DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **10**, Number **3**, April **2004**

Website: http://AIMsciences.org

pp. **589–616**

THE LIMITING DISTRIBUTION AND ERROR TERMS FOR RETURN TIMES OF DYNAMICAL SYSTEMS

Nicolai Haydn

Department of Mathematics
University of Southern California
Los Angeles, 90089-1113

Sandro Vaienti

Phymat, Université de Toulon
Centre de Physique Théorique and FRUMAM
Luminy Case 907, 13288 Marseille Cedex 09

Abstract. We develop a new framework that allows to prove that the limiting distribution of return times for a large class of mixing dynamical systems are Poisson distributed. We demonstrate our technique in several settings and obtain more general results than previously has been proven. We also obtain error estimates. For ϕ-mixing maps we obtain a close to exhausting description of return times. For (ϕ, f)-mixing maps it is shown how the separation function affects error estimates for the limiting distribution. As examples of (ϕ, f)-mixing we prove that for piecewise invertible maps and for rational maps return times are in the limit Poisson distributed.

1. **Introduction.** We study the distribution of return times for transformations to small set. Let T be a transformation on the space Ω and let μ be a probability measure on Ω. Denote by χ_A the characteristic function of a (measurable) set A and define the 'random variable'

$$\xi_A = \sum_{j=1}^{[t/\mu(\chi_A)]} \chi_A \circ T^j.$$

The value of ξ_A measures the number of times a given point returns to A within the normalised time t (the normalisation is with respect to the μ-measure of the 'return-set' A). If μ is the measure of maximal entropy for the shift transformation on a subshift of finite type, then it was shown by Pitskel [21] that the return times are in the limit Poisson distributed for cylinder sets and μ-almost every x. For equilibrium states of Hölder continuous functions, Hirata ([14], [15]) has similar results for the zeroth return time $r = 0$ using the transfer operator restricted to the complement of ε-balls in the shiftspace. Wang, Tang and Wang [24] have a result for non-homogeneous Markov chains which uses Pitskel's method which was

1991 *Mathematics Subject Classification.* 37A25, 37A50, 37F35.

Key words and phrases. Mixing dynamical systems, return times statistics, Poisson distribution, ϕ-mixing, Axiom-A systems, rational maps, piecewise expanding maps.

This work was supported by a grant from the NSF and by a grant from the Université de Toulon et du Var, France.

also used by Denker [10]. All those results don't give rates of convergence. Also note the review paper by Z Coelho [8] for some other results on the return times distribution and the review paper by Abadi and Galves [3].

For cylinder sets, Galves and Schmitt [11] have obtained rates of convergence for the zeroth order return times ($r = 0$). Hirata, Saussol and Vaienti [16] have developed a general scheme to prove that return times are in the limit Poisson distributed and applied it to a family of interval maps with a parabolic point at the origin (where the map is like $x + x^{1+\alpha}$ for some $\alpha \in (0, 1)$).

Here we develop a mechanism which allows to prove the Poisson distribution of return times and to obtain error estimates as the set A shrinks to a single point.

In the second section we prove a theorem that gives general conditions under which a sum of mutually dependent $0, 1$-valued random variables converges to the Poisson distribution and provides error terms. This quantifies a previous theorem of Sevast'yanov who originally in [22] has a version of Corollary 2 without error terms (that was the result used by B Pitskel in [21]). Theorem 1 plays a centre part in this paper to obtain the limiting distribution of entry and re-entry times of dynamical systems. (The $0, 1$-valued random variables will count returns to a given set.) The main feature of the theorem is the distinction between 'short' returns and 'long' returns and the tradeoff between the size of the set of points that have short returns and the speed of mixing that is implied by long returns.

The distribution of return times is tied to the mixing properties of the invariant measure considered. For that purpose we introduce in the third secion the (ϕ, f)-mixing property. This property is more general that the widely used ϕ-mixing property and is reminiscent of Philipp and Stout's [20] 'retarded strong mixing property'. In this way one can obtain distribution results on return times of some well studied dynamical systems that are not ϕ-mixing, e.g. rational maps, parabolic maps, piecewise expanding maps in higher dimension ....

In section four we look more closely at the return times patterns and and single out those that contain at least one short return. In order to get a good distribution it is necessary to avoid very short return times that are of the order of the length of the cylinders considered. The remaining return time patterns that contain only short returns of moderate length form a 'rare' set and is shown to be in general small. In the remainder of the section we then derive general propositions that will be used in the last sections to obtain distribution results for various mixing systems.

In section five we apply our results to ϕ-mixing systems where ϕ is algebraically or exponentially decaying. Similar results have been obtained by M Abadi [1] for α-mixing maps using the Chen-Stein method. His error estimates are better in terms of the parameter t but not as good in the dependency on the order of returns as ours are. The error terms we obtain involve the measure of a cylinder set whose length is determined by the 'shortest repeat time function' τ. All our results come in three flavours: for entry times, for return times and then we restrict to sets that have only points without 'very short' returns. The latter one we call *restricted return times* (see section 3.3) and has the advantage that the error terms are not constraint by the shortest repeat time function.

Finally, in section six we apply our method to maps that are genuinely (ϕ, f)-mixing (i.e. not ϕ-mixing): piecewise expanding maps in higher dimensions and rational maps.

In the following $c_1, c_2, \dots$ are constants that are locally used while $C_1, C_2, \dots$ indicate constants that whose values apply throughout the text.

2. **Factorial moments and mixing.** In this section we prove a technical result, Theorem 1 where we provide general conditions under which distribution of a finite set of 0, 1-valued random variables is close to Poisson, where the parameter for the approximating Poisson distribution is the expected value of the counting function ζ below. We give an explicit expression for the error term: It is determined by the independence of the given random variables (see condition (5) below) and the sparcity of short term correlations (conditions (3) and (4) below).

For (ϕ, f)-mixing maps we show in section 4 how the dynamical properties are used to satisfy the conditions (1)–(5) of Theorem 1. There we have to make a distinction between short returns and long returns. For long returns we obtain strong mixing properties, while the short returns (sets K_r and I_r) can only be estimated in a rough manner (section 4.1).

In the sections that follow we then use these results to obtain the speed of convergence for the limiting distributions for ϕ-mixing systems, some non-Markovian systems and equilibrium states for rational maps with critical points.

In the following G_r is a subset of $\mathbf{Z}^r$.

Theorem 1. *Let $\{\eta_v : v = 1, \dots, N\}$, be 0, 1-valued random variables and μ a probability measure. Put $\zeta = \sum_{v=1}^N \eta_v$, and for $\vec{v} \in G_r = \{\vec{v} \in \mathbf{Z}^r : 1 \le v_1 < v_2 < \cdots < v_r \le N\}$ let $b_{\vec{v}} = \mu(\eta_{\vec{v}})$, where $\eta_{\vec{v}} = \prod_{s=1}^r \eta_{v_s}$ (if $r = 1$ then $b_v = \mu(\eta_v)$).*

Assume that there exist sets $R_r \subset G_r$ ($r \ge 1$) and $t > 0$, $\alpha \ge 0$ so that (the numbers r', r'' satisfy $|r' - r|, |r'' - r| \le 2$) the following five assumptions are satisfied:

$$\max_{1 \le v \le N} b_v \quad \le \quad \varepsilon \tag{1}$$

$$\left| \sum_{v=1}^N b_v - t \right| \quad \le \quad \varepsilon \tag{2}$$

$$\sum_{\vec{v} \in R_r} b_{\vec{v}} \quad \le \quad \varepsilon \sum_{s=0}^{r'} \binom{r'}{s} \varepsilon^{r'-s} \frac{(\alpha t)^s}{s!} \tag{3}$$

$$\sum_{\vec{v} \in R_r} b_{v_1} \cdots b_{v_r} \quad \le \quad \varepsilon \sum_{s=0}^{r''} \binom{r''}{s} \varepsilon^{r''-s} \frac{(\alpha t)^s}{s!} \tag{4}$$

$$\left| \frac{b_{v_1} \cdots b_{v_r}}{b_{\vec{v}}} - 1 \right| \quad \le \quad \alpha^r \varepsilon \qquad \forall \ \vec{v} \in G_r \setminus R_r, \tag{5}$$

for some $\varepsilon > 0$ ($\varepsilon < 1$).

Then, for every $\alpha' > \alpha$ there exists a constant C_1 independent of ε so that for all $t > 0$ and r for which $r^2 \varepsilon / t$ is small if $k \ge 1$ and εt is small if $k = 0$ (for smallness see equation (7)):

$$\left| \mu(\mathcal{N}^k) - \frac{t^k e^{-t}}{k!} \right| \le \begin{cases} C_1 \frac{(k+t)^2}{k!} \varepsilon (\alpha' t)^{k-1} e^{\alpha' t} & \text{if} \quad k \ge 1 \\ C_1 e^{\alpha' t + 1} \varepsilon (t+1) & \text{if} \quad k = 0 \end{cases}$$

For all values of k and t one has the (weaker) bound

$$\left| \mu(\mathcal{N}^k) - \frac{t^k e^{-t}}{k!} \right| \le C_1 \varepsilon e^{2\alpha' t} t.$$

where $\mathcal{N}^k = \{y : \zeta(y) = k\}$ *is the* k*-levelset of* ζ*.*

The sets R_r are sometimes called *rare sets*.

Proof. Throughout the proof we shall assume that $r' = r'' = r$. If $r'', r' \neq r$ then there are obvious modifications below that let us arrive at the same conclusion (except the constant C_1 will be somewhat larger).

If we put $U_r = r! \sum_{\vec{v} \in G_r} b^n_{\vec{v}}$ then

$$|U_r - t^r| \leq I + II + III + IV + V,$$

where by assumption (3)

$$I = \left| U_r - r! \sum_{\vec{v} \notin R_r} b_{\vec{v}} \right| = r! \sum_{\vec{v} \in R_r} b_{\vec{v}} \leq r! \varepsilon \sum_{s=0}^{r} \left(\begin{array}{c} r \\ s \end{array} \right) \varepsilon^{r-s} \frac{(\alpha t)^s}{s!}$$

and by assumption (4)

$$II = \left| V_r - r! \sum_{\vec{v} \notin R_r} \prod_i b_{v_i} \right| \leq r! \varepsilon \sum_{s=0}^{r} \left(\begin{array}{c} r \\ s \end{array} \right) \varepsilon^{r-s} \frac{(\alpha t)^s}{s!},$$

where we put $V_r = r! \sum_{\vec{v} \in G_r} \prod_i b_{v_i}$. Moreover by assumption (2)

$$III = \left| \left(\sum_{k=0}^{N} b_k \right)^r - t^r \right| \leq r\varepsilon (t+\varepsilon)^{r-1} \leq \varepsilon \sum_{x=0}^{r} (\alpha\varepsilon)^{r+s} \frac{(\alpha t)^s}{s!} \left(\begin{array}{c} r \\ s \end{array} \right).$$

To estimate the term IV we factor out as follows

$$\left(\sum_{k=0}^{N} b_k \right)^r = r! \sum_{\vec{v} \in G_r} \prod_i b_{v_i} + \sum_{k=1}^{r-1} \sum_{\vec{v} \in H_r^k} \prod_i b_{v_i},$$

where H_r^k consists of all those unordered multi-indices $\vec{v} = (v_1, \ldots, v_r)$, $0 \leq v_j \leq N$, which have exactly $r-k$ distinct entries. We wish now to estimate the sum over each set H_r^k by the sum over the set G_{r-k} of ordered $(r-k)$-tuples. To generate all of the possible unordered r-tuples $\vec{v}$ in H_r^k, let $\vec{w} \in G_{r-k}$. There are $(r-k)!$ possible arrangements of the entries of $\vec{w}$. There are $\frac{r!}{(r-k)!k!}$ possibilities to fit any of these arrangements into the r slots of a vector $\vec{v}$ and there are $(r-k)^k$ many ways to fill the remaining k empty slots with any of the $r-k$ distinct entries of $\vec{w}$. Hence, by assumption (1)

$$\begin{aligned} \sum_{\vec{v} \in H_r^k} \prod_i b_{v_i} &\leq \frac{r!}{(r-k)!k!} (r-k)^k (\max_i b_i)^k (r-k)! \sum_{\vec{v} \in G_{r-k}} \prod_i b_{v_i} \\ &\leq \frac{r!}{(r-k)!k!} (r-k)^k \varepsilon^k V_{r-k}. \end{aligned}$$

