Exponential mixing and almost sure limit theorems in dynamical systems

Boyuan Zhao

University of St Andrews

This thesis is submitted in partial fulfilment for the degree of Doctor of Philosophy (PhD) at the University of St Andrews

October 2025

Declarations

Candidate's declaration

I, Boyuan Zhao, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 19,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in October 2021.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date: 30/10/2025 Signature of candidate: 赵博原.

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date: 04/07/2025 Signature of supervisor:

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Boyuan Zhao, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Electronic copy

No embargo on electronic copy.

Date: 30/10/2025 Signature of candidate: 赵博原.

Date: 04/07/2025 Signature of supervisor:

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Boyuan Zhao, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date: 30/10/2025 Signature of candidate: 赵博原.

Acknowledgements

General Acknowledgements

The past four or five years at St Andrews have been the highlight of my life so far and there is a long list of people I want to express my thanks to. First, my heartfelt gratitude go to the Analysis group staffs Kenneth Falconer, Jonathan Fraser, Natalia Jurga, Lars Olsen and Mike Todd, for organising the weekly seminars, inviting speakers for academic talks and leading the cordial pub trips. Especially I am grateful for Mike who 'tolerated' four years of my indecipherable mathematical thoughts, and Kenneth for scrutinising my thesis draft which was well decorated with typos and awkward expressions. I would also like to thank the (current and departed) postdocs and PhDs of our group who overlapped with my time here: Istvan Kolossvary, Luke Derry, Saeed Shaabanian, Lijian Yang; and those I had the honour to share the analysis bunker with: Raquel Couto, Alex Rutar, Amlan Banaji, Ana de Orellana, Liam Stuart, Tianyi Feng and Yunlong Xu. Special thanks to Alex, Amlan and Ana for organising reading groups during term time.

My non-academic life in St Andrews can be readily summarised by two words: running and music. It felt particularly nice and reassuring to reclaim my athletic capability after giving it up for years. I wish to thank my friends in the St Andrews athletic and XC club, for getting me through the tough walk-backs to the starting line during interval sessions, and for turning me into a proper endurance runner. Running more than five kilometres was unimaginable to me before joining the club.

And of course, my folk music community, which includes the local legends running the Tuesday Whey Pat sessions and the student society which I joined during the pandemic. I stumbled into a new world, where people from different parts of the country, or different countries, sit down under the gravity of music and converse in beautiful tunes. I want to thank the committee of the folk and trad society for successfully organising our annual folk festivals and other social events. Thanks also go to all members in the performance group which I put together in the fourth year of my PhD: Erin Dundas, Kirsty Hill, Gracie McBride, Felix Leitner-Mittersinker, Emma Montgomery and Euan Robertson. It has been a fantastic year to rehearse and perform together as a team, it was my great honour and pleasure to play alongside these wonderful musicians.

To Rica Go, Ivana Tonkova and Sebastian Mok, thank you for struggling with me through our unusual undergraduate years. To Meg Briars, Daniel Isaiah and Xuan'ang Lyu, thanks for all the inspiring conversations between us, and all other forms of mental support throughout my PhD.

To Isla Bauer, thank you for being you and being there whenever I needed you.

Lastly, I want to thank my parents, Li Li and Hui Zhao, for being my biggest fans and strongest shield, for providing me with everything I could possibly need, for their love and caring which I cannot fathom to describe with words. For raising me right and getting me through the hard times.

Funding

My research throughout my PhD is funded by the *China Scholarship Council (CSC) Programme*, number [CSC202108060177].

Publication and collaboration

This thesis is largely based on three projects completed during my PhD studies; each chapter can be read somewhat independently.

Chapter 2: [TZ] *Countable Markov shifts with exponential mixing* (with M. Todd), submitted, preprint, arxiv: 2403.02092 (2024).

Chapter 3: [Zha2] *Almost sure cover times for dynamical systems with ψ-mixing property*, submitted, preprint, arxiv: 2412.17425 (2024).

Chapter 4: [Zha1] *Closest distance between iterates of typical points*, Discrete Contin. Dyn. Syst. **44**(8) (2024) 2252-2279. arxiv: 2306.01628.

Contents

Abstract

A goal of ergodic theory is to understand the stochastic behaviours of deterministic systems. Given a measure preserving system (X, f, μ), decay of correlations ensures that for reasonable $\omega : X \to \mathbb{R}$, $(\omega(f^n x))_{n \in \mathbb{N}}$ asymptotically behaves like an *i.i.d.* process and results analogous to classical probabilistic theorems for *i.i.d.* sequences can be proved.

Decay rates for uniformly hyperbolic maps are often exponential whereas non-uniformly hyperbolic systems can have troublesome rates, *e.g.* subexponential or polynomial. A common approach to study decay of correlations is via the corresponding symbolic space, which admits the same rate of mixing. Since non-uniformly hyperbolic systems are often modelled by *countable Markov shifts* which are non-compact, it requires a more exhausting machinery to prove analogous statements for finite shifts.

This thesis will first review some thermodynamic results for subshifts of finite types (SFT) and countable Markov shifts (CMS) then focus on CMS with *strong positive recurrence* (SPR), a property shown to be equivalent to the spectral gap property, which guarantees exponential mixing rates and other desirable features. For CMS satisfying certain topological boundary conditions, we will show that SPR is characterised by the ergodic averages over periodic orbits. Examples are provided to demonstrate that our condition is rather weak.

In Chapters 3 and 4, we prove two sets of almost sure results using the Borel-Cantelli lemmas for fast mixing systems. Firstly, we show that the asymptotics of the *cover times* are almost surely quantified by the *Minkowski dimensions*, which dictate the growth of *hitting times* to geometrically small sets.

The second set of theorems shows that for a point in a topological Markov shift, the length of the longest matching substrings grows exponentially depending on the *Rényi entropy* of the Gibbs measure. Such quantitative results extend analogously to the *shortest distance problem* for interval maps.

Notation and abbreviations

Here is a list of notation used in this thesis.

- For E a set in some topological space, let $\#E$ denote the cardinality of E, and $\mathbf{1_E}$ the indicator function of E.
- The union of natural numbers with $\{0\}$ is denoted by $\mathbb{N}_0$.
- For a dynamical system f defined on a metric space (X,d), let $\mathcal{C}(X)$ be the set of continuous real-valued functions and $\mathcal{M}_f$ denotes the set of f-invariant probability measures.
- For two collections $\mathcal{P}, \mathcal{Q}$ of subsets in (X,d), let $\mathcal{P} \vee \mathcal{Q} : \{P \cap Q : P \in \mathcal{P},\ Q \in \mathcal{Q}\}$.
- When there is no confusion of the dynamics in question, the measure-theoretic entropy of a measure ν is denoted by $h(\nu)$.
- The open ball centred at x with radius $r > 0$ is denoted by $B(x,r)$.
- The *floor and ceiling* functions, $\lfloor x \rfloor$ takes the largest integer $\leq x$ and $\lceil x \rceil$ takes the smallest integer $\geq x$.
- For real numbers a, b, c, we write $a = b \pm c$ if $a \in [b-c, b+c]$.
- For two real positive sequences $\{a_n\}_n, \{b_n\}_n$, write $a_k \approx b_k$ if $\log a_k - \log b_k$ is bounded, or equivalently the ratio $\left|\frac{a_k}{b_k}\right|$ is uniformly bounded away from 0 and $+\infty$.
  Say $a_k \preceq b_k$ if there is $\{c_k\}_k$ such that $a_k \leq c_k$ for all k, and $b_k \approx c_k$. Both relations are transitive, and if b_k is summable, $a_k \preceq b_k$, then a_k is summable.
- For any function $g : X \to \mathbb{R}$ on the ambient space (X,d), let $\mathbb{E}[g]$ and $\mathrm{Var}[g]$ denote the expectation and variance of g respectively, when the relevant probability measure is clear.

The constants of the form C_i, or K_j used in the proofs in a given chapter are not inherited in later chapters, unless specified.

The following abbreviations are commonly used.

- acip: invariant probability measure absolute continuous with respect to Lebesgue measure;

- BIP: big image and preimage property;
- CI: contraction at infinity;
- CLT: central limit theorem;
- CMS: countable Markov shift;
- *i.i.d.* : independent, identically distributed;
- SGP: spectral gap property;
- SPR: strong positive recurrent;
- SFT: subshift of finite type;
- UCS: uniform contraction structure;

Chapter 1

Preliminaries

In order to understand the statistical behaviour of a dynamical system, we often study its invariant measures. There are various notions of chaotic and stochastic phenomena, but a measure's *decay of correlations*, or equivalently the *mixing properties* (these two terms will be used interchangeably throughout), is of central focus among plenty other measure theoretic properties. Given (X, d) a metric space and (f, μ) a measure preserving system on X, decay of correlations ensures that for a large class of observables $\omega : X \to \mathbb{R}$, the sequence $(\omega(f^n x))_{n\in\mathbb{N}}$ in the long run asymptotically resembles an independent, identically distributed process. If the rate of decorrelation is sufficiently fast, one can prove various results analogous to classical probabilistic theorems for sequence of *i.i.d.* random variables, *e.g.* the Law of Large Numbers, the Central Limit Theorem, *etc*. Moreover, the analysis gets particularly interesting when the observables are chosen to reflect some geometric behaviour of the system. For example, the observable evaluated at each $\{f^j x\}_{j\in\mathbb{N}}$ can decay with the distance either to a reference point $\tilde{x} \in X$, or to the initial position. The analysis of minimal times for such observable to reach some threshold value is referred to as the *hitting time* or *return time* problem respectively, and so far we know their asymptotic behaviours are closely dependent on the rates of decorrelation.

A wide range of uniformly hyperbolic systems are well understood today in terms of their statistical and stochastic properties. In particular, they very often enjoy exponential decay of correlations, whereas the non-uniformly counterparts are more difficult to handle, *e.g.* subexponential or polynomial rates. However, we can often associate a symbolic shift model to the systems in question, and analyse the invariant measures. Generally speaking, the symbolic system with the relevant measure admits the same rate of mixing as the original mapping although the relevant functional spaces may be different.

The study of *symbolic shifts* has a relatively long history compared to that of thermodynamic formalism of dynamical systems, and has been playing a crucial role in other research areas such as stochastic processes, graph theory, logic, *etc*. In this thesis all shifts are assumed to be Markovian, in the sense that the image of a partition set under the shift map is always a countable union of other partition sets; some non-Markovian shifts, *e.g.* the β-shifts, also have a sophisticated and well-developed theory but will not be discussed.

The thermodynamic formalism of finite topological Markov shifts is very well understood in terms of Gibbs measures, equilibrium states and decay of correlations. These results are nicely summarised and organised in [Bow]. Since non-uniformly hyperbolic systems are often modelled by *countable Markov shifts* which are non-compact metric spaces, it requires a more exhausting machinery to prove analogous statements for finite shifts. However, if one can show that the transfer operator associated to some potential acts on a Banach space of functions with *spectral gaps*, then exponential rates of decay of correlations are expected. That is to say, the job reduces to, in some sense, finding necessary and sufficient conditions for the existence of such a spectral gap. Important results for Markov shifts with a countable alphabet have been proved in the last thirty years due to Aaronson, Denker, Mauldin, Sarig, Úrbanski and many others. In this chapter we will do a quick review of some selected results from their works. Since many properties (especially those we care about in this thesis) of two-sided shifts can be reduced to one-sided shifts on $\mathbb{N}_0$, we restrict our discussion to this case.

1.1 Subshifts of finite type and mixing conditions

Let us start with subshifts of finite type. Let $\mathcal{A}$ be a finite alphabet, and M be an $\mathcal{A} \times \mathcal{A}$ transition matrix of $0, 1$ entries. The associated *topological Markov (sub)shift space*, denoted by Σ, is defined by

$$\Sigma := \left\{ x = (x_0, x_1, \dots) \in \mathcal{A}^{\mathbb{N}_0} : M_{x_j, x_{j+1}} = 1 \text{ for all } j \in \mathbb{N}_0 \right\},$$

where $M_{i,j}$ is the $(i, j)-$entry of the matrix M. The dynamics on Σ is the left shift $\sigma : \Sigma \to \Sigma$, given by $(\sigma x)_i = x_{i+1},$ for all $i \geq 0$. The triplet $(\Sigma, \mathcal{A}, M)$ is called a *subshift of finite type* (SFT). Denote the set of probability measures on Σ by $\mathcal{M}(\Sigma)$ and the set of σ-invariant measures by $\mathcal{M}_\sigma$.

A word of length n $\underline{w}$ is *allowable* if $M_{w_j, w_{j+1}} = 1$ for $j = 0, \dots, n-2$. Let Σ_n denote all allowable words of length n, and $\Sigma^* := \bigcup_{n \geq 1} \Sigma_n$ the set of all finite allowable words. Similarly, let $\mathcal{C}_n$ denote the set of all non-empty n-cylinders and $\mathcal{F}_n$ the sigma-algebra generated by $\mathcal{C}_n$.

Definition 1.1.1. *The* measure theoretic entropy of μ *of the subshift, denoted by* $h(\mu)$*, is given by*

$$h(\mu) = \lim_{n \to \infty} - \sum_{C \in \mathcal{C}_n} \mu(C) \log(\mu(C)). \tag{1.1.1}$$

The shift space is equipped with a symbolic metric, $d_s : \Sigma \times \Sigma \to \mathbb{R}$, given by

$$d_s(x, y) = 2^{-|x \wedge y|}, \text{ where } x \wedge y := \inf \{k \geq 0 : x_k \neq y_k\}. \tag{1.1.2}$$

Note that since $\#\mathcal{A} < \infty$, the metric space (Σ, d_s) is compact.

Definition 1.1.2. *An n-cylinder in Σ is a collection of points such that they agree on the first n symbols, denoted by square brackets,* i.e.*,*

$$[x_0, \ldots, x_{n-1}] := \{y \in \Sigma : y_0 = x_0, \ldots, y_{n-1} = x_{n-1}\}.$$

Cylinders are precisely the open balls in Σ with respect to the metric d_s. The 1-cylinders $\{[a] : a \in \mathcal{A}\}$ are called *partition sets*, and the symbolic systems studied in this thesis are Markov in the sense that the image of each partition sets under σ is a (countable) union of partition sets. Examples of non-Markov symbolic systems include Sturmian shifts, β-transformations, *etc*.

Definition 1.1.3. Σ *is* topologically transitive *if for all $a, b \in \mathcal{A}$, there exists N_{ab} such that $\sigma^{-N_{ab}}[b] \cap [a] \neq \emptyset$, and* topologically mixing *if for all $n \geq N_{ab}$, $\sigma^{-n}[b] \cap [a] \neq \emptyset$. For each pair $a, b \in \mathcal{A}$, N_{ab} can be different.*

Topological transitivity can be upgraded to topological mixing if there are coprime periodic orbits.

Proposition 1.1.4. *A topologically transitive Markov shift is topologically mixing if and only if there exist p, q coprime and periodic points $x, y \in \Sigma$ such that $x = \sigma^p x$, $y = \sigma^q y$.*

Proof. Suppose Σ is topologically mixing and fix $a \in \mathcal{A}$. Then by definition there exists $n \geq N_a$ and $x, y \in [a]$ such that $\sigma^n x = x$ and $\sigma^{n+1} y = y$.

Now suppose the Σ contains x, y periodic points with periods p, q coprime. Let $a, b \in \mathcal{A}$, by topological transitivity, there exists $N_{ax_0}, N_{x_{p-1}y_0}, N_{y_{q-1}b} \geq 1$ such that

$$[a] \cap \sigma^{-N_{ax_0}}[x_0],\ [x_{p-1}] \cap \sigma^{-N_{x_{p-1}y_0}}[y_0],\ [y_{q-1}] \cap \sigma^{-N_{y_{q-1}b}}[b] \neq \emptyset.$$

The largest integer that cannot be written as a sum of positive multiples of p, q is $(p-1)(q-1)-1$, so set

$$N_{ab} = N_{ax_0} + N_{x_{p-1}y_0} + N_{y_{q-1}b} + (p-1)(q-1),$$

then for all $j \geq N_{ab}$, $[a] \cap \sigma^{-j}[b] \neq \emptyset$. As a, b are arbitrary, this proves topological mixing. $\square$

These definitions can be compared to more general versions in topological dynamical systems language. For interval maps, transitivity roughly means there do not exist two non-empty subsets in the interval with disjoint interiors that never talk to each other under the dynamics. Topological transitivity guarantees a list of chaotic properties, *e.g.* the set of periodic points of the mapping is dense in the interval. Topological mixing implies that for any two open sets, after some finite time evolution their images always intersect, while the more important question is how fast and what proportion of the sets are saturated. Suppose μ is a *shift invariant* measure, the following measure-theoretic notions of mixing approximate the notion of independence.

Definition 1.1.5 (Mixing conditions)**.** *Let μ be a shift invariant measure on Σ, it is said to be one of the following if for all $n, m, k \in \mathbb{N}$ and $E \in \mathcal{F}_n$, $F \in \mathcal{F}_m$,*

- weakly-mixing*: $\left|\mu\left(E \cap \sigma^{-n-k} F\right) - \mu(E)\mu(F)\right| \to 0$ as $k \to \infty$;*
- α-mixing*: $\left|\mu\left(E \cap \sigma^{-n-k} F\right) - \mu(E)\mu(F)\right| \leq \alpha(k)$ for some $\alpha : \mathbb{N}_0 \to \mathbb{R}$ strictly decreasing;*
- ϕ-mixing*: $\left|\mu\left(E \cap \sigma^{-n-k} F\right) - \mu(E)\mu(F)\right| \leq \phi(k)\mu(E)$ for some $\phi : \mathbb{N}_0 \to \mathbb{R}$ strictly decreasing;*
- ψ-mixing*: $\left|\mu\left(E \cap \sigma^{-n-k} F\right) - \mu(E)\mu(F)\right| \leq \psi(k)\mu(E)\mu(F)$ for some $\psi : \mathbb{N}_0 \to \mathbb{R}$ strictly decreasing.*

It is obvious that ψ-mixing $\implies$ ϕ-mixing $\implies$ α-mixing $\implies$ weakly mixing $\implies$ ergodic. See [Bra] for more detailed discussion on mixing conditions. We will focus on systems with ψ-mixing in Chapter 3 and Chapter 4, as it is a powerful mixing condition: (1) it implies the following *quasi-Bernoulli property*, *i.e.*, for all $\underline{w}, \underline{v} \in \Sigma^*$, μ being ψ-mixing implies there exists $B = 1 + \psi(0) > 1$ such that

$$\mu([\underline{wv}]) \leq B\mu([\underline{w}])\mu([\underline{v}]), \tag{1.1.3}$$

and (2) it guarantees that the measure of an arbitrary n-cylinder decays exponentially in n.

Lemma 1.1.6. *[GalSch] If the probability measure μ is ψ-mixing with $\psi(\cdot)$ summable, there exist constants $\beta \in (0, 1)$ and $K_0 > 0$ such that $\mu(C) \leq K_0 \beta^n$ for all $C \in \mathcal{C}_n$ and all n.*

Remark 1.1.7. *The statement of the lemma holds when $\mathcal{A}$ is countably infinite as well, as the original proof given in [GalSch] does not exploit the finiteness of the alphabet.*

1.2 Equilibrium and Gibbs states for subshifts of finite type

Given a dynamical system, it is common that the system admits many, even infinitely many, invariant or ergodic measures, and since statistical behaviours depend on the measure, which measure should we choose to analyse the system? A natural answer is 'equilibrium states'. In statistical mechanics, equilibrium states are described by probability measures on topological spaces that are characterised by variational principles, maximising entropy (or the sum of entropy and an energy-like quantity). In some sense [Kel2, §1], equilibrium states should be viewed as an object on both microscopic and macroscopic scales: by invariance, equilibrium states allow one to predict the configuration of an item in the topological space for all future times, however since the number of microscopic sites are too many or simply (uncountably) infinite, we focus on the macroscopic information.

For topological shifts, an obvious question is whether the system admits equilibrium states, in such case the sum of entropy with some thermodynamic potential is minimised or maximised. Precise definitions are given below, starting from the *pressure* function.

A real-valued function $\phi : \Sigma \to \mathbb{R}$ is called a *potential*. Denote the set of continuous potentials on Σ by $\mathcal{C}(\Sigma)$. Two potentials ϕ, ϕ' are called *cohomologous* if there is $u \in \mathcal{C}(\Sigma)$ such that for all $x \in \Sigma$,

$$\phi'(x) = \phi(x) - u(x) + u(\sigma x).$$

If one of ϕ, ϕ' is a constant function then both of them are called *coboundaries*. For topologically transitive topological Markov shifts, two potentials are cohomologous if and only if the Birkhoff averages of all periodic points coincide.

Definition 1.2.1. *Let* $S_n\phi(x) := \sum_{j=0}^{n-1} \phi\left(\sigma^j x\right)$ *denote the* n*-th Birkhoff sum, and define the* partition functions *and the* topological pressure *of* ϕ *respectively by*

$$Z_n(\phi) := \sum_{C \in \mathcal{C}_n} \exp\left(\sup_{x \in C} S_n\phi(x)\right), \quad P(\phi) := \lim_{n\to\infty} \frac{1}{n} \log Z_n(\phi). \tag{1.2.1}$$

By [Bow, Lemma 1.20], $P(\phi)$ exists for all continuous ϕ. In other words, while topological entropy $h_T = P(0)$ counts the asymptotic growth of cylinders of length n, the topological pressure can be seen as its generalisation, weighted by a potential ϕ. The term pressure here perhaps should not be taken literally: here $P(\phi)$ is minus the usual value of pressure in statistical mechanics known to most physicists. It should also be stressed that, for each given ϕ, $P(\phi)$ depends only depends on the Borel structure of Σ and has nothing to do with the metric d_s.

To guarantee the value of $P(\phi)$ exists, ϕ is often assumed to behave reasonably regular, *i.e.*, continuous with respect to d_s and its variations are not extreme.

Definition 1.2.2. *Let* $\mathrm{var}_k(\phi) := \sup\{|\phi(x) - \phi(y)| : x_j = y_j,$ *for all* $0 \le j \le k-1\}$ *denote the* k-th variation *of* ϕ*, then* ϕ *is called* Hölder *or* θ*-Hölder if there exists* $c_\phi > 0$ *and* $\theta \in (0,1)$ *such that* $\mathrm{var}_k(\phi) \le c_\phi \theta^k$ *for all* $k \in \mathbb{N}_0$.

Furthermore, $P(\phi)$ satisfies the *Variational Principle*.

Theorem 1.2.3 (Variational Principle). *[Bow, §2.17]*

$$P(\phi) = \sup \left\{ h(\nu) + \int \phi d\nu : \nu \in \mathcal{M}_\sigma \right\}, \tag{1.2.2}$$

where $h(\nu)$ denotes the measure-theoretic entropy defined in (1.1.1).

In comparison to (1.2.1), the Variational Principle does not require any time-evolution information apart from invariance under the dynamics. Now it makes sense to ask if there are invariant measures realising this supremum.

Definition 1.2.4. *Given $\phi : \Sigma \to \mathbb{R}$ continuous, a measure $\mu \in \mathcal{M}_\sigma$ is called an* equilibrium state *if $h(\mu) + \int \phi d\mu = P(\phi)$.*

The statistical properties of possible equilibrium states is closely related to the properties of the *Ruelle operator* associated to ϕ, or often called the transfer operator, defined by

$$\mathcal{L}_\phi : \mathcal{C}(\Sigma) \to \mathcal{C}(\Sigma), \;\; \mathcal{L}_\phi f(x) = \sum_{y \in \sigma^{-1}x} e^{\phi(y)} f(y), \tag{1.2.3}$$

and its iterates satisfy $\mathcal{L}_\phi^n f(x) = \sum_{y \in \sigma^{-n}x} e^{S_n \phi(y)} f(y)$ for all integers $n \geq 1$. For any $\mu \in \mathcal{M}(\Sigma)$ and $g \in \mathcal{C}(\Sigma)$, $\int g \, d(\mathcal{L}_\phi^* \mu) := \int (\mathcal{L}_\phi g) d\mu$. When there is no confusion about the potential in question, denote $\mathcal{L}_\phi$ simply by $\mathcal{L}$. The following theorem is a consequence of the compactness of the functional spaces of Σ and the Schauder-Tychonoff theorem, which tells us the eigenspaces of $\mathcal{L}$.

Theorem 1.2.5. *[Rue1, Rue2] Let (Σ, σ) be topologically mixing and ϕ Hölder, then there exists $\lambda > 0$, $h \in \mathcal{C}(\Sigma)$ and a Borel measure ν such that*

$$\mathcal{L}h = \lambda h, \; \mathcal{L}^*\nu = \lambda \nu, \; \nu(h) = 1, \; \textit{and} \; \lim_{n \to \infty} \left\| \lambda^{-n} \mathcal{L}^n g - \nu(g) h \right\|_\infty = 0 \; \textit{for all} \; g \in \mathcal{C}(\Sigma).$$

In this case, $\lambda = e^{P(\phi)}$.

The terms λ, h and ν are called the eigenvalue, eigenfunction and eigenmeasure of $\mathcal{L}$ respectively. Another fruit from this theorem is the existence of a Gibbs measure for ϕ, which is defined below.

Definition 1.2.6. *Given $\phi \in \mathcal{C}(\Sigma)$, an invariant measure $\mu \in \mathcal{M}_\sigma$ is called a* Gibbs measure *for ϕ if there exists $G \geq 1$ and $P \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ and all $C \in \mathcal{C}_n$, for all $x \in C$,*

$$\frac{1}{G} \leq \frac{\mu(C)}{\exp\left(S_n \phi(x) - nP\right)} \leq G. \tag{1.2.4}$$

A simple example of Gibbs measure is a *Bernoulli measure*. Let (Σ, σ) be a fullshift on K symbols, Suppose $\mathbf{p} = \{p_k\}_{k=1}^{K}$ is a probability vector, and for each k-cylinder $[x_0, \ldots, x_{k-1}]$,

$$\mu_{\mathbf{p}}\left([x_0, \ldots, x_{k-1}]\right) = \prod_{j=0}^{k-1} p_{x_j}.$$

Then $\mu_{\mathbf{p}}$ is the Gibbs measure for $\phi : x \mapsto \log p_{x_0}$ with $G = 1$.

It should be stressed that the Gibbs definition above is given in Bowen's sense, and its form immediately suggests its importance from the dynamical point of view: the measure of each cylinder can be uniformly approximated by the ergodic average of an arbitrary point in the set.

In statistical mechanics, Gibbs states are of central importance in *equilibrium theory*, which originated from Boltzmann's work on ideal gases (see [Bol]). His ideas were adapted to other physical systems such as ferromagnets, which are often modelled by an infinite lattice, *e.g.* $\mathbb{Z}^d$. Non-rigorously speaking, given a system S and the collection of its possible states $\{s_1, \ldots, s_N\}$, if $U(s_i)$ is the total energy of S at state s_i, then the Gibbs' rule of probability distribution satisfies

$$\mathbb{P}([s_i]) = \frac{e^{-\beta U(s_i)}}{\sum_{i=1}^{N} e^{-\beta U(s_i)}},$$

for some constant β often referred to as inverse temperature. In comparison to topological shift, it is useful to think of S as a lattice with sites $0, 1, \ldots$ and each $s \in \{s_1, \ldots, s_N\}$ as $s = (x_0, x_1, \ldots)$ which means 0 is configurated in state x_0, 1 in x_1, etc.

Rigorous constructions of probability structures on infinite probability spaces of interest to us were due to Dobrushin [Dob68], Lanford and Ruelle [LanRue] in late sixties, and also frequently referred to as the Gibbs measure. Their method, the *DLR approach*, constructs a measure from the conditional expectations over the pre-images of the sigma-algebra generated by the cylinders, that is, we care about the distribution of the initial n symbols in a sequence conditioned on a fixed tail $(x_n, x_{n+1}, \ldots)$. The precise definition will be omitted here; for non-singular probability measure ν on Σ, let

$$\nu \circ \sigma(\cdot) := \sum_{a \in \mathcal{A}} \nu\left(\sigma(\cdot \cap [a])\right),$$

then if $\frac{d\nu}{d\nu\circ\sigma} = \lambda^{-1}\exp(\phi)$ (in other words ν is ϕ-conformal Remark 2.1.3), ν is a DLR measure[1]. For a topologically transitive subshift of finite type with continuous potential ϕ, Ruelle showed there exists a DLR measure. In the early 70's, Sinai showed that natural invariant measures for hyperbolic systems are DLR measures. In fact, Sinai proved that for Hölder potentials, equilibrium states are obtained from DLR states [Sin], and (1.2.4) was established by Bowen.

