Fluctuations of Birkhoff averages of return times for Axiom A diffeomorphisms

Jean-René Chazottes[1], Renaud Leplaideur[2]

[1] Centre de Physique Théorique
CNRS-Ecole polytechnique, UMR 7644
F-91128 Palaiseau Cedex
jeanrene@cpht.polytechnique.fr
[2] Laboratoire de Mathématiques
Université de Bretagne Occidentale
6, rue Victor Le Gorgeu
BP 809
F-29285 Brest Cedex
Renaud.Leplaideur@univ-brest.fr

Abstract

We study the time of nth return of orbits to some given (union of) rectangle(s) of a Markov partition of an Axiom A diffeomorphism. Namely, we prove the existence of a scaled generating function for these returns with respect to any Gibbs measure (associated to a Hölderian potential). As a by-product, we derive precise large deviation estimates and a central limit theorem for Birkhoff averages of return times. We emphasize that we look at the limiting behavior in term of number of visits (the size of the visited set is kept fixed). Our approach relies on the spectral properties of a one-parameter family of induced transfer operators on unstable leaves crossing the visited set.

1 Introduction

Let (X, f) be some dynamical system preserving a probability measure μ and pick an arbitrary Borel set $A \subset X$ of positive μ-measure. By Poincaré's recurrence Theorem, μ-almost every point $x \in A$ comes back infinitely many times in A upon iterations of f. We denote by $\mathbf{r}_A^n(x)$, $n \geq 1$, the time of the nth return of $x \in A$ to A. These times are defined by induction in the following way :

$$\begin{aligned} \mathbf{r}_A^1(x) &\stackrel{\text{def}}{=} \min\{k \geq 1,\ f^k(x) \in A\}, \\ \mathbf{r}_A^{n+1}(x) &\stackrel{\text{def}}{=} \mathbf{r}_A^n(x) + \mathbf{r}_A^1(f^{\mathbf{r}_A^n(x)}(x)). \end{aligned}$$

(For convenience we set $r_A^0 \stackrel{\text{def}}{=} 0$.) Assuming that μ is ergodic, the well-known Kač 's formula (virtually found in any textbook on ergodic theory) tells that $\int \mathbf{r}_A^1 \ d\mu = 1$. We can also introduce the successive return times (also called Poincaré cycles by Kač [13]) of A with respect to $x \in A$ by setting $\tau_A^{n+1} \stackrel{\text{def}}{=} \mathbf{r}_A^{n+1} - \mathbf{r}_A^n$, $n \geq 0$. We obviously have $\mathbf{r}_A^n = \sum_{j=1}^n \tau_A^j$. In a note [2], Birkhoff showed that

$$\lim_{n \to +\infty} \frac{\mathbf{r}_A^n(x)}{n} = \frac{1}{\mu(A)} \quad \text{for } \mu - \text{almost every } x \in A\,. \tag{1}$$

This note was so overshadowed by his subsequent proof of the Ergodic Theorem for any integrable function that it escaped notice. The 'modern' proof of (1) can be found for instance in [13, 16] (where Birkhoff's note is not cited) and it is a simple consequence of the Ergodic Theorem.

In this paper, we ask for the *fluctuations* of $\mathbf{r}_A^n/n$ around $1/\mu(A)$. It is easy to check that the random variables τ_A^n are stationary under the conditional probability measure $\mu_A \stackrel{\text{def}}{=} \mu(\cdot \cap A)/\mu(A)$ which is $f^{\tau_A^1}$-invariant. The τ_A^n's are generally not independent. Let us mention a notable exception: When A is a state of a countable state Markov chain.

We shall prove precise large deviation asymptotics and a central limit theorem for $\mathbf{r}_A^n/n$ when f is an Axiom-A diffeomorphism on a Riemannian manifold M, and μ is the equilibrium state associated to a Hölder continuous potential φ. The visited set A will be a Markov rectangle of some basic set (or a finite union of Markov rectangles).

There are a lot of recent works on return times. In most of them, "rare events" are considered, that is sets A_k such that $\mu(A_k)$ goes to zero as k tends to infinity. Typically, A_k is a cylinder set and one looks at the rescaled return times to A_k. In many dynamical systems with "sufficiently strong" mixing properties, such rescaled returns are shown to be distributed according to a Poisson law as k goes to infinity (see [1] for a recent but not up-to-date review). For a fixed set, the moments of hitting and return times are studied in [6] in the setting of 'strongly mixing' processes. We would like to emphasize that in the present work the asymptotics are taken with respect to the *number of visits* to some *fixed* set A.

Our key-result (Theorem 2.1) is the existence of a kind of "free energy" for the nth return time, as n diverges. In probabilistic terms, we prove the existence of the scaled-generating cumulant function associated to $\mathbf{r}_A^n$. We are able to analyse the properties of this function because we show it is nothing but the logarithm of the largest eingenvalue of some one-parameter family of transfer operators. These transfer operators act on the induced system on an unstable leaf of reference crossing the set A upon consideration. This construction was done in [14] for other purposes than studying return times (namely to construct equilibrium states). Once we have this free energy for successive returns and its properties, we apply two results (that are little used) to get precise large deviation estimates and a central limit theorem for the nth return time. At a more technical level, let us notice that since return times are not continuous functions, we cannot apply the so-called contraction principle of large deviation theory [10] to the empirical measure. We can neither (directly) apply the known central limit asymptotics which are established for Lipschitz continuous functions (see [21] and references therein).

Outline of the article. In Section 2, we state our main result and its consequences. Section 3 is devoted to some preparatory notions and lemmas. The proof of the main result is given in Section 4. We first handle the case when A is a single Markov rectangle. We then show how to extend the result to a finite union of rectangles. In Section 5 we derive our large deviation and central limit theorems.