With the estimate:

$$V_r = \left(\sum_{k=0}^{N} b_k \right)^r - \sum_{k=1}^{r-1} \sum_{\vec{v} \in H_r^k} \prod_i b_{v_i} \leq \left(\sum_{k=0}^{N} b_k \right)^r \leq (t+\varepsilon)^r,$$

we obtain

$$
\begin{aligned}
IV = \left| \left(\sum_{k=0}^{N} b_k \right)^r - V_r \right| &\leq \sum_{k=1}^{r-1} \frac{r!}{(r-k)!k!} (r-k)^k \varepsilon^k V_{r-k} \\
&\leq \sum_{k=1}^{r-1} \frac{r!(r-k)^k}{(r-k)!k!} r^k \varepsilon^k (t+\varepsilon)^{r-k}.
\end{aligned}
$$

To estimate the term V we proceed as follows using assumption (5) in the first inequality:

$$
\begin{aligned}
V &= \left| r! \sum_{\vec{v} \notin R_r} \prod_i b_{v_i} - r! \sum_{\vec{v} \notin R_r} b_{\vec{v}} \right| \\
&\leq r! \alpha^r \varepsilon \sum_{\vec{v} \notin R_r} \prod_i b_{v_i} \\
&\leq \alpha^r \varepsilon V_r \\
&\leq \varepsilon \alpha^r (t+\varepsilon)^r \\
&\leq r! \varepsilon \sum_{s=0}^{r} (\alpha\varepsilon)^{r-s} \frac{(\alpha t)^s}{s!}.
\end{aligned}
$$

Hence

$$
|U_r - t^r| \leq 4r! \varepsilon \sum_{s=0}^{r} \begin{pmatrix} r \\ s \end{pmatrix} \varepsilon^{r-s} \frac{(\alpha t)^s}{s!} + \sum_{k=1}^{r-1} \frac{r!(r-k)^k}{(r-k)!k!} \varepsilon^k (t+\varepsilon)^{r-k}.
$$

Let $f(z) = \sum_{k=0}^{\infty} z^k \mu(\mathcal{N}^k)$ be the generating function which we develop at $z = 1$ into a powerseries:

$$
f(z) = \sum_{r=0}^{\infty} \frac{f^{(r)}(1)}{r!} (z-1)^r = \sum_{r=0}^{\infty} \frac{(z-1)^r}{r!} \mu(\zeta^{(r)}),
$$

where $\zeta^{(r)} = \zeta(\zeta-1)\cdots(\zeta-r+1)$ is the rth factorial moment of ζ. For $x \in \mathcal{N}^k$, $k \geq r$, one has that $\zeta^{(r)}(x) = k(k-1)\cdots(k-r+1)$. For $\vec{v} \in G_r$ let us put $C_{\vec{v}} = \{x : \eta_{\vec{v}} = 1\}$ and let us observe that for any given r we have:
(i) if $x \in \mathcal{N}^k$ for some $k < r$ then $x \notin C_{\vec{v}}$, for all $\vec{v} \in G_r$,
(ii) if $x \in \mathcal{N}^k$ for $k \geq r$ then there are $\begin{pmatrix} k \\ r \end{pmatrix}$ distinct $\vec{v} \in G_r$ so that $x \in C_{\vec{v}}$.
Since $C_{\vec{v}} = \bigcup_{k=r}^{\infty} C_{\vec{v}} \cap \mathcal{N}^k$ (disjoint union) we get

$$
\begin{aligned}
\sum_{\vec{v} \in G_r} \mu(C_{\vec{v}}) &= \sum_{k=r}^{\infty} \sum_{\vec{v} \in G_r} \mu(C_{\vec{v}} \cap \mathcal{N}_k) \\
&= \sum_{k=r}^{\infty} \frac{k!}{(k-r)!r!} \mu(\mathcal{N}^k) \\
&= \frac{1}{r!} \mu(\zeta^{(r)})
\end{aligned}
$$

and therefore

$$
U_r = \mu(\zeta^{(r)}) = f^{(r)}(1).
$$

The (error) function is

$$\varphi(z)=f(z)-e^{t(z-1)}=\sum_{r=0}^{\infty}\frac{(z-1)^r}{r!}\left(f^{(r)}(1)-t^r\right).$$

We split φ into the sum $\varphi=\varphi_1+\varphi_2$ where φ_2 is the generating function for $\left(\sum_{k=0}^N b_k\right)^r - V_r$ and is estimated as follows (we put $\ell=r-k$):

$$\begin{aligned}
|\varphi_2(z)| &\le \sum_{r=1}^{\infty}\frac{|z-1|^r}{r!}\sum_{k=1}^{r-1}\frac{r!}{(r-k)!k!}(r-k)^k\varepsilon^k(t+\varepsilon)^{r-k}\\
&\le \sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty}\frac{\ell^k}{\ell!k!}|z-1|^{\ell+k}\varepsilon^k(t+\varepsilon)^{\ell}\\
&\le \sum_{\ell=1}^{\infty}\frac{|z-1|^{\ell}(t+\varepsilon)^{\ell}}{\ell!}\left(e^{|z-1|\ell\varepsilon}-1\right)\\
&= \left(e^{e^{|z-1|\varepsilon}|z-1|(t+\varepsilon)}-e^{|z-1|(t+\varepsilon)}\right).
\end{aligned}$$

The function $\varphi_1=\varphi-\varphi_2$ is (as $\alpha\ge 1$)

$$\begin{aligned}
|\varphi_1(z)| &\le \sum_{r=0}^{\infty}\frac{|z-1|^r}{r!}c_1 r!\varepsilon\sum_{k=0}^{r}\binom{r}{k}(\alpha\varepsilon)^{r-k}\frac{(\alpha t)^k}{k!}\\
&= c_1\varepsilon F(|z-1|\alpha t,|z-1|\alpha\varepsilon)
\end{aligned}$$

where we used the identity

$$F(x,y)=\sum_{r=0}^{\infty}\sum_{s=0}^{r}\binom{r}{s}y^{r-s}\frac{x^s}{s!}=\frac{1}{1-y}e^{\frac{x}{1-y}}.$$

In particular we see that φ_1 is for every value of t analytic for $|z-1|<\alpha/\varepsilon$ and moreover φ, φ_1 and φ_2 are for every value of t analytic for $|z-1|<\alpha/\varepsilon$. Let $\alpha'>\alpha$. For $\alpha|z-1|\varepsilon\le 1-\frac{\alpha}{\alpha'}$ we get ($c_2>c_1$)

$$|\varphi_1(z)|\le c_2\varepsilon e^{\alpha'|z-1|(t+\varepsilon)}$$

and

$$\begin{aligned}
|\varphi_2(z)| &\le e^{|z-1|(t+\varepsilon)}\left(e^{(e^{|z-1|\varepsilon}-1)|z-1|(t+\varepsilon)}-1\right)\\
&\le 4e^{|z-1|(t+\varepsilon)}|z-1|^2\varepsilon(t+\varepsilon),
\end{aligned}$$

provided $|z-1|^2\varepsilon(t+\varepsilon)\le 1$. A Cauchy estimate now yields ($R>0$):

$$|\varphi^{(k)}(0)|\le\frac{k!}{R^k}\left(2c_2\varepsilon e^{\alpha'(R+1)(t+\varepsilon)}+2e^{(R+1)(t+\varepsilon)}(R+1)^2\varepsilon(t+\varepsilon)\right),$$

provided, of course, that $(R+1)\varepsilon\alpha<1-\frac{\alpha}{\alpha'}$ and $(R+1)^2\varepsilon(t+\varepsilon)\le 1$. Hence, since

$$\mu(\mathcal{N}^k)=\frac{f^{(k)}(0)}{k!}=\frac{t^k}{k!}e^{-t}+\frac{\varphi^{(k)}(0)}{k!},$$

we get

$$\left|\mu(\mathcal{N}^k)-\frac{t^k}{k!}e^{-t}\right|\le\frac{\varphi^{(k)}(0)}{k!}\le c_3\frac{(R+1)^2}{R^k}e^{\alpha'(R+1)t}\varepsilon(t+1). \tag{6}$$

One can now obtain different estimates by choosing different values for R (subject to the constraint mentioned above). If $R = 1$ then we simply obtain

$$\left|\mu(\mathcal{N}^k) - \frac{t^k}{k!}e^{-t}\right| \leq c_4\varepsilon e^{2\alpha' t}(t+1)$$

for some constant c_4. A better choice of R can be obtained if for instance

$$\left(\frac{k}{\alpha' t+1}+1\right)^2 \varepsilon\alpha(t+1) < 1 - \frac{\alpha}{\alpha'} \tag{7}$$

in which case the optimal value is $R = \frac{k}{\alpha' t+1}$. Then ($c_5 > 0$)

$$\left|\mu(\mathcal{N}^k) - \frac{t^k}{k!}e^{-t}\right| \leq c_5\frac{(k+t)^2}{k^k}e^{k+\alpha' t}\varepsilon(\alpha' t+1)^{k-1}.$$

Using Stirling's formula one obtains the estimate given in the statement of the theorem. If $k = 0$ then in equation (6) we let $R \to 0$ and obtain

$$\left|\mu(\mathcal{N}^0) - e^{-t}\right| \leq \varphi^{(k)}(0) \leq c_3 e^{\alpha' t}\varepsilon(t+1).$$

$\square$

Remarks. (i) The error estimate becomes meaningless if t of the order (or larger than) $|\log k^2\varepsilon|$ because the principal term is then smaller than the error term.

(ii) If in property (5) the left hand side is equal to zero for all $\vec{v}$ then the statement of the theorem would is trivially satisfied (since then $\mu(\mathcal{N}^k) = \frac{t^k e^{-t}}{k!}$ for all k).

(iii) In property (5) the error only has to go to zero for 'most' of the multi-indices $\vec{v}$. For the remaining multi-indices (rare set) we don't require anything except that there are not too many of them (that is conditions (3) and (4). In our setting the rare set will typically consist of return time patterns $\vec{v}$ which contain a return which is 'too short'.

(iv) The error term ε is allowed to depend on t which is a parameter in the theorem.

The following corollary is a simplified version of Theorem 1.

Corollary 2. *Let $\{\eta_v : v = 1, \dots, N\}$ be an array of random $0,1$-valued variables and μ a probability measure. Put $\zeta = \sum_{v=1}^N \eta_v$, and let $b_{\vec{v}} = \mu(\eta_{\vec{v}})$ for $\vec{v} \in G_r$ (where $\eta_{\vec{v}} = \prod_{s=1}^r \eta_{v_s}$). Assume that there are there is $R_r \subset G_r$ ($r \geq 1$) and $t > 0$ so that*

$$\begin{aligned}
\max_{1\leq v\leq N} b_v &\leq \varepsilon \\
\left|\sum_{v=1}^N b_v - t\right| &\leq \varepsilon \\
\sum_{\vec{v}\in R_r}(b_{\vec{v}} + b_{v_1}\cdots b_{v_r}) &\leq \varepsilon \\
\left|\frac{b_{v_1}\cdots b_{v_r}}{b_{\vec{v}}} - 1\right| &\leq \varepsilon \qquad \forall\ \vec{v} \in G_r \setminus R_r.
\end{aligned}$$

Then there exists a constant C_1 independent of t so that for all k for which $(k^2+1)\varepsilon/t$ is small[1]*:*

$$\left|\mu(\mathcal{N}^k) - \frac{t^k e^{-t}}{k!}\right| \le \begin{cases} C_1\varepsilon e^t \frac{(k+t)^2}{k!}(t+1)^{k-1} & \text{if } k \ge 1 \\ C_1\varepsilon e^t(t+1) & \text{if } k = 0 \end{cases}$$

For all values of k and t one has the (weaker) bound

$$\left|\mu(\mathcal{N}^k) - \frac{t^k e^{-t}}{k!}\right| \le C_1\varepsilon e^{2t} t.$$

where $\mathcal{N}^k = \{y : \zeta(y) = k\}$ is the k-levelset of ζ.

3. **Properties of (ϕ, f)-mixing measures.** In this section we introduce the class of systems for which we develop our framework for obtaining distribution results on return times. The definition below generalises the 'retarded strong mixing condition' (see e.g. [20]). We consider mixing dynamical systems in which the function ϕ determines the rate of mixing while the *separation function* f specifies a lower bound for the size of the 'gap' m that is necessary to get the mixing property.

Let T be a map on a space Ω and μ a probability measure on Ω. Moreover let $\mathcal{A}$ be a measurable partition of Ω and denote by $\mathcal{A}^n = \bigvee_{j=0}^{n-1} T^{-j}\mathcal{A}$ its *n-th join* which also is a measurable partition of Ω for every $n \ge 1$. The atoms of $\mathcal{A}^n$ are called *n-cylinders*. Let us put $\mathcal{A}^* = \bigcup_n \mathcal{A}^n$ for the collection of all cylinders in Ω and put $|A|$ for the length of an n-cylinder $A \in \mathcal{A}^*$, i.e. $|A| = n$ if $A \in \mathcal{A}^n$.

