The complete $L^q$-spectrum and large deviations for return times for equilibrium states with summable potentials

M. Abadi *, J.-R. Chazottes † and S. Gallo ‡

Department of Mathematics, Universidade Federal de São Carlos (Brazil)

Abstract

Let $\{X_k\}_{k \geq 1}$ be a stationary and ergodic process with joint distribution $\mu$ where the random variables $X_k$ take values in a finite set $A$. Let $R_n$ be the first time this process repeats its first $n$ symbols of output. It is well-known that $n^{-1} \log R_n$ converges almost surely to the entropy of the process. Refined properties of $R_n$ (large deviations, multifractality, etc) are encoded in the return-time $L^q$-spectrum defined as

$$\mathcal{R}(q) = \lim_{n \to \infty} \frac{1}{n} \log \int R_n^q d\mu$$

provided the limit exists. When $\{X_k\}_{k \geq 1}$ is distributed according to the equilibrium state of a potential $\varphi$ with summable variation, we are able to compute $\mathcal{R}(q)$ for all $q \in \mathbb{R}$, namely

$$\mathcal{R}(q) = \begin{cases} P((1-q)\varphi) & \text{for } q \geq q^* \\ \sup_{\eta} \int \varphi d\eta & \text{for } q \leq q^* \end{cases}$$

*Email: leugim@ime.usp.br
†Email: chazottes@cpht.polytechnique.fr
‡Email: sandro.gallo@ufscar.br.

SG acknowledges École Polytechnique for financial support and hospitality during a two-months stay. SG was supported by FAPESP (BPE: 2017/07084-6) and CNPq (PQ 312315/2015-5 and Universal 462064/2014-0). MA and SG acknowledge the FAPESP-FCT joint project between SP-Brazil and Portugal (19805/2014).
where the supremum is taken over all shift-invariant probability measures, 
\( P((1 - q) \varphi) \) is the topological pressure of \((1 - q) \varphi\) \((q \in \mathbb{R})\), and \( q^\ast = q^\ast_\varphi \in [-1, 0) \) is the unique solution of \( P((1 - q) \varphi) = \sup_q \int \varphi \, d\eta \), as \( q \) runs through \( \mathbb{R} \). Surprisingly, this spectrum does not coincide with the hitting-time \( L^q \)-spectrum for all \( q < q^\ast \). The two spectra coincide if and only if the equilibrium state of \( \varphi \) is the measure of maximal entropy. As a by-product, we prove large deviation asymptotics for \( n^{-1} \log R_n \).

Contents

1 Introduction 3

2 Basic setting and definitions 7

2.1 The shift space, hitting times and return times 6

2.2 The return-time and the hitting-time \( L^q \)-spectra 8

2.3 The \( L^q \)-spectra of a measure and the Rényi entropy function 10

3 A result for \( \phi \)-mixing processes with summable rate 11

4 Equilibrium states for potentials with summable variation 12

4.1 Basic definitions and properties 12

4.2 Main results 14

5 Some explicit examples 17

5.1 Independent random variables 17

5.2 Markov chains 17

6 Proofs of the results of Section 3 18

6.1 Proof of Proposition 4 18

6.2 Proof of Theorem 1 20

6.2.1 First estimates 20

6.2.2 Estimations of \( I(x_n^1, q) \) and \( J(x_n^0, q) \) 22

6.2.3 End of the proof of Theorem 1 23

7 Proofs of the results of Section 4 25

7.1 Two preliminary general results 25

7.2 Proof of Proposition 5 28

7.3 Proof of Theorem 2 29

7.3.1 Control of the measure of cylinders 29

7.3.2 The key proposition 31

7.3.3 End of proof of Theorem 2 32

A Some inequalities for the upper incomplete Gamma function 33

B Proof of inequalities 35

C Proof of Theorem 1 for \( q \geq 0 \) 36
1 Introduction

Let \( \{X_k\}_{k \geq 1} \) be a stationary and ergodic process where the random variables \( X_k \) take values in a finite set \( \mathcal{A} \). We denote by \( \mu \) the joint distribution of \( \{X_k\}_{k \geq 1} \) which is a shift-invariant ergodic probability measure on \( \mathcal{A}^\mathbb{N} \) whose elements are of the form \( x = (x_1, x_2, \ldots) \) with \( x_i \in \mathcal{A} \). We are interested in the statistical properties of \( R_n(x) \), defined as the first time the process repeats its first \( n \) symbols of output, that is, the smallest \( k \geq 2 \) such that \( x_k, x_{k+1}, \ldots, x_{k+n-1} = x_1, x_2, \ldots, x_n \).

**Duality between return times and the inverse measure of cylinders.** It is well-known \([OW93]\) that
\[
\frac{1}{n} \log R_n(x) \to h(\mu), \quad \text{for } \mu\text{-almost every } x,
\]
where \( h(\mu) \) is the entropy of \( \mu \). This is a remarkable result when compared to the Shannon-McMillan-Breiman theorem which says that
\[
-\frac{1}{n} \log \mu([x^n_1]) \to h(\mu), \quad \text{for } \mu\text{-almost every } x,
\]
where \([x^n_1]\) is the \( n \)-cylinder based on \( x_1, \ldots, x_n \), that is, the set of outputs starting with the symbols \( x_1, \ldots, x_n \). Using return times, we don’t need to know \( \mu \) to estimate the entropy. Of course, in both cases, we have to assume that we have a typical output \( x = x_1, x_2, \ldots \) of the process. As a consequence of the above two results we have
\[
\log \left( R_n(x) \mu([x^n_1]) \right) = o(n), \quad \text{for } \mu\text{-almost every } x. \tag{1}
\]
We also mention Kač’s Lemma which states that, for any cylinder \([a^n_1]\),
\[
\int R_n(x) \mathrm{d}\mu_{[a^n_1]}(x) = \frac{1}{\mu([a^n_1])}
\]
where \( \mu_{[a^n_1]}(C) = \mu([a^n_1] \cap C)/\mu([a^n_1]) \), for any cylinder set \( C \). That is, on average, it takes a time \( \mu([a^n_1])^{-1} \) to come back to the cylinder \([a^n_1]\). In view of these results, it is tantalizing to replace \( R_n(x) \) by \( \mu([a^n_1])^{-1} \), for instance in the \( L^q \)-spectrum of \( \mu \) (see below). It is also tantalizing to assume that \( R_n(x) \) and \( \mu([a^n_1])^{-1} \) have the same fluctuations. For a large class of measures, it turns out that they do have the same “small” fluctuations”, but not the same ‘large deviations’, around \( h(\mu) \).

**Fluctuations of the return times.** In this paper, we are interested in the fluctuations of \( \log R_n/n \) around \( h(\mu) \). To this end, we need to assume that \( \mu \) has good mixing properties. We will mainly assume that the process is distributed according to an equilibrium state \( \mu_\phi \) of a potential with summable
variation. The latter class contains for instance finite-state Markov chains which are irreducible and aperiodic, and H"older continuous potentials whose variation decays exponentially fast.

A way to describe the fluctuations of \( \log R_n/n \) is to study the convergence in law of \( \log R_n - nh(\mu) / \sqrt{n} \) to a Gaussian distribution (central limit asymptotics). Such a convergence was first proved in [CGS99] when \( \phi \) is a H"older continuous potential, and was then extended to potentials with summable variation in [CU05]. As expected, the variance of the Gaussian distribution is the same as that of the processes \( -\log \mu_\phi([x^n]), (S_n \phi(x)/n) \) (where \( S_n \phi \) is the Birkhoff sum of \( \phi \)). Loosely speaking, this convergence means that typical fluctuations of \( \log R_n \) around \( nh(\mu_\phi) \) are of order \( \sqrt{n} \). In this paper, we are interested in the so-called large deviation regime which corresponds to fluctuations of order \( n \). Again, this was first studied in [CGS99] when \( \phi \) is a H"older continuous potential. Therein the authors proved that there exists \( u_0 > 0 \) (which is implicit) such that, for any \( u \in (0, u_0) \), one has

\[
\lim_{n \to \infty} \frac{1}{n} \log \mu_\phi \left( \frac{1}{n} \log R_n > h(\mu_\phi) + u \right) = I(h(\mu_\phi) + u)
\]

and

\[
\lim_{n \to \infty} \frac{1}{n} \log \mu_\phi \left( \frac{1}{n} \log R_n < h(\mu_\phi) - u \right) = I(h(\mu_\phi) - u),
\]

where \( I \) is a positive convex function vanishing if \( u = 0 \) which is obtained as the Legendre transform of

\[
\lim_{n \to \infty} \frac{1}{n} \log \sum_{a^n_1 \in A^n} \mu(a^n_1)^{1-q} = P((1-q)\phi), \quad q \in \mathbb{R},
\]

where \( P((1-q)\phi) \) is the topological pressure of \( (1-q)\phi \). In view of a classical theorem in large deviation theory called the G"artner-Ellis theorem (see e.g. [DZ10]), a natural route to prove such kind of result is to prove that the cumulant generating function of \( \left( \log R_n/n \right) \), namely

\[
\mathcal{R}_{\mu_\phi}(q) = \lim_{n \to \infty} \frac{1}{n} \log \int e^{q \log R_n} \, d\mu_\phi = \lim_{n \to \infty} \frac{1}{n} \log \int R_n^q \, d\mu_\phi, \quad q \in \mathbb{R}, \quad (2)
\]

does exist and has good properties like differentiability. Motivated by multifractal analysis, we call \( \mathcal{R}_{\mu_\phi}(q) \) the return-time \( L^q \)-spectrum (with respect to \( \mu_\phi \)). To obtain the above large deviation result, the authors of [CGS99] use a certain control of \( |\mu_\phi(x : R_n(x)\mu([x^n])) > t) - e^{-t} | \) as a function of \( x^n \) and \( t > 0 \) which holds only for typical cylinders, that is cylinders around aperiodic points. This allows them to prove that \( \mathcal{R}_{\mu_\phi}(q) \) exists in an interval \((q_0, q_0)\) for some implicit \( q_0 > 0 \), and coincides with \( P((1-q)\phi) \) in that interval. This equality is what we can expect if we replace \( R_n \) by \( \mu([x^n])^{-1} \), which is the natural “ansatz” mentioned above. The question is: Is this equality true outside \((q_0, q_0)\) and more importantly, what is \( q_0 \)?
Main contribution of the paper. The main result of this paper is the computation of $R_{\mu_\varphi}(q)$ for all $q \in \mathbb{R}$, and for potentials with summable variation. We prove that, if $\varphi$ is not of the form $u - u \circ \theta - \log |A|$ for some continuous function $u : A^N \to \mathbb{R}$ ($\theta$ being the shift operator), then

$$R_{\mu_\varphi}(q) = \begin{cases} P((1 - q)\varphi) & \text{for } q \geq q_\varphi^* \\ \sup_\eta \int \varphi\,d\eta & \text{for } q \leq q_\varphi^* \end{cases}$$

(3)

where the supremum is taken over all shift-invariant probability measures, and $q_\varphi^* \in (-1, 0)$ is the unique solution of $P((1 - q)\varphi) = \sup_\eta \int \varphi\,d\eta$, as $q$ runs through $\mathbb{R}$. In fact, we first establish an abstract result for $\phi$-mixing processes with a summable rate before specializing it to equilibrium states of potentials with summable variation for which we use the powerful tools of thermodynamic formalism. From this result, we deduce more precise large deviation asymptotics than in [CGS99], and cover a larger class of potentials.

In order to obtain the complete return time $L^q$-spectrum, we partition the phase space into $n$-cylinders sets. This procedure takes into account all cylinders, including the non-typical ones (those around periodic points). To cope with this situation, we need a sharp control of

$$\left| \mu_\varphi \left( T_{[x^n]} > \frac{t}{\mu_\varphi([x^n_1])} \right) - \zeta_{\mu_\varphi}([x^n_1]) e^{-\zeta_{\mu_\varphi}([x^n_1])t} \right|$$

for all $t > 0$ and $x^n_1$, where $\zeta_{\mu_\varphi}([x^n_1])$ is the “exit probability” of the cylinder $[x^n_1]$. This was achieved by Abadi and Vergne [AV09] assuming that the measure is $\phi$-mixing. This allows us to control the balance between the contribution of periodic points and that of typical points. The former is encoded by the quantity

$$\lim_n \frac{1}{n} \sum_{a^n_1} (1 - \zeta_{\mu_\varphi}([x^n_1])) \mu_\varphi([x^n_1]) .$$

The control of $\zeta_{\mu}([x^n_1])$ and the existence of this limit are an important part of this work. What we show is that this quantity determines $R_{\mu_\varphi}(q)$ for all $q < q_\varphi^*$, whereas the ansatz consisting in replacing $R_n(x)$ by $\mu_\varphi([x^n_1])^{-1}$ is indeed correct only when $q \geq q^*$.

