The limiting distribution and error terms for return times of dynamical systems

Nicolai Haydn [*] Sandro Vaienti [†]

November 21, 2010

Abstract

We develop a general framework that allows to prove that the limiting distribution of return times are Poisson distributed. The approach uses a result that connects the convergence of factorial moments to the mixing properties of transformations which often times are expressed through the decay of correlations. 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 expanding transformations to small set. Let T be an expansive transformation on the space Ω and let μ

[*]Mathematics Department, USC, Los Angeles, 90089-1113. e-mail: <nhaydn@math.usc.edu>.

[†]Centre de Physique Théorique, CNRS, Luminy Case 907, F-13288 Marseille Cedex 9 e-mail: <vaienti@cpt.univ-mrs.fr>.

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 set '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 [20] 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 ([12], [13]) has similar results for the zeroth return time $r = 0$ using the transfer operator restricted to the complement of ε-balls in the shiftspace (the argument for the higher order return times $r \geq 1$ seems to be incomplete).

For cylinder sets, Galves and Schmitt [9] have obtained rates of convergence for the zeroth order return times ($r = 0$). Hirata, Saussol and Vaienti 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 $x^{1+\alpha}$ 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. The ingredience is the following theorem which quantifies a previous result of Sevast'yanov [21]. 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 the following G_r is a subset of $\mathbf{Z}^r$. If in property (5) below we had that the left hand side were equal to zero for all $\vec{v}$ then the statement of the theorem would be trivially satisfied (since then $\mu(\mathcal{N}_n^r) = \frac{t^r e^{-t}}{r!}$ for all r). However since in property (5) the error only has to go to zero for 'most' of the multi-indices $\vec{v}$ we have to impose smallness conditions on the remaining indices for which (5) does not apply. In our setting the rare set will typically consist of return time patterns $\vec{v}$ which contain a return which is 'too short'. The conditions (3) and (4) look rather complicated but are exactly what can be shown for (some) mixing dynamical systems and still be made to work in the theorem below.

Theorem 1 *Let $\{\eta_v^n : v = 1, \dots, N(n)\}$ for $n \geq 1$ be an array of random $0, 1$-valued variables and μ a probability measure. Put $\zeta_n = \sum_{v=1}^N \eta_v^n$, and for*

$\vec v \in G_r = \{\vec v \in \mathbf{Z}^r : 1 \le v_1 < v_2 < \cdots < v_r \le N(n)\}$ let $b^n_{\vec v} = \mu(\eta^n_{\vec v})$, where $\eta^n_{\vec v} = \prod_{s=1}^r \eta^n_{v_s}$ (in particular $b^n_v = \mu(\eta^n_v)$).

Assume that there is a (monotonically to zero decreasing) sequence ε_n so that the following five assumptions are satisfied:

$$\max_{1\le v\le N} b^n_v \le \varepsilon_n, \tag{1}$$

$$\left|\sum_{v=1}^N b^n_v - t\right| \le \varepsilon_n. \tag{2}$$

Moreover assume that there exist rare sets $R_r \subset G_r$ (depending on n) ($r \ge 1$) and constants $\alpha \ge 0$ so that (the numbers r', r'' are so that $|r'-r|, |r''-r|$ are bounded)

$$\sum_{\vec v \in R_r} b^n_{\vec v} \le \varepsilon_n \sum_{s=0}^{r'} \binom{r'}{s} \varepsilon_n^{r'-s} \frac{(\alpha t)^s}{s!}, \tag{3}$$

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

$$\left|\frac{b^n_{v_1}\cdots b^n_{v_r}}{b^n_{\vec v}} - 1\right| \le \alpha^r \varepsilon_n, \tag{5}$$

for all $\vec v \in G_r \setminus R_r$.

Then there exists a constant C_1 so that for all $t > 0$, n and r for which $r^2\varepsilon_n/t$ is small (say less than 0.1) if $r \ge 1$ and $\varepsilon_n t$ is small if $r = 0$:

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

For all values of n, r and t one has the (weaker) bound

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

where $\mathcal{N}^r_n = \{y : \zeta_n(y) = r\}$ is the r-levelset of ζ_n.

Proof. Throughout the proof we shall assume that $r' = r'' = r$. If $r'', r' \ne r$ but their differences are (uniformly) bounded by some contant c_0 then there are obvious modifications below that let us arrive at the same conclusion (except the constant C_1 will have to be replaced by $C_1(c_0+1)^2$).

If we put $U_r = r! \sum_{\vec v \in G_r} b^n_{\vec v}$ then we have by assumption (3)

$$I = \left| U_r - r! \sum_{\vec v \notin R_r} b^n_{\vec v} \right| = r! \sum_{\vec v \in R_r} b^n_{\vec v} \le r! \varepsilon_n \sum_{s=0}^r \binom{r}{s} \varepsilon_n^{r-s} \frac{(\alpha t)^s}{s!}.$$

Moreover, by assumption (4)

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

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

$$III = \left| \left(\sum_{k=0}^{N} b_k^n \right)^r - t^r \right| \leq r \varepsilon_n (t + \varepsilon_n)^{r-1}.$$

Factoring out yields

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

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)

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

With the estimate:

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

we obtain

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

Since by assumption (5)

$$V = \left| r! \sum_{\vec{v} \notin R_r} \prod_i b_{v_i}^n - r! \sum_{\vec{v} \notin R_r} b_{\vec{v}}^n \right|$$

$$
\begin{aligned}
&\le r!\alpha^r \varepsilon_n \sum_{\vec{v}\notin R_r} \prod_i b^n_{v_i} \\
&\le \alpha^r \varepsilon_n V_r \\
&\le \varepsilon_n \alpha^r (t+\varepsilon_n)^r \\
&\le r!\varepsilon_n \sum_{s=0}^{r} (\alpha\varepsilon_n)^{r-s} \frac{(\alpha t)^s}{s!},
\end{aligned}
$$

we can now estimate as follows

$$
\begin{aligned}
|U_r - t^r| &\le I + II + III + IV + V \\
&\le 3r!\varepsilon_n \sum_{s=0}^{r} \binom{r}{s} \varepsilon_n^{r-s} \frac{(\alpha t)^s}{s!} + \sum_{k=1}^{r-1} \frac{r!(r-k)^k}{(r-k)!k!} \varepsilon_n^k (t+\varepsilon_n)^{r-k} + r\varepsilon_n (t+\varepsilon_n)^{r-1}.
\end{aligned}
$$

The last term can be absorbed by either of the first two sums.