We shall assume that $\mathcal{A}$ is generating, i.e. that the atoms of $\mathcal{A}^\infty$ are single points in Ω.

Definition 3. *Assume*
(i) $f : \mathcal{A}^* \to \mathbf{N}_0$ ($\mathbf{N}_0 = \{0, 1, 2, \dots\}$) so that $f(A) \ge f(B)$ if $|A| \ge |B|$, $A, B \in \mathcal{A}^*$. If U is a union of n-cylinders U_j (some n) then $f(U) = \max_j f(U_j)$.
(ii) $\phi : \mathbf{N}_0 \to \mathbf{R}^+$ is non-increasing.

*We say that the dynamical system (T, μ) is (ϕ, f)-*mixing *if*

$$\left|\mu(U \cap T^{-m-n}V) - \mu(U)\mu(V)\right| \le \phi(m)\mu(U)\mu(V)$$

for all $m \ge f(U)$, measurable V (in the σ-algebra generated by $\mathcal{A}^*$) and U which are unions of cylinders of the same length n, for all n.

In section 4 where we consider the distribution of entry and return times we consider separation functions f that depend only on the length of the cylinders, that is $f(A) = f(n)$ for $A \in \mathcal{A}^n$, $n = 1, 2, \dots$. In section 5 we discuss the entry and return times distribution for ϕ-mixing maps. This is the special case in which $f = 0$

In section 6 we then consider two cases in which the separation function is not constant: in 6.1 we discuss the return times for multidimensional piecewise continuous maps and in 6.2 we give uniform convergence results for the return times of rational maps where the separation function $f(n)$ is linear in n.

For $r \ge 1$ and (large) N denote by $G_r(N)$ the r-vectors $\vec{v} = (v_1, \dots, v_r)$ for which $1 \le v_1 < v_2 < \dots < v_r \le N$. (The set $G_r(N)$ is the intersection of a cone in $\mathbf{Z}^r$ with a ball of radius N.) Let t be a positive parameter, put $N = [t/\mu(W)]$ (the normalised time) and $W \subset \Omega$. Then the entries v_j of the vector $\vec{v} \in G_r(N)$ are the

[1] for smallness see equation (7) with $\alpha' = 1$

times at which all the points in $C_{\vec{v}} = \bigcap_{j=1}^{r} T^{-v_j} W$ hit the set W during the time interval $[1, N]$.

The following simple lemmas will be needed later on.

Lemma 4. *Let (T, μ) be (ϕ, f)-mixing, let $r > 1$ be an integer and let $W_j \subset \Omega$, be unions of n_j-cylinders, $j = 1, \dots, r$.*

Then for all 'hitting vectors' $\vec{v} \in G_r(N)$ with return times $v_{j+1} - v_j \geq f(W_j) + n_j$ ($j = 1, \dots, r-1$) one has

$$\left| \frac{\mu\left(\bigcap_{j=1}^{r} T^{-v_j} W_j\right)}{\prod_{j=1}^{r} \mu(W_j)} - 1 \right| \leq (1 + \phi(d(\vec{v}, \vec{n})))^r - 1,$$

and $d(\vec{v}, \vec{n}) = \min_k (v_{k+1} - v_k - n_k)$.

Proof. Put for $k = 1, 2, \dots, r$:

$$D_k = \bigcap_{j=k}^{r} T^{-(v_j - v_k)} W_j.$$

In particular we have $\bigcap_{j=1}^{r} T^{-v_j} W_j = T^{-v_1} D_1$ and of course $\mu\left(\bigcap_{j=1}^{r} T^{-v_j} W_j\right) = \mu(D_1)$. Also note that

$$D_k = W_k \cap T^{-(v_{k+1} - v_k)} D_{k+1}$$

and $D_r = W_r$. Hence by assumption we obtain

$$|\mu(D_k) - \mu(W_k)\mu(D_{k+1})| \leq \phi(v_{k+1} - v_k - n_k)\mu(D_{k+1})\mu(W_k).$$

Repeated application of the triangle inequality yields

$$\left| \mu\left(\bigcap_{j=1}^{r} T^{-v_j} W_j\right) - \prod_{j=1}^{r} \mu(W_j) \right| \leq ((1 + \phi(d(\vec{v}, \vec{n})))^r - 1) \prod_{j=1}^{r} \mu(W_j),$$

where we used the estimates

$$\begin{aligned} \mu(D_k) &\leq \mu(W_k)\mu(D_{k+1})\left(1 + \phi(v_{k+1} - v_k - n_k)\right) \\ &\leq \mu(W_r) \prod_{j=k}^{r-1} (1 + \phi(v_{j+1} - v_j - n_j))\mu(W_j) \\ &\leq (1 + \phi(d(\vec{v}, \vec{n})))^{r-k-1} \prod_{j=k}^{r} \mu(W_j) \end{aligned}$$

since by assumption that $v_{k+1} - v_k - n_k \geq f(W_k)$. □

The following exponential estimate has previously been shown for ϕ-mixing measures in [11] and for α-mixing measures in [2] and is an immediate consequence of Lemma 4.

Lemma 5. *There exists a $0 < \gamma < 1$ so that for all $A \in \mathcal{A}^*$:*

$$\mu(A) \leq \gamma^{|A|}.$$

This lemma has the following result as a simple consequence:

Lemma 6. *The measure μ has positive metric entropy.*

4. **Return times.** This section contains the technical results that are used in sections 5 and 6 to obtain the return times distributions. Here we determine the dynamical situations under which the conditions of Theorem 1 are satisfied. The main feature of Theorem 1 is the distinction between short return times and long return times. The parameters δ (in 4.1) and δ_n (in 4.2 and 4.3) divide the return time patterns (G_r) into those patterns which have short returns ($R_r, \tilde{R}_r$) and the remaining ones that only consist of long returns. In subsection 4.1 we show that the 'short return times patterns' overall occur infrequently. (In Proposition 10 Lemmas 7 and 8 serve to satisfy conditions (3) and (4) of Theorem 1.) In subsections 4.2 and 4.3 we then estimate the quantities in the conditions (1)–(5) in terms of dynamical variables in four different settings: entry times (Proposition 9), return times (Proposition 10), restricted entry times (Proposition 11) and restricted return times (Proposition 12).

The short returns will fall into two categories, namely those that contain very short returns (I_r) and those where all the shortest returns are of some moderate length ($K_r, \tilde{K}_r$). For that purpose we shall from now on restrict to the situation where all the sets W_j are identical and equal to some W (the return set). For a 'hitting vector' $\vec{v} \in G_r(N)$ (N a large integer) we put $C_{\vec{v}} = \bigcap_{j=1}^{r} T^{-v_j} W$. Let $\delta \geq f(W)$ and define the rare set

$$R_r(N) = \{\vec{v} \in G_r(N) : \min(v_{j+1} - v_j) < \delta\}.$$

For some $1 \leq \delta' \leq \delta$ we have the principal part of the rare set given by

$$K_r(N) = \{\vec{v} \in R_r(N) : \delta' \leq \min(v_{j+1} - v_j)\}$$

The set $K_r(N)$ will be estimated in rather general terms below, but the remaining portion

$$I_r(N) = R_r(N) \setminus K_r(N) = \{\vec{v} \in R_r(N) : \min(v_{j+1} - v_j) < \delta'\}$$

typically has to be disposed of by employing some ad hoc argument exploiting particularities of the map T.

For the return times statistics we shall use a slightly different rare set, namely

$$\tilde{R}_r(N) = \{\vec{v} \in G_{r+1}(N) : \min_j(v_{j+1} - v_j) < \delta \text{ and } v_1 = 0\}.$$

Note that $\tilde{R}$ tracks the returns of points whose origin is in W ($v_1 = 0$ implies that $x \in W$). Correspondingly the principal part is

$$\tilde{K}_r(N) = \{\vec{v} \in \tilde{R}_r(N) : \delta' \leq \min_j(v_{j+1} - v_j)\}.$$

4.1. **Short return times K_r and $\tilde{K}_r$.**

Lemma 7. *Assume (T, μ) is (ϕ, f)-mixing. Then for every union W of n-cylinders one has (for some $C_2 > 0$)*
(i) (Entry time version)

$$\sum_{\vec{v} \in K_r} \mu(C_{\vec{v}}) \leq C_2 t \mu(V) \sum_{s=0}^{r-2} \left(\begin{array}{c} r-2 \\ s \end{array} \right) \frac{(\beta t)^s}{s!} (\beta \delta \mu(V))^{r-s},$$

(ii) (Return time version)

$$\sum_{\vec{v}\in\tilde{K}_r}\mu(C_{\vec{v}})\le C_2\mu(W)\sum_{s=0}^{r-1}\left(\begin{array}{c}r-1\\ s\end{array}\right)\frac{(2\beta t)^s}{s!}(2\beta\delta\mu(V))^{r-s},$$

where $\beta>1+\phi(\min_k(v_{k+1}-v_k)-\delta'')$ *and the set* V *is a union of atoms in* $\mathcal{A}^{\delta''}$ *such that* $W\subset V$ *and* δ'' *is so that* $f(V)\le\delta'-\delta''$*.*

Proof. As in the hypothesis let W be a union of n-cylinders so that $f(W)\le\delta$.

(i) Let us first prove the first statement of the lemma. Put K_r^s for those $\vec{v}\in K_r$ where $v_{i+1}-v_i\ge\delta$ for exactly s indices $i_1,i_2,\dots,i_s$ (obviously one always has $s\le r-2$ and $i_s\le r-1$).

I. Let us now assume that $s\ge 1$ and let $i_1,i_2,\dots,i_s$ be the indices for which $v_{i_k+1}-v_{i_k}\ge\delta$ for $k=1,\dots,s$. All the other differences are $\ge\delta'$ and smaller than δ. Let V be a union of δ''-cylinders so that $V\subset W$. By assumption assumption V can be chosen so that $f(V)\le\delta'-\delta''$. Put $W_{i_1}=W_{i_2}=\cdots=W_{i_s}=W_r=W$ and $W_j=V$ for all indices j not equal to any of the i_k or r.

By our choice of δ'' we have achieved that $v_{i_k+1}-v_{i_k}\ge\delta\ge f(W)$ and $v_{j+1}-v_j\ge f(V)$ for $j\ne i_k,\ k=1,\dots,s$. This allows us to apply Lemma 4 as follows:

$$\begin{aligned}\mu\left(\bigcap_{i=1}^{r}T^{-v_i}W\right) &\le \mu\left(\bigcap_{i=1}^{r}T^{-v_i}W_i\right)\\ &\le (1+\phi(d(\vec{v},\vec{n})))^r\prod_{i=1}^{r}\mu(W_i)\\ &\le \beta^{r-1}\mu(V)^{r-s-1}\mu(W)^{s+1},\end{aligned}$$

$\beta=1+\phi(d(\vec{v},\vec{n}))$, where the components of $\vec{n}=(n_1,\dots,n_r)$ are given by $n_{i_k}=n$ for $k=1,\dots,s$ and $n_j=\delta''$ for $j\ne i_k,\ k=1,\dots,s$.

To estimate the cardinality of K_r^s let us note that the number of possibilities of $v_{i_1}<v_{i_2}\cdots<v_{i_s}<v_{i_s+1}$ (entrance times for long returns) is bounded above by $\frac{1}{(s+1)!}\left(t/\mu(W)\right)^{s+1}$ (this is the upper bound for the number of possibilities to obtain $s-1$ intervals contained in the interval $[1,t/\mu(W)]$), and each of the remaining $r-s-1$ (short) return times assume no more than δ different values. Since the indices $i_1,\dots,i_s$ can be picked in $\left(\begin{array}{c}r\\ s\end{array}\right)$ many ways, we obtain:

$$|K_r^s|\le\left(\begin{array}{c}r\\ s\end{array}\right)\frac{\delta^{r-s-1}}{(s+1)!}\left(\frac{t}{\mu(W)}\right)^{s+1}.$$

The above estimates combined yield

$$\sum_{\vec{v}\in K_r^s}\mu(C_{\vec{v}})\le\beta^{r-1}\left(\begin{array}{c}r\\ s\end{array}\right)\frac{t^{s+1}}{(s+1)!}(\delta\mu(V))^{r-s-1},$$

II. If $s=0$ then all returns are short, i.e. $v_{j+1}-v_j<\delta$ for all j. This implies $|K_r^0|\le\delta^{r-1}t/\mu(W)$ and (using Lemma 4 with $W_1=W_2=\cdots=W_{r-1}=V$ and $W_r=W$)

$$\mu\left(\bigcap_{i=1}^{r+1}T^{-v_i}W\right)\le\beta^{r-1}\mu(V)^{r-1}\mu(W),$$

$\vec{v}\in\ K_r^0$.