The waiting times $L^q$-spectrum. Another estimator of $h(\mu)$ that is worth mentioning is based on hitting times. Consider two independent realizations $x = x_1, x_2, \ldots$ and $y = y_1, y_2, \ldots$ of the process. Let $T_{x^n_1}(y)$ be the first time $x_1, \ldots, x_n$ appears in $y$. It is known [Shi96] that if $\mu$ is weak Bernoulli, then

$$\frac{1}{n} \log T_{x^n_1}(y) \to h(\mu), \quad \text{for } \mu \otimes \mu\text{-almost every } (x, y).$$

This means that $\mu_\varphi$ is not the measure of maximal entropy.
(This result is known to be false if \( \mu \) is only assumed to be ergodic [Shi96, p. 205].). We then have
\[
\log \left( T_{x_1^n}(y) \mu([x^n_1]) \right) = o(n), \quad \text{for } \mu \otimes \mu\text{-almost every } (x, y).
\]

Given a potential \( \varphi \), the analog of \( R_{\mu, \varphi}(q) \) is
\[
W_{\mu, \varphi}(q) = \lim_{n \to \infty} \frac{1}{n} \log \int \int T_{x_1^n}(y) \, d\mu \otimes \mu \otimes \varphi, \quad q \in \mathbb{R}.
\]

In [CU05] it was proved that for potentials with summable variation one has
\[
W_{\mu, \varphi}(q) = \begin{cases} 
P((1 - q)\varphi) & \text{for } q \geq -1 \\
2P(\varphi) & \text{for } q < -1.
\end{cases}
\]

It is fair to say that the expressions of \( R_{\mu, \varphi}(q) \) and \( W_{\mu, \varphi}(q) \) are both unexpected, and that it is surprising that they do not coincide on a non-trivial interval.

**More on related results.** Let us come back to large deviations for return times and comment on other results related to ours, beside [CGS99]. In [JB13], the authors obtain the following result. For a \( \phi \)-mixing process with joint distribution \( \mu \) with an exponentially decaying rate, and satisfying a property called ‘exponential rates for entropy’, there exists an implicit positive function \( I \) such that \( I(0) = 0 \) and
\[
\mu \left( \frac{1}{n} \log R_n - h(\mu) \right) > u \right) \leq 2 e^{-I(u)}, \quad n \geq N(u).
\]

In the same vein, [CRS18] considered the case of (geometric) balls in smooth dynamical systems (instead of cylinder sets which are the natural sets to look at for processes on finite alphabets). Finally, still in the context of smooth dynamical systems, let us mention that the analogue of \( R_{\mu, \varphi}(q) \) with balls instead of cylinders was considered from the viewpoint of multifractal analysis to define ‘return-time dimensions’, see [CFM+18].

**Organisation of the paper.** In Section 2 we define the return-time and the hitting-time \( L^q \)-spectra, the \( L^q \)-spectrum of a measure, and the Rényi entropy function. In Section 3 we obtain our first result which is for \( \phi \)-mixing processes with a summable rate. It gives a formula for \( R_{\mu, \varphi}(q) \) and \( W_{\mu, \varphi}(q) \), which are defined by taking, respectively, the limit inferior and the limit superior instead of the limit in (2). Under the same assumptions, we have that the exit probability \( \zeta_{\mu, \varphi}([x^n_1]) \) is uniformly bounded away from 0. This results is important on its own since the exit probability appears as a scaling parameter in the exponential approximation for the distribution of return times mentioned above. Next,

\[\text{Except, as we will demonstrate, when } \varphi = u - u \circ \theta - \log |A| \text{ for some continuous function } u : A^1 \to \mathbb{R}.\]
in Section 4 we describe the class of equilibrium states for potentials with summable variation and state our main results. In Section 5 we give a few examples where we can explicitly compute \( R_{\mu_n}(q) \). Sections 6 and 7 contain the proofs of the results of Section 3 and 4, respectively. Finally, there is an appendix which essentially contains some estimates for the upper incomplete Gamma function and some further auxiliary results.

2 Basic setting and definitions

2.1 The shift space, hitting times and return times

For any sequence \( (a_k)_{k \geq 1} \) with elements in \( \mathcal{A} \), we denote the partial sequence (‘string’) \( (a_i, a_{i+1}, \ldots, a_j) \) by \( a_i^j \), for \( i < j \). (By convention, \( a_i^i := a_i \).) In particular, \( a_\infty \) denotes the sequence \( (a_k)_{k \geq i} \).

We consider the space \( \mathcal{A}^\mathbb{N} \) of infinite sequences \( x = (x_1, x_2, \ldots) \) where \( x_i \in \mathcal{A} \), \( i \in \mathbb{N} := \{1, 2, \ldots\} \). The cylinder sets \( [a_i^j] = \{ x \in \mathcal{A}^\mathbb{N} : x_i = a_i \} \), \( i, j \in \mathbb{N} \), generate the (Borel) \( \sigma \)-algebra \( \mathcal{F} \). For a probability measure \( \mu \) we shall simply write \( \mu(a_i^j) \) instead of \( \mu([a_i^j]) \). Now define the shift \( \theta : \mathcal{A}^\mathbb{N} \to \mathcal{A}^\mathbb{N} \) by \( (\theta x)_i = x_{i+1}, i \in \mathbb{N} \). Let \( \mu \) be a shift-invariant probability measure on \( \mathcal{F} \), that is, \( \mu(\theta^{-1}B) = \mu(B) \) for each cylinder. We then consider the stationary process \( \{X_k\}_{k \geq 1} \) on the probability space \( (\mathcal{A}^\mathbb{N}, \mathcal{F}, \mu) \), where \( X_n(x) = x_n \), \( n \in \mathbb{N} \). We will use the short-hand notation \( X_i^n \) for \( (X_i, X_{i+1}, \ldots, X_j) \), where \( i < j \).

Given \( x \in \mathcal{A}^\mathbb{N} \) and \( a_i^n \in \mathcal{A}^n \), let

\[
T_{a_i^n}(x) = \inf\{ k \geq 1 : x_k^{k+n-1} = a_i^n \}.
\]

This is the first time the string \( a_i^n \) appears in \( x \). Equivalently, \( T_{a_i^n}(x) = \inf\{ k \geq 1 : \theta^k x \in [a_i^n] \} \), that is, this is the first time the orbit of \( x \) enters the cylinder \( [a_i^n] \). Now define

\[
R_n(x) = \inf\{ k \geq 2 : x_k^n = x_k^{k+n-1} \}.
\]

This is the first time that the string \( x_1^n \) reappears in \( x \). The corresponding random variable will be simply denoted by

\[
R_n = \inf\{ k \geq 2 : X_k^n = X_k^{k+n-1} \}.
\]

Finally, we define the smallest return time in a cylinder \( [a_i^n] \), namely

\[
\tau(a_i^n) = \inf_{x \in [a_i^n]} T_{a_i^n}(x).
\]

One can check that \( \tau(a_i^n) = \inf\{ k \geq 1 : [a_i^n] \cap \theta^{-k}[a_i^n] \neq \emptyset \} \). Notice that for all \( n \geq 1 \) and \( a_i^n \in \mathcal{A}^n \)

\[
1 \leq \tau(a_i^n) \leq n. \tag{4}
\]

\[\text{We endow } \mathcal{A}^\mathbb{N} \text{ with the product topology, thus it is a compact space.} \]
2.2 The return-time and the hitting-time $L^q$-spectra

For $q \in \mathbb{R}$ and $n \in \mathbb{N}$, let

$$R_n^{(q)}(x) = \frac{1}{n} \log \int R_n^q(x) \, d\mu(x) \quad (\in \mathbb{R} \cup \{+\infty\})$$

and define the upper and lower return-time $L^q$-spectra by

$$\underline{R}_\mu(q) = \limsup_n R_n^{(q)}(x) \mu(q)$$  \hspace{1cm} (5)

$$\overline{R}_\mu(q) = \liminf_n R_n^{(q)}(x) \mu(q).$$  \hspace{1cm} (6)

**Definition 2.1** (Return-time $L^q$ spectrum). When both (5) and (6) coincide for all $q \in \mathbb{R}$, this defines the return-time $L^q$-spectrum which we denote by $R_\mu(q)$.

We list several basic properties of this function. There are at least two values of $q$ for which $R_\mu(q)$ always exist, namely $q = 0, 1$:

$$R_\mu(0) = 0 \quad \text{and} \quad R_\mu(1) = \log |A|.$$  

The first equality is trivial. The second one follows from the fact that, for every cylinder $[a^n]$, $\int_{[a^n]} R_n(x) \, d\mu(x) = 1$ (this is Kac’s lemma, see e.g. [Shi96]). It can happen that, for some $n_0$ and $q_0 > 0$ (hence for any $n \geq n_0$ and $q \geq q_0$), $R_n^{(q)}(x) = +\infty$. This can be already the case for $q = 2$ because there is no reason for the second moment of the return time to a cylinder to be finite for an arbitrary invariant probability measure. Since $R_n \geq 2$, $\overline{R}_\mu(q) \geq 0$ for $q > 0$ and $\underline{R}_\mu(q) \leq 0$ for $q < 0$. Moreover, $q \mapsto \overline{R}_\mu(q)$ and $q \mapsto \underline{R}_\mu(q)$ are increasing on $\mathbb{R}$ and, using Hölder’s inequality, one can check that they are convex. Hence, if $R_\mu(q)$ and $\overline{R}_\mu(q)$ are finite for all $q$, then they are continuous functions on $\mathbb{R}$ (see e.g. [RV73]). We have the following proposition.

**Proposition 1.** If $\mu$ is ergodic then, for all $q \in \mathbb{R}$,

$$qh(\mu) \leq \overline{R}_\mu(q).$$

Moreover, $\overline{R}_\mu$ are continuous functions on $(-\infty, 0]$.

**Proof.** By Jensen’s inequality we have

$$R_n^{(q)}(x) \geq q \int \frac{1}{n} \log R_n \, d\mu.$$  

Since $(1/n) \log R_n$ converges $\mu$-almost surely to $h(\mu)$, the entropy of $\mu$ (see (12) below for the definition), we get by Fatou’s lemma

$$\overline{R}_\mu(q) \geq q h(\mu).$$

Continuity of $\overline{R}_\mu$ on $(-\infty, 0]$ follows from convexity and finiteness. \qed
Consider \( T_{x^n}(y) \) where \( x \) and \( y \) are drawn independently according to the same probability measure \( \mu \). For \( q \in \mathbb{R} \) and \( n \in \mathbb{N} \), let

\[
W_{\mu}^{(n)}(q) = \frac{1}{n} \log \int \int T_{x^n}(y)^q \, d(\mu \otimes \mu)(x, y) \quad (\in \mathbb{R} \cup \{+\infty\})
\]

and define the upper and lower hitting-time \( L^q \)-spectra by

\[
\overline{W}_{\mu}(q) = \limsup_n W_{\mu}^{(n)}(q) \quad (7)
\]

\[
\underline{W}_{\mu}(q) = \liminf_n W_{\mu}^{(n)}(q). \quad (8)
\]

**Definition 2.2 (Hitting-time \( L^q \)-spectrum).** When both (7) and (8) coincide for all \( q \in \mathbb{R} \), this defines the hitting-time \( L^q \)-spectrum which we denote by \( W_{\mu}(q) \).

It can happen that, for some \( n_0 \) and \( q_0 > 0 \) (hence for any \( n \geq n_0 \) and \( q \geq q_0 \)), \( W_{\mu}^{(n)}(q) = +\infty \). This can be already the case for \( q = 1 \) because there is no version of Kac’s lemma for hitting times. We have \( \overline{W}_{\mu}(q) \geq 0 \) for \( q > 0 \) and \( \overline{W}_{\mu}(q) \leq 0 \) for \( q < 0 \). Moreover, \( q \mapsto \overline{W}_{\mu} \) and \( q \mapsto \underline{W}_{\mu} \) are increasing on \( \mathbb{R} \) and, using Hölder’s inequality, one can check that they are convex. Hence, if \( \overline{W}_{\mu}(q) \) and \( \underline{W}_{\mu}(q) \) are finite for all \( q \), then they are continuous functions on \( \mathbb{R} \).

We have the following proposition.

**Proposition 2.** If \( \mu \) is ergodic then, for all \( q \in \mathbb{R} \),

\[
qh(\mu) \leq W_{\mu}(q).
\]

Moreover, \( \overline{W}_{\mu} \) are continuous functions on \((-\infty, 0]\).

**Proof.** By Jensen’s inequality we have

\[
W_{\mu}^{(n)}(q) \geq q \int \int \frac{1}{n} \log T_{x^n}(y) \, d(\mu \otimes \mu)(x, y).
\]

By Fatou’s lemma

\[
\overline{W}_{\mu}(q) \geq q \int \int \liminf_n \frac{1}{n} \log T_{x^n}(y) \, d(\mu \otimes \mu)(x, y).
\]

It is proved in [Shi90, p. 202] that \( \liminf_n \frac{1}{n} \log T_{x^n}(y) = h(\mu) \), for \( \mu \otimes \mu \) almost every \((x, y)\). Therefore

\[
\overline{W}_{\mu}(q) \geq qh(\mu).
\]

Continuity of \( \overline{W}_{\mu} \) on \((-\infty, 0]\) follows from convexity and finiteness, as explained above. \(\square\)
2.3 The $L^q$-spectra of a measure and the Rényi entropy function

For $q \in \mathbb{R}$ and $n \in \mathbb{N}$, let
\[
M^{(n)}_\mu(q) = \frac{1}{n} \log \int \mu(x^n)^{-q} \, d\mu(x) = \frac{1}{n} \log \sum_{a_1^n \in A^n} \mu(a_1^n)^{1-q}
\]
and define the upper and lower $L^q$ spectra of $\mu$ by
\[
\overline{M}_\mu(q) = \limsup_n M^{(n)}_\mu(q) \quad (9)
\]
\[
\underline{M}_\mu(q) = \liminf_n M^{(n)}_\mu(q) \quad (10)
\]

**Definition 2.3 (L$^q$-spectrum of a measure).** When both (9) and (10) coincide for all $q \in \mathbb{R}$, this defines the $L^q$-spectrum of $\mu$ which we denote by $M_\mu(q)$. It is easy to check that $M_\mu(q) \geq 0$ on $[0, +\infty)$ and $M_\mu(q) \leq 0$ on $(-\infty, 0]$. Note that $M_\mu(0) = 0$ and $M_\mu(1) = \log |A|$, hence $M_\mu(q) = \mathcal{R}_\mu(q)$ for $q = 0, 1$. Moreover, the functions $\overline{M}_\mu$ and $\underline{M}_\mu$ are increasing, and by Hölder’s inequality, they are convex, hence continuous if they are finite for all $q$.