Theorem 1.2.7. *[Bow, §1.4,Theorem 1.22] Let μ be such that $d\mu = hd\nu$ for h, ν in Theorem 1.2.5.*

[1] In general however, DLR measures need not to be invariant under the dynamics

Then μ is a σ-invariant probability measure with the Gibbs property with $P = P(\phi)$, and it is the unique measure in $\mathcal{M}_\sigma$ such that $h(\mu) + \int \phi d\mu = P(\phi)$.

Another important property of a Gibbs measure is that it is a fast mixing equilibrium state.

Proposition 1.2.8. *[Bow, Proposition 1.14] Let μ be the Gibbs measure in Theorem 1.2.7, then μ is exponentially ψ-mixing,* i.e.*, there exists $K_1 > 0$ and $\rho \in (0,1)$ such that for all $n, m, k \in \mathbb{N}$ and all $E \in \mathcal{C}_n$, $F \in \mathcal{C}_m$,*

$$\left| \frac{\mu(E \cap \sigma^{-n-k})}{\mu(E)\mu(F)} - 1 \right| \leq K_1 \rho^k.$$

The Gibbs measure μ also verifies a version of the Central Limit Theorem [Rat], which is beyond the scope of discussion here.

Remark 1.2.9. *In this section, the existence of an eigenmeasure for subshifts of finite type is due to the compactness of (Σ, σ), but for Markov shifts on countable alphabets those arguments may fail; instead, existence and construction of a solution to $\mathcal{L}^*\nu = \lambda\nu$ are given by a limiting procedure that produces a sequence of tight measures.*

1.3 Sarig's theorems for Gibbs measures for countable Markov shifts

Suppose $\mathcal{A}$ is countably infinite (without loss of generality we can assume $\mathcal{A} = \mathbb{N}$), and everything else remains the same, then the triplet $(\Sigma, \mathcal{A}, \sigma)$ is called a *countable Markov (sub)shift* (CMS) and (Σ, d_s) is no longer compact. As a result, Theorem 1.2.5 may even fail for $\phi = 0$ (see [Gur1]). Therefore, in order to obtain a Gibbs equilibrium state, we need to put other restrictions on the system so that (Σ, σ, ϕ) behaves like a SFT. We will recall the results regarding the entropy, equilibrium states and Gibbs states by Aaronson, Denker [AD], Gurevich [Gur2], Gurevich and Savchenko [GurSav], Mauldin and Urbański [MU] and many others. Their results lead to Sarig's work [Sar1], [Sar5] which developed a set of theorems regarding equilibrium states, the Variational Principle and characterisations of the existence and uniqueness for Gibbs measures for CMS.

First, note that the Hölder condition in the finite alphabet case is too strong in the CMS setting, *e.g.* a potential needs to be bounded from below in order to be Hölder, but only potentials unbounded from below can have finite pressure for CMS. Hence, we need some other notion of regularity of potentials for the countable alphabet case.

Definition 1.3.1. *A potential ϕ is said to be of* summable variations *if $\sum_{k\geq 2} \mathrm{var}_k(\phi) < \infty$,* weakly Hölder *(or weakly θ-Hölder) if there exists $c_\phi > 0$ and $\theta \in (0,1)$ such that $\mathrm{var}_k(\phi) \leq c_\phi \theta^k$ for all $k \geq 2$,* *and* locally Hölder *if the previous inequality holds, also for* $k = 1$.

Next, the definition of $Z_n(\phi)$ in (1.2.1) needs to be modified since for each n there may be infinitely many cylinders of depth n and the sum may easily blow up to infinity, hence Sarig introduced a new version of pressure in [Sar1].

Definition 1.3.2. *Let* $a \in \mathcal{A}$ *and set* $Z_n(\phi, a) := \sum_{\sigma^n x = x} \mathbf{1}_{[a]}(x) \exp(S_n\phi(x))$*, define the* Gurevich pressure *of* ϕ *by*

$$P_G(\phi) := \lim_{n\to\infty} \frac{1}{n} \log Z_n(\phi, a). \tag{1.3.1}$$

This quantity, introduced by Sarig [Sar1], is a generalisation of Gurevich entropy which is a special case of pressure for $\phi = 0$. For topologically mixing CMS, by [Sar1, Theorem 1] it is independent of the choice of initial symbol $a \in \mathcal{A}$, and invariant under cohomology *i.e.*, for all ϕ' cohomologous to ϕ, $P(\phi) = P(\phi')$. Similar to the Variational Principle for SFT, the Gurevich pressure for CMS can also be expressed as a supremum.

Theorem 1.3.3. *[Sar1, Theorem 2, Corollary 1, Theorem 3],[IJT, Theorem 2.10]*[2] *Let* (Σ, σ, ϕ) *be topologically mixing and* ϕ *of summable variations, then*

$$\begin{aligned}
P_G(\phi) &= \sup\{P(\phi|_Y) : Y \subseteq \Sigma \textit{ a topologically mixing finite Markov shift}\}\\
&= \sup\{P(\phi|_Y) : Y \subseteq \Sigma \textit{ compact and } \sigma^{-1}Y = Y\}\\
&= \sup\left\{h(\nu) + \int \phi\, d\nu : \nu \in \mathcal{M}_\sigma,\ \int \phi\, d\nu > -\infty\right\}\\
&= \sup\left\{h(\nu) + \int \phi\, d\nu : \nu \in \mathcal{M}_\sigma,\ \textit{and } -\int \min\{\phi, 0\}\, d\nu < \infty\right\},
\end{aligned}$$

where $P(\phi|_Y)$ *denotes the pressure of the restriction* ϕ *over the compact space* Y*. In addition, if* $\|\mathcal{L}1\|_\infty < \infty$*, then* $P_G(\phi) < \infty$.

Just as the finite alphabet case, $P_G(\phi)$ only depends on the structure of the periodic points not on the metric, so the Gurevich pressure can be viewed as the topological pressure of (Σ, σ, ϕ). We will simply write $P(\phi)$ from now on.

Definition 1.3.4. *We say* (Σ, σ) *has the* big image property*, if there exists a finite set* $B \subset \mathcal{A}$ *such that for all* $a \in \mathcal{A}$*, there is* $b \in B$ *such that* $[ab] \neq \emptyset$*. Furthermore,* (Σ, σ) *has the* big image and preimage property*,* BIP *for short, if there exists a finite set* $B \subset \mathcal{A}$ *such that for all* $a \in \mathcal{A}$*, there are* $b_1, b_2 \in B$ *such that* $[b_1 a b_2] \neq \emptyset$.

[2] The statements are proved in [Sar1] under stronger regularities on potentials but the proofs hold under summable variation assumption. See [IJT].

This is a combinatorial property and is equivalent to the *finite primitivity* notion in [MU]. CMS with the BIP property behaves like subshifts of finite type in different ways, *e.g.*

Proposition 1.3.5. *[Sar5, Corollary 1] If the BIP condition holds,*

$$P(\phi) = \lim_{n\to\infty} \frac{1}{n} \log \sum_{\sigma^n x = x} \exp\left(S_n\phi(x)\right).$$

Moreover, Gibbs measures for locally Hölder potentials are characterised by the BIP property.

Theorem 1.3.6. *Let Σ be a topologically mixing CMS and ϕ a locally Hölder[3] potential. There exists an invariant Gibbs measure if and only if Σ has the BIP property and $P(\phi) < \infty$.*

As in Theorem 1.2.5, Gibbs measures, when they exist, are constructed from the eigenfunction h and the eigenmeasure ν of the transfer operator $\mathcal{L}$ associated to ϕ, given by the generalised *Ruelle–Perron–Frobenius theorem* (see Theorem 2.1.2 in the next chapter), *i.e.*, $d\mu = hd\nu$. By [Sar1, Remark 3] this measure is unique up to multiplicative constant, and by [Sar1, Theorem 7] such a μ realises the supremum in Theorem 1.3.3 and if $\sup\phi < \infty$, it is the unique invariant equilibrium state [BS, Theorem 1]. However, there remains the caveat that an invariant measure m verifies the Gibbs property (1.2.4) with respect to a potential ϕ whilst $\int \phi dm = -\infty$, in which case the notion of equilibrium does not make sense.

Here, we prove that, just as in the SFT case, Gibbs measures for CMS are ψ-mixing. This was referred to as the *continued fraction mixing* property in terms of return-time processes for a measurable set in [ADU], and proved for Markov-fibred systems with the Schweiger property. The original proof was long and heavy since it was aiming for more general settings, so a shorter version using the transfer operator is provided here for CMS.

Lemma 1.3.7. *Under the conditions of Theorem 1.3.6, for μ the Gibbs measure with respect to a locally Hölder potential ϕ, μ is exponentially ψ-mixing.*

Proof. By [Sar1, Theorem 4, Theorem 8], $d\mu = hd\nu$ and for $\lambda = e^{P(\phi)}$, $\mathcal{L}^*\nu = \lambda\nu$ and $\mathcal{L}h = \lambda h$.

Firstly, by the locally Hölder property of ϕ, there is $M_1 > 0$ such that

$$\left| e^{S_n\phi(x) - S_n\phi(y)} - 1 \right| \leq M_1 d(x,y) \tag{1.3.2}$$

whenever x, y are in the same partition set.

[3] for existence of Gibbs measures, this can be relaxed to $\sum_{k\geq 1} \mathrm{var}_k(\phi) < \infty$ [Sar5, Theorem 1], but not for exponentially mixing properties.

Also, by the Gibbs property there is $M_2 > 0$ such that for each n-cylinder $C \in \mathcal{C}_n$, for all $x \in C$,

$$M_2^{-1}\lambda^{-n}e^{S_n\phi(x)} \leq \mu(C) \leq M_2\lambda^{-n}e^{S_n\phi(x)}. \tag{1.3.3}$$

Now define the norm (see [Sar1]) for a real-valued function f acting on Σ_A,

$$\|f\|_L := \|f\|_\infty + D_\beta f,$$

where β is the $\sigma-$algebra generated by $\{\sigma[a] : a \in \mathcal{A}\}$ and

$$D_\beta := \sup_{b\in\beta}\sup_{x\neq y\in b}\frac{|f(x)-f(y)|}{d(x,y)}.$$

The operator $\mathcal{L} : \mathrm{Lip}_{1,\beta} \to L$ where the spaces are defined by $\mathrm{Lip}_{1,\beta} := \{f : \Sigma_A \to \mathbb{R} : \|f\|_1, D_\beta f \leq \infty\}$ and $L := \{f : \Sigma_A \to \mathbb{R} : \|f\|_L < \infty\}$.

Consider $E = [e_0, e_1, \ldots, e_{n-1}] \in \mathcal{C}_n$ and $F \in \mathcal{C}^* := \bigcup_{n\geq 1}\mathcal{C}_n$, as $\mathcal{L}^*\nu = \lambda\nu$ and μ is σ invariant,

$$\begin{aligned}
&\left|\mu(E\cap\sigma^{-(n+k)}F) - \mu(E)\mu(F)\right| = \left|\mu(E\cap\sigma^{-(n+k)}F) - \mu(\sigma^{-n}E)\mu(F)\right| \\
&= \left|\int h\mathbf{1}_E\cdot\mathbf{1}_F\circ\sigma^{n+k}\,d\nu - \int h\mathbf{1}_{\sigma^{-n}E}\,d\nu\int h\mathbf{1}_F\,d\nu\right| \\
&\leq \left|\int \mathbf{1}_F\left(\lambda^{-k}\mathcal{L}^k(\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E) - h\int h\mathbf{1}_E\,d\nu\right)\,d\nu\right| \\
&\leq (\inf h)^{-1}\mu(F)\left\|\lambda^{-k}\mathcal{L}^k\left(\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)\right) - h\int\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)\,d\nu\right\|_L.
\end{aligned}$$

where the last inequality holds because $\mu(F) = \int \mathbf{1}_F h d\nu$ and h is uniformly bounded away from 0 and infinity [Sar1, Theorem 8]. The following Lasota-Yorke type inequality holds under our assumptions (see for example [AD, Theorem 1.6] or [Sar5, Corollary 3]): there are $K_\phi > 0$ and $\kappa \in (0,1)$ such that

$$\left\|\lambda^{-k}\mathcal{L}^k(\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)) - h\int\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)\,d\nu\right\|_L \leq K_\phi\kappa^k\|\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)\|_L. \tag{1.3.4}$$

+**Claim.** $\|\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)\|_L \leq M_3\mu(E)$ for some $M_3 > 0$.

Proof of claim. It is easy to see for each $E \in \mathcal{C}_n$ and $x \in \Sigma_A$, there is only one $z \in E = [e_0, e_1, \ldots, e_{n-1}]$ such that $\sigma^n z = x$, *i.e.*,

$$z = (e_0, \ldots, e_{n-1}, x_0, x_1, \ldots),$$

hence by (1.3.3), for all x,

$$\lambda^{-n}\mathcal{L}^n(h\mathbf{1}_E)(x) = \sum_{\sigma^n y = x} e^{S_n\phi(y)-nP(\phi)}(h\mathbf{1}_E)(y) \leq h(z)e^{S_n\phi(z)-nP(\phi)} \leq M_2\|h\|_\infty\mu(E),$$

and for $x, y \in [b] \in \beta$, for $z, w \in E$ allowable with $\sigma^n z = x$ and $\sigma^n w = y$.

$$\begin{aligned}\lambda^{-n}\left|\mathcal{L}^n(h\mathbf{1}_E)(x) - \mathcal{L}^n(h\mathbf{1}_E)(y)\right| &= \lambda^{-n}\left|e^{S_n\phi(z)}h(z) - e^{S_n\phi(w)}h(w)\right| \\ &\le \lambda^{-n}\left(\left|e^{S_n\phi(z)}h(z) - h(z)e^{S_n\phi(w)}\right| + \left|h(z)e^{S_n\phi(w)} - h(w)e^{S_n\phi(w)}\right|\right) \\ &\le |h(z)|M_2\mu(E)\left|1 - e^{S_n\phi(w)-S_n\phi(z)}\right| + M_2\mu(E)|h(w)|\left|1 - \frac{h(z)}{h(w)}\right|.\end{aligned}$$

By the locally Hölder property of $\log h$ and h (see [Sar1, §5]) and (1.3.2)(1.3.3), since by construction $d(z, w) = 2^{-n}d(x, y)$, the inequality above can be bounded by $\|h\|_\infty M_3\mu(E)d(x, y)$ for some constant $M_3 > 0$. □

The proof of lemma follows from the claim. □

[4]

We will prove two sets of limit theorems for asymptotic cover times and substring matching lengths in Chapter 3 and Chapter 4 respectively, for systems that are exponentially ψ-mixing in which case Gibbs measures become a natural candidate. For all systems that admit a Gibbs measure, the relevant potential has to have finite $1-$variation, whereas the formalism of equilibrium states for CMS gets more interesting once the potentials are allowed to have weaker regularities. In summable variations or weakly Hölder cases, the potential can be unbounded with $\text{var}_1(\phi) = \infty$, as a consequence equilibrium states may not (1) be exponentially mixing, or (2) be unique, or (3) even exist, depending on the *recurrent properties* of (Σ, ϕ). If a CMS does not have a Gibbs measure, the next best we can hope for that is close to the *i.i.d.* case, is *strong positively recurrent* (SPR). In this case, the system often admits an equilibrium state with exponential decay of correlations. The main goal of the next chapter is to find sufficient and necessary conditions for countable Markov shifts to be SPR.

[4] I am not sure if the locally Hölder condition in the lemma above can be relaxed to $\sum_{k\ge1}\text{var}_k(\phi) < \infty$ since the Lasota-Yorke type inequality (1.3.4) may depend on the locally Hölder property of ϕ.

Chapter 2

Strong positive recurrence

In the study of dynamical systems, an invariant ergodic measure with exponential decay of correlations is highly desirable with a simple motivation: the values of a reasonable observable along the orbit of a typical point behaves like a sequence of *i.i.d.* random variables and we have a rich collection of tools and theorems to deal with such sequences from classical probability theory. For uniformly hyperbolic systems with some mild assumptions, an equilibrium state with exponential rate of decay of correlation can be shown to exist, but that is often not the case for non-uniformly hyperbolic maps, unless we have some extra information *e.g.* the behaviour of Lyapunov exponents at a particular set of points. For example, in the literature of unimodal maps, many of them demonstrate non-uniform hyperbolicity for the geometric potential $-\log|DT|$. Moreover, for $S-$unimodal maps the *Collet-Eckmann condition*, which regulates the geometric behaviour of the image of the non-flat critical point, can be proved equivalent to a uniform lower bound of Lyapunov exponents for all periodic points, and further equivalent to the existence of an *acip* for some renormalisation of T with exponential decay rate of correlations [NS]. For more general non-geometric potentials, *e.g.* Hölder potentials, a list of conditions (see Section 2.4 below) that guarantee existence of equilibrium states have been studied since the end of 90s, and in most cases the measures in question have exponentially mixing properties.

A standard approach to prove that a Markov dynamical system has certain rate of decay of correlations is to show that the corresponding symbolic shift has the same rate of mixing. In symbolic dynamics, non-uniform hyperbolicity commonly arises from a countable alphabet. We will first review some important results of CMS regarding the modes of recurrences, with a particular focus on the notion of *strong positive recurrence* which is often equivalent to the spectral gap property. The new results in this chapter are based on ideas and results in [TZ], where a new characterisation of SPR for countable

Markov shifts is given based on some controlled boundary behaviours.

2.1 Modes of recurrences

Let $(\Sigma, \mathcal{A}, \sigma, \phi)$ be a topologically transitive countable Markov shift and $\phi \in \mathcal{SV}$, where $\mathcal{SV}$ is the set of potentials of summable variations, *i.e.*, $\sum_{k\geq 2} \mathrm{var}_k(\phi) < \infty$. The following definitions are frequently used by Sarig in his series of works [Sar1, Sar2, Sar3].

Definition 2.1.1. *Fix $a \in \mathcal{A}$ and define the return time function with respect to a by $\varphi_a(x) := \mathbf{1}_a(x) \inf\{j \geq 1 : x_j = a\}$, and*

$$Z_n^*(\phi, a) := \sum_{\sigma^n x = x} e^{S_n\phi(x)} \mathbf{1}_{\{\varphi_a = n\}}(x). \tag{2.1.1}$$

A potential ϕ on Σ is called

- recurrent *if $\sum_{n\geq 1} e^{-nP(\phi)} Z_n(\phi, a)$ diverges, and* transient *if it converges.*
- positively recurrent *if is it recurrent with $\sum_{n\geq 1} n e^{-nP(\phi)} Z_n^*(\phi, x) < \infty$, and* null-recurrent *if this sum diverges.*

The word 'recurrent' in definition above reflects the fact that the sums are taken over periodic orbits, and these quantities do not depend on the state $a \in \mathcal{A}$ for any topologically transitive CMS. The roots of recurrence properties can be found in Vere-Jones' work on Markov chains [VJ]. Heuristically speaking, a recurrent process implies the Markov chain returns to its starting state eventually while positive recurrence says this happens relatively fast; in comparison, for countable Markov shifts these notions are reflected by the behaviours of ergodic sums of periodic points.

Recurrence modes of ϕ govern the behaviour of equilibrium states (or RPF measures, in case it does not make sense to talk about equilibrium states, *e.g.* when the measure-theoretic entropy $h(\nu) = \infty$ and $\int \phi d\nu = -\infty$) in the following sense.

Theorem 2.1.2. *[Sar2],[Sar3, Theorem 2] Let Σ be a topologically mixing CMS and ϕ a potential of summable variations with $P(\phi) < \infty$. Then ϕ is recurrent if and only if there exists a conservative[1] measure ν, finite and positive on cylinder sets, and a positive continuous function h such that $\mathcal{L}^*\nu = e^{P(\phi)}\nu$, $\mathcal{L}h = e^{P(\phi)}h$, and there exists $\{a_n\}_n$ increasing such that $a_n \asymp \int_{[a]} h d\nu \sum_{k=1}^n e^{-kP(\phi)} Z_k(\phi, a)$ for all $a \in \mathcal{A}$, and for for every cylinder set C and $x \in \Sigma$,*

$$\frac{1}{a_n} \sum_{k=1}^{n} e^{-kP(\phi)} \left(\mathcal{L}^k \mathbf{1}_C\right)(x) \xrightarrow[n\to\infty]{} h(x)\nu(C).$$

In addition,

[1] A measure is conservative if all wandering sets, *i.e.*, its backward iterates are disjoint, have zero measure.

- *if ϕ is positively recurrent then $\int h d\nu < \infty$, so without loss of generality we can assume $\int h d\nu = 1$, and for every cylinder set C, $e^{-nP(\phi)}\mathcal{L}^n \mathbf{1}_C \to h\nu(C)/\int h d\nu$ uniformly on compact subsets,*
- *if ϕ is null recurrent, $\int h d\nu = \infty$, $a_n = o(n)$ and for every cylinder set C, $e^{-nP(\phi)}\mathcal{L}^n \mathbf{1}_C \to 0$ uniformly on cylinders.*

In the theorem above, h is bounded away from 0 and infinity on $[a]$ for each $a \in \mathcal{A}$, and ν, h are unique up to multiplicative constants.

Remark 2.1.3. *We also refer to the eigenmeasure of $\mathcal{L}$ associated to a potential ϕ with $P(\phi) = 0$ as the* conformal measure *of ϕ which has the following property: for $E \subset \Sigma$ such that $\sigma^n : E \to \sigma^n(E)$ is injective,*

$$\nu\left(\sigma^n(E)\right) = \int_E e^{-S_n\phi(x)} d\nu(x).$$

To sum up, if the system is null recurrent then there is an infinite RPF measure, whereas positive recurrence is a happier situation since the RPF measure is a finite equilibrium state so up to normalisation one can assume it is a probability measure. If additionally $\sup\phi < \infty$, such an equilibrium state is unique (see [BS]). However, in none of these cases we are able to confidently say at what rates the correlations between two continuous functions decay with respect to the equilibrium $d\mu = h d\nu$ *e.g.* under positive recurrence, the equilibrium can be exponentially mixing or sub-exponentially mixing [Sar4]. This is addressed in the following section.

2.2 Inducing schemes and strong positive recurrence (SPR)

Here we recall some basics of the *inducing* process in the context of countable Markov shifts. For any finite alloable word $\underline{w}$, let $|\underline{w}|$ denote its length. Fix $a \in \mathcal{A}$, the *a-induced alphabet* is

$$\mathcal{A}_a := \left\{[\underline{w}] : \underline{w} \in \Sigma^*,\ |\underline{w}| \geq 1, [\underline{w}, a] \neq \emptyset \text{ and } w_j = a \text{ iff } j = 0\right\}. \tag{2.2.1}$$

The induced shift space is $\overline{X}_a := \mathcal{A}_a^{\mathbb{N}_0}$, define the natural projection $\pi : \overline{X}_a \to \Sigma$ by $\pi\left([\underline{a}_0], [\underline{a}_1], \dots\right) = (\underline{a}_0, \underline{a}_1, \dots)$. The induced shift system is then $(\overline{X}_a, \overline{\sigma}, \overline{\phi})$ where

$$\pi \circ \overline{\sigma} = \sigma^{\varphi_a} \circ \pi, \text{ and } \overline{\phi} = \left(\sum_{j=0}^{\varphi_a - 1} \phi \circ \sigma^j\right) \circ \pi$$

Inducing schemes are always full-shifts on $\mathcal{A}_a$, and problems arising in the original system $(\Sigma, \mathcal{A}, \phi)$ due to extreme variations of ϕ on partition sets may disappear since $\text{var}_n\left(S_n\overline{\phi}\right) \leq \sum_{k\geq 2} \text{var}_k(\phi)$. In particular, if ϕ is weakly Hölder then $\overline{\phi}$ is locally Hölder, and the limit $P(\overline{\phi}) := \lim_{n\to\infty} \frac{1}{n} \log Z_n(\overline{\phi}, [\underline{a}])$ always exists and independent of $[\underline{a}] \in \mathcal{A}_a$ [Sar3, Lemma 2], and the results concerning Gibbs measures in § 1.3 are applicable.

Now we introduce the following quantities defined in [Sar3]:

$$p_a^*[\phi] := \sup\left\{p \in \mathbb{R} : P(\overline{\phi + p}) < \infty\right\}, \text{ and } \Delta_a[\phi] := \sup\left\{P(\overline{\phi + p}) : p < p_a^*[\phi]\right\} = P\left(\overline{\phi + p_a^*[\phi]}\right),$$

where $\Delta_a[\phi]$ is called the *a-discriminant* and the last equality is due to [Sar3, Proposition 3], obtained as a by-product of

$$\left| P\left(\overline{\phi + p}\right) - \log \sum_{k \geq 1} e^{kp} Z_k^*(\phi, a) \right| \leq \sum_{k \geq 2} \operatorname{var}_k(\phi). \tag{2.2.2}$$

For a topologically mixing CMS and $\phi \in \mathcal{SV}$ with $P(\phi) < \infty$, the discriminant is another indicator reflecting the recurrence modes: $\Delta_a[\phi] > 0$ implies positive recurrence and $\Delta_a[\phi] < 0$ implies transience, while $\Delta_a[\phi] = 0$ can be either positive or null recurrent [Sar3, Theorem 2]. In fact, the case $\Delta_a[\phi] > 0$ is important enough to have earned a separate label.

Definition 2.2.1. *ϕ is* strong positive recurrent*, SPR for short, if for some $a \in \mathcal{A}$, $\Delta_a[\phi] > 0$. Equivalently*[2]*, see for example [Cli, §8.5],*

$$\limsup_{n \to \infty} \frac{1}{n} \log Z_n^*(\phi, a) < P(\phi). \tag{2.2.3}$$

This is a generalisation of the *stable positivity* notion in [GurSav], in the sense that if ϕ_0 is SPR, its positive recurrence nature remains 'stable' under a small perturbation by a nice potential ϕ_1: if there is an interval in which each t has $P(\phi_0 + t\phi_1) < \infty$, then $\phi_0 + t\phi_1$ is also positively recurrent, and $t \mapsto P(\phi_0 + t\phi_1)$ is real analytic on such an interval.

Remark 2.2.2. *The notion of SPR has been applied to or generalised in other systems (perhaps non-symbolic), see for example [RV], [GST] and [BCS2]. The key idea is that there is some gap between the entropy or pressure at infinity and the topological pressure of the potential. In addition to exponential decay of correlations, SPR also implies other properties of the equilibrium such as the EKP inequality (which is proved equivalent to SPR) [RS].*

Example 2.2.1 (The renewal shift)**.** Consider the $\mathbb{N}_0 \times \mathbb{N}_0$ transition matrix M with $M_{00} = M_{0i} = M_{i,i-1} = 1$ for all $i \in \mathbb{N}$ and all other entries 0, *i.e.*,

$$M = \begin{pmatrix} 1 & 1 & 1 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The *renewal space* is defined by $\Sigma_R := \left\{x \in \mathbb{N}_0^{\mathbb{N}_0} : M_{x_i, x_{i+1}} = 1\right\}$, and the left shift dynamics on Σ_R is

[2]Another equivalent definition commonly used is: there exists p such that $P(\overline{\phi + p}) = 0$ [PZ].

represented by the following diagram.

$$\begin{tikzcd} 0 \arrow[loop left] \arrow[r, bend left] \arrow[rr, bend left] \arrow[rrr, bend left] \arrow[rrrr, bend left] & 1 \arrow[l] & 2 \arrow[l] & 3 \arrow[l] & \cdots \arrow[l] \end{tikzcd}$$

Figure 2.1: Renewal shift

Then the *renewal shift* system (Σ_R, σ) is obviously topologically mixing but only has big preimages so not BIP, and one can construct a conjugacy between the Bernoulli shift on $\{0,1\}$ and the renewal shift with all preimages of 0 removed. The renewal shift is one canonical example of CMS and models the dynamics of many interval maps with non-uniform hyperbolicity.