III. Summing over s yields

$$\begin{aligned}\sum_{\vec{v}\in K_r}\mu(C_{\vec{v}}) &= \sum_{s=0}^{r-2}\sum_{\vec{v}\in K_r^s}\mu(C_{\vec{v}})\\ &\le \frac{1}{\beta\delta}\sum_{s=0}^{r-2}\binom{r}{s}\frac{(\beta t)^{s+1}}{(s+1)!}(\beta\delta\mu(V))^{r-s-1}\\ &\le C_2 t\mu(V)\sum_{s=0}^{r-2}\binom{r-2}{s}\frac{(\beta t)^s}{s!}(\beta\delta\mu(V))^{r-s},\end{aligned}$$

with some C_2 and a slightly larger β to absorb a factor $r(r-1)$, which comes from the inequality

$$\binom{r}{s}\le r\binom{r-1}{s}\le r(r-1)\binom{r-2}{s}$$

for $s\le r-2$. This concludes the proof of the first statement.

(ii) The second inequality is proven is the same way with the obvious modifications due to the first component of the hitting vector $\vec{v}$. We split $\tilde{K}_r$ into a disjoint union of sets $\tilde{K}_r^s$, $s=0,\dots,r-1$, each of which has exactly s 'long' intervales (i.e. $\ge\delta$) and $r-s$ short intervals. For $s=0,\dots,r-1$:

$$|\tilde{K}_r^s|\le\binom{r+1}{s}\frac{\delta^{r-s}}{(s+1)!}\left(\frac{t}{\mu(W)}\right)^s,$$

and $\mu(C_{\vec{v}})\le\beta^r\mu(V)^{r-s}\mu(W)^{s+1}$, for $\vec{v}\in\tilde{K}_r^s$. As in part (i) this then yields the estimate for return times. □

Denote by

$$I_r(N)=\{\vec{v}\in R_r(N):\min(v_{j+1}-v_j)<\delta'\}$$

($\delta'>0$) the portion of very short returns within the rare set.

Lemma 8. *Let W be a measurable set in Ω. Then*

$$\left.\begin{array}{l}|R_r|\mu(W)^r\\ |\tilde{R}_r|\mu(W)^r\end{array}\right\}\le\delta\frac{\mu(W)t^{r-1}}{(r-2)!}$$

for every r.

Proof. For every vector $\vec{v}$ in R_r note that the shortest return time $\min(v_{j+1}-v_j)$ is at most δ, the position of the 'shortest' return time has $r-1$ possibilities and the remaining $r-1$ hitting times have at most $\frac{1}{(r-1)!}\left(t/\mu(W)\right)^{r-1}$ many arrangements. This leaves us with the upper bound

$$|R_r|\le\delta(r-1)\frac{1}{(r-1)!}\left(\frac{t}{\mu(W)}\right)^{r-1}.$$

The bound on the cardinality of $\tilde{R}_r$ is proven in the same way. □

Remark. Let us note that the term $\delta\mu(W)t^{r-1}/(r-2)!$ is bounded by the highest order term ($s=r-1$) in the expression $\delta\mu(W)\sum_{s=0}^r\delta\mu(W)^{r-s}t^s/s!$ which occurs in formula (4) of Theorem 1.

4.2. **Entry and return times for (ϕ, f)-mixing maps.** In the following t will always be a positive parameter and we shall denote by χ_U the characteristic function of a set U. Let A_n be an n-cylinder and define the 0,1-valued random variable $\eta_v^n = \chi_{A_n} \circ T^v$ for $v = 0, 1, \ldots, N$, where $N = [t/\mu(A_n)]$ (unless we say otherwise). In the context of studying the distribution of entry times we shall use the values $b_v^n = \mu(\eta_v^n)$ in the following Proposition 9. For $\vec{v} \in G_r(N)$ ($\vec{v} = (v_1, v_2, \ldots, v_r)$) we put

$$\eta_{\vec{v}}^n = \eta_{v_1}^n \eta_{v_2}^n \cdots \eta_{v_r}^n = (\chi_{A_n} \circ T^{v_1})(\chi_{A_n} \circ T^{v_2}) \cdots (\chi_{A_n} \circ T^{v_r})$$

for the characteristic function of $C_{\vec{v}} = \bigcap_{j=1}^r T^{-v_j} A_n$ and define the values

$$b_{\vec{v}}^n = \mu(C_{\vec{v}}).$$

For a given non-decreasing sequences of integers $\delta'_n \leq \delta_n, n = 1, 2, \ldots,$ we define the rare set $R_r(N)$ as the disjoint union of $K_r(N)$ and $I_r(N)$ where

$$\begin{aligned} K_r(N) &= \left\{ \vec{v} \in G_r(N) : \delta'_n \leq \min_j (v_{j+1} - v_j) < \delta_n \right\} \\ I_r(N) &= \left\{ \vec{v} \in G_r(N) : \min_j (v_{j+1} - v_j) < \delta'_n \right\} \end{aligned}$$

Notice that in the following Proposition 9 and 10 in the third inequality the sum is taken only over K_r, the principal part of the rare set. In Proposition 12 however we consider the full rare set. Later on we shall use Corollary 16 (which uses Proposition 12) to get bounds on the set of very short returns I_r.

Proposition 9. *Let μ be a (ϕ, f)-mixing probability measure.*

Then there exists a constant C_3 so that for every cylinder $A_n \in \mathcal{A}^n$ for which $f(A_n) \leq \delta_n - n$ and $t > 0$ one has

$$\begin{aligned} \max_{1 \leq v \leq N} b_v^n &= \mu(A_n) \\ \left| \sum_{v=1}^N b_v^n - t \right| &\leq \mu(A_n) \\ \sum_{\vec{v} \in K_r} b_{\vec{v}}^n &\leq C_3 \delta_n \mu(V_n) \sum_{s=0}^r (3\delta_n \mu(V_n))^{r-s} \frac{(3t)^s}{s!} \\ \sum_{\vec{v} \in R_r} b_{v_1}^n \cdots b_{v_r}^n &\leq \delta_n \frac{\mu(A_n) t^{r-1}}{(r-2)!} \\ \left| \frac{b_{v_1}^n \cdots b_{v_r}^n}{b_{\vec{v}}^n} - 1 \right| &\leq C_3 r \phi(\delta_n), \end{aligned}$$

where V_n a union of δ''-cylinders such that $A_n \subset V_n$ and $f(V_n) \leq \delta'_n - \delta''$.

We omit the proof which consists in checking the conditions using Lemma 7 (ii) and Lemma 8. This is technically more straightforward than the proof for Proposition 10 which we will show in detail below.

Proposition 9 has the following companion which will be used to get results on the distribution of return times and their error terms.

As before let μ be a T-invariant probability measure on Ω. For an n-cylinder A_n we then define the restricted probability measure μ_n on A_n by $\mu_n(B) = \mu(B \cap A_n)/\mu(A_n)$ for measurable B.

With t a positive parameter and $N = [t/\mu(A_n)]$ we define (for every n) the $0,1$-valued random variable $\eta_v^n = \chi_{A_n} \circ T^v$ and consider now the values

$$\begin{aligned} b_v^n &= \mu_n(\eta_v^n), \\ b_{\vec{v}}^n &= \mu_n(\eta_{\vec{v}}^n), \end{aligned}$$

where $\vec{v} \in G_r(N)$ and, as above, $\eta_{\vec{v}}^n$ is the characteristic function of $C_{\vec{v}} = \bigcap_{j=1}^r T^{-v_j} A_n$. For a given non-decreasing sequences of integers $\delta'_n \leq \delta_n$ we define the set $\tilde{K}_r$ of short (but not too short) returns by

$$\tilde{K}_r(N) = \left\{ \vec{v} \in G_{r+1}(N) : \delta'_n \leq \min\left(\min_j (v_{j+1} - v_j), v_1 \right) < \delta_n \right\}.$$

Proposition 10. *Let μ be a (ϕ, f)-mixing probability measure where $\phi(v)$ is summable.*

Then there exists a constant C_4 so that for every cylinder $A_n \in \mathcal{A}^n$ for which $f(A_n) \leq \delta_n - n$ and $t > 0$ one has:

$$\begin{aligned} \max_{\delta'_n \leq v \leq N} b_v^n &\leq C_4 \mu(V_n) \\ \left| \sum_{v=\delta'_n}^{N} b_v^n - t \right| &\leq C_4 (f(A_n) + n) \mu(V_n) \\ \sum_{\vec{v} \in \tilde{K}_r} b_{\vec{v}}^n &\leq C_4 \mu(A_n) \sum_{s=0}^{r-1} \binom{r-1}{s} (3\delta_n \mu(V_n))^{r-s} \frac{(3t)^s}{s!} \\ \sum_{\vec{v} \in \tilde{K}_r} b_{v_1}^n \cdots b_{v_r}^n &\leq C_4 \delta_n \mu(V_n) \sum_{k=0}^{r-1} (\delta_n C_4 \mu(V_n))^{r-1-k} \frac{t^k}{k!} \\ \left| \frac{b_{v_1}^n \cdots b_{v_r}^n}{b_{\vec{v}}^n} - 1 \right| &\leq 2^r \phi(\delta_n) \ \ \forall\, \vec{v} \notin \tilde{R}_r \end{aligned}$$

where V_n a union of δ''-cylinders such that $A_n \subset V_n$ and $f(V_n) \leq \delta'_n - \delta''$.

Proof. (i) To estimate

$$b_v^n = \mu_n(T^{-v} A_n) = \frac{\mu(A_n \cap T^{-v} A_n)}{\mu(A_n)},$$

we consider two cases: (a) $v \geq f(A_n) + n$ and (b) $\delta'_n \leq v < f(A_n) + n$. In the first case, $v \geq f(A_n) + n$, we use the (ϕ, f)-mixing property according to which

$$|\mu(A_n \cap T^{-v} A_n) - \mu(A_n)^2| \leq \phi(v - n) \mu(A_n)^2$$

and consequently

$$|b_v^n - \mu(A_n)| \leq \phi(v - n) \mu(A_n). \tag{8}$$

Hence ($c_1 > 0$):

$$b_v^n = \mu_n(T^{-v} A_n) \leq \mu(A_n)(1 + \phi(v - n)) \leq c_1 \mu(A_n).$$

In the second case, $\delta'_n \leq v < f(A_n) + n$, we use the set V_n chosen according to the hypothesis ($A_n \subset V_n$) and conclude in a similar way that

$$|b_v^n - \mu(V_n)| \leq \phi(v - \delta'') \mu(V_n).$$

and therefore $b_v^n \le c_1\mu(V_n)$.
(ii) Summability of the function ϕ gives us the second inequality:

$$\begin{aligned}\left|\sum_{v=\delta'_n}^{N} b_v^n - t\right| &\le (f(A_n)+n)(c_1\mu(V_n)+\mu(A_n)) + \mu(A_n)\sum_{v=f(A_n)+n}^{N}\phi(v-n)\\ &\le (1+c_1)(f(A_n)+n)\mu(V_n)+\mu(A_n)\sum_{v=0}^{\infty}\phi(v)\\ &\le c_2(f(A_n)+n)\mu(V_n).\end{aligned}$$

(iii) To obtain the third inequality we apply Lemma 7 (ii) with the parameters $\delta'=\delta'_n$, $\delta=\delta_n$, $W=A_n$, $V=V_n$ and $\tilde{K}_r$ as defined above:

$$\sum_{\vec{v}\in\tilde{K}_r} b_{\vec{v}}^n \le C_2\mu(A_n)\sum_{s=0}^{r-1}\binom{r-1}{s}(3\delta_n\mu(V_n))^{r-s}\frac{(3t)^s}{s!}$$

for all large enough n so that $\beta = 1+\phi(\delta_n - n)\le 3/2$, where V_n is as in hypothesis.
(iv) If $v_j \ge \delta_n$ then

$$b_{v_j}^n \le (1+\phi(v_j-n))\mu(A_n) \le (1+\phi(v_1-n))\mu(A_n)\le c_3\mu(A_n),$$

and otherwise ($\delta'_n v_j < \delta_n$) we use the estimate $b_{v_j}^n \le c_1\mu(V_n)$ from part (i). If the first s of the entries of $\vec{v}$ are less that $f(A_n)+n$ then we obtain similarly to Lemma 8 for $s\ge 1$:

$$\begin{aligned}\sum_{\vec{v}\in\tilde{K}_r;\, v_1,\dots,v_s<\delta_n} b_{v_1}^n b_{v_2}^n\cdots b_{v_r}^n &\le \delta_n^s c_1^s\mu(V_n)^s\mu(A_n)^{r-s}|G_{r-s}|\\ &\le (\delta_n c_1\mu(V_n))^s\frac{t^{r-s}}{(r-s)!}.\end{aligned}$$