**Proposition 3.** If $\mu$ is ergodic, then, for all $q \in \mathbb{R}$, we have
\[
qh(\mu) \leq M_\mu(q).
\]
Moreover, $\overline{M}_\mu$ are continuous on $(-\infty, 0]$. 

**Proof.** By Jensen’s inequality, $M_\mu^{(n)}(q) \geq q \frac{1}{n} \log \mu(x^n)^{-q} \, d\mu(x)$ for any $q \in \mathbb{R}$. Since $q \frac{1}{n} \log \mu(x^n)^{-q} \, d\mu(x)$ converges $\mu$ almost surely to $h(\mu)$ by the the Shannon-McMillan-Breiman theorem, we conclude by using Fatou’s lemma. Continuity of $\overline{M}_\mu$ follows from the comments made before the statement of the proposition.

Finally, for $q \neq 0$, let
\[
\mathcal{H}^{(n)}_\mu(q) = -\frac{1}{qn} \log \sum_{a_1^n} \mu(a_1^n)^{q+1}
\]
and define the upper and lower Rényi entropy functions by
\[
\overline{\mathcal{H}}_\mu(q) = \limsup_n \mathcal{H}^{(n)}_\mu(q) \quad \text{and} \quad \underline{\mathcal{H}}_\mu(q) = \liminf_n \mathcal{H}^{(n)}_\mu(q). \quad (11)
\]
When both limits coincide for all $q \neq 0$, this defines the so-called Rényi entropy function which we denote by $\mathcal{H}_\mu$. By definition, $\mathcal{H}_\mu(0) = h(\mu)$, the entropy of $\mu$, where
\[
h(\mu) = -\lim_n \frac{1}{n} \sum_{a_1^n} \mu(a_1^n) \log \mu(a_1^n). \quad (12)
\]
(This limit always exists.) Moreover, $\mathcal{H}_\mu$ is a decreasing function and we obviously have
\[
\overline{\mathcal{H}}_\mu(q) = q \underline{\mathcal{H}}_\mu(-q), \quad q \in \mathbb{R}. \quad (13)
\]
3 A result for \( \phi \)-mixing processes with summable rate

A process \( \{X_k\}_{k=1}^{\infty} \) over a finite alphabet \( \mathcal{A} \), distributed according to a probability measure \( \mu \), is \( \phi \)-mixing if there exists a sequence \( (\phi(\ell))_{\ell \geq 1} \) of positive numbers decreasing to zero such that for all integers \( n \geq 1 \) and \( \ell \geq 1 \) we have

\[
\sup_{A \in \mathcal{F}_n^\infty, \mu(A) > 0} \left| \frac{\mu(A \cap B)}{\mu(A)} - \mu(B) \right| \leq \phi(\ell)
\]  

(14)

where \( \mathcal{F}_n^\infty \) denotes the \( \sigma \)-algebra generated by \( X_n^\infty \). We shall also say that \( \mu \) is \( \phi \)-mixing. We refer the reader to [Bra05] for a survey on mixing processes. It is well-known (see e.g. [Aba01]) that for such a process there exist two strictly positive constants \( c, C \) such that

\[
\mu(a^n_1) \leq C e^{-cn}
\]

(15)

for all \( n \geq 1 \) and \( a^n_1 \in \mathcal{A}^n \).

Let \( \mu \) be a probability measure on \( \mathcal{A}^\infty \). For \( a^n_1 \in \mathcal{A}^n \), let

\[
\zeta_\mu(a^n_1) = \mu_{a^n_1}(T_{a^n_1} \neq \tau(a^n_1)) = \mu_{a^n_1}(T_{a^n_1} > \tau(a^n_1))
\]

(16)

where

\[
\mu_{a^n_1}(T_{a^n_1} > \tau(a^n_1)) := \frac{1}{\mu(a^n_1)} \mu([a^n_1] \cap \{T_{a^n_1} > \tau(a^n_1)\})
\]

Remark 1. For \( t < \mu(x^n_1)\tau(x^n_1) \) we have

\[
\mu_{x^n_1} \left( T_{x^n_1} \leq \frac{t}{\mu(x^n_1)} \right) = 0
\]

since by definition \( \mu_{x^n_1} \left( T_{x^n_1} < \tau(x^n_1) \right) = 0 \) (whence the rightmost equality in (16)).

A priori \( \zeta_\mu(a^n_1) \in [0, 1] \). However, when \( \mu \) is \( \phi \)-mixing with a summable rate, it is bounded away from 0, uniformly in \( a^n_1 \), provided that \( n \) is sufficiently large.

Proposition 4. Let \( \mu \) be a \( \phi \)-mixing measure such that \( \sum_\ell \phi(\ell) < +\infty \). Then there exist \( N \geq 1 \) and \( \zeta_- > 0 \) such that for all \( n \geq N \)

\[
\inf_{a^n_1 \in \mathcal{A}^n} \zeta_\mu(a^n_1) \geq \zeta_- .
\]

(17)

\footnote{Without loss of generality, we assume that there exists no \( a \in \mathcal{A} \) such that \( \mu(a) \in \{0, 1\} \). We draw the attention of the reader to the fact that, in [Aba01], a \( \phi \)-mixing process is called a \( \alpha \)-mixing process (and a \( \psi \)-mixing process is called a \( \phi \)-mixing process). Finally, for the first part of Lemma 1 in [Aba01], which is (15), no assumption on the rate at which \( \phi(\ell) \) goes to 0 is in fact necessary.}
This result is of independent interest in view of a theorem of Abadi and Vergne that is essential in our analysis (see Theorem 4 below). Now define
\[ \Lambda(n)(\mu) = \frac{1}{n} \log \sum_{a_i^n} (1 - \zeta_{\mu}(a_i^n)) \mu(a_i^n) \]
and
\[ \overline{\Lambda}(\mu) := \limsup_n \Lambda(n)(\mu) \quad \text{and} \quad \underline{\Lambda} := \liminf_n \Lambda(n)(\mu). \]
Note that by the previous proposition \( \underline{\Lambda} < 0. \)

We can now formulate the main result of this section.

**Theorem 1.** Let \( \mu \) be a \( \phi \)-mixing measure such that \( \sum \phi(\ell) < +\infty. \) Then
\[ R_{\mu}(q) = \begin{cases} \overline{M}_\mu(q) & \text{for } q \geq 0 \\ \overline{\Lambda}(\mu) \lor \underline{M}_\mu(-1) & \text{for } q \leq -1 \end{cases} \]
The same statement holds when \( \overline{\Lambda}, \overline{M}_\mu \) and \( \underline{\Lambda} \) are replaced by \( \overline{P}_\mu, \underline{M}_\mu \) and \( \underline{\Lambda}(\mu). \)

**Remark 2.** The Rényi entropy function (hence the \( L^q \)-spectrum of \( \mu \)) exists and is finite on \( \mathbb{R} \) under several mixing conditions which are stronger than \( \phi \)-mixing \( [LS97, HV10, AC15] \), for instance \( \psi \)-mixing with \( \psi(1) < 1. \) (We recall the definition of \( \psi \)-mixing below.) In the \( \phi \)-mixing case, the existence of the Rényi entropy function is an open question. The problem comes from cylinders whose measure decays faster than \( \exp(-cn) \), e.g., like \( \exp(-cn^{1+\delta}) \) for some \( \delta > 0. \) Of course, these ‘bad’ cylinders are untypical in the sense of the Shannon-McMillan-Breiman theorem, but we are forced to take them into account.

## 4 Equilibrium states for potentials with summable variation

We mainly refer to [Wal75] for full details on the material of this section.

### 4.1 Basic definitions and properties

Let \( \varphi : A^\mathbb{N} \to \mathbb{R} \) be a continuous function (‘potential’). The measure \( \nu \) is an equilibrium state for \( \varphi \) if
\[ h(\nu) + \int \varphi \, d\nu = \sup \left\{ h(\eta) + \int \varphi \, d\eta \right\} = P(\varphi) \quad (18) \]
where the supremum is taken over the set of shift-invariant probability measures, and \( P(\varphi) \) is the topological pressure of \( \varphi. \) (We do not give here the definition of \( P(\varphi) \). It will appear below when it is really needed.)
Under the further assumption that
\[ \sum_n \text{var}_n(\varphi) < \infty \]  
(19)
where
\[ \text{var}_n(\varphi) = \sup \{ |\varphi(x) - \varphi(y)| : x_1^n = y_1^n \}, \]
there exists a unique equilibrium state which we denote by \( \mu_\varphi \) (which is shift-invariant).

Let \( L_\varphi \) denote the transfer operator. For \( h : A^\mathbb{N} \to \mathbb{R} \) continuous it is defined by
\[ (L_\varphi h)(x) = \sum_{a \in A} e^{\varphi(ax)} h(ax) \]  
(20)
where \( ax = (a, x_1, x_2, \ldots) \). It is known that there exist a number \( \lambda > 0 \), \( f : A^\mathbb{N} \to \mathbb{R} \) continuous, and \( \nu \) a probability measure such that \( f > 0, \int f \, d\nu = 1, L_\varphi f = \lambda f, \) and \( L_\varphi^* \nu = \lambda \nu \), and \( L_\varphi^* \) is its dual which acts on probability measures. One has \( \mu_\varphi = f \nu, \) and \( \lambda = e^{P(\varphi)}. \)

Without loss of generality we can assume that \( \varphi \) is normalized in the sense that \( \lambda = 1 \) and \( h = 1 \). This means that \( L_\varphi 1 = 1 \), which reads
\[ \forall x \in A^\mathbb{N}, \quad \sum_{a \in A} e^{\varphi(ax)} = 1. \]  
(21)

We also have
\[ L_\varphi^* \mu_\varphi = \mu_\varphi. \]  
(22)
Moreover
\[ h(\mu_\varphi) + \int \varphi \, d\mu_\varphi = \sup \left\{ h(\eta) + \int \varphi \, d\eta \right\} = P(\varphi) = 0 \]  
(23)
where the supremum is taken over the set of shift-invariant probability measures.

If \( \varphi \) is not normalized, one can define
\[ \forall x \in A^\mathbb{N}, \quad g(x) = e^{\varphi(x)} \frac{f(x)}{\lambda f(Tx)}. \]  
(24)
This is a well-defined function from \( A^\mathbb{N} \) to \((0,1)\), and \( \log g \) is the normalized version of \( \varphi \) and gives rise to the same equilibrium state.

Finally, recall that if \( \varphi \) is of summable variation, the map \( q \mapsto P(q \varphi) \) is convex and continuously differentiable with
\[ P'(q \varphi) = \int \varphi \, d\mu_{q \varphi} \]  
(25)
where \( \mu_{q \varphi} \) is the unique equilibrium state for \( q \varphi \). Moreover, it is strictly decreasing, and strictly convex if and only if \( \mu_\varphi \) is not the measure of maximal entropy, that is, the equilibrium state for a potential of the form \( u - u \circ \theta - \log |A| \), where \( u : A^\mathbb{N} \to \mathbb{R} \) is continuous. (Recall that since we normalize potentials, their pressure is equal to 0, and \( P(0 + u - u \circ \theta + c) = \log |A| + c \) where \( c \in \mathbb{R} \).)

We refer to [TV99] for a proof of these facts.
4.2 Main results

We work with a (normalized) potential $\varphi$ satisfying (19). We start with a proposition which provides the critical value of $q$ below which the return-time $L^q$-spectrum turns out to be different from the $L^q$-spectrum of $\mu$.

**Proposition 5.** Let $\mu_\varphi$ be the equilibrium state for a potential $\varphi$ with summable variation. Then, the equation

$$M_{\mu_\varphi}(q) = \sup_\eta \int \varphi \, d\eta \quad (q \in \mathbb{R})$$

has a unique solution $q^*_\varphi$. Moreover, $q^*_\varphi \in [-1, 0)$, and $q^*_\varphi = -1$ if and only if $\varphi = u - u \circ \theta - \log |A|$ for some continuous function $u : \mathcal{A}^\mathbb{N} \to \mathbb{R}$.

Given a probability measure $\nu$, let

$$\gamma^+(\nu) := \lim_{n \to \infty} \frac{1}{n} \log \max_{a_1^n} \nu(a_1^n)$$

whenever the limit exists. Usually, $-\gamma^+(\nu)$ is called the ‘min-entropy’ of $\nu$.

We can now formulate the main result of this section. It is a strengthening of Theorem 1 for processes which are distributed according to the equilibrium state of a potential with summable variation. Remember that $\varphi$ is normalized, in particular $P(\varphi) = 0$.

**Theorem 2.** Let $\varphi$ be a potential with summable variation. Then $\gamma^+(\mu_\varphi)$ and $R_{\mu_\varphi}$ exist.

If $\varphi$ is not of the form $u - u \circ \theta - \log |A|$ for some continuous function $u : \mathcal{A}^\mathbb{N} \to \mathbb{R}$ (i.e., $\mu_\varphi$ is not the measure of maximal entropy), then $q^*_\varphi \in (-1, 0)$, and

$$R_{\mu_\varphi}(q) = \begin{cases} M_{\mu_\varphi}(q) = P((1 - q)\varphi) & \text{for } q \geq q^*_\varphi \\ \gamma^+(\mu_\varphi) = \sup_\eta \int \varphi \, d\eta & \text{for } q \leq q^*_\varphi \end{cases}$$

where the supremum is taken over the set of shift-invariant probability measures.