Let $\phi : \Sigma_R \to \mathbb{R}$ be a potential of summable variations, $P(\phi) \le \log \|\mathcal{L}_\phi 1\|_\infty \le \log\left(2e^{\sup\phi}\right)$. Let $(\overline{\Sigma}_R, \overline{\phi})$ be the induced shift on state 0. Note that for all n, $Z_n^*(\phi, 0) = \exp\left(\overline{\phi}(\overline{x})\right)$ where $\overline{x} = ([\underline{a}_n], [\underline{a}_n], \dots)$, and $\underline{a}_n$ is the n-length word $0, n-1, \dots, 1$, so $p_0^*[\phi] = -\limsup_n \frac{1}{n}\phi(\underline{a}_n, \underline{a}_n, \dots)$.

If ϕ is positively recurrent and $\sup\phi < \infty$, it admits a unique equilibrium state [PZ, Proposition 2.6]; if also we have $\overline{\phi}$ weakly-Hölder, by [Sar3, Theorem 5] there exists β_c such that

- $0 < \beta < \beta_c$: $\beta\phi$ is strong positive recurrent with $P(\beta\phi)$ real analytic in β, continuous but not analytic at β_c with $P(\beta_c\phi) < \infty$.
- $\beta_c < \beta < \infty$: $\beta\phi$ is transient and $P(\beta\phi)$ is linear in β.

So for renewal shifts, the one parameter family of $\{\beta\phi\}_\beta$ has a *phase transition* at β_c. For interval maps $T : \mathcal{X} \to \mathcal{X}$ that are conjugate to renewal shifts *e.g.* Manneville-Pomeau maps, this theorem implies that if the lifted potential of $-\log|DT|$ to Σ_R has finite pressure, then for some $t \le 1$, $-t\log|DT|$ has a unique equilibrium which is absolutely continuous with respect to Lebesgue, for t in some interval $(0, t_c)$ [Sar3, Proposition 1].

2.3 The spectral gap property (SGP)

Now we turn to the discussion of the *spectral gap property*. Recall that the transfer operator $\mathcal{L} = \mathcal{L}_\phi$ associated with ϕ is defined by $\mathcal{L}f(x) = \sum_{\sigma y = x} e^{\phi(y)} f(y)$. Under mild conditions, if $\mathcal{L}$ acts on a nice Banach space $\mathcal{B}$ with a spectral gap (defined below) then a finite equilibrium measure exists and enjoys many desirable mixing properties (see for example [Bal]).

Definition 2.3.1 (Spectral Gap Property (SGP))**.** *Suppose $\phi : \Sigma \to \mathbb{R}$ is weakly θ-Hölder and $P(\phi) < \infty$. Then we say ϕ has the spectral gap property (SGP) if there is a Banach space $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ of continuous functions on Σ such that*

(a) *$\mathcal{B} \subset$ dom$(\mathcal{L})$ and $\{\mathbf{1}_C : C \in \mathcal{C}_n,\, n \in \mathbb{N}\} \subset \mathcal{B}$, where $\mathcal{C}_n$ is the collection of n-cylinders and*

$$\text{dom}(\mathcal{L}) := \left\{ f : \Sigma \to \mathbb{R} : \mathcal{L}f(x) = \sum_{\sigma y = x} e^{\phi(y)} f(y) \text{ converges for all } x \in \Sigma \right\}.$$

(b) *$f \in \mathcal{B}$ implies $|f| \in \mathcal{B}$ and $\|f\|_{\mathcal{B}} \le \||f|\|_{\mathcal{B}}$,*

(c) *convergence in $\|\cdot\|_{\mathcal{B}}$ implies uniform convergence on cylinder sets,*

(d) *$\mathcal{L}(\mathcal{B}) \subset \mathcal{B}$, and $\mathcal{L} : \mathcal{B} \to \mathcal{B}$ is a bounded operator,*

(e) *The operator $\mathcal{L}$ can be decomposed into $e^{P(\phi)}P + N$, where N, P are bounded operators on $\mathcal{B}$ with $PN = NP = 0$, $P^2 = P$, $\dim(Im\, P) = 1$, and the spectral radius of N is less than $e^{P(\phi)}$,*

(f) *if g is weakly θ-Hölder, then $\mathcal{L}_{\phi+zg} : \mathcal{B} \to \mathcal{B}$ is bounded, and $z \mapsto \mathcal{L}_{\phi+zg}$ is analytic on a complex neighbourhood of 0.*

Those definitions are taken from [CS], where the authors showed the following.

Theorem 2.3.2. *[CS, Theorem 1.1] If ϕ is weakly θ-Hölder over a topologically mixing CMS, has finite supremum and satisfies the SGP defined above, then P takes the form $Pf = h \int f d\nu$, where $h \in \mathcal{B}$ is positive, and ν a measure that is finite and positive on all cylinders. The measure μ with $d\mu = h d\nu$ is in $\mathcal{M}_\sigma$ such that*

(a) *if μ has finite entropy, it is the unique equilibrium state of ϕ,*

(b) *there is $\kappa \in (0,1)$ such that for all $g \in L^\infty(\mu)$ and f bounded Hölder continuous, there exists $C(f,g) > 0$ such that for all $n \in \mathbb{N}$,*

$$\left| \int fg \circ \sigma^n d\mu - \int f d\mu \int g d\mu \right| \le C(f,g)\kappa^n,$$

(c) *the Central Limit Theorem holds (see [Sar6, Theorem 6.4]),*

(d) *if ϕ' is a bounded Hölder continuous function, then $t \mapsto P(\phi + t\phi')$ is real analytic on a neighbourhood of zero.*

So in short, SGP produces a unique equilibrium state with exponentially fast mixing behaviour, and it is not an uncommon property. For shifts with finite alphabets, it is well-known that every Hölder potential has SGP. For CMS, the set of SGP potentials is open and dense with respect to a specific class of topology (see [CS, Theorem 2.2]); but SGP may fail for different reasons, *e.g.* the potential ϕ is transient or null recurrent. It will be useful to find a necessary and sufficient condition to guarantee the spectral gap property. To this matter, Cyr and Sarig elegantly presented the following theorem.

Theorem 2.3.3. *[CS, Theorem 2.1] Suppose Σ is a topologically mixing CMS, and $\phi : \Sigma \to \mathbb{R}$ is weakly Hölder continuous with finite Gurevich pressure, then ϕ has the spectral gap property if and only if ϕ is strong positive recurrent.*

Therefore, in order to answer whether a CMS admits an equilibrium probability measure with exponential decay of correlations, it is often enough to check if (Σ, ϕ) is SPR, hence in the remaining part of this chapter, we provide a characterisation of SPR via periodic points.

2.4 Boundary behaviours of countable Markov shifts

Transience is one major obstruction to the existence of equilibrium states, and in [Cyr] it is established that transience is possible only when there are infinitely many states in $\mathcal{A}$ that get visited by 'long excursions' from a compact part of Σ. More precisely, Σ has a transient potential if there is a state $a \in \mathcal{A}$ with arbitrarily long paths back to itself that do not visit any $b \in \mathcal{A}$ twice. Paths of this kind belong to what we refer to as the *boundary* of the CMS. We can quantify the boundary behaviours via the following quantities.

Definition 2.4.1 (The $\mathcal{F}$ property)**.** *Our system Σ has the* $\mathcal{F}$-property *if for every state $a \in \mathcal{A}$ and every $n \in \mathbb{N}$, the number of periodic points in $[a]$ with period n is finite.*

The $\mathcal{F}$-property obviously fails for full-shifts.
The $\mathcal{F}$-property holds when (Σ, σ):

- is locally compact (i.e. for every $i \in \mathcal{A}$, $\sum_{j \in \mathcal{A}} M_{i,j} < \infty$ where $[M]_{ij}$ is the transition matrix);
- has $h_{top} < \infty$. Note that If ϕ is uniformly bounded from below and $P(\phi) < \infty$ then the $\mathcal{F}$-property must hold.

Definition 2.4.2 (Entropy contraction at infinity)**.** *For each n, M and q, define the set of $n+1$-cylinders*

$$B(n, M, q) := \left\{ [x_0, \dots, x_n] \in \mathcal{C}_{n+1} : x_0, x_n \le q, \, \#\{k \le n : x_k \le q\} \le \frac{n+1}{M} \right\},$$

and write $z_n(M, q) := \#B(n, M, q)$. The entropy at infinity h_∞ *as in [ITV] is defined via*

$$h_\infty(M, q) := \limsup_{n \to \infty} \frac{1}{n} \log z_n(M, q), \quad h_\infty(q) := \liminf_{M \to \infty} h_\infty(M, q),$$

$$h_\infty := \liminf_{q \to \infty} h_\infty(q).$$

In later chapters we are interested in systems with $h_\infty = 0$. A trivial example of CMS with $h_\infty = 0$ is the renewal shift given in Example 2.2.1.

If the $\mathcal{F}$-property fails then h_∞ may not make sense: suppose there are $q \in \mathcal{A}$ and $N \in \mathbb{N}$ such that there are infinitely many periodic orbits of length N intersecting $[a]$ for some $a \le q$. By pigeonhole principle some of these loops contribute, in some sense, to the boundary, but the set $B(n, M, q)$ does not see these paths if $M > N$. Another important consequence of the $\mathcal{F}$-property is that $\mathcal{M}_{\le 1}(\Sigma)$, the space of shift-invariant sub-probability (*i.e.*, the measure of Σ is in $[0,1]$) measures, is compact [IV, Theorem 1.2]

Remark 2.4.3. *Transience arises from non-compactness of the shift space, in the sense that points often escape to the 'boundary' part of the system. A CMS admits transient potentials only if its 'boundary states' is not finite. More precisely [Cyr, Theorem 2.1], if it does not have a finite uniform Rome (defined below).*

Definition 2.4.4. *Let $(\Sigma, \mathcal{A})$ be a topologically transitive CMS. For $a, b \in \mathcal{A}$, we say there is* a path of length ℓ between a and b *if there exists $\underline{w} \in \Sigma_{\ell-1}$ such that $w_j \neq a$ or b for all $j = 0, \dots, \ell-2$ and $[a\underline{w}b] \neq \emptyset$. A* finite uniform Rome *is a finite subset $\mathcal{A}' \subset \mathcal{A}$ such that for some $N \in \mathbb{N}$,*

$$\Sigma \cap (\mathcal{A} \setminus \mathcal{A}')^{\mathbb{N}_0} \text{ has no path of length } \geq N.$$

Existence of a finite uniform Rome defined in Definition 2.4.4 is 'good' for a system in the sense that it a finite pressure potential may admit equilibrium states. It also results in unusual behaviour of entropy at infinity, *i.e.*, by our definition, we can show that $h_\infty = -\infty$. Suppose there exists such a finite uniform Rome $\mathcal{A}'$. Then for all $M \geq N$ where N is given by the definition of $\mathcal{A}'$, and each q such that $\mathcal{A}' \subseteq [\leq q] := \cup_{a=0}^{q}[a]$, for all $n > NM$, if there exists some $\underline{w} \in \Sigma^*$ in $B(n, M, q)$, there must be a subword of $\underline{w}$ of length greater than N which contains no states in $[\leq q]$. However, such path does not exist by definition of a finite uniform Rome, and we have a contradiction. Hence $B(n, M, q) = \emptyset$ for all $n > MN$, which technically implies $h_\infty = -\infty$.

Definition 2.4.5 (Contraction at infinity)**.** *We define the following quantities:*

$$z_{\phi,n}(M,q) := \sup\left\{\frac{1}{n}S_n\phi(x) : x \in B(n,M,q)\right\},$$

$$\delta_{\phi,\infty}(M,q) := \limsup_{n\to\infty} z_{\phi,n}(M,q), \ \ \delta_{\phi,\infty}(q) := \liminf_{M\to\infty} \delta_{\phi,\infty}(M),$$

$$\delta_{\phi,\infty} := \liminf_{q\to\infty} \delta_{\phi,\infty}(q).$$

Then the system is said to have contraction at infinity (CI)*, if $\delta_{\phi,\infty} < P(\phi)$.*

The name contraction can be seen as the symbolic counterpart of hyperbolicity for interval maps $T : \mathcal{X} \to \mathcal{X}$, in which case the geometric potential $-\log|DT|$ decreases. The quantity h_∞ is well-studied in [ITV, Theorem 1.1, Theorem 1.4] and other recent literature, and for CMS with finite topological entropy, *i.e.*, for $P(0) < \infty$, it coincides with the measure-theoretic entropy and Buzzi's graph-theoretic entropy at infinity [Buz].

Definition 2.4.6. *For Σ a topologically transitive CMS and $\phi : \Sigma :\to \mathbb{R}$ of summable variations, the system is said to have:*

Uniform contracting structure (**UCS**) *if*

$$\chi_{per}(\phi) := \sup \left\{ \frac{1}{n} S_n\phi(x) : \sigma^n(x) = x \right\} < P(\phi);$$

Compact returns contractions (**CRC**) *if given* $q \in \mathbb{N}$*, there exist* $C_q \in \mathbb{R}$ *and* $\lambda_q > 0$ *such that if* $x \in \Sigma$ *has* $x_0, x_n \leq q$ *then*

$$S_n\phi(y) \leq C_q - n\lambda_q;$$

Contraction at infinity (**CI**) *if* $\delta_{\phi,\infty} < P(\phi)$.

The condition (UCS) should be compared with other notions of 'hyperbolicity' or 'contraction' that were defined in various ways for interval maps to show existence of equilibrium states and, in most cases, with exponential decay of correlations. We provide a partial list here below, where ϕ is often taken to be a Hölder potential.

(i) $\sup \phi - \inf \phi < h_{top}$, see for example [HK, BT1];
(ii) $\sup \phi < P(\phi)$, see [DKU];
(iii) there exists $n_0 \in \mathbb{N}$ such that $\sup_{x \in X} S_{n_0}\phi(x)/n_0 < P(\phi)$, see [IR, LiRiv2] and [LSV] (the last one involves a covering condition);
(iv) there exists $\lambda_f > 1$ such that for all $x \in [0,1]$ with $f^p(x) = x$ for some $p \in \mathbb{N}$, $|Df^p(x)| \geq \lambda_f^p$.

Our (UCS) condition is closest to (iv) which has been proved to be equivalent to the existence of an absolutely continuous probability measure with exponential decay of correlations for a class of unimodal maps, see [NS].

We first show that for a system (Σ, ϕ) with some boundary conditions, (UCS) is equivalent to (CI) meaning that in order to understand the Birkhoff averages accumulated on the paths escaping to 'infinity', it is enough to look at the behaviour of periodic points.

Theorem 2.4.7. *Let* (Σ, ϕ) *be a topologically transitive CMS with the* $\mathcal{F}$*-property,* ϕ *a potential with summable variations and the pressure* $P(\phi) < \infty$*. Then (UCS) is equivalent to (CI).*

The proof will be split into several lemmas. In the remaining part of this chapter, since $P(\phi) < \infty$ is always assumed, by subtracting a constant from ϕ we can without loss of generality take $P(\phi) = 0$. Also, take $\mathcal{A} = \mathbb{N}_0$ so we have a natural ordering of labels.

Recall that $B_\phi := \sum_{k \geq 2} \mathrm{var}_k(\phi) < \infty$. For any allowable k-word $\underline{w} = w_0, \ldots, w_{k-1} \in \Sigma_k$ such that $[w_{k-1}, w_0] \neq \emptyset$, let $(\overline{w_0, \ldots, w_{k-1}})$ denote the corresponding periodic point of period k.

Given $a, b \in \mathbb{N}$, define

$$\ell(a,b) := \min\{k : \exists \underline{w} \in \Sigma_k, w_0 = a \text{ and } \underline{w}b \in \Sigma_{k+1}\}.$$

For each pair a, b, $\ell(a,b)$ is finite by topological transitivity, hence $\ell(q) := \sup_{a,b\le q} \ell(a,b)$, is also finite.

Again by topological transitivity, for any $x \in \Sigma$ and $n \in \mathbb{N}$, there is $\ell = \ell(x_n, x_0) \geq 0$ and we can pick a finite word $\underline{w}(x_n, x_0) \in [x_n]$ of length ℓ such that the following concatenation is allowed

$$\underline{w}(x_n, x_0)x_0, \ldots, x_{n-1} \in \Sigma_{n+\ell}, \tag{2.4.1}$$

which are the first $n+\ell$ symbols of a periodic point $z = \left(\overline{\underline{w}(x_n, x_0)x_0, \ldots, x_{n-1}}\right)$.

Lemma 2.4.8. *Suppose that ϕ has summable variations. If we define*

$$\underline{C}_q(\phi) := \min_{a,b\le q} \inf_{y\in[\underline{w}(a,b)b]} \left\{S_{\ell(a,b)}\phi(y)\right\},$$

then we have $\underline{C}_q(\phi) > -\infty$.

Proof. This follows since there are finitely many words $\underline{w}(a,b)$ for $a, b \le q$ to consider and $S_{\ell(a,b)}\phi$ is bounded on each $[\underline{w}(a,b)b]$ by summable variations. □

Lemma 2.4.9. *For a topologically transitive CMS, ϕ of summable variations, (UCS) implies (CRC).*

Proof. Let $q \in \mathbb{N}$ be given and pick $\lambda_q > 0$ small enough such that $\chi_{per}(\phi) < -2\lambda_q < 0$. Suppose $x \in \Sigma$ has $x_0, x_n \le q$. Then for $z = \left(\overline{\underline{w}(x_n, x_0)x_0, \ldots, x_{n-1}}\right)$ defined as above,

$$\underline{C}_q(\phi) + S_n\phi(x) \le S_{n+\ell}\phi(z) + 2B_\phi < 2B_\phi - \left(n + \min_{a,b\le q} \ell(a,b)\right)\lambda_q$$

by Definition2.4.6, where $\underline{C}_q(\phi) > -\infty$ as in Lemma 2.4.8. Hence $S_n\phi(x) < C_q - n\lambda_q$ where $C_q = \max\left\{0, 2B_\phi - \underline{C}_q(\phi)\right\}$. □

It is straightforward that (CRC) implies (CI) since for all $x \in \Sigma$ such that $x_0, x_n \in [\le q]$, (CRC) does not care whether $x_j \in [\le q]$ or not, for $j \in [1, n-1]$. To complete the proof of Theorem 2.4.7, it only remains to show that (CI) implies (UCS) for, by the next two lemmas, non-positive potentials.

Lemma 2.4.10. *There exists $h : \Sigma \to \mathbb{R}$ bounded on each 1-cylinder such that for $\phi' := \phi + \log h - \log h \circ \sigma$, we have $\phi' \le 0$. Moreover, ϕ' has summable variations (or is weakly Hölder) if ϕ has summable variations (or is weakly Hölder).*

This is essentially [Sar3, Lemma 1], but we sketch parts of the proof here for completeness.

Proof. If ϕ is recurrent, then h is the eigenfunction of the transfer operator $\mathcal{L}$ associated to ϕ. If ϕ is transient, take $h = \sum_{n\geq 1} \mathcal{L}^n \mathbf{1}_{[a]}$. The regularity follows as in [Sar3, Lemma 1], although there the shift is assumed to be topologically mixing, h remains finite and non-positive under topological transitivity. □

We next show that ϕ' inherits (CI) from ϕ.

Lemma 2.4.11. *If $\delta_{\phi,\infty} < 0$, then $\delta_{\phi',\infty} < 0$.*

Proof. By definition, there exist $\varepsilon > 0$ and $N_\varepsilon, M_\varepsilon, q_\varepsilon$ such that

$$z_{\phi,n}(M, q_\varepsilon) < -2\varepsilon \tag{2.4.2}$$

for all $n > N_\varepsilon$ and $M > M_\varepsilon$. Then for every $n > N_\varepsilon$ large enough that $\frac{n+\ell(q_\varepsilon)}{2M} < \frac{n}{M}$, for every $x \in B$ for some $B \in B(n, 2M, q_\varepsilon)$, as in (2.4.1), there exists an allowable word $\underline{w} = \underline{w}(x_n, x_0)$ looping x_n back to x_0, and a periodic point y of period $n' = n + |\underline{w}|$, such that $[y_0, \ldots, y_{n+|\underline{w}|-1}] \in B(n', M, q_\varepsilon)$, $y = (\overline{x_0, \ldots, x_{n-1}\underline{w}})$ and by summable variations,

$$S_n\phi'(x) \leq S_{n'}\phi'(y) - \underline{C}_q(\phi') + B_{\phi'} = S_{n'}\phi(y) - \underline{C}_q(\phi') + B_{\phi'},$$

where $\underline{C}_q(\phi')$ is defined as in Lemma 2.4.8. Then (2.4.2) implies

$$S_n\phi'(x) < -2n'\varepsilon - C_q(\phi') + B_{\phi'}$$

and by choosing n large, this implies that for all $M > M_\varepsilon$, $z_{\phi',n'}(M, q_\varepsilon) < -\varepsilon$, and consequently $\delta_{\phi',\infty}(q_\varepsilon) < -\varepsilon$. Since this inequality holds for all $q > q_\varepsilon$, we conclude that $\delta_{\phi',\infty} < 0$. □

Note that this lemma also holds for any cohomologous $\psi = \phi + \xi - \xi \circ \sigma$, provided ξ has summable variations.

In the following lemma we show that we cannot have a sequence of periodic measures supported on a finite collection of partition sets such that their integrals of ϕ converge to zero (in non-normalised cases the limit is $P(\phi)$), and simultaneously converge to a probability measure. The proof requires compactness results regarding the space of sub-probability measures on Σ, in particular we say a sequence of measures $(\mu_n)_n$ *converges on cylinders* to a measure μ if for any $C \in \mathcal{C}_k$, $\mu_n(C) \to \mu(C)$ as $n \to \infty$, see [IV] for more details.

Lemma 2.4.12. *There is no $q \in \mathbb{N}$ and sequence $(x^k)_k$ of periodic points of period p_k such that $\frac{1}{p_k} S_{p_k}\phi(x^k) \to 0$ and $\nu_k([\leq q]) \to 1$ as $k \to \infty$, where $\nu_k = \frac{1}{p_k}\sum_{i=0}^{p_k-1} \delta_{\sigma^i x^k}$.*

Proof. Assume by contradiction that the lemma is false, we will show that there is an equilibrium state ν that has zero measure-theoretic entropy, which, by for example [Sar6, Theorem 5.6], is a contradiction.

Let ϕ' be as in Lemma 2.4.11, and it is easy to see $S_n\phi(x) = S_n\phi'(x)$ for all $n \in \mathbb{N}$ and all $\sigma^n x = x$. Suppose there is such a $q \in \mathbb{N}$ and sequence of periodic points as in the statement of our lemma. Note that $\int \phi' \, d\nu_k \to 0$. By [IV, Theorem 1.2], $\mathcal{M}_{\leq 1}(\Sigma)$, the space of shift-invariant sub-probability measures on Σ, is compact with respect to the convergence on cylinders topology, *i.e.*, there is $\nu \in \mathcal{M}_{\leq 1}(\Sigma)$ such that $\nu_k \to \nu$ (up to subsequences) on cylinders. Our assumption implies that ν is a probability measure. Hence, as $(\nu_k)_k$ and ν are probability measures, [IV, Lemma 3.17] implies that the convergence also holds in the weak-* topology. In particular, if we let

$$\phi'_L(x) := \begin{cases} \phi'(x) & \text{if } \phi'(x) \geq -L, \\ 0 & \text{if } \phi'(x) < -L. \end{cases}$$

Then ϕ'_L is continuous and bounded whence $\int \phi'_L \, d\nu_{n_k} \to \int \phi'_L \, d\nu$.

Claim. *Given $L > 0$, for any $\varepsilon > 0$ there exists K'_ε such that for all $k \geq K'_\varepsilon$,*

$$\left| \int \phi'_L \, d\nu_{n_k} - \int \phi' \, d\nu_{n_k} \right| < \varepsilon/4.$$

Proof of Claim. Since

$$\int \phi' \, d\nu_{n_k} = \int \phi'_L \, d\nu_{n_k} + \int_{\{\phi' < -L\}} \phi' \, d\nu_{n_k},$$

if the claim is false then there is $\varepsilon > 0$ such that for any $N \in \mathbb{N}$ we can find $k \geq N$ such that

$$\left| \int_{\{\phi' < -L\}} \phi' \, d\nu_{n_k} \right| \geq \varepsilon/4.$$

But since $\phi' \leq 0$, this means $\int \phi' \, d\nu_{n_k} \leq -\varepsilon/4$, contradicting the fact that $\int \phi' \, d\nu_{n_k} \to 0$. $\square$

Now given $L > 0$, take $K_\varepsilon \geq K'_\varepsilon$, where K'_ε is given by the claim, such that $\left|\int \phi' \, d\nu_{n_k}\right| < \varepsilon/4$, and $\left|\int \phi'_L \, d\nu - \int \phi'_L \, d\nu_{n_k}\right| < \varepsilon/2$ for all $k \geq K_\varepsilon$. Then

$$\begin{aligned} \left| \int \phi'_L \, d\nu \right| &\leq \left| \int \phi'_L \, d\nu - \int \phi'_L \, d\nu_{n_k} \right| + \left| \int \phi'_L \, d\nu_{n_k} \right| \\ &< \frac{\varepsilon}{2} + \left| \int \phi' \, d\nu_{n_k} - \int \phi'_L \, d\nu_{n_k} \right| + \left| \int \phi' \, d\nu_{n_k} \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned} \tag{2.4.3}$$

Now the Monotone Convergence Theorem implies $-\int \phi'_L \, d\nu \nearrow -\int \phi' \, d\nu = -\int \phi \, d\nu$ as $L \to \infty$.

Moreover (2.4.3) and weak* convergence of ν_{n_k} to ν imply $\left|\int \phi \, d\nu\right| < \varepsilon$ for all ε, i.e., $\int \phi \, d\nu = 0$, a contradiction. $\square$

Proposition 2.4.13. *Under the assumptions of Theorem 2.4.7, (CI) implies (UCS).*

The idea of the proof is that (UCS) must hold for periodic orbits which 'spend most of their time in a compact part' of the space. In the finite alphabet case this is trivially true, and for CMS such is assured by Lemma 2.4.12, and then (CI) ensures that orbits which 'spend significant time outside the compact part' satisfy contraction. Combining these two arguments we get (UCS).

Proof. Suppose (CI) holds for ϕ hence ϕ' and (UCS) fails. Then by definition and non-negativity of ϕ' there exists a sequence of periodic points $x^1, x^2, \ldots$ with periods $p_1, p_2, \ldots$ and with Birkhoff averages

$$s_n = \frac{1}{p_k} S_{p_k} \phi(x^n) \leq 0$$

for all n and $\lim_{n\to\infty} s_n = 0$.

By the definition of (CI) and Lemma 2.4.11, for all $\varepsilon > 0$ such that $\delta_{\phi',\infty} < -\varepsilon < 0$, there exist N_ε, $M_\varepsilon, q_\varepsilon$ such that for $n > N_\varepsilon$, $M > M_\varepsilon$, $q > q_\varepsilon$ and all x such that $[x_0, \ldots, x_{n-1}] \in B(n, M, q)$, $S_n\phi'(x) < -n\varepsilon$.

Given $q, N \in \mathbb{N}$, let $\mathcal{A}_{[\leq q],N}$ be the set of words $\underline{w} \in \Sigma_N$ such that $w_0 \leq q$ and $w_i > q$ for $i = 1, \ldots, N-1$, and $\underline{w}q' \in \Sigma_{N+1}$ for some $q' \leq q$. Let

$$\mathcal{A}_{[\leq q]} := \cup_{N\geq 1}\mathcal{A}_{[\leq q],N} \text{ and for each } \underline{w} \in \mathcal{A}_{[\leq q]},\ [\underline{w}, \leq q] := \bigcup_{q'\leq q} \{[\underline{w}q'] : \underline{w}q' \in \Sigma_{|\underline{w}|+1}\}.$$

Given $x \in \Sigma$ such that $x_0, x_n \leq q$, we can decompose $x_0, \ldots, x_{n-1} = \underline{w}_1\underline{v}_1\underline{w}_2\underline{v}_2 \ldots \underline{w}_k\underline{v}_k$, where $\underline{w}_i \in \mathcal{A}_{[\leq q]}$ and $\underline{v}_i \in \{\emptyset\} \cup (\cup_m \{1, \ldots, q\}^m)$ for all $1 \leq i \leq k$. For each x, let $\mathcal{D}(x, q)$ denote the set of words $\underline{w}_i$ in this decomposition.