If $s=0$ (no entry of $\vec{v}$ is less than $f(A_n)+n$) then

$$\sum_{\vec{v}\in R_r; v_1\ge\delta_n} b_{v_1}^n\cdots b_{v_r}^n \le |R_r|\mu(U_n)^r \le \delta_n\frac{\mu(U_n)t^{r-1}}{(r-2)!}.$$

Summing over $s=0,\dots,r$ yields (where $k=r-s$)

$$\begin{aligned}\sum_{\vec{v}\in\tilde{K}_r} b_{v_1}^n b_{v_2}^n\cdots b_{v_r}^n &\le \sum_{s=1}^{r}(\delta_n c_1\mu(V_n))^s\frac{t^{r-s}}{(r-s)!}+\delta_n\frac{\mu(U_n)t^{r-1}}{(r-2)!}\\ &\le 2\delta_n c_1\mu(V_n)\sum_{k=0}^{r-1}(\delta_n c_1\mu(V_n))^{r-1-k}\frac{t^k}{k!}.\end{aligned}$$

(v) To verify the last of the inequalities we restrict to $\vec{v}\notin\tilde{R}_r$, that is $v_{j+1}-v_j\ge\delta_n\ge f(A_n)+n$ for all j and $v_1\ge\delta_n$. Thus

$$b_{\vec{v}}^n = \mu_n(C_{\vec{v}}) = \frac{\mu(A_n\cap C_{\vec{v}})}{\mu(A_n)},$$

and by Lemma 4 we get

$$\begin{aligned}|\mu(A_n\cap C_{\vec{v}})-\mu(A_n)^{r+1}| &\le \left((1+\phi(\delta_n-n))^r-1\right)\mu(A_n)^{r+1}\\ &\le rc_4\phi(\delta_n-n)\mu(A_n)^{r+1},\end{aligned}$$

(for some $c_4 > 0$) and

$$|b^n_{\vec{v}} - \mu(A_n)^r| \le rc_4\phi(\delta_n - n)\mu(A_n)^r.$$

In order to compare $b^n_{\vec{v}}$ to the product $b^n_{v_1} \cdots b^n_{v_r}$ let us note that by equation (8) one has for $j = 1, 2, \ldots, r$:

$$|b^n_{v_j} - \mu(A_n)| \le \phi(v_j - n)\mu(A_n) \le \phi(v_1 - n)\mu(A_n),$$

and in particular $b^n_{v_j} \le c_3\mu(A_n)$. Thus

$$\begin{aligned}
|b^n_{v_1} \cdots b^n_{v_r} - \mu(A_n)^r| &\le r\left(\max_j \left|b^n_{v_j} - \mu(A_n)\right|\right)\left(\max\left(b^n_{v_1}, \ldots, b^n_{v_r}, \mu(A_n)\right)\right)^{r-1} \\
&\le r\phi(v_1 - n)c_3^{r-1}\mu(A_n)^r \\
&\le rc_3^r\phi(\delta_n - n)\mu(A_n)^r,
\end{aligned}$$

for all large enough n. By the triangle inequality

$$|b^n_{\vec{v}} - b^n_{v_1} \cdots b^n_{v_r}| \le r(c_4 + c_3^r)\phi(\delta_n - n)\mu(A_n)^r,$$

and therefore, with a slightly larger value for c_3,

$$\left|\frac{b^n_{\vec{v}}}{b^n_{v_1} \cdots b^n_{v_r}} - 1\right| \le c_3^r\phi(\delta_n).$$

Let us note that since we only consider large enough n, the number $c_3 > 1$ can be chosen arbitrarily close to 1. In particular we can assume that $c_3 < 2$.

The proof is finished if we put $C_4 = \max(1, 2c_1, c_2, 3, C_2)$. □

4.3. **Restricted entry and return times for (ϕ, f)-mixing maps.** Near periodic orbits, the Poisson statistics doesn't apply to the return times, because of the presence of persistently short returns. Pitskel [21], Hirata [14] and Abadi [1] all have given examples that show that the near periodic points the first return time is not exponentially distributed. The following results will provide us with the asymptotics of long returns to the neighbourhoods of *any* point and in particular also with the asymptotics of the first return time for periodic points. This is achieved by deleting points that return 'too soon' (of the order of the length of the cylinder) and adjusting the normalising factor for the time-interval accordingly.

We will set up the functions $\hat{\eta}^n_v$ to only counts returns when the return interval is at least of length n and to ignore all shorter ones. Let $A_n \in \mathcal{A}^n$ be an arbitrary cylinder of length n, define

$$U_n = (T^{-n}A_n) \setminus \bigcup_{j=1}^{n-1} T^{-(n-j)}A_n$$

and put $\hat{N} = [t/\mu(U_n)]$ (this is a 'non-standard' rescaling). In this way we achieve that $\tau(U_n) \ge n$. We next define the functions $\hat{\eta}^n_v$ by

$$\begin{aligned}
\hat{\eta}^n_v &= (\chi_{A_n} \circ T^v)\prod_{j=1}^{v}(1 - \chi_{A_n} \circ T^{v-j}) \qquad v = 1, 2, \ldots, n-1, \\
\hat{\eta}^n_v &= (\chi_{A_n} \circ T^v)\prod_{j=1}^{n-1}(1 - \chi_{A_n} \circ T^{v-j}) \qquad v = n, n+1, \ldots, \hat{N}
\end{aligned}$$

Note that $\hat{\eta}^n_v = \chi_{U_n} \circ T^{v-n}$ for $n \le v < \hat{N}$.

Proposition 9 has the following companion which will be used to get results on the distribution of return times and their error terms. We omit the proof which is similar.

Proposition 11. *Let μ be a (ϕ, f)-mixing probability measure.*

Then there exists a constant C_5 so that for every cylinder $A_n \in \mathcal{A}^n$ for which $f(A_n) \leq \delta_n - n$ and $t > 0$ one has

$$
\begin{aligned}
\max_{1\leq v\leq \hat{N}} \hat{b}^n_v &\leq \mu(A_n)\\
\left|\sum_{v=1}^{\hat{N}} \hat{b}^n_v - t\right| &\leq C_5 n\mu(A_n)\\
\sum_{\vec{v}\in R_r} \hat{b}^n_{\vec{v}} &\leq C_5\delta_n\mu(V_n)\sum_{s=0}^{r}(3\delta_n\mu(V_n))^{r-s}\frac{(3t)^s}{s!}\\
\sum_{\vec{v}\in R_r} \hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r} &\leq C_5\delta_n\frac{\mu(A_n)t^{r-1}}{(r-2)!}\\
\left|\frac{\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r}}{\hat{b}^n_{\vec{v}}} - 1\right| &\leq C_5 r\phi(\delta_n),
\end{aligned}
$$

where V_n a union of δ''-cylinders such that $A_n \subset V_n$ and $f(V_n) \leq n - \delta''$.

As before we define the restricted probability measure μ_n on A_n by $\mu_n(B) = \mu(B\cap A_n)/\mu(A_n)$ (for B measurable).

In the following proposition, which is the analog of Proposition 10 for the restricted returns on an adjusted time-interval, we use the values $\hat{b}^n_v = \mu_n(\hat{\eta}^n_v)$ and $\hat{b}^n_{\vec{v}} = \mu_n(\hat{\eta}_{\vec{v}})$, where $\hat{\eta}_{\vec{v}} = \hat{\eta}_{v_1}\cdots\hat{\eta}_{v_r}$ for $\vec{v} \in G_r(\hat{N})$. The rare set is as above with the obvious modification of replacing N by $\hat{N}$. We omit the proof of the following propositon since it mirrors the proof of Proposition 10.

Proposition 12. *Let μ be a (ϕ, f)-mixing probability measure where $\phi(v)$ is summable.*

Then there exists a constant C_6 so that for every $A_n \in \mathcal{A}^n$ for which $f(A_n) \leq \delta_n - 2n$ and $t > 0$:

$$
\begin{aligned}
\max_{1\leq v\leq \hat{N}} \hat{b}^n_v &\leq C_6\mu(V_n)\\
\left|\sum_{v=1}^{\hat{N}} \hat{b}^n_v - t\right| &\leq C_6(f(A_n)+n)\mu(V_n)\\
\sum_{\vec{v}\in R_r} \hat{b}^n_{\vec{v}} &\leq C_6\delta_n\mu(A_n)\sum_{s=0}^{r+1}(3\delta_n\mu(A_n))^{r+1-s}\frac{(3t)^s}{s!}\\
\sum_{\vec{v}\in R_r} \hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r} &\leq C_6\delta_n\mu(V_n)\sum_{k=0}^{r-1}(\delta_n C_6\mu(V_n))^{r-1-k}\frac{t^k}{k!}\\
\left|\frac{\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r}}{\hat{b}^n_{\vec{v}}} - 1\right| &\leq 2^r\phi(\delta_n),
\end{aligned}
$$

where V_n a union of δ''-cylinders such that $A_n \subset V_n$ and $f(V_n) \leq n - \delta''$.

Remark 1. In Proposition 12 the (ϕ, f)-mixing requirement can be weakened. It is sufficient that μ is (ϕ, f)-mixing on A_n and V_n:

$$\left|\mu(U \cap T^{-m-n}Q) - \mu(U)\mu(Q)\right| \leq \phi(m)\mu(U)\mu(Q)$$

for all measurable Q and $m \geq f(U)$, where $U = A_n, V_n$.

Remark 2. In the special case when $f = 0$ then $V_n = A_n$.

5. **Statistics of ϕ-mixing maps.** In this section we discuss classical ϕ-mixing maps. An invariant probability measure μ for the map T is called ϕ-mixing if it is (ϕ, f)-mixing for a (given) partition $\mathcal{A}$ where f is the constant 0. In various settings [11, 18] it has been shown that the measure of n-cylinders fall off geometrically, i.e. there is a constant $c_1 > 0$ so that $\mu(A) \leq e^{-nc_1}$ for all n and $A \in \mathcal{A}^n$. Since the rate of convergence of the entry and return times to the Poisson distribution depends on the decay rate of ϕ we shall prove in the first section some more general statement.

In the following A_n denotes an n-cylinder set and χ_{A_n} its characteristic function. Let μ be an invariant probability measure. For a given positive parameter value t we then define the counting function

$$\zeta_n = \sum_{k=1}^{N} \chi_{A_n} \circ T^k$$

whose value is the number of times a point hits the set A_n on the time interval $[1, N]$, where $N = [t/\mu(A_n)]$. If we denote by

$$\mathcal{N}_n^r = \{x \in \Omega : \zeta_n(x) = r\}$$

the levelset of the counting function ζ_n, then $\mu(\mathcal{N}_n^r)$ is the probability that a randomly chosen point hits A_n exactly r times on the time interval $[1, N]$. Of particular interest is when $r = 0$, in this case $\mathcal{N}_n^0 = \{x \in \Omega : \tau_{A_n}(x) > t/\mu(A_n)\}$.

We will examine two types of ϕ-mixing systems, namely those in which ϕ decays polynomially and equilibrium states on Axiom A systems for Hölder continuous functions which are ϕ mixing where ϕ decays exponentially fast. We say the measure μ is polynomially ϕ-mixing with power $p > 0$ if $\limsup_{v\to\infty} v^p\phi(v) < \infty$.

Strongly hyperbolic maps that satisfy Axiom A and have very regular behaviour (as the shadowing property and finite Markov partitions of arbitrarily small diameter) mix exponentially fast (i.e. ϕ decays exponentially). Such systems are usually studied using a symbolic description by a subshift of finite type. A good reference is the classical book by Bowen [7]. We shall study the entry and return time distribution for equilibrium states for Hölder continuous potentials.

Let $W \subset \Omega$ and define the *return time function*

$$\tau_W(x) = \min\{k \geq 1 : T^k x \in W\}.$$

τ_W measures the first entry time for points outside W and (for the first return time for points in W). This function is finite almost everywhere with respect to ergodic measures and satisfies by a theorem of Kac the identity $\int_W \tau_W(x)\, d\mu(x) = 1$ for any ergodic probability measure μ and measurable W. Let us define the *shortest return time function*:

$$\tau(A) = \min_{x\in A} \tau_A(x)$$

which measures the shortest return time within the set A (see [15, 16]). By definition $A \cap T^{-k}A = \emptyset$ for $k = 1, 2, \ldots, \tau(A) - 1$.

5.1. **Polynomially ϕ-mixing maps.** We shall prove limiting results for the entry time and return times to cylinder set. If μ is a T-invariant probability on Ω, then its restriction to an n-cylinder A_n is given by $\mu_n(B) = \mu(B \cap A_n)/\mu(A_n)$ (for all measurable B).

