RETURN TIMES AT PERIODIC POINTS IN RANDOM DYNAMICS

NICOLAI HAYDN AND MIKE TODD

Abstract. We prove a quenched limiting law for random measures on subshifts at periodic points. We consider a family of measures $\{\mu_\omega\}_{\omega\in\Omega}$, where the 'driving space' Ω is equipped with a probability measure which is invariant under a transformation θ. We assume that the fibred measures μ_ω satisfy a generalised invariance property and are ψ-mixing. We then show that for almost every ω the return times to cylinders A_n at periodic points are in the limit compound Poisson distributed for a parameter ϑ which is given by the escape rate at the periodic point.

1. Introduction

For sufficiently mixing deterministic dynamical systems the return times to shrinking sets (e.g. dynamical cylinders) around a typical point in the phase space are in the limit exponentially distributed almost surely, as shown in [A] and [AS]. Moreover, for ϕ-mixing measures it follows from [AV] that for all non-periodic points one obtains in the limit the exponential distribution for entry and return times, but that at periodic points the limiting return times distribution have a point mass at the origin. A similar distinction can be drawn for higher order returns where we know that for ψ-mixing systems return times at periodic points are in the limit compound Poisson distributed [HV2]. Assuming the ϕ-mixing property we can again conclude that higher order return times are in the limit Poisson distributed ([HP, Corollary 1]).

For random, stochastic dynamical systems, following work in [RSV] for typical points, it was shown in [RT] that the entry times distributions at periodic points show similar behaviour as in the deterministic setting if one considers a quenched limit (the annealed result then follows easily). In this case the limiting distribution has a point mass at the origin and is otherwise exponential. The relative weight of the point mass is determined by the marginal measure and applies to almost all realisations of the random dynamics. It is assumed that the marginal measures are ψ-mixing. In the present paper we consider the same setting and prove that higher order return times at periodic points are compound Poisson distributed. Again the parameter for the compound Poissonian is entirely determined by the marginal measure and applies to almost all realisations.

The perspective taken here and in the works discussed above is to see the system via dynamically defined cylinder sets, which makes it essentially a 'symbolic approach'. We note that outside this context (for example for an interval map, considering hits to balls rather than cylinders), much less is known in the random setting. However, in [AFV], a Poisson distribution was shown for first hitting times to balls in the setting of certain random dynamical systems. We note that this was for systems which were all close to a certain well-behaved system, so the randomness could be interpreted as (additive) noise. Moreover, this was an annealed law rather than a quenched one.

Date: September 9, 2016.

1.1. **Setting and conditions.** For a measure space Ω, let $\theta : \Omega \to \Omega$ be an invertible transformation preserving an ergodic probability measure $\mathbb{P}$. Set $\Sigma = \mathbb{N}^{\mathbb{N}_0}$ and let $\sigma : \Sigma \to \Sigma$ denote the shift. Let $A = \{A(\omega) = (a_{ij}(\omega)) : \omega \in \Omega\}$ be a random transition matrix, i.e., for any $\omega \in \Omega$, $A(\omega)$ is a $\mathbb{N} \times \mathbb{N}$-matrix with entries in $\{0,1\}$ such that $\omega \mapsto a_{ij}(\omega)$ is measurable for any $i \in \mathbb{N}$ and $j \in \mathbb{N}$. For any $\omega \in \Omega$ we write

$$\Sigma_\omega = \{x = (x_0, x_1, \ldots)\colon x_i \in \mathbb{N} \text{ and } a_{x_i x_{i+1}}(\theta^i\omega) = 1 \text{ for all } i \in \mathbb{N}\}$$

for the fibre over the point ω in the 'driving space' Ω. Moreover

$$\mathcal{E} = \{(\omega, x)\colon \omega \in \Omega, x \in \Sigma_\omega\} \subset \Omega \times \Sigma$$

denotes the full space on which the random system is described through a skew action. Indeed, the random dynamical system coded by the skew-product $S : \mathcal{E} \to \mathcal{E}$ given by $S(\omega, x) = (\theta\omega, \sigma x)$. While we allow infinite alphabets here, we nevertheless call S a *random subshift of finite type* (SFT). Assume that ν is an S-invariant probability measure with marginal $\mathbb{P}$ on Ω. Then we let $(\mu_\omega)_\omega$ denote its decomposition on Σ_ω (see [Ar, Section 1.4]), that is, $d\nu(\omega, x) = d\mu_\omega(x)d\mathbb{P}(\omega)$. The measures μ_ω are called the *sample measures*. Observe that $\mu_\omega(A) = 0$ if $A \cap \Sigma_\omega = \emptyset$. We denote by $\mu = \int \mu_\omega \, d\mathbb{P}$ the marginal of ν on Σ. We note that we may replace the assumption of invertibility of θ by assuming the existence of sample and marginal measures as above.

We also identify our alphabet $\mathcal{A}$ with the partition given by 1-cylinders $U(a) = \{x \in \Sigma : x_0 = a\}$. The elements of the kth join $\mathcal{A}^k = \bigvee_{j=0}^{k-1} \sigma^{-j}\mathcal{A}$, $k = 1, 2, \ldots$ are called *k-cylinders*. Put $\mathcal{A}^*$ for the forward sigma-algebra generated by $\bigcup_{j\ge1} \mathcal{A}^j$. The length $|A|$ of a cylinder set A is determined by $|A| = k$ where k is so that $A \in \mathcal{A}^k$. Note that $\mathcal{A}$ is generating, i.e. that the atoms of $\mathcal{A}^\infty$ are single points. If we denote by χ_A the characteristic function of a (measurable) set $A \subset \Sigma$ then we can define the counting function

$$\zeta_A(z) = \sum_{j=1}^{N} \chi_A \circ \sigma^j(z),$$

$z \in \Sigma$, where N is the observation time given by the invariant annealed measure μ. To wit $N = [t/\mu(A)]$ for $t > 0$ a parameter. The value of ζ_A counts the number of times a given point returns to A within the time N.

Let us make the following assumptions:

(i) The measures μ_ω are ψ-mixing: There exists a decreasing function $\psi : \mathbb{N} \to [0, \infty)$ (independent of ω) so that

$$\left|\mu_\omega(A \cap \sigma^{-n-k}B) - \mu_\omega(A)\mu_{\theta^{n+k}\omega}(B)\right| \le \psi(k)\mu_\omega(A)\mu_{\theta^{n+k}\omega}(B)$$

for all $A \in \sigma(\mathcal{A}^n)$ and $B \in \mathcal{A}^*$.

(ii) The marginal measure μ satisfies the α-mixing property:

$$\left|\mu(A \cap \sigma^{-n-k}B) - \mu(A)\mu(B)\right| \le \psi(k)$$

for all $A \in \sigma(\mathcal{A}^n)$ and $B \in \mathcal{A}^*$.

(iii) There exist $0 < \eta_0 < 1$ so that $\eta_0^n \le \mu(A)$ for all $A \in \mathcal{A}^n$, all ω and all large n.

(iv)

$$\sup_\omega \sup_{A \in \mathcal{A}} \mu_\omega(A) < 1.$$

Our main result, Theorem 7, is that under these conditions, the return times at periodic points x are compound Poissonian provided the limit $\vartheta(x) = \lim_{n\to\infty} \frac{\mu(A_{n+m}(x))}{\mu(A_n(x))}$ exists, where m is the minimal period of x and $A_n(x) \in \mathcal{A}^n$ denotes the n-cylinder that contains x. To be more precise, if we denote by ζ_n^x the counting function

$$\zeta_n^x(z) = \sum_{j=1}^{N_n} \chi_{A_n(x)} \circ \sigma^j(z)$$

with the *observation time* $N_n = \left[\frac{t}{\mu(A_n(x))}\right]$ ($t > 0$ is a parameter), then we will show that $\mu_\omega(\zeta_n^x = r)\,, r = 0, 1, 2, \dots,$ converges to the Polya-Aeppli distribution as $n \to \infty$ for $\mathbb{P}$-almost every ω.