Given $q > q_\varepsilon$, define the proportion function $\zeta(\cdot)$

$$\zeta(x^n) := \frac{1}{p_n} \sideset{}{'}\sum_{\left\{\underline{w}\in\mathcal{D}(x^n,q_\varepsilon)\cap(\bigcup_{N\geq N_\varepsilon}\mathcal{A}_{\leq q_\varepsilon,N})\right\}} |\underline{w}|. \tag{2.4.4}$$

Here $\sum'$ means that we count with multiplicity, *i.e.*, if $\underline{w}$ appears k times in the decomposition of x^n, we sum its lengths k times.

Notice that since $w_j \leq q$ if and only if $j = 0$ for each $\underline{w} \in \mathcal{A}_{[\leq q],N}$, so long as $(N+1)M_\varepsilon \geq 1$, $S_N\phi'(x) < -N\varepsilon$ for $x \in [\underline{w}]$. Hence, since we can assume that $(N_\varepsilon + 1)M_\varepsilon \geq 1$, we can show that

$\limsup_{n\to\infty}\zeta(x^n)=0$: if there exists $\eta>0$ such that $\lim_{n\to\infty}\zeta(x^n)\geq\eta$, by the non-positivity of ϕ',

$$\begin{aligned}\liminf_{n\to\infty}s_n &\leq \liminf_{n\to\infty}\frac{1}{p_n}\sideset{}{'}\sum_{\left\{\underline{w}\in\mathcal{D}(x^n,q_\varepsilon)\cap\left(\bigcup_{N\geq N_\varepsilon}\mathcal{A}_{\leq q_\varepsilon,N}\right)\right\}}\sup_{x\in[\underline{w},\leq q_\varepsilon]}S_{|\underline{w}|}\phi'(x)\\ &\leq \liminf_{n\to\infty}\frac{1}{p_n}\sideset{}{'}\sum_{\left\{\underline{w}\in\mathcal{D}(x^n,q_\varepsilon)\cap\left(\bigcup_{N\geq N_\varepsilon}\mathcal{A}_{\leq q_\varepsilon,N}\right)\right\}}-|\underline{w}|\varepsilon \qquad (2.4.5)\\ &\leq \limsup_{n\to\infty}-\zeta\left(x^n\right)\varepsilon\leq-\eta\varepsilon<0,\end{aligned}$$

contradicting our assumption that $\lim_{n\to\infty}s_n=0$.

By the $\mathcal{F}$-property and topological transitivity, $\#\{\cup_{N\leq N_\varepsilon}\mathcal{A}_{[\leq q],N}\}<\infty$, and the following quantity is also finite:

$$q'(q,N):=\min\left\{q'\in\mathbb{N}:\text{if }\underline{w}\in\cup_{N\leq N_\varepsilon}\mathcal{A}_{[\leq q],N}\text{ then }\underline{w}\in[\leq q']^{|\underline{w}|}\right\}. \qquad (2.4.6)$$

Therefore, since $\limsup_{n\to\infty}\zeta(x^n)=0$,

$$\lim_{n\to\infty}\nu_n\left([\leq q']\right)=1\text{ where }\nu_k=\frac{1}{p_k}\sum_{i=j}^{p_k-1}\delta_{\sigma^i x^k},$$

which this contradicts Lemma 2.4.12, hence (CI) implies (UCS). □

2.5 A characterisation of SPR for countable Markov shifts

For countable Markov shifts, with no information about the entropy at infinity, a usually strong condition for interval dynamics like $\sup\phi<P(\phi)$ does not even guarantee existence of an equilibrium, let alone its mixing conditions. However if we control h_∞ and δ_∞, it is possible to show the system is SPR. In this case, if the potential is weakly Hölder, by Cyr and Sarig's results, there is a unique equilibrium state with exponentially mixing behaviours. In this section, we combine the types of ideas mentioned in previous sections of this chapter to show a characterising condition for SPR. If the boundary behaviours of the CMS is 'nice', we need only to focus on the ergodic averages on periodic orbits reflected by the following quantity, $\chi_{per}(\phi):=\sup\left\{\frac{1}{n}S_n\phi(x):\sigma^n(x)=x\right\}$, and check if $\chi_{per}(\phi)<P(\phi)$. The set of theorems we are going to prove is comparable to the list in [NS, Theorem A].

The first proof ingredient is to use entropy at infinity defined in Section 2.4 (see also an equivalent notion in [Buz]) to control the asymptotic number of excursions from the compact part of the alphabet, and by Theorem 2.4.7 we can look at the global behaviours of periodic points to ensure (CI). Then together they produce some pressure gap at infinity which is similar to the idea in [RV]. The next step to show uniform contraction implies SPR involves re-inducing arguments, and can be compared to the proof of [DT, Lemma 2.17, Theorem 7.14], where the entropy at infinity for the symbolic version of the finitely branched interval maps is zero. The main theorems are presented below.

Theorem 2.5.1. *Let (Σ, ϕ) be a topologically transitive CMS with the $\mathcal{F}$-property (see Definition 2.4.1 above), ϕ a potential of summable variations satisfying (CI) and assume that $\delta_{\phi,\infty} + h_\infty < 0$. Then ϕ is SPR.*

Theorem 2.5.2. *Let (Σ, σ, ϕ) be a topologically transitive CMS with the $\mathcal{F}$-property, ϕ a potential with summable variations such that $P(\phi) < +\infty$, and entropy at infinity $h_\infty = 0$. Then (UCS) holds if and only if* (2.2.3) *holds.*

Remark 2.5.3. *Theorem 2.5.1 implies that if $h_\infty < h_{top}$, which means that the measure of maximal entropy is SPR (see [Buz, Proposition 6.1], [ITV, Proposition 2.20]), then the equilibrium state for a potential ϕ with $\sup\phi - \inf\phi < h_{top} - h_\infty$ must also be SPR, since this automatically implies $\delta_{\phi,\infty} + h_\infty < P(\phi)$. This is shown to be the case for interval maps in [BT1]. There are various cases of systems which have a coding by a countable Markov shift and where it may, in the future, be proved that the measure of maximal entropy is SPR the above idea would then apply. For example we might expect the surface diffeomorphisms considered in [BCS1] to satisfy these conditions.*

Remark 2.5.4. *Also by [CS] for topologically mixing CMS, the set of SPR potentials is open and dense in the set of weakly Hölder potentials (with respect to a sequentially defined topology). Then our theorem implies that for CMS with $h_\infty = 0$, the set of UCS is also open and dense with respect to the same topology.*

It is also worth pointing out that in most literature concerning countable Markov shifts and SPR, the default assumption on the transition matrix is topologically mixing which is slightly stronger than topologically transitive. For our results, if the system is topologically transitive but not mixing, one can use *spectral decomposition* (see the paragraph before Proposition 2.5.7) to resolve the discrepancy between different assumptions. Let us first prove Theorem 2.5.1, for which an example with $\delta_\infty + h_\infty = P(\phi)$ will be provided in Section 2.6 so our condition is sharp.

For proofs below, we again assume without loss of generality that $\mathcal{A} = \mathbb{N}_0$.

Lemma 2.5.5. *Suppose (Σ, σ, ϕ) is a topologically transitive mixing CMS with the $\mathcal{F}$-property, ϕ has summable variations with $P(\phi) < \infty$. Then for all $\varepsilon > 0$, there exists $q \in \mathbb{N}$ and $K_q \geq 0$ such that for all positive n, if $x_0, x_n \leq q$ and $x_k > q$ for all $1 \leq k \leq n-1$, then*

$$S_n\phi(x) < K_q + n\left(\delta_{\phi,\infty} + \varepsilon\right).$$

Proof. Let $\varepsilon > 0$ be given. By definition of δ_∞ there exists q such that $\delta_{\phi,\infty}(q, M) < \delta_{\phi,\infty} + \frac{\varepsilon}{2}$ for all M large. Then there exists N_ε such that for all $n > N_\varepsilon$ if $x \in \Sigma$ is such that $x_0, x_n \leq q$, but $x_k > q$ for

$i = 1, \dots, n-1$, then

$$\frac{1}{n} S_n\phi(x) < \delta_{\phi,\infty}(q, M) + \frac{\varepsilon}{2} < \delta_{\phi,\infty} + \varepsilon.$$

Since the $\mathcal{F}$-property implies that for each n the number of words of length n which start and end at $[\leq q]$ are finite, also using summable variations,

$$K_q := \max\left\{\max_{n \leq N_\varepsilon} \sup\left\{S_n\phi(x) : x_0, x_n \leq q\right\}, 0\right\}$$

is finite and satisfies the lemma. □

Given $q \in \mathbb{N}$ as in the lemma, let $Y = [\leq q]$ and define $\tau_Y : Y \to \mathbb{N} \cup \{\infty\}$ by $\tau_Y(x) := \inf\{n \geq 1 : \sigma^n(x) \in Y\}$. Then let $F : Y \to Y$ be the first return map $F = \sigma^{\tau_Y}$. Let $\mathcal{C}_p^F$ be the set of p-cylinders with respect to (Y, F), so that $Z \in \mathcal{C}_1^F$ implies that $F(Z) = [a]$ for some $a \leq q$. The topological transitivity of the original system means that there is some $J \in \mathbb{N}$ such that for any $Z, Z' \in \mathcal{C}_1^F$ there is $j \leq J$ such that $Z' \subset F^j(Z)$, which is a stronger condition than the BIP property (see [Sar5]).

Define the corresponding induced potential $\hat{\phi}_Y = \sum_{i=0}^{\tau_Y - 1} \phi \circ \sigma^i$ and note that this has summable variations (in fact $\mathrm{var}_1 \hat{\phi}_Y < \infty$), so $B_{\hat{\phi}} < \infty$. By for example [Sar3, Theorem 2], $P(\hat{\phi}) \leq 0$, so setting $\overline{\phi} = \hat{\phi} - P(\hat{\phi})$ we have a potential of zero pressure, and there is an $\overline{\phi}$-conformal measure (see Remark 2.1.3 above) m_Y and an equivalent invariant Gibbs measure μ_Y, see [Sar5, Theorem 1]; also $B_{\overline{\phi}} = B_{\hat{\phi}} < \infty$. Note that if ϕ is recurrent then $\hat{\phi} = \overline{\phi}$. We also define $S_n^F \overline{\phi}_Y = \sum_{i=0}^{n-1} \overline{\phi}_Y \circ F^i$.

Lemma 2.5.6. *There is $C_1 > 0$ and $\varepsilon > 0$ such that if $Z \in \mathcal{C}_p^F$ and for some $p \geq 1$,*

$$\sum_{i=0}^{p-1} \tau_Y(F^i(Z)) = n,$$

then

$$m_Y(Z) \leq C_1^p \exp\left(n(\delta_{\phi,\infty} + \varepsilon)\right).$$

Proof. Writing $\tau_Y(F^i(Z)) = \tau_i$, Z is an $(n+1)$-cylinder with respect to σ of the form

$$[z_0, \dots, z_{\tau_1 - 1}, z_{\tau_1}, \dots, z_{\tau_p - 1}, z_{\tau_p}]$$

where $z_0, z_{\tau_i} \leq q$ for $i = 1, \dots, p$. By conformality and Lemma 2.5.5,

$$\begin{aligned} m_Y\left([z_0]\right) &= \int_{[z_0, \dots, z_{\tau_p - 1}]} e^{-S_n\phi(x) + pP(\hat{\phi})}\, dm_Y \\ &\geq m_Y([z_0, \dots, z_{\tau_p - 1}]) e^{-\sup_{x \in Z} S_n\phi(x) + pP(\hat{\phi})}. \end{aligned}$$

Hence

$$m_Y([z_0, \dots, z_{\tau_p - 1}]) \le \exp\left(p(K_q - P(\hat{\phi})) + n(\delta_{\phi,\infty} + \varepsilon)\right) m_Y([z_0]),$$

so setting $C_1 = e^{K_q - P(\hat{\phi})}$, we are finished. □

To conclude the proof of Theorem 2.5.1, one needs to combine the Lemmas above with re-inducing arguments.

Proof of Theorem 2.5.1. The proof is similar to that of [DT, Theorem 7.14]. As we will see, by (2.2.3) it suffices to show the inducing scheme on some 1-cylinder $[a]$, that is the first return map $([a], \sigma^{\varphi_a})$ to $[a]$, has an exponential tail.

Pick $\varepsilon > 0$ such that

$$\delta_{\phi,\infty} + h_\infty < -4\varepsilon,$$

choose q satisfying Lemma 2.5.5 and such that for all large M,

$$\delta_{\phi,\infty} + h_\infty(M, q) < -3\varepsilon. \tag{2.5.1}$$

So for (Y, F) as above, which must also satisfy Lemma 2.5.6, by topological transitivity there exists N such that for all $Z \in \mathcal{C}_1^F$,

$$Y \subset \bigcup_{j=1}^{N} F^j(Z). \tag{2.5.2}$$

Pick some 1-cylinder with respect to σ, $Y_0 = [a]$, with $m_Y(Y_0) > 0$ and let m_{Y_0} be the conformal conditional measure here.

Claim. *There is some uniform constant $\beta > 0$ such that for $Z_n \in \mathcal{C}_n^F$ and N as in* (2.5.2)*,*

$$\frac{m_Y\left(x \in Z_n : F^j(x) \notin Y_0, j = n, \dots, n+N-1\right)}{m_Y(Z_n)} < e^{-\beta}. \tag{2.5.3}$$

Proof of claim. By (2.5.2), for each $b \in \mathcal{A}$ such that $[b] \subset Y$ there is some cylinder (with respect to F) $A \subset [b]$ and $0 \le k(A) \le N-1$ such that $F^{k(A)}(A) = [a]$. Denote the (finite) collection of such cylinders by $\mathcal{B}$. In particular there is some $A \in \mathcal{B}$ such that $A \subset F^n(Z_n)$. Letting $A' = F^{-n}A \cap Z_n$, it suffices to find a lower bound for $\frac{m_Y(A')}{m_Y(Z_n)}$, independent of $Z_n \in \mathcal{C}_n^F$ and $A \in \mathcal{B}$.

Then similar to the proof of Lemma 2.4.8, $\min_{A \in \mathcal{B}} \inf_{x \in A} S_{k(A)}^F \overline{\phi}_Y(x)$ is bounded from below by the finiteness of $\mathcal{B}$ and summable variations .

By conformality of m_Y, for any $C \subseteq Y$, if $F^m : C \to F^m C$ is injective, $m_Y(F^m C) = \int_C \exp(-S_m^F \overline{\phi}_Y) dm_Y$, hence

$$\begin{aligned}\frac{m_Y(A')}{m_Y(Z_n)} &\geq \frac{m_Y(A)}{m_Y(F^n Z_n)} \exp\left(-\sup_{Z_n} S_n^F\,\overline{\phi}_Y + \inf_{A'} S_n^F\,\overline{\phi}_Y\right)\\ &\geq \frac{m_Y([a])}{m_Y(Y)} e^{-B(\overline{\phi}_Y)} e^{\inf_{x\in A} S_{k(A)}^F\,\overline{\phi}_Y(x)} > 0,\end{aligned}$$

uniformly, as required. □

Claim. *For each $k \geq 1$,*

$$\frac{m_Y\left(x \in Y_0 : F^j(x) \notin Y_0, j = 1, \dots, kN\right)}{m_Y(Y_0)} < e^{-k\beta}. \tag{2.5.4}$$

Proof of Claim. This claim is proved by induction. As $Y_0 = [a]$ can be written as a union of 1-cylinders with respect to F, (2.5.3) and the fact that for all positive numbers a, b, c, d, $\frac{a+c}{b+d} \leq \max\left\{\frac{a}{b}, \frac{c}{d}\right\}$, together implies

$$\frac{m_Y\left(x \in Y_0 : F^j(x) \notin Y_0, j = 1, \dots, N\right)}{m_Y(Y_0)} < e^{-\beta}.$$

Assume inductively that for each $i \geq 1$,

$$\frac{m_Y\left(x \in Y_0 : F^j(x) \notin Y_0, j = 1, \dots, iN\right)}{m_Y(Y_0)} < e^{-i\beta}.$$

Defining the set

$$\mathcal{Z}_i := \left\{Z \in \mathcal{C}_{iN+1}^F : Z \subset Y_0,\ F^j(Z) \notin Y_0,\ j = 1, \dots, iN\right\},$$

by (2.5.3) and the inequality above:

$$\begin{aligned}&\frac{m_Y\left(x \in Y_0 : F^j(x) \notin Y_0, j = 1, \dots, (i+1)N\right)}{m_Y(Y_0)}\\ &= \frac{1}{m_Y(Y_0)} \sum_{Z \in \mathcal{Z}_i} m_Y(Z) \frac{m_Y\left(x \in Z : F^j(x) \notin Y_0, j = iN+1, \dots, (i+1)N\right)}{m_Y(Z_{i(N+1)})}\\ &< \frac{1}{m_Y(Y_0)} \sum_{Z \in \mathcal{Z}_i} m_Y(Z) e^{-\beta} \leq e^{-\beta} \frac{m_Y\left(x \in Y_0 : F^j(x) \notin Y_0,\ j = 1, \dots, iN\right)}{m_Y(Y_0)}\\ &< e^{-(i+1)\beta}.\end{aligned}$$

□

Letting $T = \gamma n$ for some $\gamma \in (0, 1)$ to be determined later, we can split the set $\{x \in Y_0 : \varphi_a(x) = n\}$

depending on whether x visits Y more or less than T times in its first n symbols, which can be written

$$\begin{aligned} m_{Y_0}\left(\{\varphi_a = n\}\right) \leq\ & m_{Y_0}\left(\left\{\varphi_a(x) = n, \sum_{j=0}^{T} \tau_Y(F^j(x)) > n\right\}\right) \\ & + m_{Y_0}\left(\left\{\varphi_a(x) = n, \sum_{j=0}^{T} \tau_Y(F^j(x)) \leq n\right\}\right) \\ =:\ & I + II. \end{aligned}$$

By (2.5.4), $I \leq \sum_{p=T}^{n} \exp\left(-\frac{p}{N}\beta\right) \leq C_2 \exp\left(-\frac{T}{N}\beta\right)$, for some $C_2 \in \mathbb{R}$. The number of n-cylinders with respect to σ which spend a proportion $\gamma \leq 1/M$ of their σ-iterates up to n in Y is no more than $\#B(n, M, q)$. Moreover, for all large n, $\#B(n, M, q) \leq C_3 e^{n(h_\infty(M,q)+\varepsilon)}$ for some $C_3 > 0$, so combining this with Lemma 2.5.6 and (2.5.1) we get

$$\begin{aligned} II &\leq C_1^T \exp\left(n(\delta_{\phi,\infty} + \varepsilon)\right) \#B(n, M, q) \\ &\leq C_1^T C_3 \exp\left(n(\delta_{\phi,\infty} + h_\infty(M, q) + 2\varepsilon)\right) \leq C_1^T C_3 \exp\left(-n\varepsilon\right). \end{aligned}$$

Then choosing $\gamma = \min\left\{\frac{1}{M}, \frac{\varepsilon}{2\log C_1}\right\}$, both I and II are exponentially small so that

$$\limsup_{n\to\infty} \frac{1}{n} \log m_{Y_0}(\{\varphi_a = n\}) < 0.$$

As m_{Y_0} is conformal,

$$m_{Y_0}(\{\varphi_a = n\}) \asymp \sum_{\sigma^n x = x, \varphi_a(x) = n} e^{S_n\phi(x) - i(x)P(\hat{\phi})} \geq Z_n^*(\phi, a),$$

where $i(x)$ corresponds to the number of hits to Y before Y_0. Hence (2.2.3) holds and the system is strong positive recurrent. □

Theorem 2.5.1 means that (UCS) implies SPR. For the other direction of Theorem 2.5.2, it suffices to prove the statement under topological mixing since we can use *spectral decomposition*, a tool to reduce arguments on topologically transitive to topologically mixing.

Briefly speaking, if (Σ, σ) is a topologically transitive CMS but not topologically mixing, by Proposition 1.1.4 there exists

$$p := \gcd\{n : \exists a \in \mathcal{A}, x \in [a] \text{ s.t. } \varphi_a(x) = n\} > 1,$$

called the *period of* Σ. The alphabet is divided into p equivalence classes $\{\mathcal{A}_1 \ldots, \mathcal{A}_{p-1}\}$ and $\Sigma = \biguplus_{i=0}^{p-1} \Sigma_i$, $\Sigma_i = \{x \in \Sigma : x_0 \in \mathcal{A}_i\}$. Then $(\Sigma_i, \phi_p, \sigma^p)$ is conjugate to a topological mixing CMS whose alphabet is given by the first return words for $a \in \mathcal{A}_i$, and most statements (especially those in this chapter) proved for $(\Sigma_i, \phi_p, \sigma^p)$ remain valid for the original CMS. For more detailed discussion, see for example [RS, §2.2,§6].

Proposition 2.5.7. *Under the assumptions of Theorem 2.5.2* i.e.*, topologically transitive, $\mathcal{F}$ property holds and $h_\infty = 0$, SPR implies (UCS).*

Proof. By SPR there is $a \in \mathcal{A}$ such that $\Delta_a[\phi] > 0$. First by [Sar3, Lemma 3], $P(\phi) = 0$ implies the induced pressure on $[a]$, $P(\overline{\phi})$, is zero, and SPR implies that there exists $\varepsilon_a > 0$ such that

$$P(\overline{\phi + 2\varepsilon_a}) < \infty. \tag{2.5.5}$$

Moreover, as in (2.2.3), there exists $N_a \in \mathbb{N}$ such that for all $n > N_a$, all x such that $\varphi_a(x) = n$,

$$\frac{1}{n} S_n\phi(x) < -\varepsilon_a. \tag{2.5.6}$$

Suppose by contradiction that $\chi_{per}(\phi) = \sup_{n\geq 1} \sup \{S_n\phi(x) : \sigma^n x = x\} = 0$; take ϕ' as in Lemma 2.4.11 and $\chi_{per}(\phi') = \chi_{per}(\phi)$. Then similar to the proof of Lemma 2.4.12 there exists a sequence of periodic points $x^1, x^2, \ldots$, with periods $p_1, p_2, \ldots$ and Birkhoff averages

$$s_n = \frac{1}{p_n} S_{p_n}\phi(x^n) = \frac{1}{p_n} S_{p_n}\phi'(x^n) > -\varepsilon_a \text{ and } \lim_{n\to\infty} s_n = 0. \tag{2.5.7}$$

Case 1. Suppose there exists $x \in \{x^1, x^2 \ldots\}$ such that $\forall k \geq 0$, $x_k \neq a$. Then as in (2.4.1), by topological transitivity, there are words $\underline{v}$, $\underline{w}$ of length $\ell_1 = \ell(a, x_0)$ and $\ell_2 = \ell(x_{n-1}, a)$ respectively such that $v_0 = a$, $\underline{v}x_0 \in \Sigma_{|\underline{v}|+1}$, $x_{n-1}\underline{w} \in \Sigma_{|\underline{w}|+1}$, $\underline{w}a \in \Sigma_{|\underline{w}|+1}$, hence

$$\underline{v}x_0, \ldots, x_{n-1}\underline{w} \in \Sigma_{\ell_1+n+\ell_2}.$$

Moreover, for each $k \in \mathbb{N}$ and $n_k = kn + \ell_1 + \ell_2$ there is a periodic point $y(k) \in [a]$ with $\varphi_a(y(k)) = n_k$ of the form:

$$y(k) = \left(\overline{\underline{v}\,(x_0, \ldots, x_{n-1})^k\, \underline{w}}\right)$$

where $(x_0, \ldots, x_{n-1})^k$ means the string is repeated k times. By summable variations, there exists a constant $C > 0$ such that for all k,

$$Z^*_{n_k}(\phi, a) \geq \exp\left(S_{n_k}\phi(y(k))\right) \geq \exp\left(C - kn\varepsilon_a\right).$$

Then as in [Sar3, (5)], $\left|P(\overline{\phi + p}) - \log \sum_{k\geq 1} e^{kp} Z^*_k(\phi, a)\right| \leq B_\phi$, therefore,

$$\infty = \log \sum_{k=1}^{\infty} e^{n_k\varepsilon_a + C} e^{S_{n_k}\phi(y(k))} \leq C + \log \sum_{n=1}^{\infty} e^{n\varepsilon_a} Z^*_n(\phi, a) \leq P(\overline{\phi + \varepsilon_a}) + B_\phi + C,$$

which is a contradiction to (2.5.5) since $B_\phi < \infty$.

Case 2. Now suppose all $x \in \{x^1, x^2, \dots\}$ contain state a. Without loss of generality one can suppose $x_0^i = a$ for all i by periodicity. Recall from (2.2.1) that the induced alphabet $\mathcal{A}_a \subset \Sigma^*$ consists of cylinders of the form $[\underline{w}]$ where $w_i = a$ if and only if $i = 0$, and moreover $\underline{w}a \in \Sigma_{|\underline{w}|+1}$, *i.e.*, each $\underline{w}$ is a *first return word* to a.

For all n,

$$x^n = (\overline{\underline{w}_0 \dots \underline{w}_{k_n-1}}) \text{ for some } k_n \geq 1,\ \underline{w}_i \in \mathcal{A}_a \text{ and } \sum_{j=0}^{k_n-1} |\underline{w}_j| = p_n;$$

that is, each x^n can be decomposed into several first return words.

As $S_m\phi(x) = S_m\phi'(x)$ for any periodic point with period m, non-positivity of ϕ' and (2.5.6) imply that for all first return words $\underline{w}$ with length longer than N_a,

$$\sup_{x \in [\underline{w}]} S_{|\underline{w}|}\phi'(x) < -|\underline{w}|\varepsilon_a.$$

Letting $\mathcal{A}_{a,>k} := \{\underline{w} \in \mathcal{A}_a : |\underline{w}| > k\}$, re-define the proportion function similarly to (2.4.4),

$$\zeta : \{x^1, \dots\} \to [0,1], \quad \zeta(x^n) = \frac{1}{p_n} \sideset{}{'}\sum_{\underline{w} \in \left\{x_0^n \dots x_{p_n-1}^n \cap \mathcal{A}_{a,>N_a}\right\}} |\underline{w}|,$$

where Σ' again means that we count with multiplicity. Then repeating (2.4.5) with $\varepsilon = \varepsilon_a$ this definition ensures $\lim_{n\to\infty} \zeta(x^n) = 0$ since otherwise we contradict the property (2.5.7) of our periodic points. By the $\mathcal{F}$-property, we can define the function q_a by

$$q_a(N) := \min\left\{q \in \mathbb{N} : \text{if } \underline{w} \in \mathcal{A}_{a,\leq N}, \text{ then } w_i \leq q \text{ for } i = 0, \dots, |\underline{w}| - 1\right\}.$$

The sequence of probability measures

$$\nu_n = \frac{1}{p_k} \sum_{j=0}^{p_n-1} \delta_{\sigma^j x^n}$$

satisfies $\lim_n \nu_n([\leq q_a(N_a)]) = 1$. But since $\lim_{k\to\infty} s_{n_k} = 0$, we have a contradiction to Lemma 2.4.12, hence such sequence of periodic points does not exist. □

This concludes the proof of Theorem 2.5.2.

2.6 Bouquet examples

The conditions for Theorem 2.5.1,2.5.2 are weak, and so our results are applicable to a wide range of CMS. In this section, a special type of CMS is constructed for which our theory applies as well as exhibiting edge cases to demonstrate the sharpness of our results.

Our examples take the form of 'bouquet' Markov graphs, see [Ru1, Example 2.9][Ru2], some of them which inspired this section come from codings for dynamical systems, particularly in the case of interval maps $f : [0,1] \to [0,1]$. For some subset $Y \subset [0,1]$ with return time function $\varphi = \varphi_Y : Y \to \mathbb{N} \cup \{\infty\}$, the inducing scheme $F = f^{\varphi}$ defines a Markov map on Y, i.e., there is a partition $\{Y_i\}_i$ such that $\varphi|_{Y_i}$ is some constant φ_i and $F(Y_i)$ is a collection of elements of this partition. We can associate bouquet Markov graphs with shift dynamics, a potential $\phi : I \to [-\infty, \infty]$ is then lifted to the symbolic model. For example, such a coding can be done for general multimodal maps of the interval, as shown in, for example, [BT2, Theorem 3], or more classical and specific inducing schemes like those given in [BLS] (which include Collet-Eckmann maps).