If $\varphi = u - u \circ \theta - \log |A|$ for some continuous function $u : \mathcal{A}^\mathbb{N} \to \mathbb{R}$, then

$$R_{\mu_\varphi}(q) = \begin{cases} q \log |A| & \text{for } q \geq -1 \\ -\log |A| & \text{for } q < -1 \end{cases}$$

**Remark 3.** Observe that $R_{\mu_\varphi}$ is not differentiable at $q^*_\mu$. We have

$$\lim_{q \searrow q^*_\varphi} R'_{\mu_\varphi}(q) = -\int \varphi \, d\mu((1-q^*_\varphi)\varphi)$$

and

$$\lim_{q \nearrow q^*_\varphi} R'_{\mu_\varphi}(q) = 0.$$

**Remark 4.** Proposition 5 and Theorem 2 remain valid if we replace the full shift $\mathcal{A}^\mathbb{N}$ by a topologically mixing subshift of finite type $\mathcal{O} \subset \mathcal{A}^\mathbb{N}$. Indeed, the properties of the equilibrium states we consider are in fact valid in that case, see [Wal75]. Moreover, it is easy to check that there exists $m \geq 1$ such that, for all $n$ and all $a_1^n \in \mathcal{A}^n$, $\tau(a_1^n) \leq n + m$, which is a slight modification of [4].
Figure 1: The three functions coincide on the interval \([q^*_\varphi, +\infty)\). We assume that \(\varphi \neq u - u \circ \theta - \log |A|\).

Let us now compare this result with the hitting-time \(L^q\)-spectrum (see Definition 2.2). Let \(\mu\) be the equilibrium state for a potential \(\varphi\) with summable variation. Then, it was proved in [CU05] that

\[
W_{\mu, \varphi}(q) = \begin{cases} 
M_{\mu, \varphi}(q) = P((1 - q)\varphi) & \text{if } q \geq -1 \\
P(2\varphi) & \text{if } q < -1.
\end{cases}
\]

Comparing with Theorem 2, we see that if \(\varphi\) is not of the form \(u - u \circ \theta - \log |A|\), then \(R_{\mu, \varphi} \neq W_{\mu, \varphi}\) on the interval \((-\infty, q^*_\varphi) \supseteq (-\infty, -1)\). The fact that \(P(2\varphi) < \sup\eta \int \varphi \, d\eta\) follows from the proof of Proposition 5 when we prove that \(q^*_\varphi > -1\).

We have \(W_{\mu, \varphi} = R_{\mu, \varphi}\) if and only if \(\mu\) is the measure of maximal entropy (that is, \(\varphi = u - u \circ \theta - \log |A|\)), which is equivalent to \(q^*_\varphi = -1\). In this case, \(W_{\mu, \varphi}\) and \(R_{\mu, \varphi}\) are piecewise affine (see (28)).

Our last theorem is about large deviations for \(n^{-1} \log R_n\). By Theorem 2 we have \(R_{\mu, \varphi}(q) = P((1 - q)\varphi)\) for all \(q \in [q^*_\varphi, +\infty)\). By the properties of \(P(q\varphi)\) recalled in Subsection 4.1 we know that, on that interval, \(R_{\mu, \varphi}\) is strictly convex, hence its derivative is strictly increasing. Therefore, the equation

\[
R'_{\mu, \varphi}(q) = u
\]
has a unique solution \( \hat{q}_u \) for each \( u \in I \) where

\[
I := \left( -\int \varphi \, d\mu_{(1-q^*_\varphi)} - \inf \int \varphi \, d\eta \right).
\]

This interval is the set of values taken by \( R'_\mu \varphi(q) \) as \( q \) runs through \((q^*_\varphi, +\infty)\). Note that \( R'_\mu \varphi(0) = h(\mu_\varphi) \) lies inside this interval. Now let

\[
J_\varphi(u) = \begin{cases} 
  u \hat{q}_u - R'_\mu \varphi(\hat{q}_u) & \text{if } u \in I, \\
  +\infty & \text{if } u \geq -\inf \int \varphi \, d\eta.
\end{cases}
\]

On \( I \), this function is positive, strictly convex and vanishes only at \( h(\mu_\varphi) \).

**Theorem 3.** Let \( \varphi \) be a potential with summable variation. We assume that \( \varphi \) is not of the form \( u - u \circ \theta - \log |A| \) for some continuous function \( u : A^N \to \mathbb{R} \).

Then we have the following.

For all \( u \geq 0 \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \mu_\varphi \left( x : \frac{1}{n} \log R_n(x) > h(\mu_\varphi) + u \right) = -J_\varphi(h(\mu_\varphi) + u).
\]

For all \( u \in \left[ 0, h(\mu_\varphi) + \int \varphi \, d\mu_{(1-q^*_\varphi)} \right] \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \log \mu_\varphi \left( x : \frac{1}{n} \log R_n(x) < h(\mu_\varphi) - u \right) = -J_\varphi(h(\mu_\varphi) - u).
\]

**Proof.** Since \( q \mapsto R'_\mu \varphi(q) \) is not differentiable at \( q^*_\varphi \), we cannot apply directly Gârtner-Ellis theorem [DZ10]. The result follows from a theorem proved in [PS75] which can be seen as a variant of Gârtner-Ellis theorem.

Let us comment on the above theorem. It describes completely the large deviations of \( n^{-1} \log R_n \) above \( h(\mu_\varphi) \). However, it does not say anything when \( n^{-1} \log R_n(x) - h(\mu_\varphi) \leq -\int \varphi \, d\mu_{(1-q^*_\varphi)} \). Compare with the large deviations for \(-n^{-1} \log \mu_\varphi(x^*_{\eta}) \). In that case, the rate function is

\[
J_\varphi(u) = \begin{cases} 
  u \hat{q}_u - R'_\mu \varphi(\hat{q}_u) & \text{if } u \in \left( -\sup \int \varphi \, d\eta, -\inf \int \varphi \, d\eta \right), \\
  +\infty & \text{otherwise}
\end{cases}
\]

and we have

\[
I \subseteq \left( -\sup \int \varphi \, d\eta, -\inf \int \varphi \, d\eta \right).
\]

So, we do not know whether or not Theorem 3 fully describes the large deviations of \( n^{-1} \log R_n \) below \( h(\mu_\varphi) \). Indeed, we do not know if the event \( \{ n^{-1} \log R_n < h(\mu_\varphi) - u \} \) for \( u \in \left( -\sup \int \varphi \, d\eta, -\int \varphi \, d\mu_{(1-q^*_\varphi)} \right) \) has a probability decaying at exponential rate. In any case, our approach cannot work to tackle this problem.
5 Some explicit examples

5.1 Independent random variables

The return-time and the hitting-time spectra are non-trivial even when \( \mu \) is a product measure. Take for instance \( A = \{0, 1\} \) and let \( \mu = m^N \) where \( m \) is a Bernoulli measure on \( A \) with parameter \( p_1 \neq \frac{1}{2} \). This corresponds to a potential \( \varphi \) which is locally constant on the cylinders \([0]\) and \([1]\). We can identify it with a function from \( A \) to \( \mathbb{R} \) such that \( \varphi(1) = \log p_1 \). To be concrete, let us take \( p_1 = \frac{1}{3} \). Then it is easy to verify that

\[
M_\mu(q) = P((1-q)\varphi) = \log \left( \left( \frac{2}{3} \right)^{1-q} + \left( \frac{1}{3} \right)^{1-q} \right)
\]

and

\[
M_\mu(-1) = P(2\varphi) = \log \frac{5}{9} \quad \text{and} \quad \gamma^+(\mu) = \log \frac{2}{3}
\]

whence \( P(2\varphi) < \gamma^+(\mu) \), as expected. It is easy to solve numerically equation (26) to find

\[
q^*_\mu \approx -0.672814.
\]

Remark 5. We can go from any \( \mu \) which is not the measure of maximal entropy (case \( p_1 = 1/2 \)) by letting \( p_1 \) tend to \( 1/2 \) and one finds that \( M_\mu(q) \to q \log 2 \), \( M_\mu(-1) = P(2\varphi) = -\log 2 = \lim_{p_1 \to 1/2} \gamma^+(\mu) \), and \( \lim_{p_1 \to 1/2} q^*_\mu = -1 \).

5.2 Markov chains

Recall that if a potential \( \varphi \) depends only on the first two symbols, that is, \( \varphi(x) = \varphi(x_1, x_2) \), then the corresponding process is a Markov chain. For Markov chains on \( A = \{1, \ldots, K\} \) with matrix \( P = (P(a, b))_{a, b \in A} \), a well-known result [Szp93] for instance] states that

\[\gamma^+_\mu = \max_{1 \leq \ell \leq K} \max_{a_1^\ell \in C_\ell} \frac{1}{\ell} \log \prod_{i=1}^{\ell} P(a_i, a_{i+1}) \quad (29)\]

where \( C_\ell \) is the set of cycles of distinct symbols of \( A \), with the convention that \( a_{\ell+1} = a_1 \) (circuits). On the other hand, it is well-known [Szp93] that

\[M_\mu(q) = \log \lambda_{1-q} \]

where \( \lambda_\ell \) is the largest eigenvalue of the matrix \( ((P(a, b))_\ell)_{a, b \in A} \). This means that, in principle, everything is explicit for the Markov case. In practice, calculations are intractable even with some innocent-looking examples. Let us restrict to binary Markov chains \((A = \{0, 1\})\) which enjoy reversibility. In this case (29) simplifies to

\[\gamma^+(\mu) = \max_{i,j} \frac{1}{2} \log P(i, j) P(j, i). \quad (30)\]
(See for instance [KV16].) If we further assume symmetry, that is \( P(1,1) = P(0,0) \), then we obtain

\[
\mathcal{M}_\mu(q) = \log \left( P(0,0)^{1-q} + P(0,1)^{1-q} \right)
\]

and \( \gamma^+(\mu) = \max\{ \log P(0,0), \log P(0,1) \} \). If we want to go beyond the symmetric case, the explicit expression of \( \mathcal{M}(q) \) gets cumbersome. As an illustration, consider the case \( P(0,0) = 0.2 \) and \( P(1,1) = 0.6 \). Then

\[
\mathcal{M}_\mu(q) = \log \left( \frac{3^{-q}}{10} \sqrt{8^{-q}(32 \cdot 225^q - 12 \cdot 600^q + 8^q \cdot (15^q + 3 \cdot 5^q)^2)} \right) + \frac{3^{-q}}{10} (15^q + 3 \cdot 5^q) .
\]

From (30) we easily obtain

\[
\gamma^+(\mu) = \log(0.6).
\]

The solution of equation (26) can be found numerically:

\[
q^*_\mu \approx -0.870750.
\]

6 Proofs of the results of Section 3

6.1 Proof of Proposition 4

Since \( \mu(x^n_1)^{-1} \geq c^{-1}_{15} e^{e^{15^n} n} \), there exists \( n_0 \geq 1 \) such that\n
\[
\forall n \geq n_0, \forall x^n_1, \mu(x^n_1) > n \geq \tau(x^n_1) .
\]

Since \( (\mu x^n_1(T x^n_1 > j))_{j \geq 1} \) is a non-increasing sequence, we thus have

\[
\zeta_\mu(x^n_1) = \mu x^n_1(T x^n_1 > \tau(x^n_1)) \geq \frac{1}{\mu(x^n_1) - \tau(x^n_1)} \sum_{j=\tau(x^n_1)}^{\mu(x^n_1) - 1} \mu x^n_1(T x^n_1 > j)
\]

\[
\geq \frac{1}{\mu(x^n_1) - 1} \sum_{j=\tau(x^n_1)}^{\mu(x^n_1) - 1} \mu x^n_1(T x^n_1 > j)
\]

\[
\geq \sum_{j=\tau(x^n_1)}^{\mu(x^n_1) - 1} \mu([x^n_1] \cap \{ T x^n_1 > j \})
\]

(31)

where the third inequality is actually an equality. Now, using stationarity, observe that

\[
\mu([x^n_1] \cap \{ T x^n_1 > j \}) = \mu(T x^n_1 > j) - \mu([x^n_1] \cap \{ T x^n_1 > j \})
\]

\[
= \mu(T x^n_1 > j) - \mu(T x^n_1 > j + 1)
\]

\[
= \mu(T x^n_1 = j + 1) .
\]
Therefore by (31) we obtain
\[
\zeta(\mu(x^n_1) \geq \mu(T_{x^n_1} \leq \tau(x^n_1)) \geq \mu(T_{x^n_1} \leq \mu(x^n_1)^{-1}) - \mu(T_{x^n_1} \leq \mu(x^n_1)^{-1})
\]
\[
\geq \mu(T_{x^n_1} \leq \mu(x^n_1)^{-1}) - \tau(x^n_1)\mu(x^n_1) \geq \mu(T_{x^n_1} \leq \mu(x^n_1)^{-1}) - C_{15} n e^{-C_{15} n}
\]
where the second and third inequalities are actually equalities. The second one is trivial. For the third one, write
\[
\mu(T_{x^n_1} \leq \tau(x^n_1)) = \sum_{j=1}^{\tau(x^n_1)} \mu(T_{x^n_1} = j) = \tau(x^n_1)\mu(x^n_1).
\]