2 Statement of results

We refer the reader to the book of Bowen (see [3]) for the precise definitions of Axiom-A diffeomorphisms, equilibrium states and basic sets.

Assumptions. Throughout we assume that f is a C^2 Axiom-A diffeomorphism on a compact Riemannian manifold M. Let Ω be a basic set for f and $\varphi : M \to \mathbb{R}$ be a Hölder continuous function. We denote by μ the (unique) equilibrium state associated to φ on Ω. Finally, $\mathcal{R} = \{R_i\}$ denotes a finite Markov partition of Ω (into more than one rectangle). Let $A \subsetneq \Omega$ be some finite union of atoms of the partition $\mathcal{R}$.

The main result that we are going to prove is the following :

Theorem 2.1. *Under the above assumptions, there exists a real number $\alpha_0 = \alpha_0(A) > 0$ such that for every $\alpha < \alpha_0$,*

$$\Psi(\alpha) \stackrel{\text{def}}{=} \lim_{n\to+\infty} \frac{1}{n} \log \mathbb{E}_{\mu_A}\left[e^{\alpha \mathbf{r}_A^n}\right] < +\infty \ .$$

Moreover the map $\alpha \mapsto \Psi(\alpha)$ has the following properties:
1. The map $z \mapsto \Psi(z)$ is analytic in a complex neighborhood of $]-\infty, \alpha_0[$.
2. It is strictly convex increasing on $]-\infty, \alpha_0[$.

We shall show that from this theorem we can deduce the following result.

Theorem 2.2 (Large deviations). *Under the above assumptions, we have the following estimates, for every* $u \in (0, \infty)$

$$\lim_{n\to\infty} \frac{1}{n} \log\ \mu_A \left\{ \frac{\mathbf{r}_A^n}{n} \geq \frac{1}{\mu(A)} + u \right\} = \inf_{\alpha<\alpha_0} \left\{ -\left(\frac{1}{\mu(A)} + u \right) \alpha + \Psi(\alpha) \right\}$$

and for every $0 < u < 1/\mu(A)$

$$\lim_{n\to\infty} \frac{1}{n} \log\ \mu_A \left\{ \frac{\mathbf{r}_A^n}{n} \leq \frac{1}{\mu(A)} - u \right\} = \inf_{\alpha<\alpha_0} \left\{ -\left(\frac{1}{\mu(A)} - u \right) \alpha + \Psi(\alpha) \right\}$$

where $\alpha_0 > 0$ *is the same as in Theorem 2.1.*

We also obtain a central limit theorem by using Theorem 2.1 and applying a result due to Bryc [5]. Notice that in general it is impossible to deduce a central limit theorem from a large deviation principle assuming only that the cumulant generating function is twice differentiable at the origin. Even real-analyticity is not enough (see a counterexample in [5]).

Theorem 2.3 (Central limit theorem). *Under the above assumptions*

$$\lim_{n\to\infty} \mu_A \left\{ \frac{\mathbf{r}_A^n - n/\mu(A)}{\sigma_A \sqrt{n}} \leq t \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{\xi^2}{2}}\ d\xi \tag{2}$$

where

$$\sigma_A^2 = \Psi''(0) = \lim_{n\to\infty} \frac{1}{n} \int \left(\mathbf{r}_A^n - \frac{n}{\mu(A)} \right)^2 d\mu_A \in \,]0, +\infty[\,. \tag{3}$$

Remarks

1. α_0 has an explicit expression: it is the difference of the topological pressures of the system and the topological pressure of the system obtained by 'removing A' from the phase space, see formula (12) below.

2. It is simple to see that the values assumed by Ψ' when α ranges from $-\infty$ to α_0 give the values of the possible deviations u around the mean $1/\mu(A)$. It will be easy to check that $\Psi'(\alpha)$ runs from 0 to $+\infty$. The fact that fluctuations above $1/\mu(A)$ can be arbitrary large is due to the fact that there are points which are typical for (invariant) measures giving arbitrary small weight to A. Notice that $\Psi'(0) = \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}_{\mu_A}[\mathbf{r}_A^n] = \frac{1}{\mu(A)}$. Indeed,

$$\frac{1}{n}\mathbb{E}_{\mu_A} \left(\sum_{j=1}^{n} \tau_A^j \right) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}_{\mu_A}(\tau_A^j) = \mathbb{E}_{\mu_A}(\tau_A^1) = \frac{1}{\mu(A)}\,.$$

The last equality is Kač formula and the fact that $\mathbb{E}_{\mu_A}(\tau_A^j) = \mathbb{E}_{\mu_A}(\tau_A^1)$ for all $j \in \mathbb{N}$ is established in [13] or [16].

3 Preparatory lemmas

In this section, we briefly recall the relevant results from [14] which are useful for the proof of Theorem 2.1 and derive a few lemmas.

3.1 Notations

Let us recall that μ is the unique equilibrium state associated to the potential φ, that is, we have

$$h_\mu(f) + \int \varphi \, d\mu = \sup_\nu \left(h_\nu(f) + \int \varphi \, d\nu \right) = P_{top}(\varphi, \Omega) \tag{4}$$

where the supremum is taken over the set of f-invariant probability measures on Ω. As usual, $h_\nu(f)$ denotes the entropy of the measure ν and $P_{top}(\varphi,\Omega)$ the topological pressure on Ω associated to the potential φ. Let $N \geq 2$ be the number of proper rectangles of the Markov partition $\mathcal{R}$ and $\mathcal{A}$ the $N \times N$-transition matrix defined as

$$a_{ij} = 1 \quad \text{if} \quad f^{-1}(\overset{\circ}{R}_j) \cap \overset{\circ}{R}_i \neq \emptyset$$
$$a_{ij} = 0 \quad \text{otherwise.}$$

Let Σ be the set of sequences $\underline{x} = \{x_n\}_{n\in\mathbb{Z}}$ such that for every n, x_n belongs to $\{1,\dots,N\}$ and $a_{x_n x_{n+1}} = 1$. If σ denotes the shift map on Σ, there exists some canonical map π from Σ onto Ω such that the following diagram commutes:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \pi \downarrow & \circlearrowleft & \downarrow \pi \\ \Omega & \xrightarrow{f} & \Omega \end{array}$$

As the map π is also Hölder continuous, the map $\widetilde{\varphi}$ defined by

$$\widetilde{\varphi} \stackrel{\text{def}}{=} \varphi \circ \pi$$

is again Hölder continuous, and there exists a unique equilibrium state $\widetilde{\mu}$ for the dynamical system (Σ, σ) associated to the potential $\widetilde{\varphi}$. Moreover $\widetilde{\mu} \circ \pi^{-1} = \mu$.