Theorem 13. *Let μ be a ϕ-mixing probability measure for the transformation $T : \Omega \to \Omega$ so that $\limsup_{v\to\infty} \phi(v)v^p < \infty$ for some positive p.*

Then there exists constants C_7, C_7' so that for all $A_n \in \mathcal{A}^n$ and all t, r for which $\frac{(r+t+1)^2}{t+1} n\mu(A_m)^{\frac{p}{p+1}} \le C_7'$ one has:

$$\left|\mu(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le n^{q^*} C_7 \mu(A_m)^{\frac{p}{1+p}} \begin{cases} \frac{(r+t)^2}{r!}(4t+1)^{r-1}e^{4t} & \text{if} \quad r \ge 1 \\ e^{4t}(t+1) & \text{if} \quad r = 0 \end{cases},$$

where m is such that $m \le \tau(A_m)$ and $A_n \subset A_m \in \mathcal{A}^m$ and where
(i) $q^* = 0$ if $\mu_* = \mu$ for the distribution of entry times,
(ii) $q^* = 1$ if $\mu_* = \mu_n$ for the distribution of return times where we also require that $\sum_v \phi(v) < \infty$.

Proof. We want to verify the conditions of Theorem 1 using Proposition 9. Notice that $A_m \cap T^{-j}A_m = \emptyset$ for $j = 1, \ldots, m-1$, and if we put $\delta_n' = m$ then the set V_n as defined in the hypothesis of Proposition 9 is equal to A_m as $f = 0$.

Assume that ϕ decays polynomially with power p, i.e. $\phi(v) \le c_1 v^{-p}$ for some c_1, and put $\delta_n = \mu(A_m)^{-\frac{1}{1+p}}$. Then

$$\delta_n \mu(A_m) \le \mu(A_m)^{\frac{p}{1+p}}$$

and

$$\phi(\delta_n) \le \frac{c_1}{\delta_n^p} \le c_1 \mu(A_m)^{\frac{p}{1+p}}.$$

(i) Entry times: With $\varepsilon_n = C_3 c_2 \mu(A_m)^{\frac{p}{1+p}}$ ($c_2 = \max(7, c_1)$) Proposition 9 ensures that the conditions (1)–(5) of Theorem 1 are satisfied with $\alpha = 3$. Then we choose $\alpha' = 4$.

(ii) Return times: If we put $\varepsilon_n = C_4 n c_2 \mu(A_m)^{\frac{p}{1+p}}$ then Proposition 10 ensures that the conditions (1)–(5) of Theorem 1 are satisfied with $\alpha = 3$ ($\alpha' = 4$).

Put $C_7 = C_1 c_2 \max(C_3, C_4)$ and $C_7' = 1/12c_2$, where we use equation (7) and $\varepsilon_n = nc_2\mu(A_m)^{\frac{p}{1+p}}$ to get the constraints on t and r. □

Remarks. (I) The smallness of $\frac{r^2}{t}\mu(A_n)^{\frac{p}{1+p}}$ (and $t\mu(A_n)^{\frac{p}{1+p}}$ if $r = 0$) in the theorem is given by equation (7) for the appropriate ε_n.
(II) Theorem 13 covers the special case when ϕ is summable (see in particular [11]), i.e. $\sum_v \phi(v) < \infty$, which implies that $\phi(v) \le c_1/v$ (for $v > 0$). Lemma 13 thus can be applied to the case when $p = 1$ ($p^* = 1/2$) and gives us the following error terms:

$$\left|\mu(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_7 \mu(A_m)^{1/2} \begin{cases} \frac{(r+t)^2}{r!}(4t)^{r-1}e^{4t} & \text{if} \quad r \ge 1 \\ e^{4t}(t+1) & \text{if} \quad r = 0 \end{cases}.$$

□

5.2. **Mappings that satisfy Axiom A.** In the following we are looking at strongly hyperbolic maps that satisfy Axiom A.

Theorem 14. *Let $T:\Omega\to\Omega$ be a topological mixing Axiom A map on the basic set Ω and μ the (invariant) equilibrium state for a Hölder continuous (with respect to the sequence topology inherited from the Markov partition) potential f.*

Then there exists constants C_8, C_8' so that for all $A_n\in\mathcal{A}^n$ ($\mathcal{A}$ finite) and all t,r for which $\frac{(r+t+1)^2}{t+1}n^2\mu(A_m)\le C_8'$ one has:

$$\left|\mu_*(\mathcal{N}_n^r)-\frac{t^r}{r!}e^{-t}\right|\le C_8 n^{q^*+1}\mu(A_m)\begin{cases}\frac{(r+t)^2}{r!}(4t+1)^{r-1}e^{4t} & \text{if}\quad r\ge 1\\ e^{4t}(t+1) & \text{if}\quad r=0\end{cases},$$

where m is such that $m\le\tau(A_m)$, $A_n\subset A_m\in\mathcal{A}^m$, and
(i) $q^*=0$, $\mu_*=\mu$ for entry times;
(ii) $q^*=1$, $\mu_*=\mu_n$ for re-entry times.

Proof. We proceed as in the proof of Theorem 13 and use the fact that Axiom A maps are ϕ-mixing where $\phi(k)=c_1\vartheta^k$ for some positive $\vartheta<1$ and a constant c_1. By the Gibbs property [7] of μ there exists constants $c_2,c_3>0$ ($c_2<h_{top}$) so that $\mu(A_m)\ge c_3e^{-mc_2}\ge c_3e^{-nc_2}$ for all m and n. Put $\delta_n=qn$, where $q=1+\frac{c_2}{|\log\vartheta|}+\log\frac{c_3}{c_1}$. Then (as $V_n=A_m$)

$$\delta_n\mu(A_m)\le qn\mu(A_m)$$

and

$$\phi(\delta_n-n)\le\vartheta^{(q-1)n}\le c_3e^{-c_2n}\le\mu(A_m)$$

for all large enough n.

(i) Entry times: If we choose $\varepsilon_n=3C_3qn\mu(A_m)$ then the conditions of Theorem 1 ($\alpha'=4$) are satisfied by Proposition 9 for $\alpha=3$. This proves the first statement of the theorem.

(ii) Return times: With the choice $\varepsilon_n=3C_4qn^2\mu(A_m)$ the conditions of Theorem 1 ($\alpha'=4$) are satisfied by Proposition 10 with $\alpha=3$.

Now put $C_8=3C_1q\max(C_3,C_4)$. The constraints on t and r are given by equation (7) and $\varepsilon_n=3qn^2\mu(A_m)$, which imply that $C_8'=1/36q$. ☐

The same asymptotics and similar error terms are valid for any ϕ-mixing measure for which ϕ is exponentially fast decreasing. In the case of an Axiom A system, the Gibbs property was used to get an exponential lower for the measure of cylinder. Systems that are not Markov will in general not have this property.

5.3. **The distribution of restricted entry and return times.** The first result we prove is on the distribution and error terms for the restricted return times.

For an n-cylinder A_n let the counting functions $\hat\eta_v^n$, $v=0,1,\dots,N$ be defined as in Proposition 12, where $\hat N=[t/\mu(U_n)]$, $U_n=(T^{-n}A_n)\setminus\bigcup_{j=1}^{n-1}T^{-(n-j)}A_n$ and t is a positive parameter. Since $f=0$ we have in the setting of Proposition 12 that $V_n=A_n$. Consider the restricted counting function $\hat\zeta_n=\sum_{v=1}^{\hat N}\hat\eta_v^n$ and its r-levelsets $\hat{\mathcal{N}}_n^r=\{x\in\Omega:\hat\zeta_n(x)=r\}$.

Theorem 15. *Let μ be a probability measure on Ω which is ϕ-mixing and invariant with respect to a map T and a partition $\mathcal{A}$. Assume that ϕ is summable.*

Then there exists a constant C_9, C_9' so that for all $A_n\in\mathcal{A}^n$ and all t,r for which $\frac{(r+t+1)^2}{t+1}n^{q^*+1}\mu(A_m)^{p^*}\le C_9'$ one has ($\mu_*=\mu$ for entry times and $\mu_*=\mu_n$ for *return times):*

$$\left|\mu_*(\hat{\mathcal{N}}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_9 n^{q^*+1}\mu(A_n)^{p^*} \left\{ \begin{array}{lll} \frac{(r+t)^2}{r!}(4t)^{r-1}e^{4t} & \textit{if} & r \ge 1 \\ e^{4t}(t+1) & \textit{if} & r = 0 \end{array} \right. ,$$

where

(i) $(q^*, p^*) = (0, \frac{p}{p+1})$ *if* ϕ *decays polynomially with power* p,

(ii) $(q^*, p^*) = (1, 1)$ *for an Axiom A system and a Hölder potential.*

Proof. Let us first note that $V_n = A_n$ since $f = 0$ and $\tau(U_n) \ge n$.

(i) If ϕ decays polynomially with power p we put $\delta_n = \mu(A_n)^{-\frac{1}{1+p}}$ and therefore obtain $\delta_n\mu(A_n) \le \mu(A_n)^{\frac{p}{1+p}}$ and $\phi(\delta_n) \le c_1\mu(A_n)^{\frac{p}{1+p}}$ for some c_1. With $\varepsilon_n = nC_5c_2\mu(A_n)^{p^*}$ $(q^*+1 = 1$, $c_2 = \max(7, c_1))$ and $\alpha = 3$, Proposition 12 implies that the conditions (1)–(5) of Theorem 1 are met. Put $C_9 = C_1c_2C_5$ and $C_9' = 1/12c_2$.

(ii) If μ is an equilibrium state on an Axiom A system for a Hölder continuous potential, then $\phi(k) = c_3\vartheta^k$ $(0 < \vartheta < 1)$ and by the Gibbs property [7] $\mu(A_n) \ge c_5e^{-nc_4}$ $(c_4, c_5 > 0)$ for all n. With $\delta_n = qn$, where $q = 1 + \frac{c_4}{|\log\vartheta|} + \log\frac{c_5}{c_3}$ we obtain (as $V_n = A_n$) that $\delta_n\mu(A_n) \le qn\mu(A_n)$ and $\phi(\delta_n - n) \le \mu(A_n)$ for all n. With $\varepsilon_n = 3C_6qn^2\mu(A_n), \alpha = 3$ the conditions of Theorem 1 are satisfied by Proposition 12. Put $C_9 = C_1c_2C_6$ and $C_9' = 1/36q$. □

Let us now look at the distribution of the first return time τ_{A_n} which is the case $r = 0$. We obtain the following result in which the numbers q^* and p^* are as in Theorem 15.

Corollary 16. *Let* μ *be a probability measure on* Ω *which is* ϕ*-mixing and invariant with respect to a map* T *and a partition* $\mathcal{A}$*. Assume that* ϕ *is summable.*

Then there exists a constant C_{10} *so that for all* $A_n \in \mathcal{A}^n$ *and* $t \ge n\mu(U_n)$ *for which* $n^{q^*+1}\mu(A_m)^{p^*}(t+1) \le C_9'$*:*

$$\left|\mu_n\left(\left\{x \in A_n : \tau_{U_n}(x) \ge \frac{t}{\mu(U_n)}\right\}\right) - e^{-t}\right| \le C_{10}(t+1)e^{4t}n^{q^*+1}\mu(A_n)^{p^*}.$$

Remark. In the case $(q^*, p^*) = (1, 1)$ the same asymptotics and similar error terms are valid for any ϕ-mixing measure for which ϕ is exponentially fast decreasing and where the measure satisfies a Gibbs property (which applies to Axiom A systems). Systems that are not Markov will in general not have this property as for instance the piecewise expanding maps we consider in section 6.

5.4. **Convergence in measure for entry and return times for ϕ-mixing maps.** For $x \in \Omega$ let us denote by $A_n(x)$ the (not necessarily unique) n-cylinder that contains the point x. Since by Lemma 6 μ has positive entropy we have by [23][2]

$$\liminf_{n\to\infty} \frac{\tau(A_n(x))}{n} \ge 1$$

μ-almost everywhere for every ergodic T-invariant probability measure μ. In other words, let $\varepsilon > 0$ then for almost every point $x \in \Omega$ there exists finite number $N_\varepsilon(x)$ so that $\tau(A_n(x)) \ge (1-\varepsilon)n \;\forall\, n \ge N_\varepsilon(x)$. Therefore, if we put

$$\mathcal{J}_{n,\varepsilon} = \{x \in \Omega : \tau(A_n(x)) \ge (1-\varepsilon)n\},$$

[2]See also [5] for another proof of this fact. It is interesting to remark that the *shortest return time function* also has been used to define a dimension-like characteristic for a wide class of invariant sets, see for example [4, 19, 6, 12, 17].

then $\mu(\mathcal{J}_{n,\varepsilon}^c) \to 0$ as $n \to \infty$ for every positive ε. Let us recall (Lemma 5) that for ϕ-mixing maps (be they Axiom A or be it that ϕ decays polynomially) the measure of cylinders decays exponentially. We thus immediately obtain the following result in which we can choose any $\varepsilon \in (0,1)$.