For the last inequality, we used (15) and the fact that \(\tau(x^n_1) \leq n\) for any \(n \geq 1\) and \(x^n_1 \in A^n\). We now look for a lower bound for \(\mu(T_{x^n_1} \leq \mu(x^n_1)^{-1})\). Let
\[
N = \sum_{j=1}^{\mu(x^n_1)^{-1}} \sum_{i=1}^j \{X_i + n - 1 = x^n_1\} \{X_{i+n-i-1} = x^n_1\}.
\]

By Cauchy-Schwarz inequality we have
\[
\mu(T_{x^n_1} \leq \mu(x^n_1)^{-1}) \geq \frac{\mathbb{E}(N)^2}{\mathbb{E}(N^2)}
\]
(33)
since \(\{T_{x^n_1} \leq \mu(x^n_1)^{-1}\} = \{N \geq 1\}\). By stationarity \(\mathbb{E}(N) = 1\). It remains to prove that there exists \(\mathbb{E}(N^2) \leq C\) for some \(C > 0\). Expanding \(N^2\), using stationarity and \(\mathbb{E}(N) = 1\) we obtain
\[
\mathbb{E}(N^2) = \mathbb{E} \left( \sum_{i=1}^{\mu(x^n_1)^{-1}} \sum_{j=1}^{\mu(x^n_1)^{-1}} \{X_i + n - 1 = x^n_1\} \{X_{i+n+i-1} = x^n_1\} \right)
\]
\[
= 1 + 2 \sum_{j=1}^{\mu(x^n_1)^{-1}} (\mu(x^n_1)^{-1} - j) \mu(X_1^n = X_{j+n-1} = x^n_1).
\]
(34)
We now estimate the sum using $\phi$-mixing by splitting it at $j = n$. We start by the sum starting from $n$:

$$
\sum_{j=n}^{n-1} (\mu(x^n_1)^{-1} - j) \mu(X^n_1 = X^{j+n}_{j+1} = x^n_1)
$$

$$
\leq \mu(x^n_1)^2 \sum_{j=n}^{n-1} (\mu(x^n_1)^{-1} - j)
$$

$$
+ (\mu(x^n_1)^{-1} - n) \mu(x^n_1) \sum_{j=n}^{n-1} \phi(j - n + 1)
$$

$$
\leq 1 + \sum_{\ell=1}^{\mu(x^n_1)^{-1}} \phi(\ell). \quad (35)
$$

When $1 \leq j \leq n - 1$, we write

$$
\mu(X^n_1 = X^{j+n}_{j+1} = x^n_1) = \mu(X^n_1 = x^n_1, X^{n+j}_{n+1} = x^n_{n-j+1}).
$$

Now we ‘open a gap’ of size $j/2$ in the second event and use $\phi$-mixing to get

$$
\mu(X^n_1 = X^{j+n}_{j+1} = x^n_1) \leq \mu(X^n_1 = x^n_1, X^{n+[j/2]}_{n+1} = x^n_{n-[j/2]+1})
$$

$$
\leq \mu(x^n_1) \left( \mu(x^n_{n-[j/2]+1}) + \phi([j/2]) \right).
$$

Therefore

$$
\sum_{j=1}^{n-1} (\mu(x^n_1)^{-1} - j) \mu(X^n_1 = X^{j+n}_{j+1} = x^n_1) \leq \sum_{j=1}^{\lfloor n/2 \rfloor} (\rho_j + \phi(j)), \quad (36)
$$

where $\rho_j := \max_{y^n_1} \mu(y^n_1) \leq C(15)^j e^{-c(15)^j}$. To conclude, we use the assumption $\sum_{\ell} \phi(\ell) < +\infty$ and (33), (34), (35) and (36) to conclude that for all $n \geq n_0$

$$
\inf_{\alpha^n_1 \in A^n} \zeta(x^n_1) \geq \frac{1}{1 + 2 \sum_{\ell} \phi(\ell) + C(15) \sum_{\ell} e^{-\ell(15)^\ell} - C(15) n e^{-c(15)^n}}.
$$

Clearly, there exits $N \geq n_0$ such that the right-hand side of this inequality is equal to some strictly positive number that we call $\zeta_-$. The proposition is proved.

### 6.2 Proof of Theorem 1

#### 6.2.1 First estimates

We start by stating a key-result established by Abadi and Vergne.
Theorem 4 (AV09). Let \( \{X_k\}_{k \geq 1} \) be a \( \phi \)-mixing process distributed according to \( \mu \). There exists \( c, C > 0 \) such that, for any \( x \in \mathbb{R}^N \), \( n \geq 1 \) and \( t \geq \mu(x_1^n) \tau(x_1^n) \) we have

\[
\begin{align*}
|\mu_{x_1^n}(T_{x_1^n} &\leq \frac{t}{\mu(x_1^n)}) - \left[1 - \zeta_{\mu}(x_1^n) e^{\zeta_{\mu}(x_1^n) \mu(x_1^n) \tau(x_1^n) e^{-\zeta_{\mu}(x_1^n) t}}\right] | \\
&\leq 54 \epsilon(x_1^n) t e^{-\left(\zeta_{\mu}(x_1^n) - 16\epsilon(x_1^n)\right) t}
\end{align*}
\]

(37)

where

\[
\max_{x_1^n \in \mathbb{R}^n} \epsilon(x_1^n) \leq C e^{-c n} + \phi \left( \frac{n}{2} \right) .
\]

(38)

In order to prove the theorem \ref{theorem:av09} we will obtain upper and lower bounds for \( E[R_n^q] \), \( q \in \mathbb{R} \). The proof in the case \( q \geq 0 \) is very similar to the proof given in [CU09]. We give the proof in Appendix \ref{appendix:c} for the sake of completeness. So we consider only the case \( q < 0 \).

For any \( q < 0 \) we have

\[
E[R_n^{-|q|}] = \sum_{x_1^n} \mu(x_1^n) E_{x_1^n}[T_{x_1^n}^{-|q|}] = \sum_{x_1^n} \mu(x_1^n) \int_0^1 \mu_{x_1^n}(T_{x_1^n}^{-|q|} \geq s) \, ds
\]

and

\[
\begin{align*}
\sum_{x_1^n} \mu(x_1^n) \int_0^1 \mu_{x_1^n}(T_{x_1^n}^{-|q|} \geq s) \, ds &= \sum_{x_1^n} \mu(x_1^n) \int_0^1 \mu_{x_1^n}(T_{x_1^n} \leq s^{-1/|q|}) \, ds \\
&= -|q| \sum_{x_1^n} \mu(x_1^n)^{|q|+1} \int_0^{\mu(x_1^n)} t^{-|q|-1} \mu_{x_1^n}(T_{x_1^n} \leq \frac{t}{\mu(x_1^n)}) \, dt \\
&= |q| \sum_{x_1^n} \mu(x_1^n)^{|q|+1} \int_{\mu(x_1^n)}^{\infty} t^{-|q|-1} \mu_{x_1^n}(T_{x_1^n} \leq \frac{t}{\mu(x_1^n)}) \, dt.
\end{align*}
\]

(39)

We now use (37) to get

\[
\begin{align*}
|E[R_n^{-|q|}] - |q| \sum_{x_1^n} \mu(x_1^n)^{|q|+1} I(x_1^n, q)| &\leq |q| \sum_{x_1^n} \mu(x_1^n)^{|q|+1} J(x_1^n, q)
\end{align*}
\]

(40)

where

\[
I(x_1^n, q) := \int_{\mu(x_1^n) \tau(x_1^n)}^{\infty} t^{-|q|-1} \left[1 - \zeta_{\mu}(x_1^n) e^{\zeta_{\mu}(x_1^n) \mu(x_1^n) \tau(x_1^n) e^{-\zeta_{\mu}(x_1^n) t}}\right] \, dt
\]

and

\[
J(x_1^n, q) := 54 \epsilon(x_1^n) \int_{\mu(x_1^n) \tau(x_1^n)}^{\infty} t^{-|q|} e^{-\left(\zeta_{\mu}(x_1^n) - 16\epsilon(x_1^n)\right) t} \, dt.
\]

21
6.2.2 Estimations of \( I(x_1^n, q) \) and \( J(x_1^n, q) \)

We take \( n \) large enough such that there exists \( \delta > 0 \) such that for all \( x_1^n \)

\[
\zeta_\mu(x_1^n) - 16 \epsilon(x_1^n) \geq \delta > 0 .
\]

(41)

This is possible in view of [35] and Proposition [1] which tells us that \( \zeta_\mu(x_1^n) \)
is uniformly bounded away from 0 by \( \zeta_- > 0 \) for \( n \) large enough. This is why we
need to assume \( \sum \phi(t) < +\infty \) although Theorem [4] is valid without any
restriction on the speed at which \( \phi(t) \) goes to 0.

In the remaining of this section section, we will use the shorthand no tation
\( A = x_1^n \).

Bounding \( I(A, q) \).

\[
I(A, q) = \int_{\mu(A) \tau(A)}^\infty t^{-|q| - 1} \left[ 1 - \zeta_\mu(A) e^{\zeta_\mu(A) \mu(A) \tau(A)} e^{-\zeta_\mu(A) \tau(A)} \right] \, dt
\]

\[
= \int_{\mu(A) \tau(A)}^\infty t^{-|q| - 1} \, dt - \zeta_\mu(A) e^{\zeta_\mu(A) \mu(A) \tau(A)} \int_{\mu(A) \tau(A)}^\infty t^{-|q| - 1} e^{-\zeta_\mu(A) \tau(A)} \, dt
\]

\[
= \int_{\mu(A) \tau(A)}^\infty t^{-|q| - 1} \, dt - \zeta_\mu(A)^{|q| + 1} e^{\zeta_\mu(A) \mu(A) \tau(A)} \int_{\mu(A) \tau(A) \zeta_\mu(A)}^\infty u^{-|q| - 1} e^{-u} \, du
\]

\[
= \left( \mu(A) \tau(A) \right)^{-|q|} - \zeta_\mu(A)^{|q| + 1} e^{\zeta_\mu(A) \mu(A) \tau(A)} \Gamma(-|q|, \zeta_\mu(A) \mu(A) \tau(A))
\]

(42)

where, for \( s \in \mathbb{R} \) and \( x > 0 \), \( \Gamma(s, x) := \int_x^\infty u^{s - 1} e^{-u} \, du \), which is usually called
the upper incomplete gamma function. Using estimates that are collected in Section A we get the following bounds.

For \( q < -1 \), we combine (42) and (43) to obtain

\[
|q| I(A, q) \begin{cases} \geq (\mu(A) \tau(A))^{-|q|} \left[ 1 - \zeta_\mu(A) (1 - C_q \zeta_\mu(A) \mu(A) \tau(A)) \right] \\ \leq (\mu(A) \tau(A))^{-|q|} \left[ 1 - \zeta_\mu(A) (1 - C_q \zeta_\mu(A) \mu(A) \tau(A)) \right] \end{cases}
\]

(43)

where \( C_q \) and \( C_q \) depend only on \( q \) (see (44)).

For \( q \in (-1, 0) \), we combine (42) and (76) to obtain

\[
|q| I(A, q) \begin{cases} \geq \left( \mu(A) \tau(A) \right)^{-|q|} \left[ 1 - \zeta_\mu(A) \left( 1 - C_q (\zeta_\mu(A) \mu(A) \tau(A))^{|q|} \right) \right] \\ \leq \left( \mu(A) \tau(A) \right)^{-|q|} \left[ 1 - \zeta_\mu(A) \left( 1 - C_q (\zeta_\mu(A) \mu(A) \tau(A))^{|q|} \right) \right] \end{cases}
\]

(44)

Bounding \( J(A, q) \).

\[
J(A, q) = 54 \epsilon(A) \int_{\mu(A) \tau(A)}^\infty t^{-|q|} e^{-\zeta_\mu(A) \mu(A) \tau(A) - 16 \epsilon(A) \tau(A)} \, dt
\]

\[
= 54 \epsilon(A) (\zeta_\mu(A) - 16 \epsilon(A))^{-|q| - 1} \int_{\mu(A) \tau(A) \zeta_\mu(A) - 16 \epsilon(A)}^\infty t^{-|q|} e^{-t} \, dt
\]

\[
= 54 \epsilon(A) (\zeta_\mu(A) - 16 \epsilon(A))^{-|q| - 1} \Gamma(-|q| + 1, \mu(A) \tau(A) (\zeta_\mu(A) - 16 \epsilon(A))) .
\]

22
Using (72), we get for $q < -1$

$$J(A, q) \leq 54 \epsilon(A)(\mu(A)\tau(A))^{-|q|+1} \frac{e^{-\mu(A)\tau(A)(\zeta(A) - 16\epsilon(A))}}{|q| - 1} \leq \frac{54}{|q| - 1} \epsilon(A)(\mu(A)\tau(A))^{-|q|+1}$$

(45)

where we used (41). For $q \in (-1, 0)$, it is easy to check that there exists $D_q > 0$ such that

$$J(A, q) \leq D_q \epsilon(A)$$

(46)

where $D_q$ depends only on $q$.

6.2.3 End of the proof of Theorem 1

We now come back to (40).

The case $q < -1$. Using (43), (45), and Proposition 4, we obtain the following upper bound, for $n$ large enough:

$$\mathbb{E}\left[R_n^{-|q|}\right] \leq |q| \sum x_1^n \mu(x_1^n)^{1+|q|} (I(x_1^n, q) + J(x_1^n, q))$$

$$\leq \sum x_1^n (1 - \zeta(x_1^n)) \mu(x_1^n) + B_q (1 + \max a_1^n \epsilon(x_1^n)) \sum x_1^n \mu(x_1^n)^2$$

$$\leq \sum x_1^n (1 - \zeta(x_1^n)) \mu(x_1^n) + 2B_q \sum x_1^n \mu(x_1^n)^2$$

(47)

where $B_q := \max(c_q, 54|q|/(|q| - 1))$. We used the following basic facts: $\tau(x_1^n) \geq 1$ for any $x_1^n$ (whence $\tau(x_1^n)^{-|q|} \leq 1$ and $\tau(x_1^n)^{1-|q|} \leq 1$), $\max a_1^n \zeta(x_1^n) \leq 1$, and there exists $n_0 \geq 1$ such that $\max a_1^n \epsilon(x_1^n) \leq 1$ for all $n \geq n_0$ (see (38)).