Let us now form the generating function for the random variable ζ_n:

$$
f_n(z) = \sum_{k=0}^{\infty} z^k \mu(\mathcal{N}_n^k)
$$

and note that $f_n^{(k)}(0) = k!\mu(\mathcal{N}_n^k)$. In particular we get for the rth derivative that

$$
f_n^{(r)}(z) = \sum_{k=0}^{\infty} k(k-1)\cdots(k-r+1) z^{k-r} \mu(\mathcal{N}_n^k),
$$

which evaluated at $z = 1$ yields

$$
f_n^{(r)}(1) = \sum_{k=0}^{\infty} k(k-1)\cdots(k-r+1)\mu(\mathcal{N}_n^k) = \mu(\zeta_n^{(r)}),
$$

where $\zeta_n^{(r)} = \zeta_n(\zeta_n - 1)\cdots(\zeta_n - r + 1)$ is the rth factorial moment of ζ_n. If we develop $f_n(z)$ at $z = 1$ into a powerseries we get

$$
f_n(z) = \sum_{r=0}^{\infty} \frac{f_n^{(r)}(1)}{r!}(z-1)^r = \sum_{r=0}^{\infty} \frac{(z-1)^r}{r!} \mu(\zeta_n^{(r)}).
$$

For $x \in \mathcal{N}_n^k$, $k \ge r$, one has that $\zeta_n^{(r)}(x) = k(k-1)\cdots(k-r+1)$. For $\vec{v} \in G_r$ let us put $C_{\vec{v}} = \{x : \eta^n_{\vec{v}} = 1\}$ and let us observe that for any given r we have:
(i) if $x \in \mathcal{N}_n^k$ for some $k < r$ then $x \notin C_{\vec{v}}$, for all $\vec{v} \in G_r$,
(ii) if $x \in \mathcal{N}_n^k$ for $k \ge r$ then there are $\binom{k}{r}$ distinct $\vec{v} \in G_r$ so that $x \in C_{\vec{v}}$.
Since $C_{\vec{v}} = \bigcup_{k=r}^{\infty} C_{\vec{v}} \cap \mathcal{N}_n^k$ (disjoint union) we get

$$
\sum_{\vec{v}\in G_r} \mu(C_{\vec{v}}) \;=\; \sum_{k=r}^{\infty} \sum_{\vec{v}\in G_r} \mu(C_{\vec{v}} \cap \mathcal{N}_k^n)
$$

$$
\begin{aligned}
&= \sum_{k=r}^{\infty} \frac{k!}{(k-r)!r!}\mu(\mathcal{N}_n^k) \\
&= \sum_{k=r}^{\infty} \frac{1}{r!}\int_{\mathcal{N}_n^k} \zeta_n^{(r)}\, d\mu \\
&= \frac{1}{r!}\mu(\zeta_n^{(r)})
\end{aligned}
$$

and therefore

$$
U_r = \mu(\zeta_n^{(r)}) = f_n^{(r)}(1).
$$

The (error) function $\varphi_n(z) = f_n(z) - e^{t(z-1)}$ splits into the sum $\varphi = \tilde{\varphi} + \tilde{\tilde{\varphi}}$. The first part is (here we used that $\alpha \geq 1$)

$$
\begin{aligned}
|\tilde{\varphi}_n(z)| &\leq \sum_{r=0}^{\infty} \frac{|z-1|^r}{r!} c_1 r! \varepsilon_n \sum_{k=0}^{r} \left(\begin{array}{c} r \\ s \end{array}\right) (\alpha\varepsilon_n)^{r-k} \frac{(\alpha t)^k}{k!} \\
&= c_1 \varepsilon_n F(|z-1|\alpha t, |z-1|\alpha\varepsilon_n)
\end{aligned}
$$

where we used the identity

$$
F(x,y) = \sum_{r=0}^{\infty}\sum_{s=0}^{r} \left(\begin{array}{c} r \\ s \end{array}\right) y^{r-s}\frac{x^s}{s!} = \frac{1}{1-y} e^{\frac{x}{1-y}}.
$$

In particular we see that $\tilde{\varphi}$ is for every value of t analytic for $|z-1| < \alpha/\varepsilon_n$. The second part of the error function is (where we put $\ell = r - k$):

$$
\begin{aligned}
\left|\tilde{\tilde{\varphi}}_n(z)\right| &\leq \sum_{r=0}^{\infty} \frac{|z-1|^r}{r!} \sum_{k=1}^{r-1} \frac{r!}{(r-k)!k!} (r-k)^k \varepsilon_n^k (t+\varepsilon_n)^{r-k} \\
&\leq \sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty} \frac{\ell^k}{\ell!k!} |z-1|^{\ell+k} \varepsilon_n^k (t+\varepsilon_n)^{\ell} \\
&\leq \sum_{\ell=1}^{\infty} \frac{|z-1|^{\ell}(t+\varepsilon_n)^{\ell}}{\ell!} \left(e^{|z-1|\ell\varepsilon_n} - 1\right) \\
&= \left(e^{e^{|z-1|\varepsilon_n}|z-1|(t+\varepsilon_n)} - e^{|z-1|(t+\varepsilon_n)}\right).
\end{aligned}
$$

It that φ, $\tilde{\varphi}$ and $\tilde{\tilde{\varphi}}$ are for every value of t analytic for $|z-1| < \alpha/\varepsilon_n$. For $\alpha|z-1|\varepsilon_n$ small enough we get

$$
|\tilde{\varphi}(z)| \leq c_2 \varepsilon_n e^{\alpha'|z-1|(t+\varepsilon_n)},
$$

for $\alpha' > \alpha$ and for $|z-1|\alpha\varepsilon_n$, $|z-1|^2\varepsilon_n(t+\varepsilon_n)$ small enough we estimate the second term as follows

$$
\begin{aligned}
\left|\tilde{\tilde{\varphi}}_n(z)\right| &\leq e^{|z-1|(t+\varepsilon_n)} \left(e^{(e^{|z-1|\varepsilon_n}-1)|z-1|(t+\varepsilon_n)} - 1\right) \\
&\leq e^{|z-1|(t+\varepsilon_n)} \left(e^{|z-1|^2\varepsilon_n(t+\varepsilon_n)3/2} - 1\right) \\
&\leq 2e^{|z-1|(t+\varepsilon_n)} |z-1|^2 \varepsilon_n (t+\varepsilon_n).
\end{aligned}
$$