The first such result was by Hirata [Hi] for the first entry time for Axiom A systems. For random systems satisfying assumptions (i)–(iv) a similar result was then shown by Rousseau and Todd [RT] for the first entry time distribution in the quenched case. Note that, as mentioned above, for systems perturbed by additive noise, which are a particular case of our systems here, an annealed version of this result is proved in [AFV]. The additivity 'washes out' any periodic behaviour.

As in [RT], if we wish to consider shifts on countable alphabets, it is no longer reasonable to assume condition (iii), but if we drop this and strengthen condition (ii) to the assumption of the ψ-mixing property for μ, then our results still follow. We close this section by noting that the conditions on our systems here are the same as those in [RT], so the main result here also applies to all the applications given there.

1.2. **Structure of the paper.** In Section 2 we describe the compound Poisson distribution and state an auxiliary limiting result on which the proof of the main result is based. The main part of the proof of the main result consists of estimating the contributions made by short returns which are outside the periodicity of the periodic point and for which the mixing property cannot be well applied. This is done in Section 3. The compound part is determined by the periodic behaviour near the periodic point where it generates geometrically distributed immediate returns. For the long returns the mixing property comes into play and results in exponentially distributed returns between clusters of short returns whose numbers are in the limit geometrically distributed. We state our main result in Section 4 and provide examples in Section 5.

Acknowledgements. This work was begun at the Conference on Extreme Value Theory and Laws of Rare Events at Centre International de Rencontres Mathématiques (CIRM), Luminy and part of it continued at the American Institute of Mathematics (AIM). The authors thank both institutions for their hospitality.

2. Factorial moments and a limiting result

This section is used to recall a result on the approximation of the compound Poisson distribution with a geometric distribution, i.e. the Polya-Aeppli distribution. More general compound Poisson distributions were considered in [F] and more recently (e.g. [CR, BCL]) there have been efforts to approach compound Poisson distributions using the Chen-Stein method. Although the treatment in [CR] applies to more general setting, the result is far from applicable to our situation. Proposition 1 is the compound analogue of other theorems for the plain Poisson distribution as for instance in [Se, HV1].

2.1. **Compound Poisson distribution.** For a parameter $p \in [0,1)$ define the polynomials

$$P_r(t,p) = \sum_{j=1}^{r} p^{r-j}(1-p)^j \frac{t^j}{j!}\binom{r-1}{j-1},$$

$r = 1, 2, \dots,$ where $P_0 = 1$ ($r = 0$). The distribution $e^{-t}P_r(t,p)$, $r = 0, 1, 2, \dots$ is the *Pólya-Aeppli distribution* [JKW] which has the generating function

$$g_p(z) = e^{-t}\sum_{r=0}^{\infty} z^r P_r = e^{t\frac{z-1}{1-pz}}.$$

It has mean $\frac{t}{1-p}$ and variance $t\frac{1+p}{(1-p)^2}$. For $p = 0$: $e^{-t}P_r(t,0) = e^{-t}\frac{t^r}{r!}$ and one recovers the Poisson distribution whose generating function $g_0(z) = e^{t(z-1)}$ is analytic in the entire plane whereas for $p > 0$ the generating function $g_p(z)$ has an essential singularity at $\frac{1}{p}$. The expansion at $z_0 = 1$ yields $g_p(z) = \sum_{k=0}^{\infty}(z-1)^k Q_k$ where

$$Q_k(t,p) = \frac{1}{(1-p)^k}\sum_{j=1}^{k} p^{k-j}\frac{t^j}{j!}\binom{k-1}{j-1}$$

($Q_0 = 1$) are the factorial moments.

2.2. **Return times patterns.** Let M and $m < M$ be given integers (typically $m \ll M$) and let $N \in \mathbb{N}$ be some (large) number. For $r = 1, 2, 3, \dots$ we define the following:
(I) $G_r(N)$: We denote by $G_r(N)$ the simplex of r-vectors $\vec{v} = (v_1, \dots, v_r) \in \mathbb{N}^r$ for which $1 \le v_1 < v_2 < \dots < v_r \le N$.
(II) $G_{r,j}(N)$: We write G_r as the disjoint union $\bigcup_j G_{r,j}$ where $G_{r,j}$ consists of all $\vec{v} \in G_r$ for which we can find j indices $i_1, i_2, \dots, i_j \in \{1, 2, \dots, r\}$, $i_1 = 1$, so that $v_k - v_{k-1} \le M$ if $k \ne i_2, \dots, i_j$ and so that $v_k - v_{k-1} > M$ for all $k = i_2, \dots, i_j$.

For $\vec{v} \in G_{r,j}$ the values of v_i will be identified with returns; returns that occur within less than time M are called *immediate returns* and if the return time is $\ge M$ then we call it a *long return* (i.e. if $v_{i+1} - v_i < M$ then we say v_{i+1} is an immediate return and if $v_{i+1} - v_i \ge M$ the we call v_{i+1} a long return). That means that $G_{r,j}$ consists of all return time patterns $\vec{v}$ which have $r - j$ immediate returns that are clustered into j blocks of immediate returns and $j - 1$ long returns between those blocks. The entries v_{i_k}, $k = 1, \dots, j$, are the beginnings (heads) of the blocks (of immediate returns). We assume from now on that all short returns are multiples of m. (This reflects the periodic structure around periodic points as in condition (b) of Proposition 1.)
(III) $G_{r,j,w}(N)$: For $\vec{v} \in G_{r,j}$ the length of each block is $v_{i_{k+1}-1} - v_{i_k}$, $k = 1, \dots, j-1$. Consequently let us put $w_k = \frac{1}{m}(v_k - v_{k-1})$ for the *individual overlaps*, for $k \ne i_1, i_2, \dots, i_j$. Then $\sum_{\ell=i_k+1}^{i_{k+1}-1} w_\ell = \frac{1}{m}(v_{i_{k+1}-1} - v_{i_k})$ is the *overlap* of the kth block and $w = w(\vec{v}) = \sum_{k \ne i_1, i_2, \dots, i_j} w_k$ the *total overlap* of $\vec{v}$. If we put $G_{r,j,w} = \{\vec{v} \in G_{r,j} : w(\vec{v}) = w\}$ then $G_{r,j}$ is the disjoint union $\bigcup_w G_{r,j,w}$.
(IV) $\Delta(\vec{v})$: For $\vec{v}$ in $G_{r,j}$ we put

$$\Delta(\vec{v}) = \min\{v_{i_k} - v_{i_k-1} : k = 2, \dots, j\}$$

for the minimal distance between the 'tail' and the 'head' of successive blocks of immediate returns.

2.3. **Compound Poisson approximations.** We shall use the following result:

Proposition 1. *Let m be as above and assume that there are sequences $\{M(n) : n\}, \{N(n) : n\}$ and $0,1$-valued random variables $\rho_{j,n}$ for $j = 1, \dots, N(n)$ on some Σ. For $\vec{v} \in G_r(n)$ put $\rho_{\vec{v}} = \prod_i \rho_{v_i,n}$. Choose $\delta(n) > 0$ and define the 'rare set' $R_r(n) = \bigcup_{j=1}^r R_{r,j}$, where $R_{r,j} = \{\vec{v} \in G_{r,j} : \Delta(\vec{v}) < \delta\}$. Let μ be a probability measure on Σ which satisfies the following conditions:*
(a)

$$\sum_{\vec{v} \in G_r \setminus R_r} \mu(\rho_{\vec{v}}) \to Q_r(t,p)$$

as $n \to \infty$ for some $p \in (0,1]$.
(b)

$$\sum_{\vec{v} \in R_r} \mu(\rho_{\vec{v}}) \to 0$$

as $n \to \infty$.

