\beta)$$

for $\beta \leq 1 - \alpha - 3\varepsilon$ and $\beta \leq 1 - \alpha - 2\varepsilon(1/\alpha - 1)$. We finally end up with

$$d(U_m) = \mathcal{O}(\mu(U_m)^\beta) \tag{16}$$

for any $\beta < 1 - \alpha$, since ε is arbitrary small, which conclude the proof by applying Theorem 2.1. □

Remark 3.9. The preceding theorem shows that the critical exponent $\bar{\beta}(\alpha)$ is smaller than 1.We point out that, by using Proposition 2.3 the power $\bar{\beta}$ cannot exceed 1.

Theorem 3.10. *For μ-almost every $z \in X$, we have for any $t \geq 0$ and $K \geq 0$ and $\beta < \bar{\beta}(\alpha)$,*

$$\left| \mu_{\xi_m(z)} \left(N_{\xi_m(z)}(t) = K \right) - \frac{t^K}{K!} \exp(-t) \right| = \mathcal{O}(\mu(\xi_m(z))^{\beta/(K+1)}).$$

with the critical exponent $\bar{\beta}(\alpha) = 1 - \alpha$.

Proof. Let z be a typical point satisfying the preceding theorem and m such that $U_m = \xi_m(z)$ is maximal.

By invoking the footnote of Theorem 2.6, it will be sufficient to use the weakened α-mixing condition

$$\alpha'(M) = \mathcal{O}_L((M - m)^{\alpha - \frac{1}{\alpha}})$$

given by Lemma 3.1 to apply Theorem 2.6. Take $M = \mu(U_m)^{-\alpha}$; we thus find for $\beta < 1 - \alpha$, and by the estimation (16) and Theorem 2.6 an error of the order

$$f(K, U_m) = const[d(U_m) + \alpha'(M) + M\mu(U)] = \mathcal{O}(\mu(U_m)^\beta).$$

By applying Theorem 2.8, the error for the probability to have K successive visits is of the order $\mu(U_m)^{\beta/(K+1)}$ for all $\beta < 1 - \alpha$. □

4. Concluding Remarks

We conclude with few observations. First, the proofs for the exponential-one law and the Poisson law given in Sect. 3 for a class of non uniform hyperbolic maps, can be easily adapted, and they are even easier, to all the cases quoted in the introduction, namely: Axiom A diffeomorphisms, transitive Markov chains, expanding maps of the interval with a spectral gap and in general to all φ-mixing dynamical systems.

For such systems, an estimation for the error can also be done: following the arguments of Theorems 3.8 and 3.10, one can easily see that the critical exponent $\bar{\beta}$ is equal to 1. This supports our beliefs that: (i) the error terms of type $\mu(U)^\beta$ could be optimal and (ii) the non uniform hyperbolicity of the map reflects in the critical exponent: in that case, in fact, it should be strictly smaller than one.

Acknowledgements. We would like to thank Viviane Baladi for a careful reading of a preliminary version of this work and Bernard Schmitt for useful discussions. B.S. acknowledges the ESF for support during the workshop "Probabilistic methods in non-hyperbolic dynamics" in Warwick.

References

[Coe97] Coelho, Z.: Asymptotic laws for symbolic dynamical systems. Lectures given in Temuco, Chili, 1997

[Col96] Collet, P.: *Some ergodic properties of maps of the interval*, Dynamical systems (Temuco, 1991/1992) (Paris), Travaux en Cours, vol. 52, Paris: Hermann, 1996, pp. 55–91

[CG93] Collet, P. and Galves, A.: Statistics of close visits to the indifferent fixed point of an interval map. J. Stat. Phys. **72**, no. 3-4, 459–478 (1993)

[CFS82] Cornfeld, I.P., Fomin, S.V. and Sinaĭ, Ya.G.: *Ergodic theory*. vol. 245, New York: Springer-Verlag, 1982

[GS97] Galves, A. and Schmitt, B.: Inequalities for hitting time in mixing dynamical systems. Random Comput. Dynam. **5**, no. 4, 337–347 (1997)

[Hay98a] Haydn, N.: The distribution of the first return time for rational maps. 1998, USC

[Hay98b] Haydn, N.: Statistical properties of equilibrium states for rational maps. 1998, USC

[Hir93] Hirata, M.: Poisson law for Axiom A diffeomorphisms. Ergodic Theory Dynamical Systems **13**, no. 3, 533–556 (1993)

[Hir95] Hirata, M.: Poisson law for the dynamical systems with the "self-mixing" conditions. In: *Dynamical systems and chaos, Vol. 1 (Hachioji, 1994) (River Edge, NJ)*, River Edge, NJ: World Sci. Publishing, 1995, pp. 87–96

[HK82] Hofbauer, F. and Keller, G.: Ergodic properties of invariant measures for piecewise monotonic transformations. Math. Z. **180**, no. 1, 119–140 (1982)

[LSV97] Liverani, C., Saussol, B. and Vaienti, S.: A probabilistic approach to intermittency. Ergodic Theory Dynamical Systems (1997). To appear

[PSV98] Penné, V., Saussol, B. and Vaienti, S.: Fractal and statistical characteristics of recurrence times. To appear in Journal de Physique (Paris), 1998, Proceeding of the conference "Disorder and Chaos" (Rome, 22–24th Sept. 1997), in honour of Giovanni Paladin

[Pit91] Pitskel, B.: Poisson limit law for Markov chains. Ergodic Theory Dynamical Systems **11**, no. 3, 501–513 (1991)

[Sau98a] Saussol, B.: Absolutely continuous invariant measures for multidimensional expanding maps. 1998, Submitted

[Sau98b] Saussol, B.: Etude statistique de systèmes dynamiques dilatants. Ph.D.thesis, Université de Toulon et du Var, 1998

Communicated by Ya. G. Sinai