A Cauchy estimate now yields ($R > 0$):

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

provided, of course, that $(R+1)\varepsilon_n$ and $(R+1)^2\varepsilon_n(t+\varepsilon_n)$ are small enough. Hence, since

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

we get

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

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

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

for some constant c_4. A better choice of R can be done if $k \ge 1$ is not too large, which is whenever $k^2\varepsilon_n/t$ is small then we can use the optimal value for R, namely $R = k/t$ and obtain ($c_5 > 0$)

$$\left|\mu(\mathcal{N}_n^k) - \frac{t^k}{k!}e^{-t}\right| \le c_5\frac{(k+t)^2}{k^k}e^{k+t}\varepsilon_n t^{k-1}.$$

Using Stirling's formular 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

$$\left|\mu(\mathcal{N}_n^0) - e^{-t}\right| \le \varphi_n^{(k)}(0) \le c_3 e^t\varepsilon_n(t+1).$$

□

Remark. The error estimate for $k < c_1\sqrt{t/\varepsilon_n}$ (for some small c_1, e.g. equal to 0.1) becomes meaningless for t larger than of the order $|\log k^2\varepsilon_n|$ because the principal term becomes smaller than the error term.

Also note that the error term ε_n is allowed to depend on t which is a parameter in the theorem.

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 is an $\varepsilon \ge 0$ so that*

$$\begin{aligned} \max_{1\le v\le N} b_v &\le \varepsilon, \\ \left|\sum_{v=1}^N b_v - t\right| &\le \varepsilon, \end{aligned}$$

and suppose there is $R_r \subset G_r$ $(r \geq 1)$ so that:

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

Then there exists a constant C_1 so that for all $t > 0$ and r for which $r^2\varepsilon/t$ is small (say less than 0.1) if $r \geq 1$ and εt is small if $r = 0$:

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

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

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

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

3 Properties of (ϕ, f)-mixing measures

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 $n \geq 1$. 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) $f : \mathcal{A}^* \to \mathbf{N}_0$ so that $f(A) \geq f(B)$ if $|A| \geq |B|$, $A, B \in \mathcal{A}^*$. If C is a union of n-cylinders C_j (some n) then $f(C) = \max_j f(C_j)$.
(ii) $\phi : \mathbf{N}_0 \to \mathbf{R}^+$ is non-increasing.

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

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

for all $m \geq f(U)$, measurable V and U which are unions cylinders of the same length.

Oftentimes the function f depends only on the length of the cylinders, that is $f(A) = f(|A|)$. The function ϕ determines the rate at which the mixing occurs and the *separation function* f specifies a lower bound for the size of the gap m that is necessary to get a good mixing property. In the special case when f is constant 0 (or some other constant) then (T, μ) is traditionally called ϕ-mixing. There is a tradeoff between the decay function ϕ and the separation function f. Typically one can achieve to have ϕ decay faster at the expense of f which as a consequence will be increasing faster.

3.1 General properties

For $r \geq 1$ and (large) N denote by $G_r(N)$ the r-vectors $\vec{v} = (v_1, \dots, v_r)$ for which $1 \leq v_1 < v_2 < \cdots < v_r \leq N$. (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 iterates 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 $[1, N]$.

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

$$\begin{aligned} \left| \mu\left(\bigcap_{j=1}^r T^{-v_j} W_j\right) - \prod_{j=1}^r \mu(W_j) \right| &\leq \sum_{k=1}^{r-1} |\mu(D_k) - \mu(W_k)\mu(D_{k+1})| \prod_{j=1}^{k-1} \mu(W_j) \\ &\leq \sum_{k=1}^{r-1} \phi(v_{k+1} - v_k - n_k)\mu(D_{k+1}) \prod_{j=1}^{k} \mu(W_j) \end{aligned}$$

$$
\begin{aligned}
&\leq \prod_{j=1}^{r} \mu(W_j)\phi(d(\vec{v},\vec{n})) \sum_{k=0}^{r-1} (1+\phi(d(\vec{v},\vec{n})))^k \\
&= ((1+\phi(d(\vec{v},\vec{n})))^r - 1) \prod_{j=1}^{r} \mu(W_j),
\end{aligned}
$$

where we used the identity $x\sum_{k=0}^{r-1}(1+x)^k = (1+x)^r - 1$ and the estimates

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

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 [9] and for α-mixing measures in [1].

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

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

Proof. If $A \in \mathcal{A}^*$ and $n = |A|$, then $A = \bigcap_{j=0}^{n-1} T^{-j}A_j$ for some $A_j \in \mathcal{A}$, $j = 0, \ldots, n-1$. Let $m \geq 1 + \max_{B\in\mathcal{A}} f(B)$ (large enough) be such that

$$
\gamma_0 = (1+\phi(m-1))\max_{B\in\mathcal{A}}\mu(B) < 1.
$$

If we put $r = [\frac{n}{m}]$, then we obtain by Lemma 4:

$$
\mu(A) \leq \mu\left(\bigcap_{i=0}^{r-1} T^{-im}A_{im}\right) \leq (1+\phi(m-1))^r \prod_{i=0}^{r-1}\mu(A_{im}) \leq \gamma_0^r,
$$

which proves the lemma with e.g. $\gamma = \gamma_0^{1/2m}$. □

The remaining lemmas in this section will be used to estimate the size of the rare set. 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

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

Correspondingly the principal part is

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

Lemma 6 *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)

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

(ii) (Return time version)

$$\sum_{\vec{v} \in \tilde{K}_r} \mu(C_{\vec{v}}) \leq C_2 \mu(W) \sum_{s=0}^{r-1} \binom{r-1}{s} \frac{(2\beta t)^s}{s!} (2\beta \delta \mu(V))^{r-s},$$

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 $f(V) \leq \delta' - \delta''$.

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

(i) Let us first prove the first statement of the lemma. Put K_r^s for those $\vec{v} \in K_r$ where $v_{i+1} - v_i \geq \delta$ for exactly s indices $i_1, i_2, \ldots, i_s$ (obviously one always has $s \leq r-2$ and $i_s \leq r-1$).
I. Let us now assume that $s \geq 1$ and let $i_1, i_2, \ldots, i_s$ be the indices for which $v_{i_k+1} - v_{i_k} \geq \delta$ for $k = 1, \ldots, s$. All the other differences are $\geq \delta'$ and smaller than δ. Let δ'' be so that V is a union of δ''-cylinders and $f(V) \leq \delta' - \delta''$. 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 $v_{i_k+1} - v_{i_k} \geq \delta \geq f(W)$ and $v_{j+1} - v_j \geq f(V)$ for $j \neq i_k$, $k = 1, \ldots, s$. 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 $\vec{n}=(n_1,\ldots,n_r)$ are given by $n_{i_k}=n$ for $k=1,\ldots,s$ and $n_j=\delta''$ for $j\ne i_k,\ k=1,\ldots,s$.