The topological pressure associated to $\widetilde{\varphi}$, $P_{top}(\widetilde{\varphi}, \sigma)$, is equal to $P_{top}(\varphi, \Omega)$. The cylinder set $[i_0, \dots, i_n] \subset \Sigma$, $i_j \in \{1, ..., N\}$, $n \geq 0$, is the set of points $\underline{x}$ such that $x_j = i_j$ (for every $0 \leq j \leq n$).

Let g be the first return map in A:

$$\begin{array}{rcl} g : A & \longrightarrow & A \\ x & \longmapsto & f^{\mathbf{r}^1_A(x)}(x). \end{array}$$

If x is a point in Ω, we denote, as usually, by $W^u(x)$, $W^u_{loc}(x)$, $W^s(x)$, $W^s_{loc}(x)$ the unstable and stable global and local manifolds. Local means that the length is equal to some expansive constant, ε_0. For every set R of small diameter (smaller than ε_0) we set

$$W^i(x,R) \stackrel{\text{def}}{=} W^i_{loc}(x) \cap R, \text{ for } i = u, s.$$

We will assume that the diameter of $\mathcal{R}$ is smaller than ε_0.

3.2 The subsystem (F, g_F)

For the sake of definiteness, we set $A = \pi[1]$. We denote by F some fixed unstable leaf in A; namely we have

$$F = W^u(x_0, A)$$

for some fixed point x_0 in $\overset{\circ}{A}$. The system of local coordinates gives a projection π_F from A onto F. This projection is Hölder continuous. We denote by $g_F : F \circlearrowleft$ the map $\pi_F \circ g$. For x in Ω and x' in $W^s(x)$, we set

$$\omega(x,x') = \sum_{k=0}^{+\infty} \varphi \circ f^k(x) - \varphi \circ f^k(x') \ .$$

The map φ is Hölder continuous, and so, by contraction on the stable leaves, the previous series converges. For x in F we set $\omega(x) = \omega(g(x), g_F(x))$, and

$$\Phi(x) = \sum_{k=0}^{\mathbf{r}^1_A(x)-1} \varphi \circ f^k(x) + \omega(x) \ .$$

This function is defined on a set of full measure with respect to any g_F-invariant measure. A simple computation gives the following lemma

Lemma 3.1. *There exists a constant $C_\omega > 0$ such that*

$$\sup_{x\in F} |\omega| \le C_\omega.$$

The inverses branches of g_F define the family of n-cylinders: for x in F we set

$$C_n(x) \stackrel{\text{def}}{=} f^{-\mathbf{r}^n_A(x)}(W^u(g^n(x), A)).$$

The n-cylinders are well-defined except for the points in F which do not return infinitely many times in R and for the points which belong to the orbit of the boundary $\partial\mathcal{R}$ of the partition. These two sets of points will have null-measure for all the measures we are going to consider. Hence, every n-cylinder is a compact set and the collection of the n-cylinders defines a partition of F (up to the boundary and points which come back less than n times), which refines the partition into $(n-1)$-cylinders. An important property is that $g^n_F(C_n(x)) = F$ for every n-cylinder. This is a consequence of the Markov property of the partition. This allows us to define the set of preimages of some point x in F by

g_F^n, denoted by $Pre_n(x)$. Hence, every n-cylinder contains exactly one element of $Pre_n(x)$, for every x in F. We define the Perron-Frobenius-Ruelle operator

$$\mathcal{L}_S(T)(x) = \sum_{y \in Pre_1(x)} e^{\Phi(y) - \mathbf{r}_A^1(y)S} T(y)$$

where $T : F \to \mathbb{R}$, $x \in F$ and S is a real parameter. It is proved in [14] (def 4.2 and prop 4.3) that there exists a critical value S_c, with $S_c \leq P_{top}(\varphi, \Omega)$, such that

$$\mathcal{L}_S(\mathbb{1}_F)(x) < +\infty \quad \text{for every } S > S_c \text{ and for every } x \in F\,. \tag{5}$$

S_c is of course defined as the smallest real number with this property. It is also proved that for every x in F and for $S = P_{top}(\varphi, \Omega)$, we still have $\mathcal{L}_S(\mathbb{1}_F)(x) < +\infty$. A part of the work in the next subsection will be to prove that in fact S_c is strictly smaller than the topological pressure $P_{top}(\varphi, \Omega)$ and indeed equal to the topological pressure when one removes transitions from or to A. Here again, the Markov property of the partition and the hyperbolic structure lead to the next lemma:

Lemma 3.2. *There exists a positive constant C_φ which does not depend on S such that for all $x, y \in F$, $S > S_c$ and integer n*

$$\frac{1}{C_\varphi}\mathcal{L}_S^n(\mathbb{1}_F)(x) \leq \mathcal{L}_S^n(\mathbb{1}_F)(y) \leq C_\varphi \mathcal{L}_S^n(\mathbb{1}_F)(x).$$