Bouquet setup

Following [Ru1, Ru2], let $a : \mathbb{N} \to \mathbb{N}_0$ with $a(1) = 1$. We define our set of vertices as

$$V := \{r\} \cup \bigcup_{n=1}^{\infty} \left\{ v_k^{n,i} : 1 \le i \le a(n),\ 1 \le k \le n-1 \right\},$$

where all vertices with distinct labels above are distinct vertices. We call r the *root*. For notational convenience write $v_0^{n,i} = v_n^{n,i} = r$. Then the only allowed transitions in our Markov graph are $v_k^{n,i} \mapsto v_{k+1}^{n,i}$ for $0 \le k \le n-1$. This defines a *bouquet* of loops: with $a(n)$ disjoint *simple* loops (from r back to r) of length n. The resulting shift space which we refer to as a *bouquet shift* is $\Sigma = \Sigma_V$: it has $a(n)$ periodic cycles of period n. The topological entropy of Σ_V is given by the formula

$$h_{top} = \limsup_{n\to\infty} \frac{1}{n} \log p(n),$$

where $p(n)$ is the number of length n loops starting and ending at the root, and $\limsup$ can be replaced with $\lim$ since the shift is mixing.

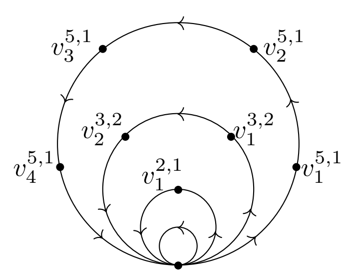

Figure 2.2: Case $a(n) = 1$ for each prime number n, and 0 otherwise. Picture credit to M. Todd.

Below we will make various choices of $(a(n))_n$ and potentials $\phi : \Sigma_V \to \mathbb{R}$. Our analysis will be via first returns to $[r]$. Note that [Ru1, Ru2] were concerned with measures of maximal entropy (in which case we set $\phi \equiv -h_{top}(\sigma)$ so that $P(\phi) = 0$), rather than the more general setting of equilibrium states that we are interested in here.

Lemma 2.6.1. $h_\infty = \limsup_{n\to\infty} \frac{1}{n} \log a(n)$.

Proof. Suppose that $\limsup_{n\to\infty} \frac{1}{n} \log a(n) = \log \lambda$, which we may assume is finite, as otherwise the conclusion is immediate since we can prove (in the next paragraph) $h_\infty \geq \log \lambda$. Then for $\varepsilon > 0$ there is $C > 0$ such that for an infinite sequence of $n_k \in \mathbb{N}$,

$$\frac{1}{C}\lambda^{n_k(1-\varepsilon)} \leq a(n_k) \leq C\lambda^{n_k(1+\varepsilon)},$$

and indeed the upper bound holds for all $n_k \in \mathbb{N}$.

We first show that $h_\infty \geq \limsup_{n\to\infty} \frac{1}{n} \log a(n)$. Fix $M, q \in \mathbb{N}$. If n is large enough so that $(n+1)/M \geq 1$, then any of the simple loops of length n that only intersects $[\leq q]$ at the root r is in $B(n, M, q)$. In other words, $z_n(M, q) \geq a(n) \geq \frac{1}{C}\lambda^{n(1-\varepsilon)}$. Taking logs then divide by n we find that $h(M, q) \geq \log \lambda$, hence the claimed lower bound holds.

For the upper bound, fix M, q and notice that for each n, if $[x_0, \ldots, x_{n-1}] \in B(n, M, q)$ then it contains no more than n/M disjoint simple loops. In other words, it visits the root r at most n/M times. Then, as for all $M > 2$ there is $\binom{n}{k} \leq \binom{n}{n/M}$ for all $k \leq n/M$,

$$\begin{aligned}\#B(n, M, q) &\leq \sum_{\substack{i_1+\cdots+i_k=n \\ k\leq n/M}} a(i_1)\ldots a(i_k) \leq \sum_{k=1}^{n/M} \binom{n}{k} C^k \lambda^{n(1+\varepsilon)} \\ &\leq C^{n/M}\frac{n}{M}\lambda^{n(1+\varepsilon)}\binom{n}{n/M} \leq C^{n/M}\frac{n}{M}\lambda^{n(1+\varepsilon)}\left(\frac{n\cdot e}{n/M}\right)^{n/M}.\end{aligned}$$

Therefore, $h_\infty(M, q) \leq \frac{1}{M}(\log C + \log M + 1) + (1+\varepsilon)\log\lambda$. This upper bound is independent of q, hence $h_\infty = \lim_{M\to\infty} h_\infty(M, q) \leq (1+\varepsilon)\log\lambda$. As $\varepsilon > 0$ was arbitrary, $h_\infty \leq \log\lambda$ as claimed.

□

2.6.1 UCS is a weak condition

Here we will use a simple set of examples to compare (UCS) with other conditions of this type.

Set $a(n) = 1$ for all n; notice that since the number of compositions for each number $n \in \mathbb{N}$ is 2^{n-1}, the topological entropy of this bouquet system is $\log 2$ and by [BBG, Theorem 6.4] Σ_V is almost isomorphic to the renewal shift defined in Example 2.2.1. Set $\phi|_{[rv_1^{n,1}]} = -n\log 2$ and $\phi = 0$ otherwise. Note that this potential is *Markov*, in the sense that for all x, $\phi(x) = \phi(x_0, x_1)$ so the induced potential (with respect to $[r]$) $\overline{\phi} : [r] \to \mathbb{R}$ is Bernoulli *i.e.*, $\overline{\phi}(\overline{x}) = \overline{\phi}([\underline{x}_0])$, and takes the value $-n \log 2$ on the vertex corresponding to the loop of length n. Then

$$P(\overline{\phi}) = \log \left(\sum_{n \geq 1} \frac{1}{2^n} \right) = 0.$$

Since, moreover, $\sum_{n \geq 1} \frac{n}{2^n} < \infty$, ϕ is positive recurrent and has $P(\phi) = 0$.

This system (Σ_V, σ, ϕ) clearly satisfies (UCS) since for the periodic point x_n of period n, $\frac{1}{n} S_n \phi(x_n) = -\log 2$. On the other hand, the hyperbolicity condition as in [IR, LiRiv1] fails since for any n, there is a point $y_n \in [v_1^{n,1} v_2^{n,1} \cdots v_{n-1}^{n,1}]$ such that $S_n \phi(y_n) = 0$. Finally, regarding the conditions of [LSV], this would require $\sum_{C \in \mathcal{C}_1} \sup_{x \in C} e^{\phi(x)} < \infty$ as well as a condition like hyperbolicity to hold, both of which fail here.

Remark 2.6.2. *One can modify ϕ to be uniformly bounded,* e.g. *putting weight $-2 \log 2$ on $n/2$ of the vertices in the loop of length n (suitably adjusting for when n is odd).*

2.6.2 An example showing the sharpness of Theorem 2.5.1

Here we give a class of examples where (UCS) holds, but $\delta_{\phi,\infty} + h_\infty = 0$ and (2.2.3) fails, so that the condition $\delta_{\phi,\infty} + h_\infty < 0$ in Theorem 2.5.1 is necessary.

Let $a(n) = 2^n$ and $C, \beta > 0$ to be chosen later. Now define $\phi|_{[rr]} = \log C$, $\phi|_{[rv_1^{n,i}]} = \log C - n \log 2 - \beta \log n$ and $\phi = 0$ otherwise (as in Remark 2.6.2 we could also spread this potential out if desired).

First observe that $Z_n(\phi, [r]) \geq C 2^n 2^{-n} n^{-\beta}$, so $P(\phi) \geq 0$.

Taking the first return map to $[r]$ the induced potential $\overline{\phi}$ corresponding to loops of length n takes the value $\log C - n \log 2 - \beta \log n$. Then

$$P(\overline{\phi}) = \log \left(C \sum_n a(n) e^{-n \log 2 - \beta \log n} \right) = \log \left(C \zeta(\beta) \right)$$

where ζ denotes the Riemann zeta function $\zeta(s) = \sum_{n \geq 1} n^{-s}$. We use the ideas of Hofbauer and Keller presented in [IT, Section 4.1], generalised to this setting (see also the ideas of [Ru1, Table 1]).

(a) If $\beta > 1$ and we choose $C = 1/\zeta(\beta)$ then the pressure of the induced system is zero, ϕ is recurrent and $P(\phi) = 0$.

(b) If $\beta > 1$ and $C > 1/\zeta(\beta)$, or $\beta \in (0, 1)$, then the pressure of the induced system is positive and this is not interesting for our purposes (note this would imply $P(\phi) > 0$).

(c) If $\beta > 1$ and $C < 1/\zeta(\beta)$ then ϕ is transient and $P(\phi) = 0$.

We will now assume that we are in case (a).

Since

$$C\sum_n na(n)e^{-n\log 2-\beta\log n}=C\sum_n n^{1-\beta},$$

the system is positive recurrent, and we have an equilibrium state μ_ϕ here, if $\beta>2$ (if $\beta\in(1,2]$ then ϕ is null recurrent); moreover there is a conformal measure m_ϕ. Since $h_{top}(\sigma)$ must satisfy

$$1=\sum_n a(n)e^{-nh_{top}(\sigma)}=\sum_n 2^n e^{-nh_{top}(\sigma)},$$

we see that $h_{top}(\sigma)=\log 4$.

The fact that $h_\infty=\log 2$ follows from Lemma 2.6.1. We next show that $\delta_{\phi,\infty}=-\log 2$. That $z_{n,\phi}(M,q)\geq -\log 2+\frac{1}{n}(\log C+\beta\log n)$ for $n+1>M$ and $n>n_q$ is immediate from the definition, so $z_{n,\phi}(M,q)\geq -\log 2$. For the upper bound, the proof is similar to, though simpler than, that of Lemma 2.6.1: if we consider $v\in\mathcal{G}(q)$ as defined there, then for $x\in[v]$, $\frac{1}{|v|}S_{|v|}\phi(x)\leq -\log 2$ and since the finite behaviour contributed by any prefixes and suffixes disappears in the limit, $\delta_{\phi,\infty}=-\log 2$ and so $\delta_{\phi,\infty}+h_\infty=0$.

We see here that $Z_n^*(\phi,r)=C/n^\beta$ so SPR fails. Hence Theorem 2.5.1 is sharp in the sense that we can satisfy (2.2.3), but if $\delta_{\phi,\infty}+h_\infty<0$ does not hold then (2.2.3) may fail. Note also that if ϕ was null recurrent or, as in case (c) above, transient, we would also fail these conditions in a more dramatic way.

2.6.3 Relation of bouquets to inducing schemes and general shifts

At the beginning of this section we described interval maps (I,f) with an inducing scheme $(Y,\tau,F=f^\tau)$ such that Y is a countably infinite union of disjoint subsets $\bigcup_i Y_i$ and $\tau|_{Y_i}$ is constant. If we have $F(Y_i)=Y$ for all i, which is the case for the examples mentioned above, then we identify Y with the root r. Suppose $a(n)=\#\{Y_i\subset Y:\tau|_{Y_i}=n\}$, then for each such Y_i in the set we associate a loop $r\mapsto v_1^{n,i_j}\mapsto v_2^{n,i_j}\mapsto\cdots\mapsto v_{n-1}^{n,i_j}\mapsto r$, for $i_j=1,\ldots,a(n)$.

We can project a sequence $(x_0,x_1,\ldots)\in\Sigma$ to $x\in I$ by a projection π as follows. Suppose that $x\in Y$ has $F^\ell(x)\in[rv_1^{n_\ell,i_\ell}]$ for all $\ell\geq 0$ for some n_ℓ,i_ℓ. Then there will be a corresponding sequence $(x_0,x_1,\ldots)\in\Sigma$ given by $(r,v_1^{n_0,i_0},v_2^{n_0,i_0},\ldots,v_{n_0-1}^{n_0,i_0},r,v_1^{n_1,i_1},\ldots)$. So let $\pi(x_0,x_1,\ldots)=x$ here, if $x_0=v_k^{n,i_j}$ for $k>1$ then consider $y\in Y$ the projection of the sequence $(r,v_1^{n,i_j},\ldots,v_{k-1}^{n,i_j},x_0,x_1,\ldots)$ and let $\pi(x_0,x_1,\ldots)=f^k(y)$.

If $\phi:I\to[-\infty,\infty]$ is a potential, then this lifts to a potential on the bouquet shift $\phi\circ\pi$. The regularity of the lifted potential depends on the regularity of the original one and the choice of inducing scheme. For some specific cases of multimodal maps where $\phi = -\log|Df|$ and there is an inducing scheme so that ϕ lifts to a potential of summable variation, see for example [BLS, Proposition 4.1] which considers multimodal maps with different rates of growth of derivative along critical orbits. In this case Collet-Eckmann maps yield symbolic models satisfying (UCS) along with our other equivalent properties, while non-Collet-Eckmann maps fail all of these.

We can extend a version of the coding used above to any topologically transitive CMS (Σ, σ): we can pick a 1-cylinder and take first returns to it and then use the induced system to recode the system via a bouquet with the root being the 1-cylinder selected. Hence the bouquet setup captures the behaviour of *any* topologically transitive CMS.

Chapter 3

Almost sure limit theorems for cover times

In this section we put the thermodynamics of CMS aside and focus on an almost sure convergence problem for interval maps. Let X be a compact interval and for some closed subset $\Lambda \subseteq X$, suppose $f : \Lambda \to \Lambda$ is topologically transitive. By transitivity, for every point x in the repeller (see (3.2.1) for definition) that is not a preimage of some periodic point, its orbit in the long-term will saturate the repeller. It is then sensible to ask the following question: given $r > 0$ small, what is the time/number of iterates needed for the orbit of x under f reaching a resolution of r? In other words, we care about the following quantity, to which we refer to as the *r-cover time of x*:

$$\tau_r(x) = \tau(x) := \inf \left\{ k \geq 0 : \text{ for all } y \in \Lambda, \text{ there exists } j \leq k \text{ s.t. } d\left(f^j(x), y\right) < r \right\}.$$

The name *cover* comes from the trivial observation that $\left\{ B\left(f^j(x), r\right) \right\}_{j=0}^{\tau_r(x)}$ forms an r-cover of Λ. Cover times were also studied for strong Markov processes, in such context they are interpreted as the minimum time for a process $\{X_n\}_n$ to have visited all of a finite subset in the state space. We start our discussion with a quick review on results for expected cover times for Brownian motions and interval maps, then move on to the almost sure convergence rate of asymptotic cover times.

3.1 Expected cover times

An important quantitative result for cover times in stochastic process context was obtained by Matthews [Mat] for Brownian motion on $\Sigma^d := \left\{ \text{surface of the unit sphere } S^d \subset \mathbb{R}^d \right\}$. In particular, the author considered two separable quantities: $C_1(r,d)$ the time taken for the geodesic balls of radius r, which the author referred to as 'caps', around the points of the Brownian motion to cover S^d, and $C_2(r,d)$ similarly defined but considering also the reflection of the points about the origin, hence there is $\mathbb{E}[C_1(r,d)] = 2\mathbb{E}[C_2(r,d)]$. For p the scale parameter of Brownian motion, and all $d \geq 4$, the expectations of $C_1(r,d)$ and $C_2(r,d)$ are sharp estimates were given, for the special case $d=3$,

$$\frac{4}{p} \leq \liminf_{r\to 0} \frac{\mathbb{E}[C_1(r,d)]}{\log\log(r^{-1})} \leq \limsup_{r\to 0} \frac{\mathbb{E}[C_1(r,d)]}{\log\log(r^{-1})} \leq \frac{16}{p},$$

and the upper and lower bounds are halved for $\mathbb{E}[C_2(r,d)]$.

The proof consists of two steps: (1) calculating the expected time $\mathbb{E}[\tau_{\mathcal{C}^d}]$ for a strongly Markov process to visit all components in a finite collection $\mathcal{C}$ of subsets in the state space, (2) and approximate the number of balls of diameter r needed to cover Σ^d. In particular, the first part involves what the author refers to as an 'auxiliary randomisation', which assigns a product measure that returns the probability of 'first hitting' each set in such a collection $\mathcal{C}$ in a particular order, and using conditional expectation to calculate $\mathbb{E}[\tau_{\mathcal{C}}]$. The expected cover time for $\mathcal{C}^d$ turns out only to depend on the logarithmic order of $\#\mathcal{C}^d$, hence the approximation in step (2) need not to be too accurate.

Such a method is transferable to the cover time problem for interval maps on their repellers, and chaos games on the attractors of iterated function systems (IFS). An iterated function system is a collection $\mathcal{F} = \{f_i\}_{i\in\mathcal{I}}$of contraction maps on a subset of $\mathbb{R}$, and it is known that there always exists a unique non-empty compact set K such that $K = \bigcup_{i\in\mathcal{I}} f_i(K)$. Then the chaos game, a process first described by [Bar], refers to applying maps from $\{f_i\}_{i\in\mathcal{I}}$ to an $x \in K$ repeatedly on the left: let $\omega = (i_0, i_1, \dots) \in \mathcal{I}^{\mathbb{N}_0}$, the orbit of x is defined by $\mathcal{O}(x) = \left\{f_{i_{n-1}} \circ \cdots \circ f_{i_0}(x)\right\}_{n\in\mathbb{N}}$, this allows us to talk about cover times.

In these settings, although the process is no longer a Brownian motion hence not necessarily independent at each iteration, the lack of independence can be resolved by fast mixing properties. For cover times in chaos games that has the ψ-mixing property, the upper and lower bounds were given for expected cover times in [BJK, Theorem 2.2]. For interval dynamics, [JT] considered cover times for the same class of maps as in [BDT] whose transfer operators have nice properties. Their main results showed that there exists $\varepsilon > 0$ such that for all r small,

$$\frac{1}{M_\mu(r/\varepsilon)} \preceq \mathbb{E}_\mu[\tau_r] \preceq \frac{-\log r}{M_\mu(\varepsilon r)}, \tag{3.1.1}$$

where μ is an f-invariant measure and $M_\mu(r) = \min_{x\in supp(\mu)} \mu\left(B(x,r)\right)$. $M_\mu(r)$ is also used to define the following quantities.

Definition 3.1.1. *The* upper and lower Minkowski dimensions *of μ are defined respectively by*

$$\overline{\dim}_M(\mu) := \limsup_{r\to 0} \frac{\log M_\mu(r)}{\log r}, \quad \underline{\dim}_M(\mu) := \liminf_{r\to 0} \frac{\log M_\mu(r)}{\log r}, \tag{3.1.2}$$

and write $\dim_M(\mu)$ *when they coincide.*

These dimension-like quantities reflect the decay rate of the minimal μ-measure ball at scale r, and they are closely related to the box-counting dimension of the ambient space (see [FFK] for more details). Alternatively, in the language of L^q dimensions for measures, Minkowski dimensions are the $L^{-\infty}$ dimensions. We are interested in the Minkowski dimensions of μ because they govern the asymptotic behaviour of hitting times associated to the balls which are most 'unlikely' to be visited at small scales. If $\dim_M(\mu) < \infty$, then (3.1.1) can be re-written in terms of $\dim_M(\mu)$; if μ is Ahlfors regular, *i.e.*, there exists $s_f > 0$ such that for all $r > 0$ and $x \in \operatorname{supp}(\mu)$, $\mu(B(x,r))$ is comparable to r^{s_f}, so (3.1.1) can be rewritten in terms of s_f.

3.2 Results on almost sure cover times

In addition to expectational results on τ_r, an almost sure asymptotic limit law was proved in [BJK], which established a connection between τ_r and $\dim_M(\mu)$. Similar statements can be obtained as well for *piecewise expanding Markov maps* described below.

Let $\mathcal{A}$ be a finite or countable index set, and $\mathcal{P} = \{P_a\}_{a\in\mathcal{A}}$ a collection of subintervals in $[0,1]$ with disjoint interiors in $[0,1]$. We say $f : \cup_{a\in\mathcal{A}} P_a \to [0,1]$ is a *Markov map* if for any $a \in \mathcal{A}$, $f([a])$ is a union of elements in $\mathcal{P} = \{P_a\}_{a\in\mathcal{A}}$ and $f_a := f|_{P_a}$ is injective, continuous and monotone. The map f is further said to be *piecewise expanding* if there is a uniform constant $\gamma > 1$ such that for all $a \in \mathcal{A}$, $|Df_a| \geq \gamma$.

The *repeller* of f, denoted by Λ, is the collection of points with all their forward iterates contained in $\mathcal{P}$, namely

$$\Lambda := \left\{ x \in \mathcal{X} : f^k(x) \in \bigcup_{a\in\mathcal{A}} P_a \text{ for all } k \geq 0 \right\}. \tag{3.2.1}$$

We study the dynamics of $f : \Lambda \to \Lambda$, together with an ergodic invariant measure μ supported on Λ. There is a shift system associated to f: let M be an $\mathcal{A} \times \mathcal{A}$ matrix such that $M_{ab} = 1$ if and only if $f(P_a) \cap P_b \neq \emptyset$, and 0 otherwise. The map f is *topologically transitive* if for all $a, b \in \mathcal{A}$, there exists k such that $M^k_{ab} > 0$. Let Σ denote the space of all *infinite allowable words*, *i.e.*,

$$\Sigma := \left\{ x = (x_0, x_1, \dots) \in \mathcal{A}^{\mathbb{N}_0} : M_{x_k, x_{k+1}} = 1, \forall\, k \geq 0 \right\}.$$

Define the projection map $\pi : \Sigma \to \Lambda$ by

$$x = \pi(x_0, x_1, \dots) \text{ if and only if } x \in \bigcap_{i=0}^{\infty} f^{-i} P_{x_i}.$$

The dynamics on Σ is the left shift $\sigma : \Sigma \to \Sigma$ given by $\sigma(x_0, x_1, \dots,) = (x_1, x_2, \dots)$, and π defines a semi-conjugacy $f \circ \pi = \pi \circ \sigma$. The corresponding symbolic measure $\tilde{\mu}$ of μ exists due to the Markov structure of f and is given by $\mu = \pi_* \tilde{\mu}$, *i.e.*, for all Borel-measurable sets $B \in \mathcal{B}([0,1])$, $\mu(B) = \tilde{\mu}\left(\pi^{-1} B\right)$.

For two partitions $\mathcal{P}$ and $\mathcal{Q}$, we define $\mathcal{P} \vee \mathcal{Q} := \{P \cap Q : P \in \mathcal{P} \text{ and } Q \in \mathcal{Q}\}$. Set $\mathcal{P}^n := \bigvee_{j=0}^{n-1} f^{-j} \mathcal{P}$, then each $P \in \mathcal{P}^n$ corresponds to an n*-cylinder* in Σ: for any $\underline{w} \in \Sigma^*$ of length n,

$$\pi[w_0, w_1, \dots, w_{n-1}] = \bigcap_{j=0}^{n-1} f^{-j} P_{w_j} =: P_{\underline{w}}.$$

We say the measure μ is *exponentially ψ-mixing* if $\tilde{\mu}$ is ψ-mixing (see Definition 1.1.5) with

$$\psi(n) \leq C_1 e^{-\rho n} \tag{3.2.2}$$

for some $C_1, \rho > 0$.

Given the setting above, the first theorem is presented below, similar to [BJK].

Theorem 3.2.1. *Let (f, μ) be a finite measure preserving system on a compact interval in $\mathbb{R}$. Assume that f is topologically transitive, Markov and piecewise expanding. If $\overline{\dim}_M(\mu) < \infty$, then for μ-a.e. x in the repeller,*

$$\limsup_{r \to 0} \frac{\log \tau_r(x)}{-\log r} \geq \overline{\dim}_M(\mu), \quad \liminf_{r \to 0} \frac{\log \tau_r(x)}{-\log r} \geq \underline{\dim}_M(\mu).$$

If μ is exponentially ψ-mixing , then for μ-almost every $x \in \Lambda$, the inequalities above are improved to

$$\limsup_{r \to 0} \frac{\log \tau_r(x)}{-\log r} = \overline{\dim}_M(\mu), \quad \liminf_{r \to 0} \frac{\log \tau_r(x)}{-\log r} = \underline{\dim}_M(\mu).$$

As in [BJK, Theorem 1.1] and [JT, Theorem 2.2], this theorem is particularly useful when $\overline{\dim}_M(\mu)$ and $\underline{\dim}_M(\mu)$ are finite (and preferably non-zero), which is true more often than not for finite IFS and finitely branched interval maps.

Remark 3.2.2. *Systems with* $\dim_M(\mu)$*, or at least $\overline{\dim}_M(\mu) < \infty$, are fairly common. For example, if μ is* doubling, i.e.*, there exists constant $D > 0$ such that for all $x \in$ supp(μ) and $r > 0$, $D\mu(B(x,r)) \geq \mu(B(x,2r)) > 0$, then $\overline{\dim}_M(\mu) < \infty$.*

Proof. For each $n \in \mathbb{N}$ let $x_n \in \text{supp}(\mu)$ be such that $\mu(B(x_n, 2^{-n})) = M_\mu(2^{-n})$, then by the doubling property,

$$M_\mu\left(2^{-n}\right)=\mu\left(B\left(x_n,2^{-n}\right)\right)\geq D^{-1}\mu\left(B\left(x_n,2^{-n+1}\right)\right)\geq D^{-1}M_\mu\left(2^{-n+1}\right)=D^{-1}\mu\left(B(x_{n-1}2^{-n+1})\right),$$

and iterating this one gets $M_\mu(2^{-n})\geq D^{-n+1}M_\mu(1/2)$, in other words

$$\frac{\log M_\mu(2^{-n})}{-n\log 2}\leq\frac{-(n-1)\log D+\log M_\mu(1/2)}{-n\log 2}.$$

As for all $r>0$, there is unique $n\in\mathbb{N}$ such that $2^{-n}<r\leq 2^{-n+1}$, and $\frac{\log 2^{-n}}{\log 2^{-n+1}}=1$,

$$\limsup_{r\to 0}\frac{\log M_\mu(r)}{\log r}=\limsup_{n\to\infty}\frac{\log M_\mu\left(2^{-n}\right)}{-n\log 2}\leq\frac{\log D}{\log 2}<\infty.$$ □

However, Minkowski dimensions are not always finite due to non-doubling behaviours, or generally more extreme decay of $M_\mu(r)$, especially when the associated symbolic shift is a countable Markov shift as discussed in Chapter 2. In this case the natural choice of the exponentially mixing measure is either not doubling or the ball of minimum measure decays stretched-exponentially as $r\to 0$ (see Example 3.3.2 below). Therefore a new notion of Minkowski dimension needs to be introduced.

Definition 3.2.3. *Define the upper and lower* stretched Minkowski dimensions *by*

$$\overline{\dim}^s_M(\mu):=\limsup_{r\to 0}\frac{\log|\log M_\mu(r)|}{-\log r},\quad \underline{\dim}^s_M(\mu):=\liminf_{r\to 0}\frac{\log|\log M_\mu(r)|}{-\log r}.$$

Those quantities should be of independent interest. Our second theorem below deals with almost sure cover times for systems in which $M_\mu(r)$ decays at stretched-exponential rates.

Theorem 3.2.4. *Let (f,μ) be a measure preserving system on $[0,1]$. Suppose f is topologically transitive, Markov and piecewise expanding. If $\overline{\dim}_M(\mu)=\infty$, but $0<\underline{\dim}^s_M(\mu),\overline{\dim}^s_M(\mu)<\infty$, then for μ-almost every $x\in\Lambda$,*

$$\liminf_{r\to 0}\frac{\log\log\tau_r(x)}{-\log r}\geq\underline{\dim}^s_M(\mu),\quad \limsup_{r\to 0}\frac{\log\log\tau_r(x)}{-\log r}\geq\overline{\dim}^s_M(\mu) \tag{3.2.3}$$

If (f,μ) is exponentially ψ-mixing, then for μ-almost every $x\in\Lambda$,

$$\liminf_{r\to 0}\frac{\log\log\tau_r(x)}{-\log r}=\underline{\dim}^s_M(\mu),\quad \limsup_{r\to 0}\frac{\log\log\tau_r(x)}{-\log r}=\overline{\dim}^s_M(\mu). \tag{3.2.4}$$

We first discuss some applications of our main theorem.