Then for every r

$$\mu(\zeta_n = r) \to e^{-t} P_r(t,p)$$

as $n \to \infty$, where $\zeta_n = \sum_{j=1}^{N(n)} \rho_{j,n}$.

Proof. The result follows by the moment method (see for example [Bi, Section 30]) that $\mu(\zeta_n^{(r)})$ converges to $Q_r(t,p)$ for each r, where $\zeta_n^{(r)} = \zeta_n(\zeta_n - 1) \cdots (\zeta_n - r + 1)$ is the factorial moment. Since $\mu(\zeta_n^{(r)}) = \sum_{\vec{v} \in G_r} \mu(\rho_{\vec{v}})$ we obtain by assumptions (a) and (b) that

$$\mu(\zeta_n^{(r)}) = \sum_{\vec{v} \in G_r \setminus R_r} \mu(\rho_{\vec{v}}) + \sum_{\vec{v} \in R_r} \mu(\rho_{\vec{v}}) \to Q_r(t,p)$$

since the second term goes to zero and the first term converges to Q_r. ∎

In the following we will apply this proposition to situations that typically arise in dynamical systems. There the stationarity condition (a) of the proposition is implied by the invariance of the measure. The random variables ρ_j will be the indicator function of a cylinder set pulled back under the jth iterate of the map. Condition (a) is then implied by the mixing property. The more difficult condition to satisfy is (b) because it involves 'short range' interaction over which one has little control and which requires more delicate estimates (see Lemma 4 below).

3. ψ-mixing measures and the rare set

In this section we only assume Assumption (i), that is the measures μ_ω are ψ-mixing i.e. satisfy

$$\left|\mu_\omega(U \cap \sigma^{-m-n} V) - \mu_\omega(U)\mu_{\theta^{m+n}\omega}(V)\right| \le \psi(m)\mu_\omega(U)\mu_{\theta^{m+n}\omega}(V)$$

for all $U \in \sigma(\mathcal{A}^n)$, $V \in \sigma(\mathcal{A}^*)$ and for all $m, n \ge 0$, where $\psi(m) \to 0$ (and ψ is independent of ω).

For instance equilibrium states for Hölder continuous potentials on Axiom A systems (which include subshifts of finite type) or on the Julia set of hyperbolic rational maps are ψ-mixing [DU].

For $r \geq 1$ and (large) $\tau \in \mathbb{N}$ let as above $G_r(N)$ be the r-vectors $\vec{v} = (v_1, \dots, v_r) \in \mathbb{Z}^r$ for which $1 \leq v_1 < v_2 < \cdots < v_r \leq N$. Let t be a positive parameter, $W \subset \Sigma$ and put $\tau = [t/\mu(W)]$ be the normalised time. Then the entries v_j of the vector $\vec{v} \in G_r(N)$ are the times at which all the points in $C_{\vec{v}} = \bigcap_{j=1}^r \sigma^{-v_j} W$ hit the set W during the time interval $[1, N]$. The following lemma, a random version of [HV1, Lemma 4], is immediate.

Lemma 2. *Let (σ, μ_ω) be ψ-mixing. For $r > 1$ let $n_i \geq 1, i = 1, \dots, r-1$, be given numbers and $\vec{n} = (n_1, \dots, n_{r-1})$. Let $W_i \in \sigma(\mathcal{A}^{n_i})$ and assume that $\vec{v} \in G_r(N)$ is such that $v_{i+1} - v_i \geq n_i$ $(i = 1, \dots, r-1)$. Then*

$$\left| \frac{\mu_\omega\left(\bigcap_{i=1}^r \sigma^{-v_i} W_i\right)}{\prod_{i=1}^r \mu_{\theta^{v_i}\omega}(W_i)} - 1 \right| \leq (1 + \psi(d(\vec{v}, \vec{n})))^{r-1} - 1,$$

and $d(\vec{v}, \vec{n}) = \min_{1 \leq i \leq r-1}(v_{i+1} - v_i - n_i)$.

Remark 3. *As in* [RT, Lemma 2.1]*, and similarly to the above lemma, under conditions (i) and (iv) there exists $0 < \eta_1 < 1$ so that for all large n*

$$\mu_\omega(A) \leq \eta_1^n$$

for all $A \in \mathcal{A}^n$, and $\mathbb{P}$-a.e. ω.

3.1. **Estimate of the rare set.** Next we will estimate the size of the rare set. As before we put $C_{\vec{v}} = \bigcap_{k=1}^r \sigma^{-v_k} W$ for $\vec{v} \in G_r(N)$ where $W \in \sigma(\mathcal{A}^n)$ for some n. Let $\delta \geq 0$ and put

$$R_{r,j}(N) = \{\vec{v} \in G_{r,j}(N) : \min_k(v_{i_k+1} - v_{i_k} - n) < \delta\},$$

where the values $v_{i_1}, \dots, v_{i_j}$ are the beginnings of the j blocks of immediate returns (notation as in section 2.2 (II)). Then we put $R_r = \bigcup_j R_{r,j}$.

Lemma 4. *Let the class of measures μ_ω be ψ-mixing. Let $\{A_n \in \mathcal{A}_n : n\}$ be a sequence of cylinders and $\{M_n < n : n\}$ a sequence of integers so that for all large n, $A_n \cap \sigma^{-\ell} A_n \neq \emptyset$ for $\ell < M = M_n$ implies that ℓ is a multiple of some given integer m.*

Then there exists a constant K_1 so that

$$\sum_{\vec{v} \in R_r} \mu_\omega(C_{\vec{v}}) \leq K_1 \gamma^{r-1} \sum_{j=2}^r \sum_{s=1}^{j-1} \binom{j-1}{s-1} (\delta \eta_1^M)^{j-s} \frac{\mu_\omega(\zeta_n^x)^s}{s!} \binom{r-1}{j-1} (\alpha' \eta_1^m)^{r-j},$$

for $\delta = \delta_n > n$ and R_r as above, where $C_{\vec{v}} = \bigcap_{k=1}^r \sigma^{-v_k} A_n$, $\gamma = 1 + \psi(\delta - n)$ and $\alpha' > 1 + \psi(0)$.

Proof. Put $R_{r,j}^s$ for those $\vec{v} \in R_{r,j}$ for which $v_{i+1} - v_i \geq \delta$ for $s-1$ indices $i_1, \dots, i_{s-1}$ and $i_s = r$ $(s \leq j-1)$ indicate the tails of the blocks which are followed by a large gap. Similarly we put $R_{r,j,u}^s$ for the set $R_{r,j}^s \cap R_{r,j,u}$. We consider two separate cases: (A) $s \geq 2$ and (B) $s = 1$.