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 $i_1,\ldots,i_s$ 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

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

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

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

for a larger C_2 if necessary. □

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 7 *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.

3.2 Entry and return times for (ϕ, f)-mixing maps

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 *Hirata-Vaienti return time function*

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

which measures the shortest return time within the set A (see [13, 14]). By definition $A \cap T^{-k}A = \emptyset$ for $k = 1, 2, \dots, \tau(A) - 1$.

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 $v = 0, 1, \dots, N$, 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 10. For $\vec{v} \in G_r(N)$ ($\vec{v} = (v_1, v_2, \dots, v_r)$) 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 $\delta_n' \leq \delta_n, n = 1, 2, \dots$, 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 10 and 9 in the third inequality the sum is taken only over K_r, the principal part of the rare set. In Proposition 11 however we consider the full rare set. Later on we shall use Corollary 15 (which uses Proposition 11) to get bounds on the set of very short returns I_r.

Proposition 8 *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

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

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

Proof. (i), (ii) By invariance of the measure μ we have

$$b_v^n = \mu(\eta_v^n) = \mu(T^{-v}A_n) = \mu(A_n)$$

for all v and therefore

$$\left|\sum_{v=1}^N b_v^n - t\right| \leq |N\mu(A_n) - t| \leq \mu(A_n).$$

This proves the first two statements of the proposition.
(iii) We can assume that $v_{j+1} - v_j \geq m$ for all j (because otherwise the set $C_{\vec{v}}$ is empty) and apply Lemma 6 (i) to the case when $\delta' = \delta'_n$, $\delta = \delta_n$. We obtain the following estimate

$$\sum_{\vec{v}\in K_r} \mu(C_{\vec{v}}) \;\leq\; 3\delta_n\mu(V_n)\sum_{s=0}^r (3\delta_n\mu(V_n))^{r-s}\frac{(3t)^s}{s!},$$

where V_n is as in the hypothesis (2) and where the value of β is bounded by $1 + \phi(\min_j(v_{j+1} - v_j)) \leq 3/2$.
(iv) The fourth inequality is easily verified using Lemma 7:

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

(v) To verify the last inequality we use Lemma 4 to obtain

$$\begin{aligned}
|\mu(C_{\vec{v}}) - \mu(A_n)^r| &\leq ((1+\phi(\delta_n - n))^r - 1)\,\mu(A_n)^r \\
&\leq C_3 r\phi(\delta_n)\mu(A_n)^r,
\end{aligned}$$

for some constant C_3 ($C_3 \geq 3$), and therefore (for large n)

$$\left| \frac{b^n_{\vec{v}}}{b^n_{v_1} \cdots b^n_{v_r}} - 1 \right| \leq C_3 r \phi(\delta_n).$$

$\square$

Proposition 10 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 $\eta^n_v = \chi_{A_n} \circ T^v$ and consider now the values

$$\begin{aligned} b^n_v &= \mu_n(\eta^n_v), \\ b^n_{\vec{v}} &= \mu_n(\eta^n_{\vec{v}}), \end{aligned}$$

where $\vec{v} \in G_r(N)$ and, as above, $\eta^n_{\vec{v}}$ 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 9 *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:

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

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

$$|b_v^n - \mu(A_n)| \leq \phi(v-n)\mu(A_n). \tag{7}$$

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 $b_v^n \leq c_1\mu(V_n)$.

(ii) Summability of the function ϕ gives us the second inequality:

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

(iii) To obtain the third inequality we apply Lemma 6 (ii) with the parameters $\delta' = \delta'_n$, $\delta = \delta_n$, $W = A_n$, $V = V_n$ and $\tilde{K}_r$ as defined above:

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

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

(iv) If $v_j \geq \delta_n$ then

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

and otherwise ($\delta'_n v_j < \delta_n$) we use the estimate $b_{v_j}^n \leq c_1\mu(V_n)$ 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 7:

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

Summing over $s=1,\dots,r$ yields (where $k=r-s$)

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

(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

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

and by Lemma 4 we get

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

(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 $j=1,2,\dots,r$:

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

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

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

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 $C_4=\max(1,c_1,c_2,3,C_2)$. □

$(,f)$-MIXING MAPS

3.3 Restricted entry and return times for (ϕ,f)-mixing maps

The following results will provide us with the asymptotics of long returns to the neighbourhoods of periodic orbits and in particular also with the asymptotics of the first return time for all points. 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

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

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

In the following proposition, which is the analog of Proposition 9 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 $\hat{N}$.

Proposition 10 *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 $f(A_n)\le\delta_n-n$ and $t>0$ one has

$$\begin{aligned}\max_{1\le v\le\hat{N}}\hat{b}^n_v &\le \mu(A_n)\\ \left|\sum_{v=1}^{\hat{N}}\hat{b}^n_v-t\right| &\le C_5 n\mu(A_n)\end{aligned}$$

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

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

Proof. (i) We have to consider two cases: (a) $1\le v<n$ and (b) $n\le v$. In the first case, $1\le v<n$, we get

$$
\hat{b}^n_v=\mu(\hat{\eta}^n_v)\le\mu(A_n),
$$

since $\hat{\eta}^n_v(x)=0$ if $x\in A_n$. In the second case, $v\ge n$, we get by invariance of μ

$$
\hat{b}^n_v=\mu(\hat{\eta}^n_v)=\mu(T^{-(v-n)}U_n)=\mu(U_n)\le\mu(A_n).
$$

(ii) With the estimates of $\hat{b}^n_v$ from part (i):

$$
\begin{aligned}
\left|\sum_{v=1}^{\hat{N}}\hat{b}^n_v-t\right| &\le n\mu(A_n)+(\hat{N}-n)\mu(U_n)-t\\
&\le n(\mu(A_n)+\mu(U_n))+\mu(U_n)\\
&\le c_1 n\mu(A_n).
\end{aligned}
$$