Corollary 17. *Let μ be a ϕ-mixing probability measure.*

Then there exist $\sigma < 1$, C_{11}, C'_{11} and a sequence of sets $\mathcal{J}_n \subset \Omega$ for which $\mu(\mathcal{J}_n) \to 1$ so that for all $x \in \mathcal{J}_n$ and all t, r for which $\frac{(r+t+1)^2}{t+1} n^2\sigma^n \le C'_{11}$ one has

$$\left|\mu_*(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_{11}\sigma^n \begin{cases} \frac{(r+t)^2}{r!}(4t)^{r-1}e^{4t} & \text{if } \ r \ge 1 \\ e^{4t}(t+1) & \text{if } \ r = 0 \end{cases},$$

where μ_* is either μ or the measure μ_n restricted to $A_n(x)$.

The value of $\sigma < 1$ is so that $n^2\mu(A_{[(1-\varepsilon)n]}(x))^{p^*} \le \text{const.}\sigma^n$ for all n.

Remarks: (I) Good distribution results always require that one avoids short return times. For instance Abadi [1] uses a similar setting in which he puts $\varepsilon = \frac{1}{2}$.

(II) Also note that if in Corollary 17 we consider the restricted return times (i.e. use $\hat{\mathcal{N}}_n^r$ instead of $\mathcal{N}_n^r$) then the resulting error estimates are true uniformly in $x \in \Omega$.

6. **Maps that are (ϕ, f)-mixing but not ϕ-mixing.** In this section we discuss some systems that exhibit mixing behaviour similar to that of the previous section but without the uniformity present there. Now, f is not necessarily equal to 0 (or a constant).

6.1. **Piecewise continuous maps.** In this section we use results on some systems that have been studied by various people and in particular by Paccaut [18] in his PhD thesis. Let M be a compact manifold, $T : M \to M$ a piecewise continuous and invertible transformation which one-to-one on the atoms of a partition $\mathcal{A}$.
(I) We assume that the partition is sufficiently regular, i.e. $\mathcal{A}$ is generating, every atom in $\mathcal{A}^*$ has only finitely many components, and for every open $U \subset M$ there is a k so that $M = T^k(U \setminus \partial\mathcal{A})$.
(II) The positive potential function $g : M \to \mathbf{R}^+$ satisfies the following bounded distortion property

$$0 < \limsup_{n\to\infty} \frac{1}{n} \log \max_{A\in\mathcal{A}^n} \sup_{x,y\in A} \left|\frac{g(y)}{g(x)} - 1\right| < 1,$$

(IV) The function also satisfies $P(g, T|_{\partial\mathcal{A}}) < P(g, T)$ (P is the pressure function).
(V) On the boundary we have the following[3]:

$$\limsup_{q\to\infty} \frac{1}{q} \log S(q) \le P(g, T)) - \log\vartheta,$$

where

$$S(q) = \sum_n \vartheta^n \sum_{A\in\mathcal{A}^n,\ \bar{A}\cap\partial T^q\mathcal{A}^q\neq\emptyset} \sup_A g_n$$

($g_n = g \cdot g \circ T \cdot g \circ T^2 \cdots g \circ T^{n-1}$ is the n-th Birkhoff product of g).

[3]For one-dimensional maps Paccaut showed that this condition (V) is implied by the previous four conditions (I)–(IV).

Then it has been proven by Paccaut [18] that there exists a unique (T-invariant) equilibrium state μ for g and $0 < \rho < 1$ so that

$$|\mu(G(H \circ T^k) - \mu(G)\mu(H)| \leq c_1 \rho^k \|G\|_\vartheta \|H\|_{L^1}, \tag{9}$$

(c_1 is some constant) for all L^1-functions H and G in the function space V_ϑ which consists of all functions χ whose ϑ-variation

$$\mathrm{var}_\vartheta = \sum_{k=1}^{\infty} \vartheta^k \sum_{A \in \mathcal{A}^k} \sup_A g_k \,\mathrm{osc}_A f$$

are bounded ($\vartheta < 1$), where $\mathrm{osc}_A f = \sup_{x,y \in A} |f(x) - f(y)|$.

Let $\mathcal{L}$ be the transfer operator with the weight function g. Then $\mathcal{L}$ has a unique positive eigenfunction h and a unique eigenfunctional ν which, if properly normalised, give the equilibrium state $\mu = h\nu$.

Now let A_n be an n-cylinder and let us estimate the ϑ-variation of its characteristic function χ_{A_n}. One has $\mathrm{osc}_U \chi_{A_n} \leq 1$ for every cylinder $U \in \mathcal{A}^k$, $k = 1, \dots, n-1$ and $\mathrm{osc}_U \chi_{A_n} = 0$ for every k-cylinder when $k \geq n$. Hence

$$\mathrm{var}_\vartheta \chi_{A_n} \leq \sum_{k=1}^{n-1} \vartheta^k \sup_{A_k} g_k \leq \sum_{k=1}^{n-1} \vartheta^k |g|_\infty \leq \kappa^n$$

for some constant $\kappa > 1$, where the k-cylinders A_k are so that $A_n \subset A_k$. If A_n has positive measure then we define

$$f(A_n) = \left[2 \frac{\log(\mu(A_n)\kappa^n)}{\log \rho} \right].$$

One sees that for $A \subset B$, $|A| \geq |B|$, $A, B \in \mathcal{A}^*$ then $f(A) \geq f(B)$. Hence f defines a separation function on $\mathcal{A}^*$ and we have by equation (9)

$$|\mu(A_n \cap T^{-k-n} V) - \mu(A_n)\mu(V)| \leq c_1 \rho^{k/2} \mu(A_n)\mu(V),$$

for all measurable $V \subset M$ and $k \geq f(A_n)$. In other words, μ is (ϕ, f)-mixing with f and $\phi(k) = \rho^{k/2}$. Clearly ϕ is summable. If μ satisfies a Gibbs inequality then $f(A) \leq c_2 |A|$ for some c_2 and all $A \in \mathcal{A}^*$.

Theorem 18. *Let T be a piecewise invertible maps as above and μ an equilibrium state.*

There exists a constant C_{12} so that for all $A_n \in \mathcal{A}^n$ and for all r, t for which $\frac{(r+t+1)^2}{t+1} n^{q^*+1} \epsilon(A_m) \leq 1/16$:

$$\left| \mu_*(\mathcal{N}_n^r) - \frac{t^r}{r!} e^{-t} \right| \leq C_{12} n^{q^*} \epsilon(A_m) \begin{cases} \frac{(r+t)^2}{r!} (4t)^{r-1} e^{4t} & \text{if } r \geq 1 \\ e^{4t}(t+1) & \text{if } r = 0 \end{cases}$$

(entry times: $q^* = 0, \mu_* = \mu$; return times: $q^* = 1, \mu_* = \mu_n$), where
(1) m is such that $m \leq \tau(A_m)$ and $A_n \subset A_m \in \mathcal{A}^m$,
(2) V_n a union of δ''-cylinders such that $A_m \subset V_n$ and $f(V_n) \leq m - \delta''$,
(3) $\epsilon(A_n) = \max((n + f(A_n)\mu(V_n), \rho^{(n+f(A_n))/2})$.

Proof. Let V_n be as in the hypothesis and put

$$\delta_n = \max\left(n + f(A_n), \frac{\log \mu(V_n)}{2 \log \rho} \right).$$

Then $\phi(\delta_n) = \rho^{\delta_n/2} \leq \mu(V_n)$.

(i) Entry times: If we choose $\varepsilon_n = \epsilon(A_n)$ and $\alpha = 3$ the conditions of Theorem 1 ($\alpha' = 4$) are satisfied by Proposition 9.

(ii) Return times: With the choice $\varepsilon_n = n\epsilon(A_n)$ and $\alpha = 3$ the conditions of Theorem 1 ($\alpha' = 4$) are satisfied by Proposition 10. □

Let $h = \limsup_{n\to\infty} \frac{\log|\mathcal{A}^n|}{n}$ denote the topological entropy of T. Let $0 < \sigma' \le e^{-3h}$ and put $\mathcal{J}_n^c = \bigcup_{A\in\mathcal{A}^n, \mu(A)\le\sigma'^n} A$. Then

$$\mu(\mathcal{J}_n^c) \le \sigma'^n |\mathcal{A}^n| \le e^{-hn}$$

for all large enough n. For $x \in \mathcal{J}_n$ one has $\mu(A_n(x)) \ge \sigma'^n$ which allows us to estimate the separation function: $f(A_n) \le 2n\frac{\log\sigma'\kappa}{\log\rho}$ (one can now read off the value of c_2 above).

Let us now examine the distribution of first return times. In order to apply Proposition 12 we put $\delta'' = \left[n\frac{\log\rho}{\log\rho\sigma'^2\kappa^2}\right]$. Thus $\delta_n = n\frac{\log\rho\sigma'^2\kappa^2}{\log\rho}$ and consequently we can use Theorem 1 with the error term

$$\varepsilon(A_n(x)) \le n\frac{\log\rho\sigma'^2\kappa^2}{\log\rho}\mu(A_{\delta''}(x)) \le \mu(A_n(x))^p,$$

where $p \ge \log\sigma/\log\sigma'$. Let $\hat{\mathcal{N}}_n^r$ be the level sets of the function $\hat{\zeta}_n$ which counts the restricted returns to the set $A_n(x)$ up to time $t/\mu(U_n(x))$, where $U_n = (T^{-n}A_n)\setminus \bigcup_{j=1}^{n-1} T^{j-n}A_n$. To emphasise the dependency on x let us denote the conditional measure on $A_n(x)$ by $\mu_{A_n(x)}$. We thus obtain:

Theorem 19. *For some C_{13}, all $x \in \mathcal{J}_n$, n large enough and all t, r for which $\frac{(r+t+1)^2}{t+1} n\mu(A_n(x))^p < (4C_{13})^{-1}$ one has:*

$$\left|\mu_{A_n(x)}(\hat{\mathcal{N}}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_{13} n\mu(A_n(x))^p \begin{cases} \frac{(r+t)^2}{r!}(4t)^{r-1}e^{4t} & \text{if } \ r \ge 1 \\ e^{4t}(t+1) & \text{if } \ r = 0 \end{cases}.$$

The distribution of the first return time is given by the case $r = 0$.

Corollary 20. *For $x \in \mathcal{J}_n$, $\mu(\mathcal{J}_n) \ge 1 - e^{-hn}$, and all n large enough and $t \ge n\mu(U_n(x))$ for which $n^2(t+1)\mu(A_n(x))^p < 1/4C_{13}$ is small:*

$$\left|\mu_{A_n(x)}\left(\left\{y \in A_n(x) : \tau_{U_n(x)}(y) \ge \frac{t}{\mu(U_n(x))}\right\}\right) - e^{-t}\right| \le C_{13}(t+1)e^{4t}\mu(A_n(x))^p,$$

6.2. **Rational Maps.** Let T be a rational map of degree at least 2 and J its Julia set. Assume that we executed appropriate branch cuts on the Riemann sphere so that we can define univalent inverse branches S_n of T^n on J for all $n \ge 1$. Put $\mathcal{A}^n = \{\varphi(J) : \varphi \in S_n\}$ (n-cylinders). Note that the diameters of the elements in $\mathcal{A}^n$ go to zero as $n \to \infty$. Moreover, $\mathcal{A}^n$ is not the join of a partition, yet they have all the properties we require.

Let f be a Hölder continuous function on J so that $P(f) > \sup f$ ($P(f)$ is the pressure of f), let μ be its unique equilibrium state on J and $\zeta_n = \sum_{j=1}^N \chi_{A_n} \circ T^{-j}$ the 'counting function' which measures the number of times a given point returns to the n-cylinder A_n within the normalised time $N = [t/\mu(A_n)]$. Although μ is not a Gibbs measure we showed in [13] that for almost every x

$$\mu(\mathcal{N}_n^r) \to \frac{t^r}{r!}e^{-t},$$

as $n \to \infty$, where $\mathcal{N}_n^r = \{y \in \Omega : \zeta_n(y) = r\}$ are the r-levelsets of ζ_n. We are now able to considerably sharpen the result on the convergence and give explicit error bounds as well as provide the limiting distribution for the return times.