(A) Assume $s \geq 2$ and $i'_1, i'_2, \dots, i'_j$ be the j tails of blocks $(i'_j = r)$ which are characterised by $v_{i'_k+1} - v_{i'_k} \geq M$ for $k = 1, \dots, j-1$ (and $v_{i'_j} = v_r$). We have $\{i_k : k\} \subset \{i'_k : k\}$ where the $j-s$ many indices in $\{i'_k : k\} \setminus \{i_k : k\}$ mark the gaps which are $\geq M$ and smaller than δ. Moreover, the remaining $r-j$ return times are immediate short returns of lengths $\in [m, M)$. Let us put

$$\begin{aligned} W_{i'_k} &= A_m \cap \sigma^{-m} A_n \cap \sigma^{-2m} A_n \cap \cdots \cap \sigma^{-(u_k-1)m} A_n \cap \sigma^{-u_k m} A_n \\ &= A_m \cap \sigma^{-m} A_m \cap \sigma^{-2m} A_m \cap \cdots \cap \sigma^{-(u_k-1)m} A_m \cap \sigma^{-u_k m} A_n, \end{aligned}$$

where u_k is the overlap for the kth block. The ψ-mixing property yields the following estimate

$$\mu_\omega(W_{i'_k}) \leq \alpha_2^{u_k m} \left(\prod_{\ell=0}^{u_k} \mu_{\theta^{\ell m}\omega}(A_m) \right) \mu_{\theta^{u_k m}\omega}(A_n) \leq (\alpha_2 \eta_1^m)^{u_k} \mu_{\theta^{u_k m}\omega}(A_n),$$

where $\alpha_2 = 1 + \psi(0)$ and where we used that (see the Remark 3) $\mu_{\omega'}(A_m) \leq \eta_1^m$ for any ω'. By Lemma 2

$$\begin{aligned} \mu_\omega\left(C_{\vec{v}}\right) &\leq \mu_\omega\left(\bigcap_{i=k}^{j} \sigma^{-(v_{i'_k} - u_k)} W_{i'_k} \right) \\ &\leq \gamma^{s-1} \alpha_2^{j-s} \prod_{i=1}^{j} \mu_{\theta^{v_{i'_k} - u_k}\omega}(W_{i'_k}) \\ &\leq \gamma^{s-1} \alpha_2^{j-s} (\eta_1^M)^{j-s} (\alpha_2 \eta_1^m)^u \prod_{k=1}^{s} \mu_{\theta^{v_{i_k}}\omega}(A_n), \end{aligned}$$

where $\gamma = 1 + \psi(\delta - n)$, and the components of $\vec{n} = (n_1, \ldots, n_r)$ as in Lemma 2 are given by $n_{i_k} = n$ for $k = 1, \ldots, s$ (for the long returns between clusters, i.e., $> \delta$) and $n_i = M$ for $i \neq i_k,\ k = 1, \ldots, s$, where $u = \sum_i u_i$ is the total overlap. We have used that $\mu_\omega(A_M) \leq \eta_1^M$ for any ω.

To count the number of return times vectors, note that there are $\binom{r-1}{j-1}$ many possibilities to choose the j positions $i'_1, \ldots, i'_j$ of the returns $> M$. Out of those we can pick in $\binom{j-1}{s-1}$ many ways the long return ($\geq \delta$) positions $i_1, \ldots, i_s$. Moreover, each choice allows for $< \delta^{j-s}$ many ways to fill in the actual $j - s$ many intermediate return times (between M and δ).

For every fixed set of j returns larger than M and for a fixed value of overlaps u there are $\binom{u-1}{r-j-1}$ many ways to distribute the u overlaps into the remaining $r - j$ many returns which are shorter than M.

For each fixed set of long ($\geq \delta$) return times $v_{i_1}, \ldots, v_{i_s}$ and given value of overlaps u there are consequently

$$\binom{j-1}{s-1} \binom{r-1}{j-1} \delta^{j-s} \binom{u-1}{r-j-1}$$

many possibilities. We thus obtain:

$$\begin{aligned} &\sum_{\vec{v} \in R^s_{r,j,u}} \mu_\omega(C_{\vec{v}}) \\ &\quad \leq \binom{j-1}{s-1} \binom{r-1}{j-1} \delta^{j-s} \binom{u-1}{r-j-1} \gamma^{s-1} (\eta_1^M)^{j-s} (\alpha_2 \eta_1^m)^u \sum_{v_{i_1} < \cdots < v_{i_s} \leq N} \prod_{k=1}^{s} \mu_{\theta^{v_{i_k}}\omega}(A_n) \\ &\quad \leq \binom{j-1}{s-1} \binom{r-1}{j-1} \delta^{j-s} \binom{u-1}{r-j-1} \gamma^{s-1} (\eta_1^M)^{j-s} (\alpha_2 \eta_1^m)^u \frac{1}{s!} \left(\sum_{i=1}^{N} \mu_{\theta^i\omega}(A_n) \right)^s. \end{aligned}$$

Therefore, since $\mu_\omega(\zeta_n^x) = \sum_{i=1}^N \mu_{\theta^i\omega}(A_n)$,

$$\sum_{\vec{v}\in R^s_{r,j,u}} \mu_\omega(C_{\vec{v}}) \le \gamma^{s-1}\binom{j-1}{s-1}\binom{r-1}{j-1}(\delta\eta_1^M)^{j-s}\binom{u-1}{r-j-1}(\alpha_2\eta_1^m)^u\frac{\mu_\omega(\zeta_n^x)^s}{s!}.$$

(B) If $s=1$ then all returns between blocks are less than δ for all k. In the same way as above we obtain

$$\sum_{\vec{v}\in R^1_{r,j,u}} \mu_\omega(C_{\vec{v}}) \le (\delta\eta_1^M)^{j-1}\binom{r-1}{j-1}\binom{u-1}{r-j-1}(\alpha_2\eta_1^m)^u\mu_\omega(\zeta_n^x).$$

Summing over s and using the estimates from (A) and (B) yields

$$\begin{aligned}
\sum_{\vec{v}\in R_r} \mu_\omega(C_{\vec{v}}) &= \sum_j\sum_{s=1}^{j-1}\sum_{u=r-j}^{\infty}\sum_{\vec{v}\in R^s_{r,j,u}}\mu_\omega(C_{\vec{v}}) \\
&\le \sum_{j=2}^r \gamma^{s-1}\sum_{s=1}^{j-1}\binom{j-1}{s-1}\frac{\mu_\omega(\zeta_n^x)^s}{s!}(\delta\eta_1^M)^{j-s}\binom{r-1}{j-1}\sum_{u=r-j}^{\infty}\binom{u-1}{r-j-1}(\alpha_2\eta_1^m)^u \\
&\le \sum_{j=2}^r \gamma^{s-1}\sum_{s=1}^{j-1}\binom{j-1}{s-1}\frac{\mu_\omega(\zeta_n^x)^s}{s!}(\delta\eta_1^M)^{j-s}\binom{r-1}{j-1}\left(\frac{\alpha_2\eta_1^m}{1-\alpha_2\eta_1^m}\right)^{r-j}
\end{aligned}$$

(as $\sum_{u=q}^\infty \binom{u-1}{q-1}x^u = \left(\frac{x}{1-x}\right)^q$). The lemma now follows since $\frac{\alpha_2\eta_1^m}{1-\alpha_2\eta_1^m} \le \alpha'\eta_1^m$ with an α' slightly larger than α_2. ∎

4. Distribution near periodic points for ψ-mixing measures

We will also need the almost sure convergence of $\zeta_n^x = \sum_{j=1}^{N_n}\chi_{A_n}\circ\sigma^j$ (where $N_n = [t/\mu(A_n(x))]$) which is proved in [RSV, Lemma 9]. The following lemma requires the Assumption (iii).

Lemma 5. *If there is $q > 2\frac{\log\eta_1}{\log\eta_0}$ such that $\psi(k)k^q \to 0$ as $k\to\infty$, then $\mu_\omega(\zeta_n^x)\to t$ for $\mathbb{P}$-almost every ω.*

Let us put $\zeta_{n,u}^x = \sum_{k=0}^N \chi_{A_{n+mu}}\circ\sigma^k$, where $N = \frac{t}{\mu(A_n)}$. We will assume that the limit

$$\vartheta(x) = \lim_{n\to\infty}\frac{\mu(A_{n+m}(x))}{\mu(A_n(x))} \tag{1}$$

exists. Then $\mu(\zeta_{n,u}^x) = N\mu(A_{n+mu}) = \frac{\mu(A_{n+mu})}{\mu(A_n)}t$ converges to $\vartheta^u t$ as $n\to\infty$. By the same argument as in Lemma 5 we conclude the following result of which Lemma 5 is the special case $u=0$.