(iii) We can assume that $v_{j+1}-v_j\ge n$ for all j (because otherwise the set $C_{\vec{v}}$ is empty) and apply Lemma 6 (i) to the case when $\delta'=\delta'_n$, $\delta=\delta_n$. We obtain the following estimate

$$
\sum_{\vec{v}\in K_r}\mu(C_{\vec{v}}) \le C_3 t\mu(V_n)\sum_{s=0}^{r-2}\begin{pmatrix} r-2\\ s\end{pmatrix}\frac{(2t)^s}{s!}(2\delta\mu(V_n))^{r-s}
$$

where V_n is as in the hypothesis and where the value of β is bounded by $1+\phi(\min_j(v_{j+1}-v_j))\le 2$.
(iv) If $v_j<\delta_n$ we use the estimate from part (i) $\hat{b}^n_{v_j}\le\mu(A_n)$ and otherwise if $v_j\ge n$ then $\hat{b}^n_{v_j}=\mu(U_n)$. If the first entry of $\vec{v}$ is less that n then we obtain similarly to Lemma 7:

$$
\begin{aligned}
\sum_{\vec{v}\in\tilde{R}_r;\, v_1<n}\hat{b}^n_{v_1}\hat{b}^n_{v_2}\cdots\hat{b}^n_{v_r} &\le n\mu(A_n)\mu(U_n)^{r-1}|G_{r-1}|\\
&\le n\mu(A_n)\frac{t^{r-1}}{(r-2)!}.
\end{aligned}
$$

If none of the entries of $\vec{v}$ is less than n then we get similarly to Lemma 7

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

(v) To verify the last inequality we use Lemma 4 to obtain (here $v_{j+1}-v_j \ge \delta_n \ge f(A_n)+2n$ as $\vec{v}\notin R_r$)

$$\begin{aligned}|\mu(C_{\vec{v}})-\mu(U_n)^r| &\le ((1+\phi(\delta_n-n))^r-1)\,\mu(U_n)^r\\ &\le C_5 r\phi(\delta_n)\mu(U_n)^r,\end{aligned}$$

for some constant $C_5\ge 3$, and therefore (for large n)

$$\left|\frac{\hat{b}^n_{\vec{v}}}{\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r}}-1\right| \le C_5 r\phi(\delta_n).$$

□

Proposition 10 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 $\eta^n_v=\chi_{A_n}\circ T^v$ and consider now the values

$$\begin{aligned}b^n_v &= \mu_n(\eta^n_v),\\ b^n_{\vec{v}} &= \mu_n(\eta^n_{\vec{v}}),\end{aligned}$$

where $\vec{v}\in G_r(N)$ and, as above, $\eta^n_{\vec{v}}$ 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 $\delta'_n\le\delta_n$ we define the set $\tilde{K}_r$ of short (but not too short) returns by

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

Proposition 11 *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 $f(A_n)\le\delta_n-2n$ and $t>0$:

$$\begin{aligned}\max_{1\le v\le\hat{N}}\hat{b}^n_v &\le C_6\mu(V_n)\\ \left|\sum_{v=1}^{\hat{N}}\hat{b}^n_v-t\right| &\le C_6(f(A_n)+n)\mu(V_n)\end{aligned}$$

$$\begin{aligned}
\sum_{\vec{v}\in R_r}\hat{b}^n_{\vec{v}} &\le C_6\delta_n\mu(A_n)\sum_{s=0}^{r+1}(3\delta_n\mu(A_n))^{r+1-s}\frac{(3t)^s}{s!}\\
\sum_{\vec{v}\in R_r}\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r} &\le C_6\delta_n\mu(V_n)\sum_{k=0}^{r-1}(\delta_nC_6\mu(V_n))^{r-1-k}\frac{t^k}{k!}\\
\left|\frac{\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r}}{\hat{b}^n_{\vec{v}}}-1\right| &\le 2^r\phi(\delta_n),
\end{aligned}$$

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

Proof. Let us first note that U_n is a union of $2n$-cylinders and similarly that for $0\le v<n$ the functions $\hat{\eta}^n_v$ are characteristic functions on sets $U_n(v)$ which are unions of $(n+v)$-cylinders.
(i) We have to consider three cases, namely (a) $1\le v<n$, (b) $n\le v<f(A_n)+2n$ and (c) $f(A_n)+2n\le v\le\hat{N}$. In the first case $1\le v<n$ and by definition $\hat{\eta}^n_v\chi_{A_n}=0$. Thus $\hat{b}^n_v=0$. In the second case, $v=n,\dots,f(A_n)+2n-1$, the fact that $U_n\subset T^{-n}A_n$ yields

$$\mu(A_n)\hat{b}^n_v=\mu(\chi_{A_n}\hat{\eta}^n_v)=\mu(A_n\cap T^{-(v-n)}U_n)\le\mu(V_n\cap T^{-v}A_n).$$

The ϕ-mixing property compares the last term to $\mu(V_n)\mu(A_n)$:

$$|\mu(V_n\cap T^{-v}A_n)-\mu(V_n)\mu(A_n)|\le\phi(v-\delta'')\mu(V_n)\mu(A_n)$$

and therefore $\hat{b}^n_v\le\mu(V_n)(1+\phi(v-\delta''))$. In the third case, $v\ge f(A_n)+2n$, one has

$$|\mu(\chi_{A_n}\hat{\eta}^n_v)-\mu(A_n)\mu(U_n)|\le\phi(v-2n)\mu(A_n)\mu(U_n)$$

and (note that $\mu_n(\hat{\eta}^n_v)=\mu(\chi_{A_n}\hat{\eta}^n_v)/\mu(A_n)$)

$$|\hat{b}^n_v-\mu(U_n)|=\left|\frac{\mu(\chi_{A_n}\hat{\eta}^n_v)}{\mu(A_n)}-\mu(U_n)\right|\le\phi(v-2n)\mu(U_n).$$

Hence for $f(A_n)+2n\le v\le\hat{N}$

$$\hat{b}^n_v=\mu_n(T^{-(v-n)}U_n)\le\mu(U_n)1+\phi(v-2n)\mu(U_n)\le c_1\mu(U_n). \tag{8}$$

(ii) Summability of the function ϕ gives the second inequality:

$$\begin{aligned}
\left|\sum_{v=1}^{\hat{N}}\hat{b}^n_v-t\right| &\le (f(A_n)+2n)(\mu(A_n)+c_1\mu(V_n))+\mu(U_n)\sum_{v=f(A_n)+2n}^{\hat{N}}\phi(v-n)\\
&\le (1+c_1)(f(A_n)+n)\mu(V_n)+\mu(A_n)\sum_{v=0}^{\infty}\phi(v)\\
&\le c_2(f(A_n)+n)\mu(V_n).
\end{aligned}$$