Using (43) and (45), we have the following lower bound:

$$\mathbb{E}\left[R_n^{-|q|}\right] \geq n^{-|q|} \left[ \sum x_1^n (1 - \zeta(x_1^n)) \mu(x_1^n) + n\left( C_q \zeta_2^2 - \frac{54|q| \max a_1^n \epsilon(x_1^n)}{|q| - 1} \right) \sum x_1^n \mu(x_1^n)^2 \right].$$

(48)

We used the fact that $\tau(x_1^n) \leq n$. Note that because of (38) the prefactor of $\sum x_1^n \mu(x_1^n)^2$ is positive for $n$ large enough.
From (47) we deduce that
\[ \limsup_n \frac{1}{n} \log \mathbb{E}[R_n^{-|q|}] \leq \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1). \]
From (48) we deduce that
\[ \limsup_n \frac{1}{n} \log \mathbb{E}[R_n^{-|q|}] \geq \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1). \]
Hence
\[ \limsup_n \frac{1}{n} \log \mathbb{E}[R_n^{-|q|}] = \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1). \]

We proceed in the same way for the limit inferior. The statement of Theorem 1 for \( q < -1 \) is thus proved.

The case \( q \in (-1, 0) \). Using (44) and (46), and using the same basic facts as above, we get for \( n \) large enough
\[ \mathbb{E}[R_n^{-|q|}] \leq \sum_{x_1^n} (1 - \zeta_\mu(x_1^n)) \mu(x_1^n) + \max(c'_q, D_q) \sum_{x_1^n} \mu(x_1^n)^{1+|q|}. \]

Using (44), (46), and Proposition 4, we get the following lower bound:
\[ \mathbb{E}[R_n^{-|q|}] \geq n^{-|q|} \sum_{x_1^n} (1 - \zeta_\mu(x_1^n)) \mu(x_1^n) \]
\[ + (C_q \zeta_\mu^{1+|q|} - |q| D_q \max \epsilon(x_1^n)) \sum_{x_1^n} \mu(x_1^n)^{1+|q|}. \]

Proceeding as in the case \( q < -1 \), we obtain the statement of Theorem 1 for \( q \in (-1, 0) \).

The case \( q = -1 \). This case is treated by a continuity argument. By definition, for each \( n \), \( q \mapsto \mathcal{R}_\mu^{(n)}(q) \) is monotonically increasing in \( q \), thus, for any \( \epsilon > 0 \), we have
\[ \mathcal{R}_\mu^{(n)}(-1 - \epsilon) \leq \mathcal{R}_\mu^{(n)}(-1) \leq \mathcal{R}_\mu^{(n)}(-1 + \epsilon). \]

Using what we obtained in the cases \( q < -1 \) and \( q \in (-1, 0) \), we get
\[ \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1) \leq \limsup_n \mathcal{R}_\mu^{(n)}(-1) \leq \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1 + \epsilon). \]

Since \( \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1 + \epsilon) \to \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1) \) as \( \epsilon \to 0 \) (this follows from the discussion after Definition 2.3), we must have \( \mathcal{R}_\mu^{(n)}(-1) = \overline{\Lambda}(\mu) \lor \overline{M}_\mu(-1) \) as claimed. The limit inferior follows identically.

The proof of Theorem 1 is now complete.

\[^5\text{We use the elementary fact that, if } (a_n) \text{ and } (b_n) \text{ are sequences of positive real numbers, then} \limsup_n \frac{1}{n} \log(a_n + b_n) = (\limsup_n \frac{1}{n} \log a_n) \lor (\limsup_n \frac{1}{n} \log b_n). \text{ The same holds for the limit inferior.}\]
7 Proofs of the results of Section 4

7.1 Two preliminary general results

We consider a larger class of equilibrium states than the one considered in Section 4 (see again [Wal75]). Let $g : \mathcal{A}^\mathbb{N} \to (0, 1)$ be a continuous function such that

$$\forall x \in \mathcal{A}^\mathbb{N}, \quad \sum_{a \in \mathcal{A}} g(ax) = 1.$$  \hfill (49)

Letting $\varphi = \log g$, the previous condition can be rewritten as

$$L^* \varphi 1 = 1.$$  \hfill (50)

By an abstract compactness argument, there exists at least one probability measure $\mu$ such that

$$L^* \varphi \mu = \mu$$  \hfill (50)

and it is known that (50) holds if and only if $\mu$ is an equilibrium state for $\varphi$:

$$h(\mu) + \int \varphi \, d\mu = \sup \left\{ h(\eta) + \int \varphi \, d\eta \right\} = P(\varphi) = 0.$$  

In general, there are several equilibrium states for a given $\varphi$. A basic result is that, for any $n \geq 1$, any cylinder $[a^n_1]$ and any $x \in [a^n_1]$,

$$e^{-n \varepsilon_n} \leq \frac{\mu(a^n_1)}{\exp \left( \sum_{k=1}^{n} \varphi(x^\infty_k) \right)} \leq e^{n \varepsilon_n}$$ \hfill (51)

where

$$\varepsilon_n = \frac{1}{n} \sum_{k=1}^{n} \text{var}_k(\varphi).$$  \hfill (52)

Of course, $\varepsilon_n \to 0$ by Cesàro lemma. (We provide a proof of (51) in Appendix B since we couldn’t find a reference.) As a consequence, for all $n, m \geq 1$ and any pair $a^n_1 \in \mathcal{A}^n, b^m_1 \in \mathcal{A}^m$, we have

$$e^{-2(n \varepsilon_n + m \varepsilon_m)} \leq \frac{\mu(a^n_1 b^m_1)}{\mu(a^n_1) \mu(b^m_1)} \leq e^{2(n \varepsilon_n + m \varepsilon_m)}.$$  \hfill (53)

Remark 6. The equilibrium states we consider are called $g$-measures in [Wal75], but for some other authors, $g$-measures are more generally defined. They are a special class of ‘chains with complete connections’. We refer to [FM05] for more informations.

The next proposition shows that $\gamma^+(\mu)$ (see (27) for the definition) satisfies a variational principle.
Proposition 6. Let $\mu$ be an equilibrium state for a potential $\varphi = \log g$ as above. Then $\gamma^+(\mu)$ exists and
\begin{equation}
\gamma^+(\mu) = \sup_\eta \int \varphi \, d\eta
\end{equation}
where the supremum (which is attained) is taken over all shift-invariant probability measures.

Proof. For each $n \geq 1$ let
\begin{equation}
s_n(\varphi) := \sup_y \sum_{k=1}^n \varphi(y_k^n).
\end{equation}
We have
\begin{equation}
s_n(\varphi) = \max_{a_1^n} \sup_{y: y_1^n = a_1^n} \sum_{k=1}^n \varphi(y_k^n) = \max_{a_1^n} \sum_{k=1}^n \varphi(a_k^n y_{n+1}^\infty).
\end{equation}
Since $\mathcal{A}_N$ is compact and $\varphi$ is continuous, for each $n$ there exists a point $z^{(n)} \in \mathcal{A}_N$ such that
\begin{equation}
s_n(\varphi) = \max_{a_1^n} \sum_{k=1}^n \varphi(a_k^n (z^{(n)})_{n+1}^\infty).
\end{equation}
Now using (51) we get
\begin{equation}
\left| \frac{1}{n} \max_{a_1^n} \log \mu(a_1^n) - \frac{1}{n} \max_{a_1^n} \sum_{k=1}^n \varphi(a_k^n x_{n+1}^\infty) \right| \leq \varepsilon_n
\end{equation}
for any choice of $x_{n+1}^\infty \in \mathcal{A}_N^\infty$ and for all $n \geq 1$. Let us take $x_{n+1}^\infty = (z^{(n)})_{n+1}^\infty$. By using (56) and (55) we thus obtain
\begin{equation}
\left| \frac{1}{n} \max_{a_1^n} \log \mu(a_1^n) - \frac{s_n(\varphi)}{n} \right| \leq \varepsilon_n, \; n \geq 1.
\end{equation}
One can check that $(s_n(\varphi))_n$ is subadditive and $\inf_m m^{-1} s_m(\varphi) \geq -\|\varphi\|_\infty$. It follows by Fekete’s lemma (see e.g. [Szp01]) that $\lim_n n^{-1} s_n(\varphi)$ exists. Since $\varepsilon_n$ goes to 0, the limit of $\left( n^{-1} \max_{a_1^n} \log \mu(a_1^n) \right)$ also exists and coincides with $\lim_n n^{-1} s_n(\varphi)$. We now use the fact that
\begin{equation}
\lim_n \frac{s_n(\varphi)}{n} = \sup_\eta \int \varphi \, d\eta.
\end{equation}
The proof is found in [Jen06, Proposition 2.1]. Hence (54) is proved. \hfill \Box

We have the following proposition about existence of the Rényi entropy function and one of its properties to be used later on.
Proposition 7. Let \( \mu \) be an equilibrium state for a potential \( \varphi = \log g \) as above. For any \( q \in \mathbb{R} \), \( M_\mu(q) \) exists and
\[
M_\mu(q) = P((1 - q)\varphi).
\]
In particular, for all \( q \in \mathbb{R} \setminus \{0\} \), one has
\[
H_\mu(q) = -\frac{P((1 + q)\varphi)}{q}.
\]
For \( q = 0 \), \( H_\mu(0) = h(\mu) \) if and only if \( \mu \) is the unique equilibrium state for \( \varphi \). Moreover
\[
H_\mu(q) \leq -\gamma^+(\mu) \quad \text{as} \quad q \to +\infty.
\]
Proof. The identity \( M_\mu(q) = P((1 - q)\varphi) \) follows easily from Definition (2.3) and (51), and the identity \( H_\mu(q) = -\frac{P((1 + q)\varphi)}{q} \) for \( q \neq 0 \) follows at once from (13). Now, suppose that \( \varphi \) has a unique equilibrium state. This means that
\[
\lim_{q \to 0} \frac{d}{dq} P(\varphi + q\varphi) \bigg|_{q=0} = \int \varphi d\mu.
\]
So we can use l'Hospital rule to conclude that \( H_\mu(q) \xrightarrow{q \to 0} -\gamma^+(\mu) \) by (23).
To prove the last statement, we use the variational principle (18) with \( q\varphi \), \( q \in \mathbb{R} \) instead of \( \varphi \), to get
\[
P(q\varphi) \geq h(\eta) + q \int \varphi d\eta
\]
for any shift-invariant probability measure \( \eta \). Hence, for any \( q > 0 \) we get
\[
\frac{P(q\varphi)}{q} \geq \int \varphi d\eta + \frac{h(\eta)}{q}.
\]
Hence
\[
\liminf_{q \to +\infty} \frac{P(q\varphi)}{q} \geq \int \varphi d\eta
\]
and taking \( \eta \) to be a maximizing measure for \( \varphi \) we obtain
\[
\liminf_{q \to +\infty} \frac{P(q\varphi)}{q} \geq \sup_{\eta} \int \varphi d\eta.
\]
(By compactness of the set of shift-invariant probability measures, there exists at least one shift-invariant measure maximizing \( \int \varphi d\eta \).) We now use the following formula giving \( P(q\varphi) \):
\[
P(q\varphi) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{a_n^1} e^{qS_n\varphi(x(a_n^1))}
\]
where, for each $a^n_1$, $x(a^n_1)$ is an arbitrary point in $[a^n_1]$. It can be easily deduced from [Bow08 Lemma 1.20]. Now, for any $q > 0$, we have the trivial bound

$$\frac{1}{n} \log \sum_{a^n_1} e^{q S_n \varphi(x(a^n_1))} \leq q \frac{1}{n} \sup_y S_n \varphi(y) + \log |A|.$$  

Hence, by taking the limit $n \to \infty$ on both sides, and using (57), we have for any $q > 0$

$$\frac{P(q \varphi)}{q} \leq \sup_\eta \int \varphi \, d\eta + \frac{\log |A|}{q}.$$  

It remains to take the limit $q \to +\infty$ to get

$$\limsup_{q \to +\infty} \frac{P(q \varphi)}{q} \leq \sup_\eta \int \varphi \, d\eta.$$  

Combining this inequality with (59) we thus proved that

$$\lim_{q \to +\infty} \frac{P(q \varphi)}{q} = \sup_\eta \int \varphi \, d\eta. \quad (61)$$  

To finish the proof, recall that $\mathcal{H}_\mu(q) = -\frac{1}{q} M_\mu(-q)$ and use Proposition 5 to conclude. □

### 7.2 Proof of Proposition 5

The map $q \mapsto M_{\mu_\varphi}(q)$ is a bijection from $\mathbb{R}$ to $\mathbb{R}$. This can be easily checked using the identity $M_{\mu_\varphi}(q) = P((1 - q) \varphi)$ and the facts on the pressure function recalled in Subsection 4.4 together with the fact that $M_{\mu_\varphi}(q) \to +\infty$ as $q \to +\infty$ and $M_{\mu_\varphi}(q) \to -\infty$ as $q \to -\infty$. This is because

$$M_{\mu_\varphi}(q) \sim (-\inf \int \varphi \, d\eta) \, q \quad \text{as } q \to +\infty$$

which follows from (61). It is easy to check that, as $q \to -\infty$, one has to replace the infimum by a supremum.