There exists some quasi-metric η on F such that for every $S > S_c$, the operator $\mathcal{L}_S$ is a quasi-compact operator on the Banach space $\mathcal{C}_\eta$ of Lipschitz-continuous function (for the quasi-metric η). We recall that the norm $\|\cdot\|_\eta$ on $\mathcal{C}_\eta$ is defined by

$$\|\phi\|_\eta = \|\phi\|_\infty + \sup_{x \neq x'} \frac{|\phi(x) - \phi(x')|}{\eta(x, x')}\,.$$

The quasi-metric η is chosen such that the ϑ-Hölder continuous functions on F (where ϑ is the Hölder coefficient of φ) are η-Lipschitz continuous functions. The Schauder-Tychonov theorem holds for the adjoint-operator $\mathcal{L}_S^*$ and $S > S_c$, thus there exists some probability measure m_S such that

$$\mathcal{L}_S^*(m_S) = \lambda_S m_S \tag{6}$$

where $\lambda_S > 0$. The Ionescu-Tulcea & Marinescu theorem (see [12]) holds for the operator $\mathcal{L}_S$ and $S > S_c c$, and $\mathcal{L}_S$ is quasi-compact and λ_S is an eigenvalue for $\mathcal{L}_S$. There is a function H_S in $\mathcal{C}_\eta$ such that

$$\mathcal{L}_S H_S = \lambda_S H_S \quad \text{and} \quad \int H_S \, dm_S = 1.$$

This function satisfies

$$\frac{1}{C_0} \leq H_S \leq C_0 \tag{7}$$

where $C_0 > 0$. A consequence of Lemma 3.2 is that C_0 *does not depend* on S.

The measure ν_S defined by

$$d\nu_S = H_S dm_S$$

is an equilibrium state for (F, g_F) associated to $\Phi(\cdot) - S\mathbf{r}_A^1(\cdot)$.
Now, the mixing properties of f implies ergodicity of ν_S. Hence we have uniqueness for H_S and m_S (thus for ν_S too); moreover there exist $p = p(S)$ complex numbers of modulus one, $1 = \lambda(1), \ldots, \lambda(p)$ such that

$$\mathcal{L}_S = \sum_{i=1}^{p} \lambda_S \lambda(i) \Psi_i + \lambda_S \widetilde{\Psi} \qquad (8)$$

where the Ψ_i are linear projectors defined on $\mathcal{C}_\eta$ with finite rank, $\widetilde{\Psi}$ has a spectral radius strictly smaller than 1, and all the kernels of these operators contain the images of the others (see [**?**] for this more restrictive decomposition).
A crucial fact is that for $S = P_{top}(\varphi, \Omega)$, we have $\lambda_S = 1$ and ν_S is the projection on F of the measure μ_A (see [14] Subsection 6.2 and 6.3).

3.3 Computation of S_c.

We can remove the set A from the Markov partition to define a new subshift of finite type in the following way. For the sake of definiteness, we assume that the first line and the first column of $\mathcal{A}$ encode the transitions to and from A. We denote by $\mathcal{A}'$ the $(N-1) \times (N-1)$ matrix obtained by removing from $\mathcal{A}$ the first line and the first column. Σ' will denote the subset in Σ of all sequences $\underline{x} = (x_n)$ such that $a'_{x_n x_{n+1}} = 1$. For convenience we assume that the matrix $\mathcal{A}'$ is aperiodic; if this is not the case, we can restrict our work to the irreducible components (see exercices page 55 in [15]). The map $\widetilde{\varphi}$ can be restricted to Σ', thus the dynamical system (Σ', σ) admits exactly one equilibrium state, $\widetilde{\mu}'$, with topological pressure $P_{top}(\widetilde{\varphi}, \Sigma')$.

Lemma 3.3. *With the previous notations,* $P_{top}(\widetilde{\varphi}, \Sigma') < P_{top}(\widetilde{\varphi}, \Sigma)$.

The proof can be found in [9]. Another proof based on relative entropy can be found in [7].

We can now prove the main lemma of this subsection:

Lemma 3.4. *The critical value* S_c*, defined in* (5)*, is equal to* $P_{top}(\widetilde{\varphi}, \Sigma')$.

Proof. To ease notations we set $P' = P_{top}(\widetilde{\varphi}, \Sigma')$ throughout this proof. Let x be in F. By lemma 3.2 we have just to prove that for every $S > P'$, $\mathcal{L}_S(\mathbb{1}_F)(x) < +\infty$ and for $S = P'$, $\mathcal{L}_S(\mathbb{1}_F)(x) = +\infty$. We can choose x such that it does not belong to the set $\bigcup_{\mathbb{Z}} f^{-n}\partial\mathcal{R}$. Therefore $\pi_F^{-1}(x)$ is a single point $\underline{x}$ in Σ; this also holds for every y in $Pre_1(x)$.

Let n be some integer. The set, $Pre_1^n(x)$, of points y in $Pre_1(x)$ such that $\mathbf{r}_A^1(y) = n$ (which is well defined because $f^k(x)$ never belongs to the boundary of the Markov partition) is a (ε_0, n)-separated set of points: all these points belong to $F \subset W^u_{loc}(x)$ and all their images by f^n belong to $W^s_{loc}(x)$. If we also assume that ε_0 is small enough, the same argument proves that the set $f(Pre^n_1(x))$ is $(\varepsilon_0, (n-2))$-separated (but not necessarily maximal). Therefore, it follows from [19] that there exists some constant κ_1, independent of S and n, such that

$$\sum_{y \in Pre^n_1(x)} e^{\Phi(y)} e^{-nS} \leq e^{\kappa_1 + C_\omega} e^{(n-2)P' - nS}. \tag{9}$$

Because the matrix $\mathcal{A}'$ is aperiodic, there exists some integer K such that ${\mathcal{A}'}^K$ has only positive entries. We can assume that n is strictly greater than $2K+2$. Let us consider the set E'_{n-2K-2} of all words in Σ' with length $n+1-2K-2$. We pick in each element of E'_{n-2K-2} exactly one point (of Σ'); this collection of points is denoted by I_n. As before, there exists some constant κ_2 (independent of n and S) such that

$$\sum_{\underline{y} \in I_n} e^{S_{n-2K-1}(\widetilde{\varphi})(\underline{y})} \geq e^{\kappa_2 + (n-2K-1)P'}, \tag{10}$$

where $S_{n-2K-1}(\widetilde{\varphi})(\cdot)$ means $\widetilde{\varphi}(\cdot) + \widetilde{\varphi} \circ \sigma(\cdot) + \dots + \widetilde{\varphi} \circ \sigma^{n-2K-2}(\cdot)$.