Corollary 6. *If $\psi(k)k^q \to 0$ as $k\to\infty$ for some $q > 2\frac{\log\eta_1}{\log\eta_0}$ and the limit $\vartheta(x) = \lim_{n\to\infty}\frac{\mu(A_{n+m}(x))}{\mu(A_n(x))}$ exists, then*

$$\mu_\omega(\zeta_{n,u}) \to \vartheta^u t$$

as $n\to\infty$ for $\mathbb{P}$-almost every ω.

Although for a periodic point x with (minimal) period m the limit $\lim_{\ell\to\infty}\frac{1}{\ell}|\log\mu(A_{\ell m}(x))|$ always exists (see Lemma 7 of [HV2]), we cannot necessarily conclude that the limit $\vartheta=\lim_{n\to\infty}\frac{\mu(A_{n+m}(x))}{\mu(A_n(x))}$ exists.

For $t>0$ and integers n we put ζ_n^t for the counting function $\sum_{j=0}^{N_n}\chi_{A_n(x))}\circ\sigma^j$ with the observation time $N_n=[t/\mu(A_n(x))]$ (where x is periodic with minimal period m). For equilibrium states for Hölder continuous potentials f (with zero pressure) on Axiom A systems, Hirata [Hi] has shown that $\vartheta(x)=\exp\sum_{j=1}^m f(\sigma^j x)$ for periodic points x with minimal period m, see Example 5.3.

In order to satisfy the assumptions of Proposition 1 we put $\gamma=\alpha$, $\gamma_1=\alpha\delta_n\eta^M$ and $\gamma_2=\alpha\eta^m$

Theorem 7. *Suppose that we have a random SFT driven by an invertible ergodic measure preserving system $(\Omega,\theta,\mathbb{P})$ with marginal measure μ and satisfying conditions (i)–(iv), where the function ϕ is such that there is $q>2\frac{\log\eta_1}{\log\eta_0}$ with $\psi(k)k^q\to0$ as $k\to\infty$. Let $x\in\Sigma$ a periodic point with minimal period m and assume the limit defining ϑ exists and let $\Theta=1-\vartheta$.*

Then

$$\mu_\omega(\zeta_n^x=r)\longrightarrow e^{-t}P_r(\Theta t,\vartheta)$$

as $n\to\infty$ for $\mathbb{P}$-a.e. ω.

Proof. We use Proposition 1 and have to verify conditions (a) and (b).

Let $\vec v\in G_{r,j}\setminus R_{r,j}$ and let $v_{i_1},v_{i_2},\dots,v_{i_j}$ be the heads of the blocks of short returns, that is $i_1=1$ and $v_{i_k}-v_{i_k-1}\ge\delta$ for $k=2,\dots,j$. Moreover $v_\ell-v_{\ell-1}\le M$ for $\ell\notin\{i_1,\dots,i_j\}$. By Lemma 2 ($\Delta$ is as defined in section 2.2)

$$\left|\mu_\omega(C_{\vec v})-\prod_{k=1}^j\mu_{\theta^{v_{i_k}}\omega}(A_{n+mu_k})\right|\le((1+\psi(\Delta(\vec v)-n))^j-1)\prod_{k=1}^j\mu_{\theta^{v_{i_k}}\omega}(A_{n+mu_k}),$$

where u_k is the overlap of the kth block beginning with v_{i_k}. Notice that for the kth cluster one has

$$A_{n+mu_k}=\bigcap_{\ell=i_k}^{i_{k+1}-1}\sigma^{-v_\ell}A_n.$$

Thus

$$\mu_\omega(C_{\vec v})=(1+\mathcal{O}(j\psi(\delta)))\prod_{k=1}^j\mu_{\theta^{v_{i_k}}\omega}(A_{n+mu_k})$$

and summing over the total overlaps (see (III)) yields for the principal term

$$S_r(n):=\sum_{\vec v\in G_r\setminus R_r}\mu_\omega(C_{\vec v})=\sum_{j=1}^r(1+\mathcal{O}(j\psi(\delta)))\sum_{u\ge r-j}\sum_{\vec v\in G_{r,j,u}\setminus R_{r,j,u}}\prod_{k=1}^j\mu_{\theta^{v_{i_k}}\omega}(A_{n+mu_k}).$$

As there are $\binom{r-1}{j-1}$ many ways to choose the indices $i_1,\dots,i_j$, we can now write for the last sum on the RHS:

$$\sum_{\vec v\in G_{r,j,u}\setminus R_{r,j,u}}\prod_{k=1}^j\mu_{\theta^{v_{i_k}}\omega}(A_{n+mu_k})=\binom{r-1}{j-1}\sum_{u_1+\cdots+u_j=u}\sum_{\vec w\in G_j}\prod_{k=1}^j\mu_{\theta^{w_i}\omega}(A_{n+mu_k})+E_{r,j,u},$$

where the error splits into two parts: $E_{r,j,u} = E'_{r,j,u} + E''_{r,j,u}$. We now estimate the two last error terms separately, where (a) $E'_{r,j,u}$ accounts for the over counts of short returns on the RHS and (b) $E''_{r,j,u}$ accounts for the contribution made by $\vec{v} \in R_{r,j,u}$ which are included on the RHS but are excluded on the LHS.

The summation over $u_1, \dots, u_j$ is such that the total overlap u has been divided into $r-j$ non-empty sections for the short returns and then are clustered into the j clusters where some of the u_k might be zero which happens when there is no short return in the associated cluster (i.e. $i_{k+1} = i_k + 1$).

(a): The error term $E'_{r,j,u}$ accounts for the over counting of those return combinations that do not occur since not all u_k are included in our sums. For $\vec{v} \in G_r$ the short overlaps cannot be larger than n which limits the vales and multiplicities for the overlaps u_k. On the RHS however we impose no such restriction and thus have to correct for it with the error term $E'_{r,j,u}$. Since every overlap u_k has to be generated by at least one of the $r-j$ short returns each of which in its turn is bounded by M we therefore obtain for $E'_{r,j}(n) = \sum_j E'_{r,j,u}$ the following upper bound

$$
\begin{aligned}
|E'_{r,j}(n)| &\le \sum_{u \ge r-j} \sum_{\substack{u_1+\dots+u_j=u \\ \min_k u_k \ge \frac{M}{m}}} \frac{1}{j!} \prod_{k=1}^{j} \mu_\omega(\zeta^x_{n,u_k}) \\
&\le \sum_{u \ge \frac{M}{m}} \sum_{u_1+\dots+u_j=u} \frac{1}{j!} \prod_{k=1}^{j} \mu_\omega(\zeta^x_{n,u_k}) \\
&\le c_1 t^j \sum_{u \ge \frac{M}{m}} \binom{u+j-1}{j-1} \vartheta^u
\end{aligned}
$$

for n large enough. Hence $E'_r(n) = \sum_j \binom{r-1}{j-1} E'_{r,j} \longrightarrow 0$ as $n \to \infty$ (as $\frac{M}{m} > \frac{n}{2m} \to \infty$) for almost every ω as the sum on the RHS is a tail sum of $\sum_{u=0}^{\infty} \binom{u+j-1}{j-1} \vartheta^u = (1-\vartheta)^{-j}$.