Since $\varphi < 0$, we know that $q^*_\varphi < 0$ since $M_{\mu_\varphi}(q) \geq 0$ for $q \geq 0$ and $M_{\mu_\varphi}(q) \leq 0$ for $q \leq 0$. Now, using Proposition 7 and the variational principle (18) twice (first for $2 \varphi$, then for $\varphi$) we obtain (recall that $P(\varphi) = 0$)

$$M_{\mu_\varphi}(-1) = P(2 \varphi) = h(\mu_{2 \varphi}) + 2 \int \varphi \, d\mu_{2 \varphi}$$

$$= h(\mu_{2 \varphi}) + \int \varphi \, d\mu_{2 \varphi} + \int \varphi \, d\mu_{2 \varphi}$$

$$\leq P(\varphi) + \int \varphi \, d\mu_{2 \varphi} \leq \int \varphi \, d\mu_{2 \varphi}$$

$$\leq \sup_\eta \int \varphi \, d\eta.$$
Hence \( q^* \in [-1, 0) \) since \( q \mapsto M_{\mu, \varphi}(q) \) is strictly increasing. We now analyse the ‘critical case’, that is, \( q^* = -1 \). If \( \varphi = u - u \circ \theta - \log |A| \) where \( u : \mathcal{A}^N \to \mathbb{R} \) is continuous, the the equation \( M_{\mu, \varphi}(q) = \sup \int \varphi \, d\eta \) boils down to \( q \log |A| = -\log |A| \), whence \( q^* = -1 \). We now prove the converse. To this end, we need a few properties of the map \( q \mapsto H_{\mu, \varphi}(q) \) on \((0, +\infty)\). Using (58) and (25), and then the variational principle (18) twice, we get

\[
H'_{\mu, \varphi}(q) = \frac{1}{q^2} \left( P((1 + q)\varphi) - q \int \varphi \, d\mu_{(1+q)\varphi} \right) \\
= \frac{1}{q^2} \left( h(\mu_{(1+q)\varphi}) + \int \varphi \, d\mu_{(1+q)\varphi} \right) \\
\leq \frac{P(\varphi)}{q^2} = 0 .
\] (62)

Hence, \( H'_{\mu, \varphi}(q) \leq 0 \) on \((0, +\infty)\). Moreover, \( H_{\mu, \varphi}(q) \) decreases to \(-\gamma^+(\mu)\) as \( q \to +\infty \), in particular \( \lim_{q \to +\infty} H_{\mu, \varphi}(q) = 0 \) (last statement of Proposition 7).

Now, the condition \( q^* > -1 \) is equivalent to \( M_{\mu, \varphi}(-1) < \sup \eta \int \varphi \, d\eta \), which in turn is equivalent to \( H(\mu_{\varphi}) > -\gamma^+(\mu_{\varphi}) \) (by [13] and Proposition 5). For the latter inequality to hold, it is sufficient that \( H'(\mu_{\varphi}) \) decreases to \(-\gamma^+(\mu_{\varphi})\) as \( q \to +\infty \), that is, if and only if \( \mu_{\varphi} = \mu_{\varphi} \) (since \( \mu_{\varphi} \) is the unique equilibrium state of \( \varphi \)), that is, if and only if there exists a continuous function \( u : \mathcal{A}^N \to \mathbb{R} \) and \( c \in \mathbb{R} \) such that

\[
2\varphi = \varphi + u - u \circ \theta + c
\]
which is equivalent to

\[
\varphi = u - u \circ \theta + c .
\]

Since \( P(\varphi) = 0 \), one must have \( c = -\log |A| \). Therefore, if \( \varphi \) is not of the form \( u - u \circ \theta - \log |A| \), then \( H'(\mu_{\varphi}) < 0 \), which means that \( M_{\mu_{\varphi}}(-1) < \sup \eta \int \varphi \, d\eta \). The proof of the proposition is complete.

7.3 Proof of Theorem [2]

We want to apply Theorem [1]. The main point is to prove that \( \overline{\Lambda}(\mu_{\varphi}) = \Lambda(\mu_{\varphi}) = \gamma^+(\mu_{\varphi}) \) when \( \mu_{\varphi} \) is the equilibrium state of \( \varphi \). We first recall basic estimates to control the measure of cylinders.

7.3.1 Control of the measure of cylinders

Since we assume that \( \varphi \) is of summable variation, we can strengthen the following way. There exists a constant \( C = C(\varphi) > 1 \) such that for any \( n \geq 1 \), any cylinder \([a^n_i] \) and any \( x \in [a^n_i] \),

\[
C^{-1} \leq \frac{\mu_{\varphi}(a^n_i)}{\exp \left( \sum_{k=1}^{n} \varphi(x_k^n) \right)} \leq C .
\] (63)
This implies that there exists a constant $D = D(\varphi) > 1$ such that, for all $n, m \geq 1$ and any pair $a^n \in A^n$, $b^m \in A^m$ we have

$$D^{-1} \leq \frac{\mu_{\varphi}(a^n b^m)}{\mu_{\varphi}(a^n) \mu_{\varphi}(b^m)} \leq D. \quad (64)$$

### 7.3.2 The key proposition

**Proposition 8.** Let $\varphi$ be a potential with summable variation. Then $\Lambda(\mu_{\varphi})$ exists and

$$\Lambda(\mu_{\varphi}) = \gamma^+(\mu_{\varphi}).$$

**Proof.** Recall that

$$\zeta_{\mu_{\varphi}}(a^n_1) = \mu_{\varphi, a^n_1} (T_{a^n_1} \neq \tau(a^n_1)) = \mu_{\varphi, a^n_1} (T_{a^n_1} > \tau(a^n_1)).$$

Since $a^n_1 a^{0, \tau(a^n_1)+1} = a_1^{\tau(a^n_1)} a^n_1$ we have

$$(1 - \zeta_{\mu_{\varphi}}(a^n_1)) \mu_{\varphi}(a^n_1) = \mu_{\varphi}(a_1^{\tau(a^n_1)} a^n_1)$$

hence

$$\Lambda(\mu_{\varphi}) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{a^n_1} \mu_{\varphi}(a_1^{\tau(a^n_1)} a^n_1)$$

provided the limit exists. We will prove that the limit superior $\overline{\Lambda}(\mu_{\varphi})$ (resp. limit inferior $\underline{\Lambda}(\mu_{\varphi})$) is upper bounded (resp. lower bounded) by $\gamma^+(\mu_{\varphi})$.

To prove that $\overline{\Lambda}(\mu_{\varphi}) \leq \gamma^+(\mu_{\varphi})$, we first use (64) to get

$$\sum_{a^n_1} \mu_{\varphi}(a_1^{\tau(a^n_1)} a^n_1) \leq D \sum_{a^n_1} \mu_{\varphi}(a_1^{\tau(a^n_1)}) \mu_{\varphi}(a^n_1)$$

$$\leq D \max_{a^n_1} \mu_{\varphi}(a^n_1) \sum_{a^n_1} \mu_{\varphi}(a_1^{\tau(a^n_1)}). \quad (65)$$

Partitioning according to the values of $\tau(a^n_1)$

$$\sum_{a^n_1} \mu_{\varphi}(a_1^{\tau(a^n_1)}) = \sum_{i=1}^{n} \sum_{\tau(a^n_1) = i} \mu_{\varphi}(a^n_1).$$

Now, observe that

$$\sum_{\tau(a^n_1) = i} \mu_{\varphi}(a^n_1) = \mu_{\varphi}(\{x^n: \tau(x^n) = i\}).$$

This implies in particular that

$$\sum_{a^n_1} \mu_{\varphi}(a_1^{\tau(a^n_1)}) \leq n.$$
Coming back to (65) we conclude by Proposition 6 that
\[ \Lambda(\mu_\varphi) \leq \limsup_n \frac{1}{n} \log \left( D_n \max_{a_1^n} \mu_\varphi(a_1^n) \right) = \gamma^+(\mu_\varphi). \]

We now show that \( \Lambda(\mu_\varphi) \geq \gamma^+(\mu_\varphi). \) Using (64) we have
\[
\sum_{a_1^n} \mu_\varphi(a_1^n a_1^n) \geq D^{-1} \sum_{a_1^n} \mu_\varphi(a_1^n) \mu_\varphi(a_1^n).
\]

We need the following lemma whose proof is given right after the present proof.

**Lemma 1.** Let \( \varphi \) be a potential with summable variation. Then there exists a sequence of strings \((A_n)_{n \geq 1}, A_n \in \mathcal{A}^n, \) such that
\[
\lim_{n} \frac{1}{n} \log \mu_\varphi(A_n) = \gamma^+(\mu_\varphi) \quad \text{and} \quad \lim_{n} \frac{\tau(A_n)}{n} = 0.
\]

For any \( n \geq 1 \) and any string \( x_1^n, \) let us introduce the notation \( \Upsilon(x_1^n) = x_1^{\tau(x_1^n)} \) which is the prefix of \( x_1^n \) of size \( \tau(x_1^n). \) We have the following obvious lower bound
\[
\frac{1}{n} \log \sum_{a_1^n} \mu_\varphi(a_1^n a_1^n) \geq \frac{1}{n} \log \left( D^{-1} \mu_\varphi(\Upsilon(A_n)) \mu_\varphi(A_n) \right).
\]

We now use (63). For any point \( x \in A_n, \) and using the fact that \( \varphi(x) \geq -\|\varphi\|_\infty > -\infty \) (since \( \varphi \) is continuous and \( \mathcal{A}^n \) is compact), we obtain
\[
\frac{1}{n} \log \left( D^{-1} \mu_\varphi(\Upsilon(A_n)) \mu_\varphi(A_n) \right) \geq \frac{\log(D^{-1}C^{-1})}{n} + \frac{1}{n} \sum_{k=1}^{\tau(A_n)} \varphi(x_k^n) + \frac{1}{n} \log \mu_\varphi(A_n)
\]
\[
\geq \frac{\log(D^{-1}C^{-1})}{n} - \|\varphi\|_\infty \frac{\tau(A_n)}{n} + \frac{1}{n} \log \mu_\varphi(A_n).
\]

Therefore by Lemma 1 we get
\[ \Delta(\mu_\varphi) \geq \liminf_n \frac{1}{n} \log \mu_\varphi(A_n) = \gamma^+(\mu_\varphi) \]
and the proof is finished.

**Proof of Lemma 1.** We know that \( \gamma^+(\mu_\varphi) \) exists by Proposition 6. This means that there exists a sequence of strings \( B_1, B_2, \ldots, \) with \( B_i \in \mathcal{A}^i, i \geq 1, \) such that
\[ \lim_i \frac{1}{i} \log \mu_\varphi(B_i) = \gamma^+(\mu_\varphi). \]
Now, let \((k_i)_{i \geq 1}\) be a diverging sequence of positive integers. Then, for each \(i \geq 1\), consider the string \(B_i^{k_i}\) obtained by concatenating \(k_i\) times the string \(B_i\):

\[
B_i^{k_i} = B_i \ldots B_i. \quad \text{\(k_i\) times}
\]

Using (64) we have

\[
\mu_\varphi(B_i)^{k_i} D^{-k_i} \leq \mu_\varphi(B_i^{k_i}) \leq \mu_\varphi(B_i)^{k_i} D^{k_i}. \quad (67)
\]

For any \(n \geq 1\), consider the unique integer \(i_n\) such that \(n \in [ik_i, (i+1)k_i-1]\) (we omit the subscript \(n\) of \(i_n\) to alleviate notations). We write \(r = r(i, n) := n - ik_i\) and let \(A_n = B_i^{k_i} B_r(i)\) where \(B_r(i)\) is the beginning (or prefix) of size \(r(i)\) of \(B_i\):

\[
A_n = B_i \ldots B_i B_r(i). \quad \text{\(k_i\) times}
\]

Therefore

\[
\frac{\tau(A_n)}{n} \leq \frac{i}{ik_i + r(i)} \xrightarrow{n \to \infty} 0
\]

since \(i\) (and therefore \(k_i\)) diverges as \(n \to \infty\).

To prove that \(\lim_{n \to \infty} \frac{1}{n} \log \mu_\varphi(A_n) = \gamma^+(\mu_\varphi)\), observe that

\[
\frac{\log \mu_\varphi(B_i^{k_i+1})}{n} \leq \frac{\log \mu_\varphi(A_n)}{n} \leq \frac{\log \mu_\varphi(B_i^{k_i})}{n}.
\]

which gives, using (67),

\[
\frac{\log \mu_\varphi(B_i^{k_i+1} D^{k_i+1})}{ik_i + r} \leq \frac{\log \mu_\varphi(A_n)}{n} \leq \frac{\log \mu_\varphi(B_i^{k_i} D^{k_i})}{ik_i + r}. \quad (68)
\]

Now it is enough to observe that both sides of (68) converge to \(\gamma^+(\mu_\varphi)\). Observe that the RHS equals

\[
\frac{1}{ik_i + r} \left( \log \mu_\varphi(B_i^{k_i}) + k_i \log D \right) = \frac{k_i}{ik_i + r} \left( \frac{1}{i} \log \mu_\varphi(B_i) + \frac{\log D}{i} \right).
\]

To conclude, recall that \(\frac{1}{i} \log \mu_\varphi(B_i) \to \gamma^+(\mu_\varphi)\), whereas \(k_i(k_i + \frac{1}{i})^{-1} \to 1\). The limit of the LHS of (68) follows similarly. This concludes the proof of the lemma.