We are now going to prove that points $\underline{y}$ in I_n can be chosen in such a way that all the $f^{-K-1}(\pi(\underline{y})) = \pi(\sigma^{-K-1}(\underline{y}))$ are in $Pre^n_1(x)$.

We pick some z in $Pre^n_1(x)$ and denote by $\underline{z}$ its preimage by π; we also set $\underline{z} = (z_n)_{n \in \mathbb{Z}}$. Hence z_1 and z_{n-1} must be different from 1. Because of our choice of K, for each W_{n-2K-1} in E'_{n-2K-2}, there exist two words w_K and w'_K of length K in Σ' such that:

- the first letter of w_K is z_1,
- the last letter of w'_K is z_{n-1},
- the word $(w_K W_{n-2K-1} w'_K)$ of length $n-1$ is admissible for Σ'.

Therefore the point

$$\underline{Z}(W_{n-2K-1}) \stackrel{\text{def}}{=} \dots, z_{-1}, z_0, (w_K, W_{n-2K-1}, w'_K), z_n, z_{n+1}, \dots$$

(with the initial position in z_0) belongs to Σ, $\pi(\underline{Z}(W_{n-2K-1}))$ is in $Pre^n_1(x)$ and $\sigma^{K+1}(\underline{Z}(W_{n-2K-1}))$ is in the cylinder W_{n-2K-1}. The sum

$$\sum_{y \in Pre^n_1(x)} e^{\Phi(y)} e^{-nS}$$

is greater than the same sum but restricted to the $\pi(\underline{Z}(W_{n-2K-1}))$'s, thus by (10) proves that there exists some constant κ_3, which does not depend on S and n, such that

$$\sum_{y \in Pre^n_1(x)} e^{\Phi(y)} e^{-nS} \geq e^{\kappa_3 - C_\omega} e^{(n-2K-1)P' - nS}. \tag{11}$$

As $\mathcal{L}_S(\mathbb{1}_F)(x) = \sum_n \sum_{y \in Pre_1^n(x)} e^{\Phi(y)} e^{-nS}$, (9) and (11) prove that $\mathcal{L}_S(\mathbb{1}_F)(x)$ converges for every $S > P'$ and diverges to $+\infty$ for every $S \leq P'$. The lemma is thus proved. $\square$

4 Proof of Theorem 2.1

Let us start with the case when A is a single rectangle of the Markov partition. At the end of this section, we briefly sketch the modifications that are necessary to handle the case when A is a finite union of rectangles.

For notational convenience, let us set

$$\alpha_0 = P_{top}(\widetilde{\varphi}, \Sigma) - S_c = P_{top}(\widetilde{\varphi}, \Sigma) - P_{top}(\widetilde{\varphi}, \Sigma') \,. \tag{12}$$

We shall also use the notation $S_j(\phi)$ to abbreviate $\phi + \phi \circ f + \cdots + \phi \circ f^{j-1}$. Let α be in $]0, \alpha_0[$. For convenience we will write P in subscript instead of $P_{top}(\varphi, \Omega)$. We start by observing that if we pick any point on the reference unstable leaf F, then all points lying in its stable leaf have the same return times as this point. That is why we can reduce the problem to the unstable leaf F. We have

$$\begin{aligned} \mathbb{E}_{\mu_A}\left[e^{\alpha \mathbf{r}_A^n}\right] &= \int_F e^{\alpha \mathbf{r}_A^n(x)} H_P(x) dm_P(x) = \int_F \mathcal{L}_P^n(e^{\alpha \mathbf{r}_A^n} H_P)(x) dm_P(x) \\ &= \int_F \sum_{y \in Pre_n(x)} e^{S_{\mathbf{r}_A^n(y)}(\varphi)(y) + \omega(g^n(y)) - (P-\alpha)\mathbf{r}_A^n(y)} H_P(y)\, dm_P(x)\,. \end{aligned} \tag{13}$$

We have used the fact that the measure $H_P dm_P$ is the projection of $d\mu_A$, see the end of Subsection 3.2. We first remark that for $\alpha < \alpha_0$ we have $P - \alpha > S_c$. Hence, using lemmas 3.1-3.2, the Markov property and the properties of the function H_P one obtains the existence of some constant C, which depends only on φ and f such that for every x in F,

$$\begin{aligned} e^{-C} \mathcal{L}_{P-\alpha}^n(\mathbb{1}_F)(x) &\leq \\ \int_F \sum_{y \in Pre_n(x)} & e^{S_{\mathbf{r}_A^n(y)}(\varphi)(y) + \omega(g^n(y)) - (P-\alpha)\mathbf{r}_A^n(y)} H_P(y)\, dm_P(x) \\ &\leq e^{C} \mathcal{L}_{P-\alpha}^n(\mathbb{1}_F)(x). \end{aligned} \tag{14}$$

If we integrate this double inequality with respect to the measure $m_{P-\alpha}$, we obtain the following estimate:

$$\log \lambda_{P-\alpha} - \frac{C}{n} \leq \frac{1}{n} \log \mathbb{E}_{\mu_A}\left[e^{\alpha \mathbf{r}_A^n}\right] \leq \log \lambda_{P-\alpha} + \frac{C}{n}\,. \tag{15}$$