(iii) The fact that $U_n \subset T^{-n}A_n$ and $U_n(v) \subset T^{-v}A_n$ implies that $\hat{\eta}^n_{\vec{v}}, \vec{v} \in G_r(\hat{N})$, is the characteristic function of a set which is contained in $\bigcap_{j=1}^r T^{-v_j}A_n$. Since $U_n \cap T^{-j}U_n = \emptyset$ and $U_n(v) \cap T^{-j}U_n = \emptyset$ for $j < n$ and $v < n$ we can employ Lemma 6 (ii) with $\delta' = n$, $\delta = \delta_n$, $\delta'' = n$, $V = A_n$ and $\tilde{K}_r = \tilde{R}_r$ yields

$$\sum_{\vec{v}\in\tilde{R}_r} \hat{b}^n_{\vec{v}} \leq 7\delta_n\mu(A_n)\sum_{s=0}^{r+1}(3\delta_n\mu(A_n))^{r+1-s}\frac{(3t)^s}{s!}$$

for all large enough n so that $\beta = 1 + \phi(\delta_n - n) \leq 3/2$.
(iv) If $v_j \geq \delta_n$ then by part (i) of the proof

$$\hat{b}^n_{v_j} \leq c_3\mu(U_n),$$

and otherwise ($v_j < \delta_n$) we use the estimate $\hat{b}^n_{v_j} \leq c_1 \max(\mu(V_n), \tilde{\varepsilon})$ from part (i). If the first s of the entries of $\vec{v}$ are less that $f(A_n) + 2n \leq \delta_n$ then we obtain similarly to Lemma 7 for $s \geq 1$:

$$\begin{aligned}\sum_{\vec{v}\in\tilde{R}_r;\, v_1,\ldots,v_s<\delta_n} \hat{b}^n_{v_1}\hat{b}^n_{v_2}\cdots\hat{b}^n_{v_r} &\leq \delta_n^s c_3^s\mu(V_n)^s\mu(U_n)^{r-s}|G_{r-s}| \\ &\leq (\delta_n c_1\mu(V_n))^s\frac{t^{r-s}}{(r-s)!}.\end{aligned}$$

If $s = 0$ (no entry of $\vec{v}$ is less than $f(A_n) + 2n$) then

$$\sum_{\vec{v}\in R_r; v_1\geq\delta_n} \hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r} \leq |R_r|\mu(U_n)^r \leq \delta_n\frac{\mu(U_n)t^{r-1}}{(r-2)!}.$$

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

$$\begin{aligned}\sum_{\vec{v}\in\tilde{R}_r} \hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r} &\leq \sum_{s=1}^{r}(\delta_n c_3\mu(V_n))^s\frac{t^{r-s}}{(r-s)!} \\ &\leq \delta_n c_3\mu(V_n)\sum_{k=0}^{r-1}(\delta_n c_3\mu(V_n))^{r-1-k}\frac{t^k}{k!}.\end{aligned}$$

The fourth inequality now follows since the binomial coefficient is ≥ 1.
(v) To verify the last of the inequalities we restrict to $\vec{v} \notin \tilde{R}_r$, that is $v_{j+1} - v_j > \delta_n \geq f(A_n) + n$ for all j and $v_1 \geq \delta_n$. Since

$$\hat{b}^n_{\vec{v}} = \mu_n(\hat{C}_{\vec{v}}) = \frac{\mu(A_n \cap \hat{C}_{\vec{v}})}{\mu(A_n)},$$

where $\hat{C}_{\vec{v}} = \bigcap_{j=1}^r T^{-v_j}U_n$, by Lemma 4

$$\begin{aligned}|\mu(A_n \cap \hat{C}_{\vec{v}}) - \mu(A_n)\mu(U_n)^r| &\leq ((1 + \phi(\delta_n - n))^r - 1)\,\mu(A_n)\mu(U_n)^r \\ &\leq rc_4\phi(\delta_n - n)\mu(A_n)\mu(U_n)^r,\end{aligned}$$

and therefore

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

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

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

and in particular $\hat{b}^n_{v_j} \le c_1\mu(U_n)$. Hence

$$\begin{aligned} |\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r} - \mu(U_n)^r| &\le r\left(\max_j\left|\hat{b}^n_{v_j} - \mu(U_n)\right|\right)\left(\max\left(\hat{b}^n_{v_1},\dots,\hat{b}^n_{v_r},\mu(U_n)\right)\right)^{r-1} \\ &\le r\phi(v_1 - n)c_1^{r-1}\mu(U_n)^r \\ &\le rc_1^r\phi(\delta_n - n)\mu(U_n)^r, \end{aligned}$$

for all large enough n. By the triangle inequality

$$|\hat{b}^n_{\vec{v}} - \hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r}| \le r(c_4 + c_1^r)\phi(\delta_n - n)\mu(U_n)^r,$$

and therefore, with a slightly larger value for c_1,

$$\left|\frac{\hat{b}^n_{\vec{v}}}{\hat{b}^n_{v_1}\cdots\hat{b}^n_{v_r}} - 1\right| \le c_1^r\phi(\delta_n).$$

Let us note that since we only consider large enough n, the number $c_1 > 1$ can be chosen arbitrarily close to 1. In particular we can assume that $c_1 < 2$. An appropriate choice for C_6 finishes the proof. □

Remark 1. In Proposition 11 the ϕ-mixing requirement can be weakened. It is sufficient that μ is ϕ-mixing on A_n and V_n:

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

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

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

4 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 [9, 19] it has been shown that the measure of n-cylinders fall off geometrically, i.e. there is a constant $c_1 > 0$ so that $\mu(A) \le e^{-nc_1}$ 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 whall 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

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

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

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

the levelset of the counting function ξ_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 the Axiom A properties and have very regular behaviour as the shadowing property and finite Markov partitions of arbitrarily small diameter. Such systems are usually studied using a symbolic description by a subshift of finite type. A good reference is the classical book by Bowen [4]. We shall study the entry and return time distribution for equilibrium states for Hölder continuous potentials.