3.3 Examples

Theorem 3.2.1 and Theorem 3.2.4 are applicable to the following systems.

Example 3.3.1. Finitely branched Gibbs-Markov maps: let f be a topologically transitive piecewise expanding Markov map with $\mathcal{A}$ finite. Let $h := \frac{dLeb}{dLeb \circ f}$, then f is a *Gibbs-Markov map* if

- (Distortion) $\log h \circ \pi|_a$ is Lipschitz with respect to the symbolic metric d_s for all $a \in \mathcal{A}$.
- (Big image property) There exists $B_f >$ such that $Leb\,(f(\pi[a])) > B_f$ for all $a \in \mathcal{A}$.

The corresponding symbolic measure $\tilde{\mu}$ is a Gibbs measure as in Definition 1.2.6. For maps of this kind, the Gibbs measure μ (which is also the unique equilibrium measure for the potential $-\log|Df|$) is doubling (cf [Dol, Appendix 3]) so by Remark 3.2.2 $\dim_M(\mu)$ exists and is finite, then by Theorem 3.2.1,

$$\lim_{r\to 0} \frac{\log \tau_r(x)}{-\log r} = \dim_M(\mu)$$

for μ-a.e. x in the repeller of f.

In the next example, when $r \to 0$ at polynomial rate, $M_\mu(r)$ decays exponentially, hence $\overline{\dim}_M(\mu)$ is infinite so the stretched Minkowski dimensions are needed for computing a finite limit.

Example 3.3.2. Similar to [JT, Example 7.4], consider the following class of infinitely full-branched maps: pick $\kappa > 1$ and set $c = \zeta(\kappa) = \sum_{n\in\mathbb{N}} \frac{1}{n^\kappa}$. Let $a_0 = 0$, $a_j = \sum_{j=1}^{n} \frac{1}{cj^\kappa}$ and define f by

$$\forall\, n \in \mathbb{N},\ f(x) = cn^\kappa(x - a_{n-1}) \text{ for } x \in [a_{n-1}, a_n) =: P_n.$$

Then f is an infinitely full-branched affine map, and we can associate this map with a full-shift system on $\mathbb{N}$: $x = \pi(i_0, i_1, \dots)$ if for all $j \geq 1$, $f^j(x) \in P_{i_j}$.
Let $\omega > 1$ and construct $\tilde{\mu}$ the finite Bernoulli measure by

$$\tilde{\mu}([i_0, \dots, i_{n-1}]) = \prod_{j=0}^{n-1} \omega^{-i_j},$$

so the push-forward measure $\mu = \pi_* \tilde{\mu}$ has $\mu(P_n) = \omega^{-n}$.

Proposition 3.3.1. *For (f, μ) defined in Example 3.3.2, $\overline{\dim}_M(\mu) = \infty$, but $\dim_M^s(\mu) = \frac{1}{\kappa-1}$.*

Proof. For each $r > 0$, as the measure of P_j decays exponentially while their diameter only decays polynomially, the r-ball of minimum measure is found near 1. In particular, along the sequence $r_n = \frac{1}{2c}\sum_{j\geq n} j^{-\kappa} \approx \frac{1}{2c(\kappa-1)n^{\kappa-1}}$, the ball that realises $M_\mu(r_n)$ is contained in $\bigcup_{j=n}^{\infty} P_j$, hence

$$\omega^{-n} \leq M_\mu(r_n) \leq \frac{\omega^{-n}}{1 - \omega^{-1}}.$$

Therefore

$$\overline{\dim}_M(\mu) \geq \limsup_{n\to\infty} \frac{n \log \omega}{(\kappa - 1)\log n} = \infty,$$

whereas for all n,

$$\frac{\log n+\log\left(\log\omega+\frac{\log(1-1/\omega)}{n}\right)}{(\kappa-1)\log n+\log(2c(\kappa-1))}\le\frac{\log|\log M_\mu(r_n)|}{-\log r_n}\le\frac{\log n+\log\log\omega}{(\kappa-1)\log n+\log(2c(\kappa-1))}.$$

As for all $r>0$, there is unique $n\in\mathbb{N}$ such that $r_{n+1}\le r<r_n$ while $\lim_{n\to\infty}\frac{\log r_{n+1}}{\log r_n}=1$, we can conclude with $\dim_M^s(\mu)=\frac{1}{\kappa-1}$. □

As in [JT, Example 7.4] it is very difficult for the system to cover small neighbourhoods of 1 so Theorem 3.2.1 says $\limsup_{r\to0}\frac{\log\tau_r(x)}{-\log r}\ge\overline{\dim}_M(\mu)=\infty$, but since $\tilde{\mu}$ is Bernoulli hence ψ-mixing, Theorem 3.2.4 implies

$$\lim_{r\to0}\frac{\log\log\tau_r(x)}{-\log r}=\frac{1}{\kappa-1}\qquad\text{for } \mu\text{-a.e. } x.$$

3.4 Proof of Theorem 3.2.4

The proofs in this section are adapted from those of [BJK, Proposition 3.1, 3.2]. We will only demonstrate the proofs for Theorem 3.2.4, *i.e.*, the asymptotics are determined by stretched Minkowski dimensions; the proofs for Theorem 3.2.1 are obtained verbatim by replacing all stretched exponential sequences in the proofs below by some exponential sequence, *e.g.* for a given constant $s\in\mathbb{R}$, $e^{\pm n^s}$ will be replaced by $2^{\pm ns}$.

Remark 3.4.1. *Assuming the conditions of Theorem 3.2.4, we will prove that the statements hold along a subsequence $r_n=n^{-1}$ that has: for each $r>0$ there is a unique $n\in\mathbb{N}$ with $r_{n+1}<r\le r_n$ while $\lim_{n\to\infty}\frac{\log r_{n+1}}{\log r_n}=1$ (if $\overline{\dim}_M(\mu)$ or $\underline{\dim}_M(\mu)$ are finite we simply choose $r_n=2^{-n}$ instead). Since $\log\tau_r(x)$ is increasing as $r\to0$,*

$$\limsup_{n\to\infty}\frac{\log\log\tau_{r_n}(x)}{-\log r_n}=\limsup_{r\to0}\frac{\log\log\tau_r(x)}{-\log r}.$$

The same argument applies similarly to the $\liminf$*'s.*

Proof of the inequalities (3.2.4)

Assuming the inequalities in (3.2.3), we first prove the set of inequalities (3.2.4) which requires the exponentially ψ-mixing condition.

Proposition 3.4.2. *Assume that (f,μ) is exponentially ψ-mixing, and the upper stretched Minkowski dimension of μ, $\overline{\dim}_M^s(\mu)$, is finite. Then for μ-almost every $x\in\Lambda$,*

$$\limsup_{n\to\infty}\frac{\log\log\tau_r(x)}{-\log r}\le\overline{\dim}_M^s(\mu).$$

Proof. Let $\varepsilon > 0$, and for simplicity denote $\overline{\alpha} := \overline{\dim}^s_M(\mu)$.

For any finite k-word $\mathbf{i} = x_0, \ldots, x_{k-1} \in \Sigma_k$, let $\mathbf{i}^- = x_0, \ldots, x_{k-2}$, *i.e.*, $\mathbf{i}$ dropping the last letter. Recall that for each $\mathbf{i} \in \Sigma^*$, $P_{\mathbf{i}} = \pi[\mathbf{i}]$, and we define

$$\mathcal{W}_r := \{\mathbf{i} \in \Sigma^* : \text{diam}(P_{\mathbf{i}}) \le r < \text{diam}(P_{\mathbf{i}^-})\}.$$

By expansion, for each $n \in \mathbb{N}$, the lengths of the words in $\mathcal{W}_{n^{-1}}$ are bounded from above, hence we can define

$$L(n) := \frac{\log n}{\log \gamma} + 1 \ge \max\{|\mathbf{i}| : \mathbf{i} \in \mathcal{W}_{n^{-1}}\}.$$

Given $y \in [0,1]$ and $r > 0$ such that $B(y,r) \subset \text{supp}(\mu)$, define the corresponding symbolic balls by

$$\tilde{B}(y,r) := \{[\mathbf{i}] : \mathbf{i} \in \mathcal{W}_r, P_{\mathbf{i}} \cap B(y,r) \ne \emptyset\}.$$

If for some $x \in P_{\mathbf{i}}$, $[\mathbf{i}] \in \tilde{B}(y,r)$, then $d(x,y) \le r + diam(P_{\mathbf{i}}) \le 2r$. In other words, for all y and all $r > 0$,

$$B(y,r) \subset \pi\tilde{B}(y,r) \subset B(y,2r). \tag{3.4.1}$$

Let $\mathcal{Q}_n$ be a cover of Λ with balls of radius $r_n = 1/2n$ with disjoint interior, denote the collection of their centres by $\mathcal{Y}_n$, and $\#\mathcal{Q}_n = \#\mathcal{Y}_n \le n$. Let $\tau(\mathcal{Q}_n, x)$ be the minimum time for the orbit of x to have visited each element of $\mathcal{Q}_n$ at least once,

$$\tau(\mathcal{Q}_n, x) := \min\left\{k \in \mathbb{N} : \text{ for all } Q \in \mathcal{Q}_n, \text{ there exists } 0 \le j \le k : f^j(x) \in Q\right\}.$$

Then $\tau_{1/n}(x) \le \tau(\mathcal{Q}_n, x)$ for all n and all x since for all $y \in \Lambda$, there is $Q \in \mathcal{Q}_n$ and $j \le \tau(\mathcal{Q}_n, x)$ such that $f^j(x) \in Q$ and $y \in Q$ hence $d(f^j(x), y) \le 1/n$. Let $\varepsilon > 0$ be an arbitrary number and for each $k \in \mathbb{N}$, set $L'(k) = \lceil L(k) + \frac{1}{\rho}\left(k^{\overline{\alpha}+\varepsilon} + \log C_1\right)\rceil$ where C_1, ρ were given in Definition 1.1.5 and $\lceil t \rceil$ takes the least integer no smaller than t. We have

$$\begin{aligned}
&\mu\left(x : \tau_{1/n}(x) > e^{n^{\overline{\alpha}+\varepsilon}} L'(4n)\right) \le \mu\left(x : \tau(\mathcal{Q}_n, x) > e^{n^{\overline{\alpha}+\varepsilon}} L'(4n)\right)\\
&= \mu\left(x : \exists y \in \mathcal{Y}_n : f^j(x) \notin B(y, 1/2n), \forall j \le e^{n^{\overline{\alpha}+\varepsilon}} L'(4n)\right)\\
&\le \mu\left(x : \exists y \in \mathcal{Y}_n : f^{jL'(4n)}(x) \notin B(y, 1/2n), \forall j \le e^{n^{\overline{\alpha}+\varepsilon}}\right)\\
&= \mu\left(\bigcup_{y \in \mathcal{Y}_n} \bigcap_{j=1}^{e^{n^{\overline{\alpha}+\varepsilon}}} \left(f^{-jL'(4n)} B(y, 1/2n)\right)^c\right) \le \sum_{y \in \mathcal{Y}_n} \mu\left(\bigcap_{j=1}^{e^{n^{\overline{\alpha}+\varepsilon}}} \left(f^{-jL'(4n)} B(y, 1/2n)\right)^c\right).
\end{aligned} \tag{3.4.2}$$

A cylinder $[\mathbf{i}]$ in $\tilde{B}(y, 1/4n)$ has depth at most $L(4n)$, then by our choice of $L'(4n)$ and the exponentially

ψ-mixing property of $\tilde{\mu}$,

$$\mu(\tilde{B}(y,1/4n)\cap f^{-L'(4n)}\tilde{B}(y,1/4n))\leq(1+\exp(-((4n)^{\overline{\alpha}+\varepsilon})+\log C_1)\mu(\tilde{B}(y,1/4n)).$$

Similar calculations holds for $\mu\left(\bigcap_{j=1}^{e^{n^{\overline{\alpha}+\varepsilon}}}\left(f^{-jL'(4n)}B(y,1/2n)\right)^c\right)$ since the compliment of $\tilde{B}(y,1/4n)$ can be written as a countable union of cylinders of depths no greater then $L(4n)$.

By (3.4.1), for all z and all $r>0$,

$$\begin{aligned}
\sum_{y\in\mathcal{Y}_n}\mu\left(\bigcap_{j=1}^{e^{n^{\overline{\alpha}+\varepsilon}}}\left(f^{-jL'(4n)}B(y,1/2n)\right)^c\right)&\leq\sum_{y\in\mathcal{Y}_n}\tilde{\mu}\left(\bigcap_{j=1}^{e^{n^{\overline{\alpha}+\varepsilon}}}\left(\sigma^{-jL'(4n)}\tilde{B}\left(y,1/4n\right)\right)^c\right)\\
&\leq\left(1+\psi\left(\frac{1}{\rho}\left((4n)^{\overline{\alpha}+\varepsilon}+\log C_1\right)\right)\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}\sum_{y\in\mathcal{Y}_n}\left(1-\tilde{\mu}\left(\tilde{B}\left(y,\frac{1}{4n}\right)\right)\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}\\
&\leq\left(1+e^{-n^{\overline{\alpha}+\varepsilon}}\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}\sum_{y\in\mathcal{Y}_n}\left(1-\mu\left(B\left(y,\frac{1}{4n}\right)\right)\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}.
\end{aligned}\tag{3.4.3}$$

By definition of $\overline{\alpha}$, for all n large such that $\frac{\varepsilon}{4}\log n\geq\left(\overline{\alpha}+\frac{\varepsilon}{4}\right)\log 4$, we have

$$\log\left(-\log M_\mu\left(\frac{1}{4n}\right)\right)\leq(\overline{\alpha}+\varepsilon/4)(\log 4n)\leq(\overline{\alpha}+\varepsilon/2)\log n.$$

So for all $y\in$ supp(μ) and all n large enough,

$$\mu\left(B\left(y,\frac{1}{4n}\right)\right)\geq e^{-n^{\overline{\alpha}+\varepsilon/2}}\geq\frac{e^{n^{\varepsilon/2}}}{e^{n^{\overline{\alpha}+\varepsilon}}}.$$

As for all $u\in\mathbb{R}$ and all large k, $(1+\frac{u}{k})^k\approx e^u$, combining (3.4.2) and (3.4.3), for some uniform constant $C_2>0$,

$$\begin{aligned}
\mu\left(x:\tau_{1/n}(x)>e^{n^{\overline{\alpha}+\varepsilon}}L'(4n)\right)&\leq\left(1+e^{-n^{\overline{\alpha}+\varepsilon}}\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}\sum_{y\in\mathcal{Y}_{k+1}}\left(1-e^{-n^{\overline{\alpha}+\varepsilon/2}}\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}\\
&\leq\left(1+e^{-n^{\overline{\alpha}+\varepsilon}}\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}n\left(1-\frac{e^{n^{\varepsilon/2}}}{e^{n^{\overline{\alpha}+\varepsilon}}}\right)^{e^{n^{\overline{\alpha}+\varepsilon}}}\leq C_2\exp\left(\log n-e^{n^{\varepsilon/2}}\right).
\end{aligned}$$

The last term is clearly summable over n, then by Borel Cantelli, for all n large enough $\tau_{1/n}(x)\leq e^{n^{\overline{\alpha}+\varepsilon}}L'(4n)$. Since $\log L'(4n)\approx(\overline{\alpha}+\varepsilon)\log n\ll n^{\overline{\alpha}+\varepsilon}$, we have for $\mu-$a.e. $x\in\Lambda$,

$$\limsup_{n\to\infty}\frac{\log\log\tau_{1/n}(x)}{\log n}\leq\limsup_{n\to\infty}\frac{\log\log\left(e^{n^{\overline{\alpha}+\varepsilon}}L'(4n)\right)}{\log n}\leq\overline{\alpha}+\varepsilon.$$

By Remark 3.4.1 this upper bound for lim sup holds for all sequences decreasing to 0, and as $\varepsilon>0$ was arbitrary, we can conclude that for μ-a.e. $x \in \Lambda$,

$$\limsup_{r\to 0} \frac{\log\log \tau_r(x)}{-\log r} = \limsup_{n\to\infty} \frac{\log\log \tau_{1/n}}{\log n} \leq \overline{\alpha}.$$

□

Proposition 3.4.3. *Suppose (f,μ) is exponentially ψ-mixing and the lower stretched Minkowski dimension of μ, $\underline{\dim}^s_M(\mu)$, is finite, then for μ-a.e. $x \in \Lambda$,*

$$\liminf_{r\to 0} \frac{\log\log \tau_r(x)}{-\log r} \leq \underline{\dim}^s_M(\mu).$$

Proof. Again for simplicity, denote $\underline{\alpha} := \underline{\dim}^s_M(\mu)$. Let $\varepsilon > 0$, then by definition of liminf there is a subsequence $n_k \to \infty$ such that for all k,

$$\frac{\log(-\log M_\mu(1/n_k))}{\log n_k} \leq \underline{\alpha} + \varepsilon.$$

Then repeating the proof of Proposition 3.4.2 by replacing n by n_k everywhere, one gets that for $\mu-$almost every x,

$$\liminf_{k\to\infty} \frac{\log\log \tau_{1/n_k}(x)}{\log n_k} \leq \underline{\alpha} + \varepsilon.$$

Again send $\varepsilon \to 0$, and use the fact that liminf over the entire sequence is no greater than the liminf along any subsequence, the proposition is proved. □

Proof of the inequalities (3.2.3)

Next, we show the almost sure lower bounds which do not require the ψ-mixing property of μ.

Proposition 3.4.4. *For μ-almost every $x \in \Lambda$,*

$$\liminf_{r\to 0} \frac{\log\log \tau_r(x)}{-\log r} \geq \underline{\dim}^s_M(\mu).$$

Proof. We continue to use the notation $\underline{\alpha} = \underline{\dim}^s_M(\mu)$. Let $\varepsilon > 0$ be arbitrary, and by definition of $\underline{\alpha}$ for all large n there exists $y_n \in \mathrm{supp}(\mu)$ such that $\mu(B(y_n, 1/n)) \leq e^{-n^{\underline{\alpha}-\varepsilon}}$. Let

$$T(x,y,r) := \inf\left\{j \geq 0 : f^j(x) \in B(y,r)\right\},$$

so for all $n \in \mathbb{N}$ and all x, $\tau_{1/n}(x) \geq T(x, y_n, 1/n)$. Then by invariance,

$$\mu\left(x : \tau_{1/n}(x) < e^{n^{\underline{\alpha}-\varepsilon}}/n^2\right) \leq \mu\left(x : T(x, y_n, 1/n) < e^{n^{\underline{\alpha}-\varepsilon}}/n^2\right)$$
$$= \mu\left(x : \exists\, 0 \leq j < e^{n^{\underline{\alpha}-\varepsilon}}/n^2 :\ f^j(x) \in B(y_n, 1/n)\right) \leq \sum_{j=0}^{e^{n^{\underline{\alpha}-\varepsilon}}/n^2-1} \mu\left(x : f^j(x) \in B(y_n, 1/n)\right)$$
$$= \sum_{j=0}^{e^{n^{\underline{\alpha}-\varepsilon}}/n^2-1} \mu\left(f^{-j}B\left(y_n, \frac{1}{n}\right)\right) \leq \frac{e^{n^{\underline{\alpha}-\varepsilon}}}{n^2} e^{-n^{\overline{\alpha}-\varepsilon}} = \frac{1}{n^2},$$

which is summable. By Borel-Cantelli, since $2\log n \ll n^{\underline{\alpha}-\varepsilon}$, for μ-almost every x

$$\liminf_{n\to\infty} \frac{\log\log \tau_{1/n}(x)}{\log n} \geq \underline{\alpha} - \varepsilon.$$

As $\varepsilon > 0$ is arbitrarily small, the proposition is proved. □

In a similar way to Proposition 3.4.2 and Proposition 3.4.3,

Proposition 3.4.5. *For μ-almost every $x \in \Lambda$,*

$$\limsup_{r\to 0} \frac{\log\log \tau_r(x)}{-\log r} \geq \overline{\dim}^s_M(\mu).$$

Proof. Let $\varepsilon > 0$, then by definition of limsup there exists a subsequence $\{n_k\}_k \to \infty$ such that for all k,

$$\frac{\log\log\left(-M_\mu(1/n_k)\right)}{\log n_k} \geq \overline{\alpha} - \varepsilon.$$

Then repeating the proof of Proposition 3.4.4 along $\{n_k\}_k$, one gets that for μ-almost every x:

$$\limsup_{k\to\infty} \frac{\log\log \tau_{1/n_k}(x)}{\log n_k} \geq \overline{\alpha} - \varepsilon.$$

As ε was arbitrary,

$$\limsup_{r\to 0} \frac{\log\log \tau_r(x)}{-\log r} \geq \limsup_{k\to\infty} \frac{\log\log \tau_{1/n_k}(x)}{\log n_k} \geq \overline{\alpha}$$

for μ-a.e. $x \in \Lambda$. □

3.5 A non-mixing example: irrational rotations

The proof of Proposition 3.4.2 requires an exponentially ψ-mixing rate which is rather strong (see Definition 1.1.5). It is natural to ask if the same asymptotic growth in Theorem 3.2.4 remains valid under different mixing conditions, e.g. exponentially ϕ-mixing and α-mixing, or even polynomial ψ-mixing. Although these questions are unresolved, in this section we will show in Theorem 3.5.4 that the $\limsup$ and $\liminf$ of the asymptotic growth rate can differ if the system is not mixing.

Let $\theta \in (0,1)$ be an irrational number and define the rotation map $T = T_\theta : [0,1) \to [0,1)$, $T(x) = x + \theta \pmod 1$. Denote the one-dimensional Lebesgue measure on $[0,1)$ by μ, then (T,μ) is an ergodic probability preserving system with $\dim_M(\mu) = 1$.

Definition 3.5.1. *For a given irrational number θ, the* type of T_θ *is given by the following number*

$$\eta = \eta(\theta) := \sup \left\{ \xi : \liminf_{n\to\infty} n^{\xi} \|n\theta\| = 0 \right\},$$

where for every $r \in \mathbb{R}$, $\|r\| = \min_{n\in\mathbb{Z}} |r-n|$.

The fact that (T_θ, μ) is non-mixing for all irrational $\theta \in (0,1)$ is standard, which can be deduced by computing the following:
let $[a,b) \subset [0,1)$, and set $F := \mathbf{1}_{[a,b)}$. For each θ there exists $n_k \to \infty$ such that $\|n_k\theta\| \to 0$, hence

$$\int F \circ T_\theta^{n_k} \cdot F d\mu = ||b-a| - \|n_k\theta\|| \to |b-a|$$

as $k \to \infty$, hence $\lim_{n\to\infty} \int F \circ T_\theta^n \cdot F d\mu \neq (\int F d\mu)^2$, hence μ is non-mixing.

Remark 3.5.2. *(See [Khi]) For every $\theta \in (0,1)$ irrational, $\eta(\theta) \geq 1$ and $\eta(\theta) = 1$ almost everywhere, but for all real numbers $v \in (1,\infty]$, there exist irrational numbers with $\eta(\theta) = v$. The Liouville numbers have $\eta(\theta) = \infty$.*

For any irrational number $\theta \in (0,1)$ there is a unique infinite continued fraction expansion

$$\theta = [a_1, a_2, \dots] := \frac{1}{a_1 + \frac{1}{a_2 + \dots}},$$

where $a_i \geq 1$ for all $i \geq 1$. Set $p_0 = 0$ and $q_0 = 1$, and for each $i \geq 1$, choose $p_i, q_i \in \mathbb{N}$ coprime such that

$$\frac{p_i}{q_i} = [a_1, \dots, a_i] = \frac{1}{a_1 + \frac{1}{\dots \frac{1}{a_i}}}.$$

Definition 3.5.3. *The term a_i is called the i-th* partial quotient *and p_i/q_i the i-th* convergent. *In particular, (see [Khi])*

$$\eta(\theta) = \limsup_{n\to\infty} \frac{\log q_{n+1}}{\log q_n}.$$

The almost sure cover time for an irrational rotation is given by the theorem below.

Theorem 3.5.4. *For any irrational rotation T_θ, for μ-a.e. x,*

$$\liminf_{r\to 0} \frac{\log \tau_r(x)}{-\log r} = \dim_M(\mu) = 1 \leq \eta(\theta) = \limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \tag{3.5.1}$$

Remark 3.5.5. *In fact, by the nature of rotation and cover time,* (3.5.1) *all μ-almost-every statements in this section can be upgraded to for every $x \in [0,1)$.*

By Remark 3.5.2, there exist irrational rotations such that the asymptotic cover time does not converge. The proof of this theorem relies on the algebraic properties of $\eta(\theta)$. For simplicity, we fix θ and write η from now on.

Lemma 3.5.6. *[KS, Fact 1, Lemma 7]*

(a) $q_{i+2} = a_{i+2}q_{i+1} + q_i$ *and* $p_{i+2} = a_{i+2}p_{i+1} + p_i$.
(b) $1/(2q_{i+1}) \leq 1/(q_{i+1} + q_i) < \|q_i\theta\| < 1/q_{i+1}$ *for* $i \geq 1$.
(c) If $0 < j < q_{i+1}$, *then* $\|j\theta\| \geq \|q_i\theta\|$.
(d) For each $\varepsilon > 0$, *there exists a uniform* $C_\varepsilon > 0$ *such that for all* $j \in \mathbb{N}$, $j^{\eta+\varepsilon}\|j\theta\| > C_\varepsilon$.

We use ideas and results from [KS, Proposition 6, Proposition 10] to prove the following propositions.

Proposition 3.5.7. *For μ-a.e. x,*

$$\limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \geq \eta. \tag{3.5.2}$$

Proof. First it is easy to see that for all $r > 0$ and all $x, y \in [0,1)$, by the nature of rotation, $\tau_r(x) = \tau_r(y)$. In particular, $\tau_r(x) = \tau_r(Tx)$, hence the function $x \mapsto \limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r}$ is T invariant therefore constant μ-a.e. by ergodicity of μ.

By [KS, Proposition 10], for almost every x, y

$$\limsup_{r\to 0} \frac{\log W_{B(y,r)}(x)}{-\log r} \geq \eta,$$

where $W_E(x) := \inf\{n \geq 1 : T^n x \in E\}$ denotes the waiting time of x before visiting E. Hence there exists a set of strictly positive measure consisting of points that satisfy

$$\limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \geq \limsup_{r\to 0} \frac{\log W_{B(y,r)}(x)}{-\log r} \geq \eta,$$

since for all $y \in [0,1)$, $\tau_r(x) \geq W_{B(y,r)}(x)$. As $\limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r}$ is μ-a.e. constant, the inequality above holds for μ-a.e. x hence the proposition is proved. $\square$

Proposition 3.5.8. *For μ-a.e. x,*

$$\limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \leq \eta.$$

Proof. Let $\mathcal{Q}_n := \{[2^{-n}j, 2^{-n}(j+1)) : j = 0, \ldots, 2^n - 1\}$ and set $\tau(\mathcal{Q}_n, x)$ as the minimum time for x to have visited each element of $\mathcal{Q}_n$. Again, we have $\tau_{2^{-n+1}}(x) \leq \tau(\mathcal{Q}_n, x)$ for all x. By Lemma 3.5.6 (a) and (c), $\{\|q_i\theta\|\}_i$ is a decreasing sequence, and for each $n \in \mathbb{N}$ there exists a minimal j such that $\|q_j\theta\| < 2^{-n} \leq \|q_{j-1}\theta\|$, write $j = j_n$.