This proves that for every $\alpha < \alpha_0$

$$\lim_{n \to +\infty} \frac{1}{n} \log \mathbb{E}_{\mu_A}\left[e^{\alpha \mathbf{r}_A^n}\right] = \log \lambda_{P-\alpha} < +\infty\,.$$

Let us set $\Psi(\alpha) \stackrel{\text{def}}{=} \log \lambda_{P-\alpha}$. We now have to prove that the function Ψ is analytic in some complex neighborhood of $]-\infty, \alpha_0[$. Analyticity of Ψ is equivalent to the analyticity of the map $S \mapsto \log \lambda_S$ in some complex neighborhood of $]S_c, \infty[$. We want to use a perturbation theorem due to Hennion and Hervé [11]. To this end, we first have to check that $\mathcal{L}_S$ has $p = p(S)$ dominating simple eigenvalues (see [11], III.2). This is a consequence of ergodicity of the system (F, g_F, ν_S). In fact, the proof of the proposition 4.11 in [4] can be adapted to our case; hence the projectors Ψ_i in (8) have rank one, and the $\lambda(i)$'s all satisfy

$$\lambda(i)^p = 1 \; .$$

Thus the operator $\mathcal{L}_S$ has p dominating simple eigenvalues.

Lemma 4.1. *Let $\mathcal{OP}_{\mathcal{C}_\eta}$ denote the set of linear bounded operators on $\mathcal{C}_\eta$. Then the map $z \mapsto \mathcal{L}_z$ is analytic map from $\{z \in \mathbb{C}, \Re(z) > S_c\}$ to $\mathcal{OP}_{\mathcal{C}_\eta}$, uniformly on compact sets.*

Proof. For any z in $\mathbb{C}$ with $\Re(z) > S_c$, we set (by extension):

$$\mathcal{L}_z(\phi)(x) = \sum_{y \in Pre_1(x)} e^{\Phi(y) - \mathbf{r}_A^1(y) z} \phi(y)$$

which can also be written

$$\sum_{n=1}^{+\infty} \left(\sum_{y \in Pre_1(x),\ \mathbf{r}_A^1(y) = n} e^{\Phi(y)} \phi(y) \right) e^{-nz} \, .$$

Let us introduce

$$K_m(z)(\phi)(x) = \sum_{n=1}^{m} \left(\sum_{y \in Pre_1(x),\ \mathbf{r}_A^1(y) = n} e^{\Phi(y)} \phi(y) \right) e^{-nz}.$$

We are going to prove that the sequence $(K_m)(\cdot)$ converges to $\mathcal{L}_{\cdot}$, when m goes to $+\infty$, as analytic functions from $\Re(z) > S_c$ to $\mathcal{OP}_{\mathcal{C}_\eta}$, uniformly on compact sets. Let us fix some compact set Γ in $\{\Re(z) > S_c\}$ and pick z in Γ. Let ϕ be some function in $\mathcal{C}_\eta$, with $\|\phi\|_\eta = 1$. We want to compute

$$\|K_m(z)(\phi) - \mathcal{L}_z(\phi)\|_\eta.$$

There exists $S > S_c$ such that for every z in Γ, $\Re(z) > S$, which proves that the series $K_m(z)(\phi)(x)$ is normally convergent (in z) and uniformly convergent in x to $\mathcal{L}_z(\phi)(x)$ when $m \to \infty$. Hence $K_m(z)$ is uniformly convergent to $\mathcal{L}_z$ in Γ for the norm $\|\cdot\|_\infty$ on $\mathcal{C}_\eta$. Let y and y' be two points in F with $y' \in C_1(y)$. Then we have

$$e^{\Phi(y)}\phi(y) - e^{\Phi(y')}\phi(y') = e^{\Phi(y)}\phi(y) - e^{\Phi(y)}\phi(y') + e^{\Phi(y)}\phi(y') - e^{\Phi(y')}\phi(y').$$

Therefore, the Lipschitz properties of the functions ϕ and Φ (for the quasi-metric η) and the expansion on the unstable leaves imply that there exists some constant C which depends only on f and ϕ such that for every n, and for every x and x' in F,

$$\left| \sum_{y\in Pre_1(x),\ \mathbf{r}_A^1(y)=n} \left(e^{\Phi(y)}\phi(y) - e^{\Phi(y')}\phi(y') \right) \right| \leq C\eta(x,x'), \tag{16}$$

where for each y in $Pre_1(x)$, y' is the preimage of x' in the 1-cylinder $C_1(y)$. Inequality (16) and convergence for the norm $\|\cdot\|_\infty$ imply that the series $K_m(z)$ is uniformly convergent to $\mathcal{L}_z$ in Γ for the norm $\|\cdot\|_\eta$. This proves that $z \mapsto \mathcal{L}_z$ is analytic. □

By Theorem III.8 from [11], for every $S > S_c$, there exists some open disc in $\mathbb{C}$ centered at S, such that for every z in this disk, $\mathcal{L}_z$ has $p(S)$ dominating simple eigenvalues, $\lambda^1(z), \ldots, \lambda^p(z)$ and the maps $z \mapsto \lambda^i(z)$ are analytic. Because the map log is analytic in $\mathbb{C} \setminus \mathbb{R}_-$, we can conclude that the map $z \mapsto \log \lambda_z$ is analytic in a complex neighborhood of $]S_c, +\infty[$.

It remains to prove strict convexity of $\alpha \mapsto \Psi(\alpha)$ on $]-\infty, \alpha_0[$. Convexity is obvious (by Hölder inequality). Strict convexity is equivalent to strict convexity of the map $S \mapsto \log \lambda_S$ on $]S_c, +\infty[$. We will assume that the map $S \mapsto \log \lambda_S$ is not strictly convex and will arrive to a contradiction with the fact that A is a proper rectangle of the Markov partition.