7.3.3 End of proof of Theorem 2

An equilibrium state for a potential satisfying \(\int \psi(\ell)\) is \(\psi\)-mixing, which means that there exists a sequence \((\psi(\ell))\) of positive numbers decreasing to zero such that for all integers \(n \geq 1\) and \(\ell \geq 1\) we have

\[
\sup_{A \in \mathcal{F}_n^+, B \in \mathcal{F}_{n+\ell}^\infty} \left| \frac{\mu_\varphi(A \cap B)}{\mu_\varphi(A) \mu_\varphi(B)} - 1 \right| \leq \psi(\ell). \quad (69)
\]
Clearly, $\psi$-mixing implies $\phi$-mixing (see (14)). Property (69) is what is actually proved in [Wal75] (see the proof of Theorem 3.2 therein). For instance, $\psi(\ell)$ goes to 0 exponentially fast for Hölder continuous potentials [Bow08]. Actually, we will not need to know at which speed $\psi(\ell)$ decreases to 0.

We only deal with the case $q < 0$ since the case $q > 0$ is already treated in [CU05]. We cannot apply directly Theorem 1. In fact, everything works except that Proposition 4 has to be replaced by Proposition 9, see below. Indeed, Theorem 1 is restricted to $\phi$-mixing processes such that $\sum \phi(\ell) < +\infty$ only to guarantee a uniform control on $\zeta_\mu(x^n_\tau)$ (Proposition 4). Then, we use Propositions 7, 8, 6, and 5.

**Proposition 9.** Let $\mu_\varphi$ be the equilibrium state for a potential $\varphi$ with summable variation. Then there exists $\hat{\zeta}_- > 0$ such that

$$\inf_{a_1^n \in A^n} \zeta_{\mu_\varphi}(a_1^n) \geq \hat{\zeta}_- .$$

**Proof.** We have the obvious following inclusion: There exists $a \neq x_1$ such that

$$[x_1^n] \cap \{T_{x_1^n} \tau(x_1^n)\} \supseteq [x_1^n a]$$

hence

$$\mu_{x_1^n}(T_{x_1^n} \tau(x_1^n)) \geq \frac{\mu([x_1^n a])}{\mu(x_1^n)} .$$

Now, using (51), we get that for any $y \in [x_1^n a]$

$$\mu([x_1^n a]) \geq e^{-(n+1)\varepsilon_n \varphi(y_\infty^n)} \quad \text{and} \quad \mu(x_1^n) \leq e^{n \varepsilon_n} e^{\sum_{k=1}^n \varphi(y_k^n)} .$$

Using (52), it follows that

$$\zeta_{\mu_\varphi}(x_1^n) \geq \exp \left(-2 \sum_{k=1}^{n+1} \varphi_k(\varphi) e^{\varphi(y_{n+1}^\infty)}\right)$$

$$\geq \exp \left(-2 \sum_{k=1}^{\infty} \varphi_k(\varphi) e^{\inf \varphi} =: \hat{\zeta}_- > 0 .\right)$$

The proposition is proved. \(\square\)

**A Some inequalities for the upper incomplete Gamma function**

The upper incomplete Gamma function is defined by

$$\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} \, dt, x > 0, s \in \mathbb{R} .$$
We are interested in $s < 0$ and finite $x$. As a matter of fact, $x$ will be vanishingly small. We will use the following well-known recursive equation (integration by parts)

$$s \Gamma(s, x) = \Gamma(s + 1, x) - x^s e^{-x}.$$ 

In particular, for $s < 0$, we have

$$\Gamma(s, x) = \frac{\Gamma(s + 1, x) - x^s e^{-x}}{s}$$

(71)

which is positive. We now distinguish two cases.

- $s < -1$. In view of (71), we need to bound $\Gamma(r, x)$ where $r = s + 1 < 0$.

  On the one hand, we obviously have

  $$\Gamma(r, x) \leq e^{-x} \int_{x}^{\infty} t^{r-1} dt = \frac{x^{-|r|} e^{-x}}{|r|} = C_r x^{-|r|} e^{-x}.$$ 

  (72)

  On the other hand, using the well-known continued fraction representation of $\Gamma(r, x)$ [Wal73, p. 356], we deduce that

  $$\Gamma(r, x) \geq \frac{x^{-|r|} e^{-x}}{x + 1 + |r|} \geq c_r x^{-|r|} e^{-x}.$$ 

  (73)

  where $c_r = 1/(2 + |r|)$.

  Thus, using (71), (72) and (73), one has

  $$\frac{c_s x^{-|s|+1} e^{-x} - x^{-|s|} e^{-x}}{-|s|} \leq \Gamma(-|s|, x) \leq \frac{C_s x^{-|s|+1} e^{-x} - x^{-|s|} e^{-x}}{-|s|}$$

  where $c_s = (|s| + 2)^{-1}$ and $C_s = (|s + 1|)^{-1}$. Therefore

  $$x^{-|s|} e^{-x}(1 - c_s x) \leq |s| \Gamma(-|s|, x) \leq x^{-|s|} e^{-x}(1 - C_s x).$$ 

  (74)

- $s \in (-1, 0)$. Using (71) we need to bound $\Gamma(s + 1, x)$ which has positive first argument $s + 1 =: r \in (0, 1)$. To get a lower bound, we write

  $$\Gamma(r, x) = \int_{0}^{\infty} (u + x)^{r-1} e^{-u-x} du = e^{-x} E[(E + x)^{r-1}]$$

  where $E$ denotes a exponential random variable with parameter 1. Then, since $0 < r < 1$, it follows by Jensen’s inequality that

  $$\Gamma(r, x) \geq e^{-x}(1 + x)^{r-1}.$$ 

To get an upper bound, we write

$$\Gamma(r, x) = \int_{0}^{\infty} (u + x)^{r-1} e^{-u-x} du$$

$$= e^{-x} \Gamma(r) \int_{0}^{\infty} \left(1 + \frac{x}{u}\right)^{r-1} u^{r-1} e^{-u} \frac{1}{\Gamma(r)} du$$

$$= e^{-x} \Gamma(r) E \left[(1 + \frac{x}{G})^{r-1}\right]$$

34
where $G$ is a random variable with distribution Gamma($r, 1$). Then, using Jensen’s inequality, we get (since $\text{E}(G) = r$)

$$
\Gamma(r, x) \leq e^{-x} \Gamma(r) \left(1 + \frac{x}{r}\right)^{r-1}.
$$

Hence we get

$$
e^{-x} 2^{r-2} \leq \Gamma(r, x) \leq e^{-x} \Gamma(r).
$$

Thus, using once more (71)

$$
\frac{2s^{-1} e^{-x} - x^s e^{-x}}{s} \leq \Gamma(-|s|, x) \leq \frac{\Gamma(s+1) e^{-x} - x^s e^{-x}}{s}
$$

which yields

$$
x^{-|s|} e^{-x} (1 - c'_q x^{|s|}) \leq |s| \Gamma(-|s|, x) \leq x^{-|s|} e^{-x} (1 - C'_q x^{|s|})
$$

where $c'_q := 2^{|q|} - 1$ and $C'_q := \Gamma(|q| + 1)$.

### B Proof of inequalities (51)

We first prove the following identity: For $a^n_1 \in A^n$ and $x \in A^N$ we have

$$
\mu(a^n_2) = \int_{[a^n_1]} e^{-\varphi(y)} \, d\mu(y). \tag{77}
$$

Indeed we have

$$
\mu(a^n_2) = \int_{[a^n_2]} 1_{[a^n_2]}(y) \, d\mu(y) = \int \left(\sum_{a \in A} 1_{[a^n_1]}(ay) e^{-\varphi(ay)}\right) \, d\mu(y)
$$

$$
= \int \left(\sum_{a \in A} e^{\varphi(ay)} 1_{[a^n_1]}(ay) e^{-\varphi(ay)}\right) \, d\mu(y)
$$

$$
= \int \left(L_{\varphi^*} 1_{[a^n_1]} e^{-\varphi}\right)(y) \, d\mu(y)
$$

$$
= \int 1_{[a^n_1]}(y) e^{-\varphi(y)} \, d(L_{\varphi^*} \mu)(y)
$$

$$
= \int 1_{[a^n_1]}(y) e^{-\varphi(y)} \, d\mu(y).
$$

We used (20) and (22). Since $|\varphi(x) - \varphi(y)| \leq \text{var}_n(\varphi)$, (71) implies

$$
e^{-\text{var}_n(\varphi)} \leq \frac{\mu(a^n_1)}{\mu(a^n_2)} e^{-\varphi(x)} \leq e^{-\text{var}_n(\varphi)}.
$$
In the same way we can prove that
\[ e^{-\text{var}_n-1(\varphi)} \leq \frac{\mu(a_n)}{\mu(a_0)} e^{-\varphi(\theta x)} \leq e^{-\text{var}_{n-1}(\varphi)} \]
\[ \vdots \]
\[ e^{-\text{var}_1(\varphi)} \leq \mu(a_n) e^{-\varphi(\theta^{n-1}x)} \leq e^{-\text{var}_1(\varphi)}. \]

Multiplying the above inequalities yields
\[ e^{-\sum_{k=1}^n \text{var}_k(\varphi)} \leq \frac{\mu(a_n)}{\exp\left(\sum_{k=1}^n \varphi(x_k^\infty)\right)} \leq e^{\sum_{k=1}^n \text{var}_k(\varphi)}. \]

Therefore (51) is proved.

C Proof of Theorem 1 for \( q \geq 0 \)

The case \( q = 0 \) is trivial. For any \( q > 0 \) we have
\[ \mathbb{E}[R_q^n] = q \sum_{x_n^1} \mu(x_n^1)^{1-q} \int_0^{\infty} t^{q-1} \mu_{x_n^1} \left( T_{x_n^1} > \frac{t}{\mu(x_n^1)} \right) \, dt. \quad (78) \]

In order to use (78) we will need the following facts:

- By Proposition 4 for all \( a_n^1 \) we have \( \zeta_{\mu}(a_n^1) \geq \zeta > 0 \), and by definition \( \zeta_{\mu}(a_n^1) \leq 1. \)
- \( \epsilon(x_n^1) \) converges uniformly to 0.
- From the two preceding items, there exists a constant \( \rho > 0 \) such that for large enough \( n \), \( \zeta_{\mu}(x_n^1) - 16\epsilon(x_n^1) \geq \rho \) uniformly in \( x_n^1. \)
- For \( n \) large enough we have \( \zeta_{\mu}(x_n^1)\mu(x_n^1)\tau(x_n^1) \leq 1 \) uniformly in \( x_n^1 \) while obviously \( \zeta_{\mu}(x_n^1)\mu(x_n^1)\tau(x_n^1) \geq 0. \)

Using these facts together with (37) we obtain that, for \( n \) large enough, any \( x_n^1 \) and any \( t > 0 \)
\[ 0 < \zeta - e^{-t - 54t\epsilon(x_n^1)} e^{-h_{1t}} \leq \mu_{x_n^1}(T_{x_n^1} > \frac{t}{\mu(x_n^1)}) \leq e^{-\zeta - t} + 54t\epsilon(x_n^1) e^{-h_{1t}}. \]

This implies that the integral in (78) is bounded between two positive constants, and therefore there exist \( c, C > 0 \) such that
\[ c \sum_{x_n^1} \mu(x_n^1)^{1-q} \leq \mathbb{E}[R_q^n] \leq C \sum_{x_n^1} \mu(x_n^1)^{1-q}. \]

This concludes the proof of the theorem for \( q \geq 0. \)
References

[Aba01] Miguel Abadi. Exponential approximation for hitting times in mixing processes. *Math. Phys. Electron. J.*, 7:Paper 2, 19 pp. (electronic), 2001.

[AC15] Miguel Abadi and Liliam Cardeño. Rényi entropies and large deviations for the first match function. *IEEE Trans. Inform. Theory*, 61(4):1629–1639, 2015.

[AV09] Miguel Abadi and Nicolas Vergne. Sharp error terms for return time statistics under mixing conditions. *J. Theoret. Probab.*, 22(1):18–37, 2009.

[Bow08] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, revised edition, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.

[Bra05] Richard C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2:107–144, 2005. Update of, and a supplement to, the 1986 original.

[CFM+18] Théophile Caby, Davide Faranda, Giorgio Mantica, Sandro Vaienti, and Pascal Yiou. Generalized dimensions, large deviations and the distribution of rare events. 2018.

[CGS99] Pierre Collet, Antonio Galves, and Bernard Schmitt. Repetition times for Gibbsian sources. *Nonlinearity*, 12(4):1225–1237, 1999.

[CRS18] Adriana Coutinho, Jérôme Rousseau, and Benoît Saussol. Large deviation for return times. published in NL, 2018.

[CU05] Jean-René Chazottes and Edgardo Ugalde. Entropy estimation and fluctuations of hitting and recurrence times for Gibbsian sources. *Discrete Contin. Dyn. Syst. Ser. B*, 5(3):565–586, 2005.

[DZ10] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[FM05] Roberto Fernández and Grégory Maillard. Chains with complete connections: general theory, uniqueness, loss of memory and mixing properties. *J. Stat. Phys.*, 118(3-4):555–588, 2005.

[HV10] Nicolai Haydn and Sandro Vaienti. The Rényi entropy function and the large deviation of short return times. *Ergodic Theory Dynam. Systems*, 30(1):159–179, 2010.

[JB13] Siddharth Jain and Rakesh Kumar Bansal. On large deviation property of recurrence times. *2013 IEEE International Symposium on Information Theory*, pages 2880–2884, 2013.
...