By [KS, Proposition 6] for all n, there is $\mu\left(W_{[0,2^{-n})} > q_{j_n} + q_{j_n-1}\right) = 0$. Notice that for all $a, b \in [0,1)$,

$$\mu\{W_{[a,a+b)}(x) = k\} = \mu\left\{\{x : W_{[0,b)}(x) = k\} + a\right\} = \mu\left\{W_{[0,b)}(x) = k\right\}, \tag{3.5.3}$$

as $\mu = Leb$ is translation invariant. Then by (3.5.3)

$$\begin{aligned}
&\mu\left\{\tau\left(\mathcal{Q}_n, x\right) > q_{j_n} + q_{j_n-1}\right\} = \mu\left\{x :\ \forall Q \in \mathcal{Q}_n :\ W_Q(x) > q_{j_n} + q_{j_n-1}\right\} \\
&= \mu\left(x : \bigcup_{Q\in\mathcal{Q}_n} \{W_Q(x) > q_{j_n-1} + q_{j_n}\}\right) \leq \sum_{Q\in\mathcal{Q}_n} \mu\left(W_Q > q_{j_n-1} + q_{j_n}\right) \\
&= \sum_{j=0}^{2^n-1} \mu\left(W_{[2^{-n}j, 2^{-n}(j+1))} > q_{j_n} + q_{j_n-1}\right) = \sum_{j=0}^{2^n-1} \mu\left(W_{[0,2^{-n})} > q_{j_n} + q_{j_n-1}\right) = 0.
\end{aligned}$$

Hence by Borel-Cantelli, for all n large enough, $\tau_{2^{-n+1}}(x) \leq (q_{j_n} + q_{j_n-1})$ for μ-a.e $x \in [0,1)$.

Let $\varepsilon > 0$, and by Lemma 3.5.6(b) and (d) there exists C_ε such that

$$\log\left(q_{j_n} + q_{j_n-1}\right) \leq \log\left(2q_{j_n}\right) \leq \log\frac{2}{\|q_{j_n}\theta\|} \leq (\eta+\varepsilon)\log q_{j_n} + \log 2 - \log C_\varepsilon,$$

Again by Lemma 3.5.6 and our choice of j_n, for μ-a.e. x and all n large enough,

$$\log\tau_{2^{-n+1}}(x) \leq \log(q_{j_n} + q_{j_n-1}) \lesssim (\eta+\varepsilon)\log q_{j_n} \leq -(\eta+\varepsilon)\log\|q_{j_n-1}\theta\| \leq (\eta+\varepsilon)n\log 2.$$

where $a \lesssim b$ means $a \leq b$ up to a uniform constant. Hence $\limsup_{n\to\infty}\frac{\log\tau_{2^{-n}}(x)}{n\log 2} \leq \eta+\varepsilon$ for μ-almost every x. Again, since for each $r < 0$ there is a unique $n \in \mathbb{N}$ for which $2^{-n} < r \leq 2^{-n+1}$, we can apply the subsequence trick again. As $\varepsilon > 0$ is arbitrarily small, the proposition is proved. □

Proposition 3.5.9. *For μ-almost every $x \in [0,1)$,*

$$\liminf_{r\to 0}\frac{\log\tau_r(x)}{-\log r} = 1.$$

Proof. Let $\varepsilon > 0$ and using the same arguments in the last proof, *i.e.*, cover time is greater than the hitting time of the ball of smallest measure at scale r, along the sequence $r_n = 2^{-(n+1)}$, one gets that for all $[a - r_n, a + r_n) \subset [0,1)$,

$$\begin{aligned}
&\sum_{n\geq 1}\mu\left(\tau_{r_n}(x) < 2^{n(1-\varepsilon)}\right) \leq \sum_{n\geq 1}\mu\left(x : W_{[a-2^{-n-1}, a+2^{-n-1})}(x) < 2^{n(1-\varepsilon)}\right) \\
&\leq \sum_{n\geq 1}\sum_{k=0}^{2^{n(1-\varepsilon)}}\mu\left(T^{-k}[a-2^{-n-1}, a+2^{-n-1})\right) = \sum_{n\geq 1}2^{n(1-\varepsilon)}2^{-n} = \sum_{n\geq 1}2^{-\varepsilon n} < \infty.
\end{aligned}$$

For each r there is a unique n such that $r_n < r \le r_{n-1}$ and $\lim_n \frac{\log r_n}{\log r_{n-1}} = 1$, so by Borel-Cantelli,

$$\liminf_{r\to 0} \frac{\log \tau_r(x)}{-\log r} = \liminf_{n\to\infty} \frac{\log \tau_{2^{-n}}(x)}{n\log 2} \ge 1-\varepsilon, \tag{3.5.4}$$

and as ε is arbitrarily small, the lower bound for the $\liminf$ is proved.

For the upper bound of liminf, recall that $\tau(\mathcal{Q}_n, x) \ge \tau_{2^{-n}}(x)$, we can repeat the proof of Proposition 3.5.8, apart from that this time we choose $\{2^{-n_i}\}_i$ according to $\{q_i\}_{i\in\mathbb{N}}$: for each i, choose $n_i \in \mathbb{N}$ to be the smallest number such that

$$\|q_{i+1}\theta\| < 2^{-n_i} \le \|q_i\theta\|.$$

Then, as in the proof of Proposition 3.5.8,

$$\mu\left(\tau\left(\mathcal{Q}_{n_i}, x\right) > q_{i+1} + q_i\right) \le \sum_{Q\in\mathcal{Q}_{n_i}} \mu\left(W_Q > q_{i+1} + q_i\right) = 0.$$

Again by Lemma 3.5.6 (b), $q_{i+1}+q_i \le 2q_{i+1} \le \frac{2}{\|q_i\theta\|} < 2^{n_i+1}$ by our choice of n_i, so $\lim_{i\to\infty} \frac{\log(q_i+q_{i+1})}{n_i \log 2} \le 1$, therefore for μ-a.e. x,

$$\liminf_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \le \liminf_{i\to\infty} \frac{\log \tau_{2^{-n_i}}(x)}{n_i \log 2} \le \liminf_{i\to\infty} \frac{\log \tau\left(\mathcal{Q}_{n_i}, x\right)}{n_i\log 2} \le 1.$$

Combining this with (3.5.4) $\liminf_{r\to 0} \frac{\log \tau_r(x)}{-\log r} = 1$ for μ almost every x. □

Theorem 3.5.4 is obtained by combining the three propositions above, and we have therefore shown that for irrational rotations, which have no mixing properties, the asymptotic limit may not converge to the Minkowski dimension dim(μ).

3.6 Cover time for flows

In this section we prove an analogous almost sure limit regarding cover times for the same class of flows that was discussed in [RT, §4].

Let $\{f_t\}_t$ be a flow on a metric space $(\mathcal{X}, d)$ preserving an ergodic probability measure ν, *i.e.*, $\nu\left(f_t^{-1}A\right) = \nu(A)$ for every $t \ge 0$ and A measurable. Recall that a point $x \in \mathcal{X}$ is called non-wandering if for every open neighbourhood $U \ni x$ and all $T > 0$, there exists $|t| > T$ such that $f_t(U) \cap U \ne \emptyset$. Let Ω denote the non-wandering set and define the cover time of x at scale r by

$$\tau_r(x) := \inf\left\{T > 0 : \forall y \in \Omega, \exists t \le T : d(f_t(x), y) < r\right\}.$$

We will assume the existence of a Poincaré section $Y \subset \mathcal{X}$ with $\overline{R} := \int R_Y d\nu < \infty$, where $R_Y(x)$

denotes the first hitting time to Y, *i.e.*, $R_Y(x) := \inf\{t > 0 : f_t(x) \in Y\}$. Define the Poincaré map by (Y, F, μ) where $F = f_{R_Y}$ and let μ be the induced measure on Y. Additionally, assume the following conditions:

(H1) $\dim_M(\mu)$ exists and is finite for (F, μ),

(H2) (Y, F, μ) is Gibbs-Markov so Theorem 3.2.1 is applicable for μ-almost every $y \in Y$.

(H3) $\{f_t\}_t$ has bounded speed: there exists $K > 0$ such that for all $s \in \mathbb{R}$ and $t > 0$, $d(f_s(x), f_{s+t}(x)) < Kt$.

(H4) $\{f_t\}_t$ is topologically transitive and there exists $T_1 > 0$ such that

$$\bigcup_{0<t\leq T_1} f_t(Y) = \mathcal{X}. \tag{3.6.1}$$

(H5) There exists

$$C_f := \sup\{\text{diam}(f_t(I))/\text{diam}(I) : I \text{ a connected component in } Y, 0 < t \leq T_1\} \in (0, \infty).$$

Remark 3.6.1. *The last condition is satisfied when (H3) holds and the flow is, for example, Lipschitz,* i.e., *there exists $L > 0$ such that for all $x, y \in \mathcal{X}$,*

$$d(f_t(x), f_t(y)) \leq L^t d(x, y).$$

Theorem 3.6.2. *Let (f_t, ν) be a measure preserving flow satisfying conditions (H1)-(H5). Then for ν-almost every $x \in \Omega$,*

$$\liminf_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \geq \underline{\dim}_M(\nu) - 1. \tag{3.6.2}$$

Furthermore, if $\overline{\dim}_M(\nu) = \dim_M(\mu) + 1$,

$$\limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \leq \overline{\dim}_M(\mu) = \overline{\dim}_M(\nu) - 1 \quad \nu\text{-a.e.} \tag{3.6.3}$$

Proof of (3.6.2). This proof is analogous to those of Proposition 3.4.3 and [RT, Theorem 4.1].

Fix some $y \in \Omega$ and $r > 0$ and consider the random variable

$$S_{T,r}(x) := \int_0^T \mathbf{1}_{B(y,r)}(f_t(x))dt.$$

Observe that by the bounded speed property, for all $T > r/K$,

$$\{x : \exists 0 \leq t \leq T \text{ s.t. } d(f_t(x), y)) < r\} \subset \{S_{2T,2r}(x) > r/K\},$$

since if $d(f_s(x), y) < r$ for some s, then for all $t < r/K$, $d(f_{t+s}(x), y) < 2r$. Also set

$$T(x, y, r) := \inf\{t \geq 0 : f_t(x) \in B(y, r)\},$$

and similarly for all $r>0$ and all x,z, $\tau_r(x)\ge T(x,y,r)$.

Let $\varepsilon>0$ be arbitrary and by definition of $\underline{\alpha}$ for all large $n\in\mathbb{N}$ there exists $y_n\in\Omega$ such that $\nu(B(y_n,2^{-n}))\le 2^{-n(\underline{\alpha}-\varepsilon)}$. By Markov's inequality, for some $\mathcal{T}_n>0$ to be decided later,

$$\begin{aligned}&\nu\left(x:\tau_{2^{-n}}(x)<\mathcal{T}_n\right)\le\nu\left(x:T(x,y_n,2^{-n})<\mathcal{T}_n\right)=\nu\left(x:\exists\,0\le t<\mathcal{T}_n:\ f_t(x)\in B(y_n,2^{-n})\right)\\&\le\nu\left(x:S_{2\mathcal{T}_n,2^{-n+1}}(x)>2^{-n}/K\right)\le K2^n\int_0^{2\mathcal{T}_n}\int\mathbf{1}_{B(y_{n-1},2^{-n+1})}(f_t(x))d\nu(x)dt\\&\le K2^{n+1}\mathcal{T}_n\nu(B(y_{n-1},2^{-n+1}))\le 4K\mathcal{T}_n2^{-(n-1)(\underline{\alpha}-\varepsilon-1)}.\end{aligned}$$

Choosing $\mathcal{T}_n=2^{(n-1)(\underline{\alpha}-\varepsilon-1)}/n^2$, the last term above is summable along n hence by Borel-Cantelli, for ν-almost every x

$$\liminf_{r\to0}\frac{\log\tau_r(x)}{-\log r}\ge\liminf_{n\to\infty}\frac{\log\mathcal{T}_n}{n\log 2}=\underline{\alpha}-1-\varepsilon.$$

Since $\varepsilon>0$ was arbitrarily small the lower bound is $\underline{\alpha}-1$, and by Remark 3.4.1 the proposition is proved. $\square$

Note that the proof of lower bound is independent of the existence or mixing properties of the Poincaré map (Y,F,μ). For upper bound, we first prove that the cover time of the Poincaré F in Y is comparable to the cover time of the flow.

Lemma 3.6.3. *Define*

$$\tau_r^F(x):=\min\{n\in\mathbb{N}_0:\forall y\in Y,\,\exists 0\le j\le n:d(y,F^jx)<r\}.$$

There exists $\lambda=\frac{1}{C_f}$ for C_f defined in (H5) such that $\tau_r(x)\le T_1+\sum_{j=0}^{\tau^F_{\lambda r}(x)}R_Y(F^jx)$.

Proof. This is adapted from the proof of [JT, Lemma 6.4] and [RT, Theorem 2.1]. F is by assumption Gibbs-Markov so one can find $\mathcal{P}(r)$, a natural partition of Y using cylinder sets with respect to F, such that for each $P\in\mathcal{P}(r)$: (a) diam$(P)\le r/C_f$, and (b) for all $0<t\le T_1$, $f_t(P)$ is connected. Suppose $\tau^F_{r/C_f}(x)=k$, then the orbit $\{x,F(x),\dots,F^k(x)\}$ must have visited every element of $\mathcal{P}$. By (3.6.1) for each $y\in\Omega$ there is $P\in\mathcal{P}(r)$ and $0<s\le T_1$ such that $y\in f_s(P)$ and hence there exists $j\le k$ such that $d\left(f_s(F^j(x)),y\right)\le C_f|P|<r$. Then set $\lambda=1/C_f$ the lemma is proved. $\square$

Proof of (3.6.3). Now assume $\overline{\dim}_M(\nu)=\dim_M(\mu)+1$. Let $\xi>0$ be arbitrary and define the sets

$$U_{\xi,N}:=\left\{x\in Y:|R_n(x)-n\overline{R}|\le\xi n,\forall n\ge N\right\},$$

where $R_n(x)=\sum_{j=0}^{n-1}R_Y(F^j(x))$. By ergodicity, $\lim_N\mu(U_{\xi,N})=1$ so for N large, $\nu(U_{\xi,N})>0$ hence by invariance,

$$\lim_{N\to\infty} \nu\left(\bigcup_{t=0}^{\xi N} f_{-t}(U_{\xi,N})\right) = 1. \tag{3.6.4}$$

Let $\varepsilon > 0$ be arbitrary. By (3.6.4) one can pick N^* such that for each ν typical $x \in \mathcal{X}$ there is some $t^* \le \xi N^*$ such that $f_{t^*}(x) \in Y$. By Theorem 3.2.1 applied to the Poincaré map and Lemma 3.6.3, for all sufficiently small $r > 0$ we have the following two inequalities,

$$\frac{\log \tau^F_{\lambda r}(f_{t^*}x)}{-\log \lambda r} \le \dim_M(\mu) + \varepsilon, \quad \frac{\log\left(\tau_r(x) - T_1\right)}{-\log r} \le \frac{\log\left((\overline{R}+\xi)\tau^F_{\lambda r}(f_{t^*}x)\right)}{-\log r}.$$

Then as $\lambda, \overline{R}$ are constants and ε is arbitrary, for ν-almost every x,

$$\limsup_{r\to 0} \frac{\log \tau_r(x)}{-\log r} \le \dim_M(\mu) = \overline{\dim}_M(\nu) - 1. \qquad \square$$

3.6.1 Example: suspension flow over topological Markov shifts

In this section, we give an example of a flow for which $\dim_M(\nu) = \dim_M(\mu) + 1$ is satisfied, so (3.6.3) is applicable.

Consider two-sided Markov subshift of finite type $(\Sigma, \sigma, \phi, \mu)$: $\mathcal{A}$ a finite alphabet and $M = [M_{ij}]_{\mathcal{A}\times\mathcal{A}}$ transition matrix,

$$\Sigma := \left\{x = (\dots, x_{-1}, x_0, x_1, \dots) \in \mathcal{A}^{\mathbb{Z}} : \text{ for all } j \in \mathbb{Z},\, x_j \in \mathcal{A} \text{ and } M_{x_j, x_{j+1}} = 1\right\},$$

σ the usual left shift, ϕ a Hölder potential and μ the unique Gibbs measure with respect to ϕ. We assume that $\dim_M(\mu) \in (0, \infty)$. The natural symbolic metric on Σ is $d_s(x, y) = 2^{-x\wedge y}$, where

$$x \wedge y = \sup\{k \ge 0 : x_j = y_j,\, \forall |j| < k\}. \tag{3.6.5}$$

An n-cylinder in this setting is given by $\left[x_{-(n-1)}, \dots, x_0, \dots, x_{n-1}\right] := \{y \in \Sigma,\, y_j = x_j, \forall |j| < n\}$, and it is a well-known fact that balls in Σ are precisely the cylinder sets. The left-shift map σ is bi-Lipschitz with Lipschitz constant $L = 2$. For more detailed description of the shift space, see [Bow, §1].

Let $\varphi \in L^1(\mu)$ be a positive Lipschitz function, define the space

$$Y_\varphi := \{(x, s) \in \Sigma \times \mathbb{R}_{\ge 0} : 0 \le s \le \varphi(x)\} / \sim$$

where $(x, \varphi(x)) \sim (\sigma(x), 0)$ for all $x \in I$. The suspension flow Ψ over σ is the function which acts on Y_φ by

$$\Psi_t(x, s) = (\sigma^k(x), v),$$

where $k, v \ge 0$ are determined by $s + t = v + \sum_{j=0}^{k-1} \varphi(\sigma^j(x))$. The invariant measure ν for the flow Ψ

on Y_φ satisfies the following: for every $g : Y_\varphi \to \mathbb{R}$ continuous,

$$\int g d\nu = \frac{1}{\int_\Sigma \varphi d\mu} \int_\Sigma \int_0^{\varphi(x)} g(x,s) ds d\mu(x). \tag{3.6.6}$$

The standard metric on Y_φ is the *Bowen-Walters distance* d_Y (see for example [BW]). Define an alternative metric d_π on Y_φ: for all $(x,s),(y,t) \in Y_\varphi$,

$$d_\pi((x,s),(y,t)) := \min \begin{Bmatrix} d(x,y) + |s-t|, \\ d(\sigma x, y) + \varphi(x) - s + t, \\ d(x, \sigma y) + \varphi(y) - t + s \end{Bmatrix}, \tag{3.6.7}$$

the following proposition says d_π is comparable to the Bowen-Walters distance.

Proposition 3.6.4. *[BS, Proposition 17] There exists $c = c_\pi$ such that*

$$c^{-1} d_\pi((x_1,t_1),(x_2,t_2)) \le d_Y((x_1,t_1),(x_2,t_2)) \le c\, d_\pi((x_1,t_1),(x_2,t_2)).$$

Then the Minkowski dimension of the flow-invariant measure ν is given by the following proposition.

Proposition 3.6.5. *For (μ) the Gibbs measure with respect to ϕ on the two-sided subshift and ν the flow invariant measure,* $\dim_M(\nu) = \dim_M(\mu) + 1$.

Proof. The proof is based on the proof of [RT, Theorem 4.3] for correlation dimensions.

By Proposition 3.6.4 for all $r > 0$,

$$(B(x, r/2c) \times (s - r/2c, s + r/2c)) \cap Y \subset B_Y((x,s), r)$$

where B_Y denotes the ball with respect to the metric d_Y, and set $\overline{\varphi} = \int_\Sigma \varphi d\mu$. Then for all $(x,s) \in Y_\varphi$,

$$\nu(B_Y((x,s),r)) \ge \nu(B(x, r/2c) \times \left(s - \frac{r}{2c}, s + \frac{r}{2c}\right),$$

$$\frac{\log \nu(B_Y((x,s),r))}{\log r} \le \frac{\log \left(\frac{r}{c\overline{\varphi}} \mu\left(B(x, \frac{r}{2c})\right)\right)}{\log r}.$$

Hence $\overline{\dim}_M(\nu) = \limsup_{r\to 0} \frac{\log \min_{(x,s)\in supp(\nu)} \nu(B_Y((x,s),r)}{\log r} \le \dim_M(\mu) + 1$.

For the lower bound, define

$$B_1 := B(x, cr) \times (s - cr, s + cr), \quad B_2 := B(\sigma x, cr) \times [0, cr),$$

$$B_3 := \left\{(y,t) : y \in B(\sigma^{-1} x, 2cr), \text{ and } \varphi(y) - cr \le t \le \varphi(y)\right\}.$$

Then as in the proof of [RT, Theorem 4.3], $B_Y((x,s),r) \subset (B_1 \cup B_2 \cup B_3) \cap Y_\varphi$.

For all $r > 0$ and $(x,s) \in Y_\varphi$ by (3.6.6), and as μ is σ, σ^{-1} invariant,

$$\nu(B_1 \cap Y_\varphi) = 2cr\mu(B(x,cr))/\overline{\varphi}, \quad \nu(B_2, Y_\varphi) \leq cr\mu(B(x,cr))/\overline{\varphi}$$
$$\nu(B_3 \cap Y_\varphi) \leq cr\mu(\sigma^{-1}B(x,2cr))/\overline{\varphi} = cr\mu(B(x,2cr))/\overline{\varphi}.$$

Therefore

$$\nu(B_Y((x,s),r) \leq \frac{1}{\overline{\varphi}}\left(3r\mu(B(x,cr)) + cr\mu(B(x,2cr))\right),$$

which is enough to conclude that $\underline{\dim}_M(\nu) \geq \dim_M(\mu) + 1$. Combining with the upper bound above we obtain $\dim_M(\nu) = \dim_M(\mu) + 1$. $\square$

Thus by Theorem 3.6.2, for ν the invariant measure of the suspension flow, for ν-almost every $(x,t) \in Y_\varphi$,

$$\lim_{r\to 0} \frac{\log \tau_r(x,t)}{-\log r} = \dim_M(\nu) - 1 = \dim_M(\mu),$$

where μ is the Gibbs measure of the two-sided subshift (Σ, σ, ϕ). We will revisit suspension flows in the next chapter for the shortest distance problem.

Chapter 4

Limit theorems for shortest distance problems

In this chapter we show another set of almost sure asymptotic convergence for systems with a ψ-mixing measure. In fact, the natural candidate of such a measure is a Gibbs measure. We first deal with symbolic dynamics. Consider a topological Markov shift (Σ, σ) on an (at most) countably infinite alphabet $\mathcal{A}$ with respect to a transition matrix M equipped with the natural symbolic metric d_s. The following quantity is of interest: for each $x \in \Sigma$,

$$M_n(x) = \max\{k : \exists\, 0 \leq i < j \leq n-1\colon x_i, \dots, x_{i+k-1} = x_j, \dots, x_{j+k-1}\}. \tag{4.0.1}$$

$M_n(x)$ counts the maximum length of self-repeating subwords in the first n symbols, and in the *longest common substring matching problem* we study the asymptotic growth of $M_n(x)$ as $n \to \infty$. We will first show that for topological Markov shifts with a Gibbs probability measure μ, $M_n(x)/\log n$ converges μ-almost surely, and then apply similar techniques to show an analogous statement for interval dynamics.

To give a motivation of the problem, let us first consider a primitive model: coin tosses. One can ask what is the probability of two identical coins coinciding for K consecutive times, and how large can such K be. This was solved by Rényi [Ren] with a formula we now call an Erdös-Rényi law. Another motivation for the substring matching problem comes from matching nucleotide sequences in DNA which is a shift on four symbols; nucleotide sequences in DNA transcription correspond to the amino acid chain which ultimately determines the structure of the protein produced by an organism's cells. The longer the matchings in substrings of DNA, the more similarities in the protein structure of the organism. Early results were established in the 80s by Arratia and Waterman's work [AW]. They considered the length of the longest common substring among two *i.i.d.* sequences $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ with different distributions taking letters in a finite alphabet. The longest matching subsequence with and without shift are defined respectively by

$$M_n(X,Y) := \sup\left\{k : X_{i+m} = Y_{j+m} \text{ for all } m = 1 \text{ to } k \text{ and } 1 \leq i,j \leq n-k\right\},$$
$$R_n(X,Y) := \sup\left\{k : X_{i+m} = Y_{i+m} \text{ for all } m = 1 \text{ to } k \text{ and } 1 \leq i \leq n-k\right\},$$

and they satisfy an Erdős-Rényi law

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{M_n}{\log n} = \frac{2}{\log 1/p}\right) = 1, \;\; \mathbb{P}\left(\lim_{n\to\infty} \frac{M_n}{R_n} = 2\right) = 1, \tag{4.0.2}$$

where $p = \mathbb{P}(X_1 = Y_1)$ is the *collision probability*. The quantity $-\log \mathbb{P}(X_1 = Y_1)$ is often called the *collision entropy* or *Rényi entropy*. Analogous convergence results as (4.0.2) are generalised to Markov processes and matching with 'scores', see [DKZa],[DKZb].

The same problem can easily be translated to the topological shifts context: for (Σ, σ) a countable or finite Markov subshift with an invariant probability measure μ, we study the $\mu \times \mu$-a.e. growth of

$$M_n(x,y) := \sup\left\{k : \exists\, 0 \leq i,j \leq n-1 \colon x_i, \dots, x_{i+k-1} = y_j, \dots, y_{j+k-1}\right\}.$$

In [BLR], the authors show that for $\mu \times \mu$-almost every $(x,y) \in \Sigma \times \Sigma$,

$$\limsup_{n\to\infty} \frac{M_n(x,y)}{\log n} \leq \frac{2}{\underline{H}_2},$$

and if the measure μ is α-mixing (see Definition 1.1.5) with exponential decay or ψ-mixing with polynomial decay, then

$$\liminf_{n\to\infty} \frac{M_n(x,y)}{\log n} \geq \frac{2}{\overline{H}_2}, \;\; \mu \times \mu\text{-a.e.}$$

The quantities $\overline{H}_2, \underline{H}_2$ (see Definition 4.1.1) are called the upper and lower *Rényi entropies*, which are generalisations of the collision entropy $\log 1/p$ in (4.0.2). This almost sure result is later generalised for orbits generated by k μ-typical points, for $k \in \mathbb{N}\backslash\{1\}$ [BR21], and random shift systems [GRS]. The question is, does an analogous almost sure convergence hold for single points orbit case, *i.e.*, substitute $M_n(x,y)$ by $M_n(x)$ as defined in (4.0.1).

The analogous problem for $M_n(x)$ is more difficult due to lack of independence and *short returns*, the latter in symbolic context refers to the *overlapping* phenomena which will be discussed soon. For subshifts of finite type, Collet et al in [CGR] applied first and second-moment analysis to the counting random variable $N(x,n,r_n)$, which counts the number of matches of strings of length r_n among the first n iterates in x. They then showed that for H_2 the *Rényi entropy* of a Gibbs measure μ, for all $\varepsilon > 0$,

$$\lim_{n\to+\infty} \mu\left(\left|\frac{M_n(x)}{\log n} - \frac{2}{H_2}\right| > \varepsilon\right) = 0.$$

That is, $\frac{M_n(x)}{\log n}$ converges to $\frac{2}{H_2}$ in probability for typical x. Then one may ask if this result can be improved to almost sure convergence, or if the convergence remains valid when the alphabet is countably infinite. The answer is given in Theorem 4.2.1.

4.1 Rényi entropy

Let $\mathcal{A}$ be a countable or finite alphabet and Σ a topological Markov subshift defined by the transition matrix $M = [M_{ij}]_{\mathcal{A}\times\mathcal{A}}$. Recall that $\mathcal{C}_n$ denotes the set of n-cylinders in Σ.

Definition 4.1.1 (Rényi entropy)**.** *For each $n \in \mathbb{N}$, $t > 0$, define the quantities*

$$F_n(t) = \sum_{C\in\mathcal{C}_n} \mu(C)^{1+t}.$$

The upper and lower Rényi entropy *(with respect to the natural partition given by the alphabet $\mathcal{A}$) of the system are defined respectively by*

$$\overline{H}_2(\mu) := \limsup_{n\to\infty} \frac{\log F_n(1)}{-n}, \quad \underline{H}_2(\mu) := \liminf_{n\to\infty} \frac{\log F_n(1)}{-n},$$

and write $H_2(\mu)$ whenever these coincide. The generalised Rényi entropy function *is*

$$\mathcal{R}_\mu(t) = \liminf_{n\to+\infty} \frac{\log F_n(t)}{-tn}.$$

In the information theory context, this is also called *collision entropy*, as it reflects the probability of two *i.i.d.* random variables coinciding in value. Therefore, heuristically, the probability of a k-matching *i.e.*, $X_{m+j} = Y_{n+j}$ for $j = 1, \dots, k$, is roughly e^{-kH_2}.