4.1 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 12 *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 a constant C_7 so that for all $A_n \in \mathcal{A}^n$ and all t, r for which $\frac{r^2}{t}\mu(A_n)^{\frac{p}{1+p}}$ is small (and $t\mu(A_n)^{\frac{p}{1+p}}$ is small if $r = 0$) one has:
(i) (Distribution of entry times)

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

(ii) (Distribution of return times) if $\sum_v \phi(v) < \infty$ *then*

$$\left|\mu_n(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \leq nC_7\mu(A_m)^{\frac{p}{1+p}} \begin{cases} \frac{(r+t)^2}{r!}t^{r-1}e^t & \text{if } r \geq 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases}$$

where $m = \min(n, \tau(A_n))$ *and* $A_n \subset A_m \in \mathcal{A}^m$*.*

Proof. We want to verify the conditions of Theorem 1 using Proposition 10. 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 10 is equal to A_m as $f = 0$.

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

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

and

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

(i) With $\varepsilon_n = nc_2\mu(A_n)^{\frac{p}{1+p}}$ ($c_2 = \max(7, c_1)$) and $\alpha = 3$, Proposition 10 ensures that the conditions (1)–(5) of Theorem 1 are satisfied.

(ii) If we put $\varepsilon_n = nc_2\mu(A_m)^{\frac{p}{1+p}}$ and $\alpha = 3$ then Proposition 9 ensures that the conditions (1)–(5) of Theorem 1 are satisfied.

Put $C_7 = \max(3C_1c_1, C_1c_2)$. □

Remark. Theorem 16 covers the special case when ϕ is summable (see in particular [9]), i.e. $\sum_v \phi(v) < \infty$, which implies that $\phi(v) \leq c_1/v$ (for $v > 0$). Lemma 16 thus can be applied to the case when $p = 1$ ($p^* = 1/2$) and gives us the following error terms:

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

□

4.2 Mappings that are Axiom A

In the following we are looking at strongly hyperbolic maps that satisfy the Axiom A properties and conseqently have very regular behaviour as the shadowing property and finite Markov partitions of arbitrarily small diameter. Such systems are usually studied using a symbolic description by a subshift of finite type. A good reference is the classical book by Bowen [4].

Theorem 13 *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 potential* f *(*$\mu(\Omega) = 1$*).*

Then there exists a constant C_8 so that for all $A_n \in \mathcal{A}^n$ and all t, r for which $\frac{r^2}{t}\mu(A_m)$ is small ($t\mu(A_m)$ mall if $r = 0$) one has (for large enough m):
(i) (Distribution of entry times)

$$\left|\mu(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \leq C_8 n\mu(A_m)\begin{cases} \frac{(r+t)^2}{r!}t^{r-1}e^t & \text{if } r \geq 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

(ii) (Distribution of return times)

$$\left|\mu_n(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \leq C_8 n^2\mu(A_m)\begin{cases} \frac{(r+t)^2}{r!}t^{r-1}e^t & \text{if } r \geq 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

where $m = \min(\tau(A_n), n)$ and $A_n \subset A_m \in \mathcal{A}^m$.

Proof. We shall use 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 [4] of μ there exists a number $c_2 > 0$ so that $\mu(A_m) \geq e^{-mc_2} \geq e^{-nc_2}$ for all large enough m and n. Put $\delta_n = qn$, where $q = 1 + c_2/|\log\vartheta|$. Then (as $V_n = A_m$)

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

and

$$\phi(\delta_n - n) \leq \vartheta^{(q-1)n} \leq e^{-c_1 n} \leq \mu(A_m)$$

for all large enough n.

(i) If we choose $\varepsilon_n = 3qn\mu(A_m)$ and $\alpha = 3$ the conditions of Theorem 1 are satisfied by Proposition 10. This proves the first statement of the theorem.

(ii) With the choice $\varepsilon_n = 3qn^2\mu(A_m)$ and $\alpha = 3$ the conditions of Theorem 1 are satisfied by Proposition 9. □

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.

4.3 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 $\hat{\eta}_v^n$, $v = 0, 1, \ldots, N$ be defined as in Proposition 11, 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 11 that $V_n = A_n$. Consider the restricted counting function $\hat{\xi}_n = \sum_{v=1}^{\hat{N}} \hat{\eta}_v^n$ and its r-levelsets $\hat{\mathcal{N}}_n^r = \{x \in \Omega : \hat{\xi}_n(x) = r\}$.

Theorem 14 *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 so that for all $A_n \in \mathcal{A}^n$ and all t, r for which $\frac{r^2}{t}\mu(A_n)^{p^*}$ is small ($t\mu(A_n)^{p^*}$ small if $r = 0$) one has:

$$\left|\mu_n(\hat{\mathcal{N}}_n^r) - \frac{t^r}{r!}e^{-t}\right| \leq C_9 n^{q^*+1}\mu(A_n)^{p^*}\begin{cases} \frac{(r+t)^2}{r!}t^{r-1}e^t & \text{if } r \geq 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

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 $\tau(U_n) \geq n$.

(i) If ϕ decays polynomially with power p we put $\delta_n = \mu(A_n)^{-\frac{1}{1+p}}$ and therefore obtain $\delta_n\mu(A_n) \leq \mu(A_n)^{\frac{p}{1+p}}$ and $\phi(\delta_n) \leq c_1\mu(A_n)^{\frac{p}{1+p}}$ for some c_1. With $\varepsilon_n = nc_2\mu(A_n)^{p^*}$ ($q^* + 1 = 1$, $c_2 = \max(7, c_1)$) and $\alpha = 3$, Proposition 11 implies that the conditions (1)–(5) of Theorem 1 are met.

(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 [4] $\mu(A_n) \geq e^{-nc_4}$ ($c_4 > 0$) for all large enough n. With $\delta_n = qn$, where $q = 1 + c_4/|\log\vartheta|$ we obtain (as $V_n = A_n$) that $\delta_n\mu(A_n) \leq qn\mu(A_n)$ and $\phi(\delta_n - n) \leq \mu(A_n)$ for all large enough n. With $\varepsilon_n = 3qn^2\mu(A_n), \alpha = 3$ the conditions of Theorem 1 are satisfied by Proposition 11. □

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 14.

Corollary 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_{10} so that for all $A_n \in \mathcal{A}^n$ and $t \geq n\mu(U_n)$ for which $t\mu(A_n)^{p^*}$ is small:

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

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 5.

4.4 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. For a given point x denote by μ_n the conditional measure on the set $A_n(x)$. By [**?**]

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

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 $\tau(A_n(x)) \geq (1-\varepsilon)n \;\forall\; n \geq N_\varepsilon(x)$. Therefore, if we put

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

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.