(b) The second part is given by $E''_{r,j,u} = \sum_{\vec{v} \in R_{r,j,u}} \prod_{k=1}^{j} \mu_{\theta^{w_i}\omega}(A_{n+mu_k})$. For n large enough we obtain as at the end of the proof of Lemma 4,

$$
\begin{aligned}
|E''_r| &\le c_2 \sum_{j=2}^{r} \binom{r-1}{j-1} \sum_{s=1}^{j-1} \binom{j-1}{s-1} \frac{t^s}{s!} \left(\frac{\delta t}{N}\right)^{j-s} \sum_{u=r-j}^{\infty} \binom{u-1}{r-j-1} \vartheta^u \\
&\le c_2 \sum_{j=2}^{r} \binom{r-1}{j-1} \sum_{s=1}^{j-1} \binom{j-1}{s-1} \frac{t^s}{s!} \left(\frac{\delta t}{N}\right)^{j-s} \left(\frac{\vartheta}{1-\vartheta}\right)^{r-j}.
\end{aligned}
$$

Since $j - s \ge 1$ we get that $E''_r \to 0$ as $n \to \infty$ provided $\delta/N \to 0$.

Combining the estimates for these two error terms we conclude that $E_r = E'_r + E''_r \to 0$ as $n \to \infty$ provided $\delta/N \to 0$.

We thus obtain

$$
S_r(n) = \sum_j \binom{r-1}{j-1} (1 + \mathcal{O}(j\psi(\delta))) \sum_{u \ge r-j} \sum_{u_1+\dots+u_j=u} \frac{1}{j!} \prod_{k=1}^{j} \left(\sum_{w=1}^{N} \mu_{\theta^w\omega}(A_{n+mu_k}) \right) + E_r,
$$

and let $\delta \to \infty$ with $n \to \infty$. Hence we obtain for almost every ω the innermost sum is $\mu_\omega(\zeta^x_{n,u_k}) \to t\vartheta^{u_k}$ and consequently get that the principal term converges as follows:

$$\begin{aligned} S_r(n) \longrightarrow& \sum_j \binom{r-1}{j-1} \sum_{u \geq r-j} \binom{u-1}{r-j-1} \frac{1}{j!} t^j \vartheta^u \\ &= \sum_j \frac{t^j}{j!} \binom{r-1}{j-1} \left(\frac{\vartheta}{1-\vartheta}\right)^{r-j} \\ &= \frac{1}{(1-\vartheta)^r} \sum_j \frac{(\Theta t)^j}{j!} \binom{r-1}{j-1} \vartheta^{r-j}. \end{aligned}$$

The combinatorial factor $\binom{u-1}{r-j-1}$ expresses the number of ways in which u overlaps are distributed into $r-j$ short returns $\leq M$ (and which are then clustered into j clusters where some of them might be empty). This implies

$$S_r(n) \longrightarrow Q_r(\Theta t, \vartheta)$$

and thus verifies condition (a) of Proposition 1.

To verify assumption (b) we obtain

$$\sum_{\vec{v} \in R_r} \mu_\omega(C_{\vec{v}}) \leq K_1 \alpha^{r-1} \sum_{j=2}^{r} \sum_{s=1}^{j-1} \binom{j-1}{s-1} (\delta \eta^M)^{j-s} \frac{\mu_\omega(\zeta^x_n)^s}{s!} \binom{r-1}{j-1} (\alpha \eta^m)^{r-j},$$

where $\alpha = 1 + \psi(0)$. Hence condition (II) of Proposition 1 is satisfied since $j - s \geq 1$. ∎

5. Examples

Let us note that the recurrence properties at periodic points or otherwise do not require the entropy to be finite (which is necessary in the theorem of Shannon-McMillan-Breiman for example). To make this point we will give below an example of infinite entropy

5.1. **Two-element Bernoulli shifts.**

Some classes of examples of random SFTs to which our results apply was given in [RT, Section 6], including examples in the infinite alphabet case. However, to give the reader some idea of systems to which our methods apply, we first give an elementary example (this is a simple version of [RSV, Example 19]).

Let $\Omega = \Sigma = \{0,1\}^{\mathbb{N}_0}$ and let $\mathbb{P}$ be a Gibbs measure on Ω (i.e., $\mathbb{P}$ need not be Bernoulli or Markov). In this case the fibre Σ_ω is equal to Σ for every ω (in particular the transition matrix $A = A(\omega)$ is the full matrix of 1s. Then fixing $\alpha, \beta \in (0,1)$, for $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$, let

$$p(\omega) = \begin{cases} \alpha & \text{if } \omega_0 = 0, \\ \beta & \text{if } \omega_0 = 1. \end{cases}$$

Then we can define a random Bernoulli measure by

$$\mu_\omega[x_0, \ldots, x_n] = p_{x_0}(\omega) p_{x_1}(\theta\omega) \cdots p_{x_n}(\theta^n \omega)$$

where

$$p_{x_i}(\omega) = \begin{cases} p(\omega) & \text{if } x_i = 0, \\ 1 - p(\omega) & \text{if } x_i = 1. \end{cases}$$

As shown in [RSV, Example 19], this system satisfies conditions (i)–(iv), so Theorem 7 holds. Moreover we can give a formula for the parameter ϑ explicitly: if $x \in \Sigma$ is a periodic point of period m, the Bernoulli property of our sample measures and θ-invariance of $\mathbb{P}$ allow us to compute that

$$\vartheta(x) = \int p_{x_0}(\omega)p_{x_1}(\theta\omega)\cdots p_{x_{m-1}}(\theta^{m-1}\omega)\ d\mathbb{P}(\omega).$$

5.2. **i.i.d. infinite alphabet systems.** Let us now consider a stochastic system which is a shift space Σ over a countably infinite alphabet with a Bernoulli measure which we below choose to obtain infinite entropy. The 'driving space' Ω will be the infinite product of intervals.

Let I be a measurable space with a measure m and let $\Omega = I^{\mathbb{N}_0}$ be equipped with the product measure $\mathbb{P}$. Let $\Sigma = \mathbb{N}^{\mathbb{N}_0}$ with the left shift map σ be the full shift over a countable alphabet and $\vec{p} : \Omega \to (0,1)^{\mathbb{N}}$ a function that depends only on the zeroth coordinate, i.e. $p(\omega) = p(\omega_0)$, and satisfies $\sum_{n=1}^{\infty} p_n(\omega_0) = 1$ for all ω_0, where p_n are the components of $\vec{p}$. Assume that $\sup_{\omega_0,n} p_n(\omega_0) < 1$. For every $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$ we thus obtain a Bernoulli measure μ_ω on Σ defined by

$$\mu_\omega(x_0, \ldots, x_n) = p_{x_0}(\omega)p_{x_1}(\theta\omega)\cdots p_{x_n}(\theta^n\omega).$$

Clearly, μ_ω satisfies the Assumptions (i) and (iv). The marginal measure μ is Bernoulli with weights $\bar{p}_n = \int_\Omega p_n(\omega_0)\, d\mathbb{P}(\omega)$ and consequently ψ-mixing. Assumption (iii) therefore need not be met, so the conclusions of Theorem 7 hold..

To compute ϑ, if $x \in \Sigma$ is a periodic point of minimal period m then, as before,

$$\vartheta(x) = \int_\Omega p_{x_0}(\omega)p_{x_1}(\theta\omega)\cdots p_{x_{m-1}}(\theta^{m-1}\omega)\, d\mathbb{P}(\omega) = \prod_{j=0}^{m-1} \bar{p}_{x_j} = \mu(x_0x_1\cdots x_{m-1})$$

by θ-invariance of $\mathbb{P}$ and since $\mathbb{P}(\omega_0\omega_1\cdots\omega_{m-1}) = \prod_{j=0}^{m-1}\mathbb{P}(\omega_j)$. Equivalently, if $N_s = |\{j \in \{0,\ldots,m-1\} : x_j = s\}|$ denotes the number of times the symbol $s \in \mathbb{N}$ occurs during a period of x, then

$$\vartheta(x) = \prod_s \bar{p}_s^{N_s}.$$

5.2.1. *Two-element Bernoulli revisited.* In the special case above when the two element alphabet $\{0,1\}$ is used we can equip $\Omega = \{0,1\}^{\mathbb{N}_0}$ with the Bernoulli measure with weights $p, q = 1-p$, $p, q > 0$. Then $\bar{p}_0 = \alpha p + \beta q$ and $\bar{p}_1 = (1-\alpha)p + (1-\beta)q$ and consequently

$$\vartheta = (\alpha p + \beta q)^\ell((1-\alpha)p + (1-\beta)q)^{m-\ell},$$

where ℓ is the number of times the symbol 0 occurs on the periodic string $x_0\cdots x_{m-1}$. In the deterministic case when $\alpha = \beta$ we obtain $\vartheta = \alpha^\ell(1-\alpha)^{m-\ell}$ which is in accordance with Hirata's result.