If the map $S \mapsto \lambda_S$ is not strictly convex, this means that its graph contains a straight line interval, say $\mathcal{I} \subset]S_c, \infty[$. In turn, this means that λ'_S/λ_S is constant on $\mathcal{I}$, which means that $\lambda_S = \Lambda_A e^{\gamma_A S}$ where Λ_A, γ_A are real constants *a priori* depending on A. Now we invoke the unicity of analytic continuation of analytic functions to deduce that

$$\forall z \in \mathbb{C} \quad \text{with} \quad \Re(z) > S_c, \quad \lambda_z = \Lambda_A e^{-\gamma_A z}\,.$$

Let us define the smallest return time in A as follows:

$$\tau(A) = \inf\{k \geq 1 : f^{-k}A \cap A \neq \emptyset\} = \inf\{\mathbf{r}_A^1(x) : x \in A\}\,. \tag{17}$$

Using Lemma 3.2, we readily get for every $S > S_c$

$$C_\varphi^{-1} \sum_{n\geq\tau(A)} \left(\sum_{y\in Pre_1(x), \mathbf{r}_A^1(x)=n} e^{\Phi(y)}\mathbb{I}_F(y) \right) e^{-nS} \leq$$

$$\Lambda_A e^{\gamma_A S} \leq$$

$$C_\varphi \sum_{n\geq\tau(A)} \left(\sum_{y\in Pre_1(x), \mathbf{r}_A^1(x)=n} e^{\Phi(y)}\mathbb{I}_F(y) \right) e^{-nS}\,.$$

Letting $S \to \infty$ we deduce that $\gamma_A = -\tau(A)$.

Now, observe that

$$\tau(A) = \Psi'(0) = 1/\mu(A)$$

(remember Remark 3 in Section 2); we use Kač formula to obtain

$$\tau(A) = \sum_{n=\tau(A)}^{\infty} n\mu_A\{\mathbf{r}_A^1 = n\}\,.$$

Therefore, $\mathbf{r}_A^1(x) = \tau(A)$ for μ_A-almost every x. Since $\mu_A(\mathring{A}) = 1$, the topological mixing property imposes that $\tau(A)$ must equal to one, which is absurd since the Markov partition is made of more than two rectangles.

Therefore, the function $S \mapsto \log\lambda_S$ is strictly convex on $]S_c, \infty[$, so is the function $\alpha \mapsto \Psi(\alpha) = \log\lambda_{P-\alpha}$ on $]-\infty, P - S_c[\ = \ ]-\infty, \alpha_0[$.

Extension to a finite union of rectangles

We sketch how to extend the previous proof when the set A is a union $R_{i_1} \cup \ldots R_{i_k}$ of $k \geq 2$ rectangles (of the partition $\mathcal{R}$). In each rectangle R_{i_j} we pick some unstable leaf F_j, and define F as the union of the F_j's. Up to the boundary of $\partial\mathcal{R}$, this union is a disjoint union. Thus, if we denote by π_j the projection from R_{i_j} onto F_j (namely $[., F_j]$, see [3]). This defines a map π_F from A onto $F \stackrel{\text{def}}{=} \bigcup F_j$ (up to some set with zero-measure for all the measures we are going to consider). Hence the map $g_F \stackrel{\text{def}}{=} \pi_F \circ g$ is well defined; the Markov property of the partition $\mathcal{R}$ allows again us to define the partitions of n-cylinders; but in this case we will not have $g_F^n(C_n(x)) = F$, but only $g_F^n(C_n(x)) = F_j$ for some j.

The definition of the operator $\mathcal{L}_S$ is the same, but the value of the critical S, S_c, is changing. Let $\mathcal{A}''$ be the matrix obtained from $\mathcal{A}$ by removing the lines and the columns corresponding to the subscripts i_j's. Let Σ'' be the set of sequences $\underline{x}$ such that for every n, $a''_{x_n,x_{n+1}} = 1$ and let P'' be the topological pressure for the dynamical system (Σ'', σ) associated to the potential $\widetilde{\varphi}$. Then, $S_c \leq P'' < P_{top}(\varphi, \Omega)$.

Lemma 3.2 still holds but only if we pick x and y in the same F_j. Accordingly, the proof of proposition 3.4 can be adapted.

Now, observe that formulas (13) and (14) are still valid, except that we have to split the integrals over the F_j's. Therefore (15) still holds, for some positive constant C.

However, it is important to notice that transitivity on Ω implies that for every j and j', the set of points in F_j which returns infinitely many times in $F_{j'}$ is dense. Moreover, if x is a point in $F_{j'}$ for every j, the set of preimages of x in F_j (for the map g_F) is dense. Therefore ergodicity of ν_S still holds, which is the sufficient condition to get simple dominating eigenvalues for $\mathcal{L}_S$. We can thus again apply Theorem III.8 from [11] to get analyticity of the map $S \mapsto \log\lambda_S$.

The proof of strict convexity is analogous to the previous case.

5 Proofs of Theorem 2.2 and 2.3

Proof of Theorem 2.2. To prove this theorem, we apply a result due to Plachky and Steinebach [18]. Since, by Theorem 2.1, the map Ψ is strictly convex and real-analytic on $]-\infty, \alpha_0[$, the function $\alpha \mapsto \Psi'(\alpha)$ is strictly increasing on that interval and Plachky–Steinebach's large deviations result applies.

We also observe that $\Psi'(\alpha) \to +\infty$ when $\alpha \to \alpha_0$. This is equivalent to show that $\lambda'_S/\lambda_S \to \infty$ when $S \to S_c$. On the other hand, $\Psi'(\alpha) \to 0$ when $\alpha \to -\infty$. This can be showed by using Lemmas 3.2 and 3.4, and convexity.