Corollary 16 *Let μ be a ϕ-mixing probability measure.*

Then there exists C_{11}, $\sigma < 1$ and a sequence of sets $\mathcal{J}_n \subset \Omega$ for which $\lim_{n\to\infty} \mu(\mathcal{J}_n) \to 1$ *so that for all $x \in \mathcal{J}_n$ and suitable t, r*

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

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

5 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).

5.1 Piecewise continuous maps

In this section we use results on some systems that have been studied by various people and in particular by Paccaut [19] in his PhD thesis. Let M be a compact manifold, $T : M \to M$ a piecewise invertible transformation which one-to-one on the atoms of a partition $\mathcal{A}$. We assume that the partition is sufficiently regular, i.e. that it satisfies
(i) $\mathcal{A}$ is generating,
(ii) every atom in $\mathcal{A}^*$ has only finitely many components,

(iii) For every open $U \subset M$ there is a k so that $M = T^k(U \setminus \partial\mathcal{A})$.
Moreover let $g : M \to \mathbf{R}^+$ be a positive potential function which satisfies the following bounded distortion property

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

and for which $P(g, T|_{\partial\mathcal{A}}) < P(g, T)$ (P is the pressure function). Then it has been proven by Paccaut ([?] Theorem 2) that there exists a unique (T-invariant) equilibrium state μ 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 ($\vartheta > 1$).

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

There exists a constant C_{12} so that for all $A_n \in \mathcal{A}^n$:
(i) (Distribution of entry times)

$$\left|\mu(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_{12}\epsilon(A_m)\begin{cases} \frac{(r+t)^2}{r!}t^{r-1}e^t & \text{if } r \ge 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

(ii) (Distribution of return times)

$$\left|\mu_n(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le nC_{12}\epsilon(A_m)\begin{cases} \frac{(r+t)^2}{r!}t^{r-1}e^t & \text{if } r \ge 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

for all t, r for which $\frac{r^2}{t}\epsilon(A_m)$ is small ($t\epsilon(A_m)$ small if $r = 0$) where
(1) $m = \min(\tau(A_n), n)$ and $A_m \in \mathcal{A}^m$ is such that $A_n \subset A_m$,
(2) V_n a union of δ''-cylinders such that $A_m \subset V_n$ and $f(V_n) \le m - \delta''$,
(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 $\phi(\delta_n) = \rho^{\delta_n/2} \le \mu(V_n)$.

(i) If we choose $\varepsilon_n = \epsilon(A_n)$ and $\alpha = 3$ the conditions of Theorem 1 are satisfied by Proposition 10.

(ii) With the choice $\varepsilon_n = n\epsilon(A_n)$ and $\alpha = 3$ the conditions of Theorem 1 are satisfied by Proposition 9. □

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 11 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 $\hat{\xi}_n$ 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 18 *For some C_{13}, all $x \in \mathcal{J}_n$, n large enough and all t, r for which $\frac{r^2}{t}\mu(A_n(x))^p$ is small ($t\mu(A_n(x))^p$ small if $r = 0$) one has:*

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

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

Corollary 19 *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 $t\mu(A_n(x))^p$ is small:*

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

5.2 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).

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 $\xi_n = \sum_{j=1}^N \chi_{A_n} \circ T^{-j}$ the 'counting function' which measures the number of times a given point returns to the n-cylinder A_n within the normalised time $N = [t/\mu(A_n)]$. In [11] we showed that for almost every x

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

as $n \to \infty$, where $\mathcal{N}_n^r = \{y \in \Omega : \xi_n(y) = r\}$ are the r-levelsets of ξ_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 20 *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:
(i) (Entry times)

$$\left|\mu(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_{14}\tilde{\rho}^n \begin{cases} \frac{(r+t)^2}{r!} t^{r-1} e^t & \text{if } r \ge 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

(ii) (Return times)

$$\left|\mu_n(\mathcal{N}_n^r) - \frac{t^r}{r!}e^{-t}\right| \le C_{14} n \tilde{\rho}^n \begin{cases} \frac{(r+t)^2}{r!} t^{r-1} e^t & \text{if } r \ge 1 \\ e^t(t+1) & \text{if } r = 0 \end{cases},$$

for all r, t for which $r^2\tilde{\rho}^n/t$ respectively $t\tilde{\rho}^n$ is small.

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 [11] we showed the following result:

Lemma 21 *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 22 *[10] 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 22. 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 23

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

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 $j = 1, \dots, k$, 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

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

where there are at most $|S_m|$ choices for ψ^1 and then for every $j = 1, \dots, k-1$ 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 $b^n_{\vec v} = \mu(C_{\vec v})$. 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

$$\sum_{\vec v \in I_r} b^n_{\vec v} = 0.$$

Proof of Theorem 20. 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 $b^n_v = \mu(A_\varphi)$ for all v.
(iii) The assumption of Lemma 6 (i) is satisfied if we choose $\delta' = pn$ and $\delta = (1+q)n$. According to Lemma 22 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 $\mu(V) \le \rho^{pn/(1+q)}$. This yields

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

for some $\tilde{\rho} \in (\rho^{p/(1+q)}, 1)$ and some $c_1 \ge 1$.
(iv) By Lemma 7 one has for every r:

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

(v) This is shown in Lemma 22.

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

References

[1] M Abadi: Exponential Approximation for Hitting Times in Mixing Stochastic Processes; preprint Universidade de São Paulo 1999

[2] M Abadi and A Galves: Inequalities for the occurence times of rare events in mixing processes; preprint Universidade de São Paulo 2000

[3] P Billingsley: *Probability and Measure*; Wiley 1979

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

[5] Z Coelho: Asymptotic laws for symbolic dynamical processes: Review paper 1997

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

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

[8] M Denker and M Urbanski: Ergodic theory of equilibrium states for rational maps, Nonlinearity 4 (1991), 103–134

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

[10] N T A Haydn: Convergence of the transfer operator for rational maps; Ergod. Th. & Dynam. Syst. 19 (1999), 657–669

[11] N T A Haydn: Statistical properties of equilibrium states for rational maps; J. Stat. Phys. 94 (1999), 1027–1036

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

[13] 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

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

[15] I Kontoyiannis: Asymptotic Recurrence and Waiting Times for Stationary Processes; J. Theor. Prob. 11 (1998), 795–811

[16] A Liverani, B Saussol and S Vaienti: A probabilitstic approach to intermittency

[17] R Mañé: On the Bernoulli property for rational maps; Ergod. Th. & Dynam. Syst. 5 (1985), 71–88

[18] R Mañé: Differentiable Dynamics; Springer 1982

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

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

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