5.2.2. *Infinite entropy system.* Let us now choose $I=[\varepsilon,1]$ with the normalised Lebesgue measure and $\varepsilon\in(0,1)$. One can then define a family of probability vectors $\{\vec{p}(\omega)\}_{\omega\in\Omega}$ by

$$p_n(\omega)=\frac{G(\omega_0)}{n\log^{1+\omega_0}n},$$

if $n\geq 3$, and equal to 0 if $n=1,2$, where $G(\omega_0)\in\mathbb{R}_+$ is a normalising constant for every $\omega_0\in I$. The marginal measure μ is Bernoulli with weights $\bar{p}_n=\int_\varepsilon^1\frac{G(t)}{n\log^{1+t}n}\frac{dt}{1-\varepsilon}$ where $G(t)\sim\frac{1}{t}$. Its entropy is infinite since

$$\begin{aligned}h(\mu) &= -\sum_n\int_\varepsilon^1\frac{G(t)}{n\log^{1+t}n}\frac{dt}{1-\varepsilon}\log\int_\varepsilon^1\frac{G(t)}{n\log^{1+t}n}\frac{dt}{1-\varepsilon}\\ &\geq \sum_n\int_\varepsilon^1\frac{G(t)}{n\log^{1+t}n}\frac{dt}{1-\varepsilon}\left(\log n+\log\log n-\log\int_\varepsilon^1 G(t)\frac{dt}{1-\varepsilon}\right)\\ &\geq \sum_n\int_\varepsilon^1\frac{G(t)}{n\log^{t}n}\frac{dt}{1-\varepsilon}-c_1\sum_n\int_\varepsilon^1\frac{G(t)}{n\log^{1+t}n}\frac{dt}{1-\varepsilon}=\infty\end{aligned}$$

as the second term converges, where we used that $\frac{1}{\log^{1+t}n}\leq\frac{1}{\log n}$ and $c_1=\log\int_\varepsilon^1 G(t)\frac{dt}{1-\varepsilon}$.

Again, if $x\in\Sigma$ is periodic with minimal period m then $\vartheta(x)=\bar{p}_{x_0}\bar{p}_{x_1}\cdots\bar{p}_{x_{m-1}}$.

5.3. **Equilibrium states for Axiom A systems.** It is standard to code Axiom A dynamical systems to understand their ergodic properties, see [Bo]. In this case, a Hölder potential gives rise to a Gibbs state. Here we briefly discuss the random case, going directly to the symbolic setting, and show how our results apply here. We emphasise that in our exposition, we fix the dynamics and put the randomness in the measure.

Let (Ω,θ) be a two-sided shift with an invariant measure $\mathbb{P}$. We assume (Σ,σ) is a topologically mixing subshift of finite type and $\{f_\omega\}_{\omega\in\Omega}$ a family of Hölder continuous functions on Σ whose Hölder norms are uniformly bounded. For some $\hat{\kappa}\in(0,1)$, the $\hat{\kappa}$-Hölder norm of a function $f:\Sigma\to\mathbb{R}$ is given by $\|f\|=|f|_\infty+\sup_n\hat{\kappa}^{-n}\text{var}_nf$, where $\text{var}_nf=\sup_{x_i=y_i:|i|<n}|f(x)-f(y)|$ is the n-variation of f.

By [K] there exist random equilibrium states μ_ω that satisfy the generalised invariance property $\sigma\mu_\omega=\mu_{\theta\omega}$. The fibred measures μ_ω are Gibbs with respect to f_ω. We can assume that the functions f_ω have zero pressure. We conclude from the Gibbs property that the fibred measures are ψ-mixing where ψ decays exponentially at some rate $\kappa<1$. Recall that (see e.g. [Bo]) $\mu_\omega=h_\omega\nu_\omega$ where h_ω are the normalised eigenfunction for the largest eigenvalue of the transfer operator and ν_ω are the associated eigenfunctionals which are e^{-f_ω}-conformal, i.e. if σ is one-to-one on a set $A\subset\Sigma$ then $\nu_\omega(\sigma A)=\int_A e^{-f_\omega}\,d\nu_\omega(x)$. If we replace f by the normalised function $\tilde{f}_\omega=f_\omega+\log h_\omega-\log h_\omega\circ\sigma$ then μ_ω is $e^{-\tilde{f}_\omega}$-conformal. We now assume that μ_ω is $e^{-\tilde{f}_\omega}$-conformal for every ω. Moreover, if $U(\alpha)$ is the n-cylinder which is determined by the n-word α, then

$$\mu_\omega(U(\alpha))=\int\chi_{U(\alpha)}\,d\mu_\omega=\int\mathcal{L}_\omega^n\chi_{U(\alpha)}\,d\mu_\omega=\int e^{f_\omega^n(\alpha x)}\,d\mu_\omega(x),$$

where $\mathcal{L}_\omega$ is the transfer operator for the normalised function f_ω, $f_\omega^n=\sum_{j=0}^{n-1}f\circ\sigma$ is the n-th ergodic sum of f_ω and αx is the concatenation of α and x (subject to the transition rules).

Let $\mu=\int\mu_\omega\,d\mathbb{P}(\omega)$ be the marginal measure. We now want to establish that μ is ψ-mixing and for this purpose assume that the family of functions $\{f_\omega\}_{\omega\in\Omega}$ satisfies the following additional regularity assumption with respect to ω: Assume there exists a constant K so that for every n:

$$|f_\omega-f_{\omega'}|_\infty\le K\hat\kappa^n$$

for $\omega,\omega'\in\Omega$ for which $\omega_i=\omega'_i\forall|i|\le n$. The supremum norm is over Σ.

Let α be an n-word in Σ and denote by $V=U(\alpha)\subset\Sigma$ the associated n-cylinder. Similarly for an m-word β we write $W=U(\beta)$ for the associated m-cylinder. Then

$$\begin{aligned}\mu(V\cap\sigma^{-n-k}W) &= \int_\Omega\mu_\omega(V\cap\sigma^{-n-k}W)\,d\mathbb{P}(\omega)\\ &= \int_\Omega\mu_\omega(V)\mu_{\theta^{-n-k}\omega}(W)(1+\mathcal{O}(\kappa^k))\,d\mathbb{P}(\omega)\\ &= (1+\mathcal{O}(\kappa^k))\int_\Omega\int_\Sigma e^{f^n_\omega(\alpha x)}\,\mu_\omega(x)\int_\Sigma e^{f^m_{\theta^{-n-k}\omega}(\beta y)}\,d\mu_{\theta^{-n-k}\omega}(y)\,d\mathbb{P}(\omega).\end{aligned}$$