Theorem 21. *Let T be a rational map of degree ≥ 2 and μ be an equilibrium state for Hölder continuous f (with $P(f) > \sup f$).*

Then there exists a $\tilde{\rho} \in (0,1)$ and C_{14} so that on a set of measure larger than $1 - \tilde{\rho}^n$ one has (entry times: $q^* = 0, \mu_* = \mu$; return times: $q^* = 1, \mu_* = \mu_n$):

$$\left|\mu_*(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \leq C_{14} n^{q^*} \tilde{\rho}^n \begin{cases} \frac{(r+t)^2}{r!}(3t)^{r-1}e^{3t} & \text{if} \quad r \geq 1 \\ e^{3t}(t+1) & \text{if} \quad r = 0 \end{cases},$$

for all t, r for which $\frac{(r+t+1)^2}{t+1} n\tilde{\rho}^n < (4C_{14})^{-1}$.

The univalent inverse branches S_n of T^n (with appropriate branch cuts) split into two categories, namely the uniformly exponentially contracting inverse branches S_n' and the remaining $S_n'' = S_n \setminus S_n'$ for which do not contract uniformly. In [13] we showed the following result:

Lemma 22. *There exists a C_{15}, $\sigma < 1$ and $\kappa > 1$ so that*

$$\left|\mu(W \cap T^{-k-n}V) - \mu(W)\mu(V)\right| \leq C_{15}\sigma^k\kappa^n\mu(V)\mu(W),$$

where $W = \bigcup_j A_{\varphi_j}$ for finitely many $\varphi_j \in S_n'$, $k, n > 0$ and Q measurable.

If in the last lemma we would not have to restrict to the cylinder sets of contracting branches in S_n' then (T, μ) would be (ϕ, f)-mixing, with decay function $\phi(k) = \sigma^{k/2}$ and separation function $f(A) = q|A|$, $A \in \mathcal{A}^*$, where q is an integer so that $\sigma^q\kappa < 1$. However the contributions from the non-contracting branches can still be well controlled and allows us to proceed in a way that nearly identical to the (ϕ, f)-mixing case with $f(A) = q|A|$. The following lemma is the equivalent of Lemma 4.

Lemma 23. *([13] Lemma 9) Let $\eta \in (0,1)$, $r > 1$ an integer. Then there exists a constant C_{16} and a $q > 0$ so that for all $\vec{v} = (v_1, v_2, \ldots, v_r) \in G_r$ satisfying $\min_j(v_{j+1} - v_j) \geq (1+q)n$:*

$$\left|\frac{\mu(\bigcap_{j=1}^r T^{-v_j}W_j)}{\prod_{j=1}^r \mu(W_j)} - 1\right| \leq C_{16}\eta^n,$$

for all sets $W_1, \ldots, W_r$ each of which is a union of atoms in $\mathcal{A}^n$ and for all $n \geq 1$.

Let us define the rare set and its components I_r and K_r. For $p > 0$ let us put $I_r(N) = \{\vec{v} \in G_r(N) : \min_j(v_{j+1} - v_j) \leq pn\}$, where the value of p will be determined in the next paragraph. The set $K_r(N)$ is then given by all $\vec{v} \in G_r(N)$ for which $pn < \min_j(v_{j+1} - v_j) \leq (1+q)n$, where q is as in Lemma 23. In the terminology of the previous section we use $\nu_1 = [pn] + 1$ and $\nu_2 = (1+q)n$.

Let $0 < p < 1$ be so that $d^p\sqrt{\rho} \leq 1$ where $\rho = e^{\sup f - P(f)}$. In the next lemma we show that those cylinders $A \in \mathcal{A}^n$ that return 'too soon' to themselves constitute a small set. Define

$$\mathcal{J}_n^c = \bigcup_{A \in \mathcal{A}^n} \bigcup_{m=1}^{[pn]} A \cap T^{-m}A,$$

and then put $\mathcal{J}_n$ for its complement.

Lemma 24.

$$\mu(\mathcal{J}_n^c) \le n\rho^{n/2}$$

Proof. Let and τ_φ denote the first return time to the set A_φ, $\varphi \in S_n$ and define

$$U_m = \{y \in J : \tau_\varphi(y) = m\}$$

and obtain

$$U_m \cap A_\varphi \subseteq A_\varphi \cap T^{-m} A_\varphi \subseteq \bigcup_{k=0}^{m} U_k \cap A_\varphi.$$

With $V = T^m U_m \cap A_\varphi$ we have $V = A_\varphi \cap T^m A_\varphi$. Let us write $\varphi = \psi^1 \varphi^1$, where $\psi^1 \in S_m$ and $\varphi^1 = T^m \varphi \in S_{n-m}$ (with suitable branch cuts). We proceed inductively and obtain

$$\varphi = \psi^k \psi^{k-1} \cdots \psi^1 \varphi^k,$$

where $n = mk + \ell$, $0 \le \ell < m$, $\psi^j \in S_m$ and $\varphi^k = T^{mk}\varphi \in S_\ell$. Let us note that $T^{mj}V = A_{\varphi^j} \cap A_{\varphi^{j+1}}$ for $j = 1, \ldots, k$, where $\varphi^j = T^{jm}\varphi = \psi^{j+1} \cdots \psi^1 \varphi^k$. Since $\mu(A_{\psi^k \cdots \psi^1 \varphi^k}) \le \rho^{n+m}$ we can now estimate

$$\begin{aligned} \sum_{\varphi \in S_n} \mu(U_m \cap A_\varphi) \quad &\le \sum_{\psi^1, \cdots, \psi^k \in S_m} \mu(A_{\psi^k \cdots \psi^1 \varphi^k}) \\ &\le \quad |S_m| \rho^{n+m}, \end{aligned}$$

where there are at most $|S_m|$ choices for ψ^1 and then for every $j = 1, \ldots, k-1$ the $\psi^{j+1} \in S_m$ must satisfy $T^{jm}V \subset A_{\psi^{j+1}} \cap A_{\psi^j}$. For every ψ^j we get a unique ψ^{j+1} since the sets $\psi(J \cap \mathrm{int}(\Omega_m)), \psi \in S_m$ are disjoint. Hence the last inequality, where we also used the fact that $\mu(A_{\tilde\varphi}) \le |\tilde\varphi|_\infty \le \rho^{n+m}$ for $\tilde\varphi \in S_{n+m}$.

Since by assumption $d^p\sqrt{\rho} \le 1$ we get

$$\sum_{\varphi \in S_n} \mu(U_m \cap A_\varphi) \le d^m \rho^{n+m} \le (d^p \rho^{1/2})^n \rho^{n/2} \rho^m \le \rho^{n/2},$$

and therefore

$$\mu(\mathcal{J}_n^c) \le \sum_{m=0}^{[pn]} \sum_{\varphi \in S_n} \mu(U_m \cap A_\varphi) \le n\rho^{n/2},$$

which goes to zero as n goes to infinity. □

For $\vec v \in G_r(N)$ let us put $C_{\vec v} = \bigcap_{j=1}^r T^{-v_j} A_\varphi$, $\varphi \in \mathcal{A}^n$, $N = t/\mu(A_\varphi)$. Let us put $b_{\vec v}^n = \mu(C_{\vec v})$. If we put $I_r = \{\vec v \in G_r : \min_j(v_{j+1} - v_j) < pn\}$, then the last lemma showed us that for all $x \in \mathcal{J}_n$ one has

$$\sum_{\vec v \in I_r} b_{\vec v}^n = 0.$$

Proof of Theorem 21. We are going to check on the conditions of Theorem 1. First for the entry times. We assume that $x \in \mathcal{J}_n$ which impliesthat $R_r = K_r$.
(i), (ii) By invariance of the measure $b_v^n = \mu(A_\varphi)$ for all v.
(iii) The assumption of Lemma 7 (i) is satisfied if we choose $\delta' = pn$ and $\delta = (1+q)n$. According to Lemma 23 our separation function f is given by $f(k) = (1+q)k = \delta$.

Hence $\delta'' = [pn/(1+q)]$. With this choice, V is a δ''-cylinder whose measure is $\mu(V) \leq \rho^{pn/(1+q)}$. This yields

$$\begin{aligned}\sum_{\vec{v}\in K_r} b^n_{\vec{v}} &\leq 2(1+C_{13})(1+q)n\rho^{pn/(1+q)}\sum_{s=0}^{r}(2(1+q)n\rho^{pn/(1+q)})^{r-s}\frac{(2t)^s}{s!} \\ &\leq c_1\tilde{\rho}^n\sum_{s=0}^{r}\tilde{\rho}^{n(r-s)}\frac{(2t)^s}{s!}.\end{aligned}$$

for some $\tilde{\rho} \in (\rho^{p/(1+q)}, 1)$ and some $c_1 \geq 1$.
(iv) By Lemma 8 one has for every r:

$$\sum_{\vec{v}\in K_r} b^n_{v_1}\cdots b^n_{v_r} \leq \frac{\mu(A_\varphi)t^{r-1}}{(r-2)!}.$$

(v) This is shown in Lemma 23.

Naturally $\mu(A_\varphi) \leq \tilde{\rho}^n$. Hence, if we put $\varepsilon_n = c_1\tilde{\rho}^n$ and $\alpha = 2$ then we obtain the result follows from Theorem 1 ($\alpha' = 3$). The proof of the result for the return times proceeds in a similar way with the obvious modifications (mainly in (v)). $\square$

REFERENCES

[1] M Abadi: Instantes de ocorrência de eventos raros em processos misturadores; PhD thesis 2001, Universidade de São Paulo

[2] M Abadi: Exponential Approximation for Hitting Times in Mixing Processes; Math. Phys. Elec. J. 7 No.2 (2001)

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

[4] V Afraimovich: Pesin's dimension for Poicaré recurrences; Chaos, 7 (1997) 12–20

[5] V Afraimovich, J-R Chazottes, B Saussol: Pointwise dimensions for Poincaré recurrences associated with maps and special flows; Discrete and Continuous Dynamical Systems, 9 (2003) 263–280

[6] V Afraimovich, J Schmeling, E Ugalde, J Urias: Spectra of dimensions for Poincaré recurrences, Discrete and Continuous Dynamical Systems, 6 (2000) 901–914

[7] R Bowen: *Equilibrium States for Anosov Diffeomorphism*; Springer Lecture Notes 470 (1975)

[8] Z Coelho: Asymptotic laws for symbolic dynamical processes: In *Topics in Symbolic Dynamics and Applications*, edited by F. Blanchard, A. Maass and A. Nogueira, LMS Lecture Notes Series 279, Cambridge University Press 2000, 123–165.

[9] P Collet, A Galves and B Schmitt: Fluctuations of repetition times for Gibbsian sources; Nonlinearity 12 (1999) 1225–1237

[10] M Denker: Remarks on weak limit laws for fractal sets; Progress in Probability Vol. 37, Birkhäuser 1995, 167–178

[11] A Galves and B Schmitt: Inequalities for hitting times in mixing dynamical systems; Random and Computational Dynamics 1997

[12] B Fernandez, E Ugalde, J Urias: Spectrum of dimensions for Poincaré recurrences of Markov maps; Discrete and Continuous Dynamical Systems, 8 (2002) 835–849

[13] N T A Haydn: Statistical properties of equilibrium states for rational maps; Ergo. Th. & Dynam. Sys. 20 (2000), 1371–1390

[14] M Hirata: Poisson law for Axiom A diffeomorphisms; Ergod. Th. & Dynam. Sys. 13 (1993), 533–556

[15] M Hirata: Poisson law for the dynamical systems with the "self-mixing" conditions; *Dynamical Systems and Chaos*, Vol. 1 (Worlds Sci. Publishing, River Edge, New York (1995), 87–96

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

[17] P Kurka, V Penné, S Vaienti: Dynamically defined recurrence dimension; Discrete and Continuous Dynamical Systems, 8 (2002) 137–146

[18] F Paccaut: Propriétés Statistiques de Systèmes Dynamiques Non Markovian; PhD Thesis Dijon 2000

[19] V Penné, B Saussol, S Vaienti: Dimensions for recurrence times: topological and dynamical properties; Discrete and Continuous Dynamical Systems, 5 (1999) 783–798

[20] W Philipp and W Stout: *Almost sure invariance principles for partial sums of weakly dependent random variables*; AMS Memoirs 161 (1975)

[21] B Pitskel: Poisson law for Markov chains; Ergod. Th. & Dynam. Syst. 11 (1991), 501–513

[22] B A Sevast'yanov: Poisson limit law for a scheme of sums of independent random variables; Th. Prob. Appl. 17 (1972), 695–699

[23] B Saussol, S Troubetzkoy and S Vaienti: Recurrence, dimensions and Lyapunov exponents; J. Stat. Phys. 106 (2002)

[24] H Wang, M Tang and R Wang: A Poisson limit theorem for a strongly ergodic non-homogeneous Markov chain; submitted to J. Math. Analysis and Applications.

Received October 2002; first revision February 2003; second revision July 2003.

E-mail address: nhaydn@math.usc.edu, vaienti@cpt.univ-mrs.fr