Proof of Theorem 2.3. Bryc's Theorem [5] applies. In particular, it says that the variance is equal to $\Psi''(0)$. It is easy to get formula (3) by differentiating twice Ψ. The only point to be proved is that $\sigma_A > 0$.

Now we follow the proof of Proposition 4.12, p. 63, in [17] (note that Herglotz's Theorem also applies in our setting). To this end, we just have to check that $\mathbb{E}_A(\tau_A^1\ \tau_A^j) - \mathbb{E}_A(\tau_A^1)\mathbb{E}_A(\tau_A^j)$ decreases exponentially fast to 0. Remember that in this case, the variance can be rewritten as follows

$$\sigma_A^2 = \mathbb{E}_A((\tau_A^1)^2) - \frac{1}{\mu(A)^2} + 2\sum_{j=2}^{\infty}\left(\mathbb{E}_A(\tau_A^1\ \tau_A^j) - \mathbb{E}_A(\tau_A^1)\mathbb{E}_A(\tau_A^j)\right) .$$

The exponential decreasing of the above correlations follows from the ψ-mixing property of the induced system (see [8]):

$$\left|\mathbb{E}_A(\tau_A^1\ \tau_A^j) - \mathbb{E}_A(\tau_A^1)\mathbb{E}_A(\tau_A^j)\right| \le$$

$$\sum_{p,q\in\mathbb{N}} pq\ \left|\mu_A(\tau_A^1 = p, \tau_A^j = q) - \mu_A(\tau_A^1 = p)\ \mu_A(\tau_A^j = q)\right| \le$$

$$C\left(\sum_{p\in\mathbb{N}} p\mu_A(B_p)\right)^2 \theta^j = C\mu(A)^{-2}\theta^j$$

where $C > 0$, $0 < \theta < 1$ and $B_p \stackrel{\text{def}}{=} \{\tau_A^1 = p\}$ (we used Kač formula).

We conclude that $\sigma_A^2 = 0$ if and only if $\tau_A^1 - 1/\mu(A)$ is a $L^2(\mu_A)$ coboundary with respect to g, the induced map on A. But if A is a single Markov rectangle, this is impossible. Indeed, there exists a fixed point for g which is periodic with period $\tau(A)$ for f ($\tau(A)$ is defined in (17)). Reasoning as above, this leads to a contradiction with the fact that A is a strict subset of Ω in measure. The case when A is a (finite) union of rectangles is left to the reader.

Therefore we arrive at the conclusion that the variance σ_A^2 defined in (3) is strictly positive. Theorem 2.3 has now a complete proof.

References

[1] M. Abadi, A. Galves, *Inequalities for the occurrence of rare events in mixing processes. The state of the art*, 'Inhomogeneous random systems' (Cergy-Pontoise, 2000), Markov Process. Related Fields 7 (2001), no. 1, 97–112.

[2] G.D. Birkhoff, *Proof of the recurrence theorem for strongly transitive systems*, Collected Mathematical Papers, Vol. II, Dover, 1968. Reprinted from: Proc. Nat. Acad. Sci., December, 1931, Vol. 17, pp. 650–655.

[3] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture notes in Math. **470**, Springer-Verlag, 1975.

[4] A. Broise, F. Dal'bo, M. Peigné, *Etudes spectrales des opérateurs de transferts et applications*, Astérisque **238**, SMF, 1996.

[5] W. Bryc, *A remark on the connection between the large deviation principle and the central limit theorem*, Statist. & Probab. Lett. **18**, 253–256 (1993).

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

[7] J.-R. Chazottes, E. Floriani, R. Lima, *Relative entropy and identification of Gibbs measures in dynamical systems*, J. Stat. Phys. **90**, Nos. 3/4 (1998), 697–725.

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

[9] P. Collet, A. Galves, A. Lopes, *Maximum Likelihood and Minimum Identification of Grammars*, Random Comput. Dyn. **3**, no. 4, 1995, 241–250.

[10] A. Dembo, O. Zeitouni, *Large Deviations Techniques & Applications*, Applic. Math. **38**, Springer, 1998.

[11] H. Hennion, L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Math. **1766**, Springer-Verlag, 2001.

[12] C.T. Ionescu-Tulcea, G. Marinescu, *Théorie ergodique pour des classes d'opérations non complètement continues*, Annals of Mathematics **52**, no. 1, 1950, 140–147.

[13] P.W. Kasteleyn, *Variations on a theme by Mark Kac*, J. Statist. Phys. **46** (1987), no. 5-6, 811–827.

[14] R. Leplaideur, *Local product structure for equilibrium states*, Trans. Amer. Math. Soc. **352** (2000), no. 4, 1889–1912.

[15] A. Katok and B. Hasselblatt, *Introduction to the modern Theory of Dynamical Systems*, Cambridge University Press, (1995), Encyclopedia of mathematics and its applications.

[16] S.C. Moy, *Successive recurrence times in a stationary process*, Ann. Math. Statist. **30** (1959) 1254–1257.

[17] W. Parry, Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque **187-188**, SMF, (1990).

[18] D. Plachky, J.A. Steinebach, *A theorem about probabilities of large deviations with an application to queuing theory*, Periodica Math. Hungar. **6**, 343–345 (1975).

[19] D. Ruelle, *Thermodynamic formalism for maps satisfying positive expansiveness and specification*, Nonlinearity **5**, 1992, 1223–1236.

[20] D. Ruelle, *Thermodynamic formalism*, The mathematical structures of classical equilibrium statistical mechanics, Encyclopedia of Mathematics and its Applications **5**, Addison-Wesley Publishing Co., Reading, Mass., 1978.

[21] S. Waddington, *Large deviation asymptotics for Anosov flows*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), no. 4, 445–484.