Define $\omega^{(n,k)}(\omega)\in\Omega$ by putting $\omega^{(n,k)}_i=\omega_i$ for $i\le n+\frac{k}{2}$ and $\omega^{(n,k)}_i=\omega_{i-(n+\frac{k}{2})}$ for $i>n+\frac{k}{2}$. This implies for all $x\in\Sigma$

$$f^n_\omega(\alpha x)-f^n_{\omega^{(n,k)}}(\alpha x)=\mathcal{O}(\kappa^{\frac{k}{2}}),$$

and for all y

$$f^m_{\theta^{-n-k}\omega}(\beta y)-f^m_{\theta^{-n-k}\omega^{(n,k)}}(\beta y)=\mathcal{O}(\kappa^{\frac{k}{2}}).$$

Hence

$$\begin{aligned}&\mu(V\cap\sigma^{-n-k}W)\\ &= (1+\mathcal{O}(\hat\kappa^{\frac{k}{2}}))\int_\Omega\int_\Sigma e^{f^n_{\omega^{(n,k)}}(\alpha x)}\,\mu_\omega(x)\int_\Sigma e^{f^m_{\theta^{-n-k}\omega^{(n,k)}}(\beta y)}\,d\mu_{\theta^{-n-k}\omega}(y)\,d\mathbb{P}(\omega)\\ &= (1+\mathcal{O}(\hat\kappa^{\frac{k}{2}}))\int_\Omega\int_\Sigma e^{f^n_{\omega^{(n,k)}}(\alpha x)}\,\mu_\omega(x)\,d\mathbb{P}(\omega)\int_\Omega\int_\Sigma e^{f^m_{\theta^{-n-k}\omega^{(n,k)}}(\beta y)}\,d\mu_{\theta^{-n-k}\omega}(y)\,d\mathbb{P}(\omega).\end{aligned}$$

Hence replacing the the modified $\omega^{(n,k)}$ again by ω we obtain

$$\int_\Omega\int_\Sigma e^{f^n_{\omega^{(n,k)}}(\alpha x)}\,\mu_\omega(x)\,d\mathbb{P}(\omega)=(1+\mathcal{O}(\hat\kappa^{\frac{k}{2}}))\int_\Omega\int_\Sigma e^{f^n_\omega(\alpha x)}\,\mu_\omega(x)\,d\mathbb{P}(\omega)=(1+\mathcal{O}(\hat\kappa^{\frac{k}{2}}))\mu(U(\alpha)),$$

and similarly

$$\int_\Omega\int_\Sigma e^{f^m_{\theta^{-n-k}\omega^{(n,k)}}(\beta y)}\,d\mu_{\theta^{-n-k}\omega}(y)=(1+\mathcal{O}(\hat\kappa^{\frac{k}{2}}))\mu(U(\beta)).$$

We thus obtain

$$\mu(V\cap\sigma^{-n-k}W)=(1+\mathcal{O}(\kappa'^k))\mu(V)\mu(W),$$

that is, the marginal measure μ is ψ-mixing at rate $\kappa'=\max\{\kappa,\sqrt{\hat\kappa}\}$.

Let us note that condition (iii) is satisfied since μ is ψ-mixing.

Thus, if x is a periodic point with minimal period m, then with $\alpha = x_0 \cdots x_{m-1}$,

$$\begin{aligned}
\mu(A_{n+m}(x)) &= \int_\Omega \int_\Sigma \chi_{A_{n+m}(x)}\, d\mu_\omega\, d\mathbb{P}(\omega) \\
&= \int_\Omega \int_\Sigma e^{f_\omega^m(\alpha y)} \chi_{A_n(x)}(y)\, d\mu_\omega(y)\, d\mathbb{P}(\omega) \\
&= (1+\mathcal{O}(\kappa^n)) \int_\Omega e^{f_\omega^m(x)} \int_\Sigma \mu_\omega(A_n(x)))\, d\mathbb{P}(\omega).
\end{aligned}$$

In particular the limit

$$\vartheta = \lim_{n\to\infty} \frac{\mu(A_{n+m}(x))}{\mu(A_n(x))}$$

exists and converges exponentially: $\frac{\mu(A_{n+m}(x))}{\mu(A_n(x))} = \vartheta + \mathcal{O}(\kappa^n)$. We can then choose $\delta(n)$ to be proportional to $|\log \mu(A_n(x))|$, or equivalently a multiple of n, and obtain the following result.

Theorem 8. *Let (Σ, σ) be an Axiom A system and $\{\mu_\omega\}_{\omega\in\Omega}$ be a family of equilibrium states for Hölder continuous potentials $\{f_\omega\}_{\omega\in\Omega}$ whose Hölder norms are uniformly bounded. We moreover assume that the family of functions $\{f_\omega\}_{\omega\in\Omega}$ is Hölder continuous in ω.*

If $x \in \Sigma$ is periodic with minimal period m, then the value $\vartheta = \lim_{n\to\infty} \frac{\mu(A_{n+m}(x))}{\mu(A_n(x))}$ exists. Moreover for every parameter value $t > 0$ and $r = 0, 1, \ldots$ one has

$$\mathbb{P}(\zeta_n^x = r) \to e^{-t} P_r(\Theta t, \vartheta)$$

as $n \to \infty$ ($\Theta = 1 - \vartheta$).

References

[A] M Abadi: Hitting, returning and the short correlation function; Bull. Braz. Math. Soc. New Series **37(4)** (2006), 1–17.

[AS] M Abadi and B Saussol: Hitting and returning into rare events for all alpha-mixing processes; Stoch. Proc. Appl. **121** (2011), 314–323.

[AV] M Abadi and N Vergne: Sharp Error Terms for Return Time Statistics under Mixing Conditions; J Theor Probab **22** (2009): 18–37.

[Ar] L Arnold, Random dynamical systems. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[AFV] H Aytaç, J M Freitas and S Vaienti: *Laws of rare events for deterministic and random dynamical systems;* Trans. Amer. Math. Soc. **367** (2015) 8229–8278

[BCL] A D Barbour, L H Y Chen and W L Loh: Compound Poisson approximation for nonnegative random variables via Stein's method; Ann. Probab. **20** (1992), 1843–1866.

[Bi] P. Billingsley, Probability and measure. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, 2012. Anniversary edition.

[Bo] R Bowen: *Equilibrium States for Anosov Diffeomorphism*; Springer Lecture Notes 470, Springer, New York/Berlin 1975.

[CR] L H Y Chen and M Roos: Compound Poisson approximation for unbounded function on a group, with application to large deviations; Prob. Th. & Rel. Fields **103** (1995) 515–528.

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

[F] W Feller: *An Introduction to Probability Theory and Its Applications*; Wiley 1950.

[HP] N Haydn and Y Psiloyenis: Return times distribution for Markov towers with decay of correlations; Nonlinearity **27(6)** (2014), 1323–1349.

[HV1] N Haydn and S Vaienti: The limiting distribution and error terms for return times of dynamical systems; Disc. Cont. Dyn. Syst. **10** (2004) 589–616.

[HV2] N Haydn and S Vaienti: The distribution of return times near periodic orbits; Probability Theory and Related Fields **144** (2009), 517–542.

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

[JKW] N L Johnson, S Kotz and A W Kemp: *Univariate Discrete Distributions*, 3rd ed., Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, New York, 2005.

[K] Y Kifer: Limit Theorems for Random Transformations and Processes in Random Environments; Trans. Amer. Math. Soc. **350(4)** (1998), 1481–1518.

[RT] J Rousseau and M Todd: Hitting times and periodicity in random dynamics; J. Stat. Phys. **161** (2015) 131–150.

[RSV] J Rousseau, B Saussol and P Varandas: Exponential law for random subshifts of finite type; Stochastic Process. Appl. **124** (2014), 3260–3276.

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

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

E-mail address: nhaydn@usc.edu

Mike Todd, Mathematical Institute, University of St Andrews, North Haugh, St Andrews, KY16 9SS, Scotland

E-mail address: m.todd@st-andrews.ac.uk

URL: http://www.mcs.st-and.ac.uk/~miket/