Rényi entropy does not always exist, especially when the alphabet is not finite[1]. For the finite alphabet case, Haydn and Vaienti proved in [HV, Theorem 1] that $\mathcal{R}_\mu(t)$ converges uniformly on compact subsets of $\mathbb{R}^+$ for all weakly ψ-mixing invariant measures, in particular, if μ is a Gibbs measure, $H_2(\mu) = \mathcal{R}_\mu(1) = 2P(\phi) - P(2\phi)$ where P is the topological pressure defined in (1.2.1). We now show the formula for Gibbs measures for countable Markov shifts.

Lemma 4.1.2. *Let (Σ, σ, ϕ) be a countable Markov shift with the BIP property, ϕ a locally Hölder potential such that $P(\phi) < \infty$, and μ the unique (up to multiplicative constants) Gibbs probability measure for ϕ.*

[1]For infinite alphabet Markov chains, Rényi entropy is calculated in [GGL].

Then the Rényi entropy $H_2(\mu)$ is finite and well defined for μ and is explicitly given by

$$H_2(\mu) = 2P(\phi) - P(2\phi),$$

where P is the Gurevich pressure defined in (1.3.1).

Proof. We first show that $\overline{H}_2$ and $\underline{H}_2$ coincide for μ. Recall that for a real-valued potential ϕ, $\mathrm{var}_k(\phi) := \sup\{|\phi(x) - \phi(y)| : x_i = y_i, \text{ for all } i = 1, \dots, k-1\}$. As μ is a Gibbs measure for the potential ϕ, the constant $B_1(\phi) := \sum_{k\geq 1} \mathrm{var}_k(\phi)$ must be finite. Then if $x, y \in C$ for some $C \in \mathcal{C}_n$, $|S_n\phi(x) - S_n\phi(y)| \leq B_1$. By the Gibbs property (see (1.2.4)) for G the Gibbs constant and $P = P(\phi)$, for all $n, k \in \mathbb{N}$ since every allowable word of length $n+k$ must be some concatenation of a length n word and a length k word,

$$\begin{aligned}
F_{n+k}(1) = \sum_{C\in\mathcal{C}_{n+k}} \mu(C)^2 &\leq G^2 \sum_{C\in\mathcal{C}_{n+k}} \exp\left(2\left(S_{n+k}\phi(x) - (n+k)P\right)\right) \\
&\leq G^2 \sum_{\substack{C\in\mathcal{C}_n, D\in\mathcal{C}_k \\ C\cap\sigma^{-n}D\neq\emptyset}} \exp\left(\sup_{x\in C} 2\left(S_n\phi(x) - nP\right)\right) \exp\left(\sup_{y\in D} 2(S_k\phi(y) - kP)\right) \\
&\leq G^2 e^{4B_1} \sum_{C\in\mathcal{C}_n, D\in\mathcal{C}_k} \exp\left(2\left(S_n\phi(x) - nP\right)\right) \exp\left(2\left(S_k\phi(y) - kP\right)\right) \leq G^4 e^{2B_1} F_n(1) F_k(1),
\end{aligned}$$

where x, y in the sums are simply arbitrary points in the cylinders. So $\log F_n(1)$ is almost subadditive hence $\lim_{n\to\infty} \frac{\log F_n(1)}{-n} = \sup \frac{\log F_n(1)}{-n}$ exists. Therefore, to show that the limit is finite, it suffices to find a subsequence converging to a finite constant, as every convergent subsequence of a convergent sequence must converge to the same limit.

For some $x \in \Sigma$, let $C_k(x)$ denote the unique k-cylinder containing x. Suppose x is a periodic point with period k, then for all n obviously $\mu(C_{nk}(x))^2 \leq F_{nk}(1) = \sum_{C\in\mathcal{C}_{nk}} \mu(C)^2$. As $S_{nk}\phi(x) = nS_k\phi(x)$,

$$2\left(S_k\phi(x)/k - P\right) = \liminf_{n\to\infty} \frac{-2\log G + 2\left(S_{nk}\phi(x) - nkP\right)}{nk} \leq \liminf_{n\to\infty} \frac{\log F_{nk}(1)}{nk} = \liminf_{n\to\infty} \frac{\log F_n(1)}{n}.$$

Since $|S_k\phi(x)/k|,\ P(\phi)$ are bounded, we get $\lim_n \frac{\log F_n(1)}{-n}$ is finite.

Combining the BIP property and the locally Hölder property of ϕ with [Sar1, Lemma 4] one can show that

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{C\in\mathcal{C}_n} \exp\left(\sup_{x\in C} 2S_n\phi(x)\right) \leq P(2\phi)$$

which implies

$$\limsup_{n\to\infty} \frac{1}{n} \log \sum_{C\in\mathcal{C}} \mu(C)^2 \leq \limsup_{n\to\infty} \frac{1}{n} G^2 e^{-2nP(\phi)} \sum_{C\in\mathcal{C}_n} \exp\left(\sup_{x\in C} 2S_n\phi(x)\right) \leq P(2\phi) - 2P(\phi).$$

Also for each $C \in \mathcal{C}_n$, there is at most one $x \in C$ such that $\sigma^n x = x$, thus

$$\begin{aligned}\sum_{C\in\mathcal{C}_n} \mu(C)^2 \geq & G^{-2}e^{-2nP(\phi)} \sum_{C\in\mathcal{C}_n} \exp\left(\sup_{x\in C} 2S_n(\phi(x))\right) \\ \geq & G^{-2}e^{-2nP(\phi)} \sum_{\substack{\sigma^n x = x\in C \\ C\in\mathcal{C}_n,\, C\subseteq[a]}} \exp\left(2S_n\phi(x)\right) = G^{-2}e^{-2nP(\phi)}F_n(2\phi, a).\end{aligned}$$

This implies

$$\liminf_{n\to\infty} \frac{1}{n}\log\sum_{C\in\mathcal{C}_n}\mu(C)^2 \geq P(2\phi) - 2P(\phi).$$

Then putting the inequalities for $\limsup$ and $\liminf$ together,

$$H_2 = \lim_{n\to+\infty} \frac{\log\sum_{C\in\mathcal{C}_n}\mu(C)^2}{-n} = 2P(\phi) - P(2\phi). \qquad \square$$

Remark 4.1.3. *It is also easy to see that $H_2(\nu) \leq 2h(\nu)$ for all ergodic invariant probability measure ν: for all x and all $n \in \mathbb{N}$,*

$$\frac{\log F_n(1)}{-n} \leq \frac{2\log\nu(C_n(x))}{-n}.$$

By the Shannon-McMillan-Breiman Theorem, the left hand side converges to $2h(\nu)$ for almost every x, therefore $\limsup_n \frac{\log F_n(1)}{-n} \leq h(\nu)$. So the Rényi entropy is finite whenever the measure-theoretic entropy of ν is finite.

4.2 Longest substring matching of one-point orbit

Here we present the main theorem for almost sure asymptotic growth of $M_n(x)$.

Theorem 4.2.1. *Let $(\Sigma, \sigma, \phi, \mu)$ be a topologically mixing countable (or finite) Markov subshift with the BIP property, and ϕ a locally Hölder (or Hölder) potential admitting a Gibbs measure μ. Then for μ-a.e. $x \in \Sigma$,*

$$\lim_{n\to\infty} \frac{M_n(x)}{\log n} = \frac{2}{H_2(\mu)}. \tag{4.2.1}$$

We will use first and second moment estimation methods together with the Borel-Cantelli lemma to prove separately that

$$\limsup_{n\to+\infty} \frac{M_n(x)}{\log n} \leq \frac{2}{H_2(\mu)}, \tag{4.2.2}$$

$$\liminf_{n\to+\infty} \frac{M_n(x)}{\log n} \geq \frac{2}{H_2(\mu)}. \tag{4.2.3}$$

We will continue the habit of denoting a dimension-like object in proofs by α, but this time $\alpha = \frac{H_2}{2}$. The following lemma is crucial for approximating the values of $F_n(t)$.

Lemma 4.2.2. *For a countable Markov shift $(\Sigma, \sigma, \phi, \mu)$ satisfying the assumptions of Theorem 4.2.1, we* *have $\alpha > 0$ and*

$$F_k(1) = \sum_{C \in \mathcal{C}_k} \mu(C)^2 \approx e^{-2k\alpha}, \tag{4.2.4}$$

and for each $t > 2$,

$$F_k(t-1) = \sum_{C \in \mathcal{C}_k} \mu(C)^t \preceq e^{-tk\alpha}. \tag{4.2.5}$$

Proof. Let $b_n := \max_{C \in \mathcal{C}_n} \mu(C)$, then $\sum_{C \in \mathcal{C}_n} \mu(C)^2 \leq b_n \sum_{C \in \mathcal{C}_n} \mu(C) \leq K_0 \beta^n$ where $K_0 > 0$, $\beta \in (0,1)$ were given by Lemma 1.1.6, hence

$$\limsup_{n \to \infty} \frac{\log F_n(1)}{-n} \geq \limsup_{n \to \infty} \frac{-\log b_n}{-n} \geq -\log \beta > 0.$$

The approximation formulae (4.2.4)(4.2.5) are from [CGR, Lemma 2.13]. They were originally proved for finite alphabets and the proof remains valid for countable cases if one combines with [HV, Theorem 1 (IV)], which holds whenever the relevant measure admits exponential decay of cylinder measures. □

4.2.1 Proof of Theorem 4.2.1

We first prove the upper bound which requires approximating the measure of points of certain recurrence times, and a first moment summation. The proof for lower bound is similar but involves a second moment argument.

Proof of upper bound (4.2.2). Set $\alpha = \frac{H_2(\mu)}{2}$ and

$$r_n := \frac{1}{\alpha - \varepsilon} \left(\log n + \log \log n \right).$$

As $M_n(x) = r_n$ implies the return time of some iterate of x under σ to some r_n-cylinder is strictly less than n, we need to approximate the size of short return sets in the system in order to apply Borel-Cantelli Lemmas to obtain almost everywhere statements. Hence, as in [HV] and [CGR], we intend to solve this by considering different cases of overlapping between r_n-substrings in x.

Overlapping Analysis

Let $n \in \mathbb{N}$. If r_n is not an integer, we simply take the closest integer. For each $k \in \mathbb{N}$, define the following auxiliary sets.

$$S_k(r_n) = \left\{ x \in \Sigma : \sigma^k x \in C_{r_n}(x) \right\}.$$

To put it into words, $S_k(r_n)$ is the set of points whose return time to the r_n-cylinder containing itself is k. Then by construction,

$$\begin{aligned}\mu\left(x:\ M_n(x)\geq r_n\right)&=\mu\left(\left\{x\in\Sigma:\exists\,0\leq i\leq n-1,\ 1\leq k\leq n-i-1\text{ s.t. }d_s(\sigma^i x,\sigma^{i+k}x)\leq 2^{-r_n}\right\}\right)\\&\leq\mu\left(\bigcup_{i=0}^{n-1}\bigcup_{k=1}^{n-i-1}\sigma^{-i}S_k(r_n)\right),\end{aligned}\tag{4.2.6}$$

where $d_s(\cdot,\cdot)$ is the symbolic metric defined in (1.1.2).

In order to obtain good estimates of $\mu(S_k(r_n))$, we consider three separate cases according to the values of k. Let

$$\Sigma_0:=\mu\left(\bigcup_{i=0}^{n-1}\bigcup_{k=1}^{\lfloor r_n/2\rfloor}\sigma^{-i}S_k(r_n)\right).$$

Similarly, set

$$\Sigma_1:=\mu\left(\bigcup_{i=0}^{n-1}\bigcup_{k=\lfloor r_n/2\rfloor+1}^{r_n}\sigma^{-i}S_k(r_n)\right),$$

and

$$\Sigma_2:=\mu\left(\bigcup_{i=0}^{n-1}\bigcup_{k=r_n+1}^{n-i-1}\sigma^{-i}S_k(r_n)\right).$$

Then (4.2.6) is replaced by

$$\mu(M_n\geq r_n)\leq\Sigma_0+\Sigma_1+\Sigma_2.\tag{4.2.7}$$

We will show that the measure of $\Sigma_0,\ \Sigma_1$ are insignificant compared with Σ_2.

Σ_0**: return time** $1\leq k\leq\lfloor r_n/2\rfloor$

Notation. For any finite k-subword of x starting from j, $x_j,x_{j+1},\dots,x_{j+k-1}$, write $x(j,k)$. For each $\omega\in\mathbb{N}$, $x(j,k)^\omega$ means that particular subword is repeated ω-times consecutively whenever it is allowed by the transition matrix.

Let $\omega_k=\lfloor\frac{r_n}{k}\rfloor$ and $0\leq\gamma_k<k$ so that $r_n=k\omega_k+\gamma_k$. Then if $x\in\sigma^{-i}S_k(r_n)$, $x_j=x_l$ if $j=l$ (mod k) for all $j,\ l\in[i,i+r_n+k-1]$, therefore $\sigma^i x$ has the following form:

$$\sigma^i(x)=(\underbrace{x(i,k),x(i,k),\dots,x(i,k)}_{k\text{-word repeating }\omega_k+1\text{ times}},x(i,\gamma_k),\dots)=\left(x(i,k)^{\omega_k+1},x(i,\gamma_k),\dots\right).$$

That is, a k-word $(x_i,\dots,x_{i+k-1})$ will be repeated fully for ω_k+1 times, followed by a truncated γ_k-word with the same initial symbols.

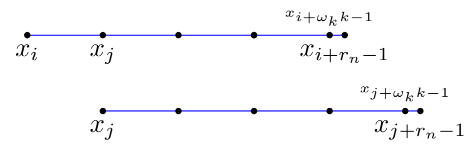

Figure 4.1: Overlapping for points in Σ_0; each segment stands for one copy of the repeated word.

Note that for each given i and all $k < k' \leq \lfloor r_n/2 \rfloor - i$, if k divides k', then $\sigma^{-i} S_k(r_n) \subset \sigma^{-i} S_{k'}(r_n)$. Since for each $k \leq \lfloor r_n/2 \rfloor$, there is some minimum $\ell_k \in [\lceil r_n/4 \rceil, \lfloor r_n/2 \rfloor]$ such that ℓ_k is a multiple of k so that $x \in \sigma^{-i} S_k(r_n) \subseteq \sigma^{-i} S_{\ell_k}(r_n)$ where the ℓ_k word is fully repeated $\omega_{\ell_k} + 1 \leq 5$ times.

Recall that μ is ψ-mixing, and so it is quasi-Bernoulli in the sense for some uniform constant $B > 1$, for any words $\underline{w}_1, \underline{w}_2$ such that $[\underline{w}_1 \underline{w}_2] \neq \emptyset$, $\mu([\underline{w}_1 \underline{w}_2]) \leq B \mu([\underline{w}_1]) \mu([\underline{w}_2])$. Therefore for each $k \in [\lfloor r_n/4 \rfloor, \lfloor r_n/2 \rfloor]$, by the quasi-Bernoulli property of ψ-mixing measures,

$$\mu\left(\{x : \sigma^{i+k}(x) \in C_{r_n}(\sigma^i x)\}\right) = \mu(S_k(r_n)) \leq \mu(S_{\ell_k}(r_n)) \preceq B^6 \sum_{C_{\ell_k} \in \mathcal{C}_\ell} \mu(C_\ell)^{\omega_{\ell_k}+1} \beta^{\gamma_k} \leq B^6 F_{\ell_k}(\omega_{\ell_k}),$$

where β is a given by Lemma 1.1.6. As $r_n \leq \ell_k(\omega_{\ell_k} + 1)$, $e^{-\alpha \ell_k \omega_k} \leq e^{-\alpha r_n}$, by definition of r_n and (4.2.5),

$$F_{\ell_k}(\omega_k) \preceq e^{-\alpha r_n} \leq \exp(-\log n - \log\log n) \leq \frac{1}{n \log n}.$$

To reduce again the redundant terms we have to sum up for Σ_0, for each $i \leq n-1$ and $1 \leq k \leq \lfloor r_n/2 \rfloor$, we can omit $\sigma^{-i} S_k(r_n)$ if $2k \leq n-i-1$, which follows from again $\sigma^{-i} S_k(r_n) \subseteq \sigma^{-i} S_{2k}(r_n)$. Considering all those discussed above, for Σ_0 we only need to consider points x such that $\sigma^i x$ has short return time *i.e.*, $k \leq r_n/2$ and $i \geq n - r_n$. As r_n is of the scale of $\log n$, we may choose n large such that $r_n \leq n^{1/2}$ so that by Lemma 4.2.2,

$$\Sigma_0 = \sum_{i \geq n-r_n} \sum_{k \geq 1} \sigma^{-i} S_k(r_n) \leq B^6 \sum_{k=1}^{\lfloor r_n/2 \rfloor} k F_{\ell_k}(\omega_{\ell_k}) \preceq B^6 r_n^2 e^{-\alpha r_n} \leq \frac{B^6 r_n^2}{n \log n} \preceq \frac{1}{\log n}. \tag{4.2.8}$$

Σ_1**: return time** $\lfloor r_n/2 \rfloor + 1 \leq k \leq r_n$

In this case, $x \in \sigma^{-i} S_k(r_n)$ implies $x_j = x_l$ if $j = l \pmod{k}$ for all $j, l \in [i, i + r_n + k - 1]$, hence $\sigma^i(x)$ has the form

$$\sigma^i(x) = (x(i, r_n - k), x(i + r_n - k, 2k - r_n), x(i, r_n - k), x(i + r_n, 2k - r_n), x(i, r_n - k), x_{i+r_n+k}, \ldots\ldots)$$

shown by the following illustration (same colour implies the same subword repeated), so the $(r_n - k)$ word starting from x_i is repeated three times, separated by two identical $(2k - r_n)$ words. Hence by

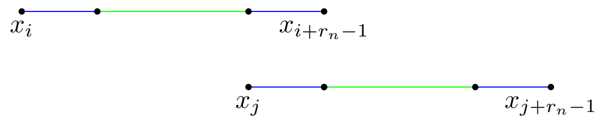

Figure 4.2: Overlapping of subwords for points in Σ_1

Lemma 4.2.2 and the quasi-Bernoulli property,

$$\mu(\sigma^{-i}S_k(r_n)) = \mu(S_k(r_n)) \leq B^6 \sum_{\substack{C\in\mathcal{C}_{r_n-k}\\ D\in\mathcal{C}_{2k-r_n}}} \mu(C)^3\mu(D)^2 = B^6 F_{r_n-k}(2)F_{2k-r_n}(1).$$

Then we obtain an upper bound for Σ_1:

$$\begin{aligned}
\Sigma_1 \leq & B^6 \sum_{k=\lfloor r_n/2\rfloor+1}^{r_n} (n-k)F_{r_n-k}(2)F_{2k-r_n}(1) \preceq B^6 \sum_{k=\lfloor r_n/2\rfloor+1}^{r_n} (n-k)e^{-\alpha 3(r_n-k)}e^{-\alpha 2(2k-r_n)} \\
= & B^6 \sum_{k=\lfloor r_n/2\rfloor+1}^{r_n} (n-k)e^{-\alpha(r_n+k)} \preceq B^6 e^{-\frac{3}{2}r_n\alpha} \sum_{k=\lfloor r_n/2\rfloor+1}^{r_n} (n-k) \preceq r_n n e^{-\frac{3}{2}\alpha r_n} \\
\leq & \frac{nr_n}{(n\log n)^{3/2}} \leq \frac{n^{1/2}n}{(n\log n)^{3/2}} \leq \frac{1}{\log n}.
\end{aligned} \tag{4.2.9}$$

Σ_2**: return time** $r_n+1 \leq k \leq n-i-1$

In this case, $k-r_n \geq 1$ and $x \in \sigma^{-i}S_k(r_n)$ implies $x(i,r_n)$, the r_n-word starting from position i of x, is repeated from the $i+k$ entry without any overlapping with itself, i.e.

$$\sigma^i(x) = (\underbrace{x(i,r_n), x(i+r_n, k-r_n), x(i,r_n)}_{r_n\text{-word repeated with } k-r_n \text{ gap}}, \ldots).$$

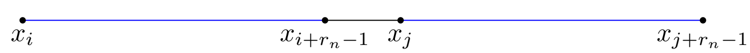

Figure 4.3: Overlapping for points in Σ_2; black segment represents an arbitrary word of length r_n-k

Then by the ψ-mixing property, the measure of such set of points is bounded by

$$\mu\left(\sigma^{-i}S_k(r_n)\right) \leq (1+\psi(k-r_n)) \sum_{C\in\mathcal{C}_{r_n}} \mu(C)^2 = (1+\psi(k-r_n))F_{r_n}(1).$$

Hence

$$\Sigma_2 \leq \sum_{k=r_n+1}^{n-1} (n-k)\mu(S_k(r_n)) \leq \sum_{k=r_n+1}^{n-1} (n-k)(1+\psi(k-r_n))F_{r_n}(1).$$

Again $(1+\psi(k))$ is monotonically decreasing in k,

$$\begin{aligned}\Sigma_2 \approx e^{-2\alpha r_n} &\sum_{k=r_n+1}^{n-1}(n-k)(1+\psi(k-r_n)) \le (1+\psi(1))e^{-2\alpha r_n}\sum_{k=r_n+1}^{n-1}(n-k)\\ \le&(1+\psi(1))e^{-2\alpha r_n}n^2 \le \frac{1+\psi(1)}{(\log n)^2} \le \frac{1+\psi(1)}{\log n}.\end{aligned} \tag{4.2.10}$$

Then, combining (4.2.7)-(4.2.10), there is some constant $K_1 > 0$ independent of n such that

$$\mu\left(\{M_n > r_n\}\right) \le K_1 \frac{1}{\log n}.$$

Using the technique in the proof of [BLR, Theorem 5], we pick a subsequence $n_k = e^{\lceil k^2 \rceil}$ so for all k large enough,

$$\mu(\{M_{n_k} > r_{n_k}\}) \le K_1 \frac{1}{k^2}.$$

Then by the Borel-Cantelli Lemma, for μ-almost every $x \in \Sigma$,

$$M_{n_k}(x) \le r_{n_k},$$

which implies that for all k large enough,

$$\frac{M_{n_k}(x)}{\log n_k} \le \frac{1}{\alpha-\varepsilon}\left(1+\frac{\log\log n_k}{\log n_k}\right).$$

Since $M_n(x)$ is non-decreasing in n for all x, for each n, there is a unique k such that $n_k \le n < n_{k+1}$. In particular,

$$\frac{\log n_k}{\log n_{k+1}}\cdot\frac{M_{n_k}(x)}{\log n_k} \le \frac{M_n(x)}{\log n} \le \frac{M_{n_{k+1}}(x)}{\log n_{k+1}}\cdot\frac{\log n_{k+1}}{\log n_k}, \tag{4.2.11}$$

As $\lim_{k\to+\infty}\frac{\log n_{k+1}}{\log n_k} = 1$ and $\lim_{k\to+\infty}\frac{\log\log n_k}{\log n_k} = 0$, taking the $\limsup$ of the inequalities above,

$$\limsup_{n\to+\infty}\frac{M_n(x)}{\log n} = \limsup_{n\to+\infty}\frac{M_{n_k}(x)}{\log n_k} \le \frac{1}{\alpha-\varepsilon}.$$

(4.2.2) is proved by since $\varepsilon > 0$ is arbitrarily small. □

Proof of lower bound (4.2.3). We apply a similar second-moment analysis as in the proof of [CGR, Theorem 4.1]. Let

$$r_n = \frac{1}{\alpha+\varepsilon}\left(\log n + \lambda\log\log n\right)$$

for some uniform constant $\lambda < 0$ to be determined later. Since $\sigma^{i+k}x \in C_{r_n}(\sigma^i x)$ if and only if $x \in \sigma^{-i}S_k(r_n)$, then we can define the random variable $\mathcal{S}_n$:

$$\mathcal{S}_n(x) := \sum_{i=0}^{n-2r_n-1}\sum_{k=2r_n}^{n-i-1}\mathbf{1}_{\{C_{r_n}(\sigma^i x)\}}(\sigma^{i+k}x) = \sum_{i=0}^{n-2r_n-1}\sum_{k=2r_n}^{n-i-1}\mathbf{1}_{\{\sigma^{-i}S_k(r_n)\}}(x), \tag{4.2.12}$$

which counts the number of times that x belongs to some $\sigma^{-i}S_k(r_n)$. As $M_n(x) < r_n$ implies for all $0 \leq i \leq n-1$, $1 \leq k \leq n-i-1$, $x \notin \sigma^{-i}S_k(r_n)$, and in particular not in the $\sigma^{-i}S_k(r_n)$ sets with $k \geq 2r_n$,

$$\{x : M_n(x) < r_n\} \subseteq \{x : S_n(x) = 0\}.$$

By the Paley-Zygmund inequality,

$$\mu(\{M_n < r_n\}) \leq \mu(\{\mathcal{S}_n = 0\}) = 1 - \mu\left(\{S_n > 0\}\right) \leq \frac{\mathrm{Var}[\mathcal{S}_n]}{\mathbb{E}[\mathcal{S}_n^2]} \leq \frac{\mathrm{Var}[\mathcal{S}_n]}{\mathbb{E}[\mathcal{S}_n]^2}. \tag{4.2.13}$$

By definition of $\sigma^{-i}S_k(r_n)$ with $k \geq 2r_n$, this set corresponds to the set of points in which an r_n-word repeats itself at least once with at least an r_n gap, therefore we have the following lower bound using the ψ-mixing property,

$$\begin{aligned}\mu\left(\{C_{r_n}(\sigma^i x) = C_{r_n}(\sigma^{i+k}x)\}\right) =& \mu\left(\sigma^{-i}S_k(r_n)\right) = \sum_{C \in \mathcal{C}_{r_n}} \mu(C \cap \sigma^{-k}C) \\ \geq& (1-\psi(k-r_n)) \sum_{C \in \mathcal{C}_{r_n}} \mu(C)^2 \geq (1-\psi(r_n))F_{r_n}(1).\end{aligned}$$

Therefore, as $\sum_{i=0}^{n-2r_n-1}\sum_{k=2r_n}^{n-i-1} 1 = (n-2r_n-1) + (n-2r_n-2) + \cdots 1$,

$$\mathbb{E}[\mathcal{S}_n] \geq \frac{1}{2}(1-\psi(r_n))(n-2r_n)^2 F_{r_n}(1). \tag{4.2.14}$$

Next, we need to consider

$$\mathbb{E}[\mathcal{S}_n^2] = \sum_{i,j=0}^{n-2r_n-1} \sum_{k=2r_n}^{n-i-1} \sum_{l=2r_n}^{n-j-1} \mu(\sigma^{-i}S_k(r_n) \cap \sigma^{-j}S_l(r_n)). \tag{4.2.15}$$

Define the index set

$$I := \left\{(i,j,k,l) \in \mathbb{N}^4 : 0 \leq i,\, j \leq n-2r_n-1, 2r_n \leq k \leq n-i-1,\, 2r_n \leq l \leq n-j-1\right\},$$

then

$$\mathbb{E}[\mathcal{S}_n^2] = \sum_{(i,j,k,l)\in I} \mu(\sigma^{-i}S_k(r_n) \cap \sigma^{-j}S_l(r_n)), \tag{4.2.16}$$

and the cardinality of I satisfies

$$\#I = \left(\sum_{i=2r_n}^{n-2r_n-1} n-i\right)\left(\sum_{j=2r_n}^{n-2r_n-1} n-j\right) \leq \frac{1}{4}(n-2r_n)^4.$$

Define the counting function by

$$\theta : I \to \mathbb{N}, \quad \theta(i,j,k,l) = \sum_{\substack{a\in\{i,i+k\}\\ b\in\{j,j+l\}}} \mathbf{1}_{(a-r_n,a+r_n)}(b),$$

i.e., it counts the occurrences that two indices in $\{i, j, i+k, j+l\}$ are r_n-close to each other. $\theta > 0$ implies there are overlaps between some r_n words, e.g. $|i-j| < r_n$ implies the r_n word $x(i, r_n)$ overlaps with the r_n word $x(j, r_n)$, and both r_n-strings are repeated later.

By our definition of $\mathcal{S}_n$, for each quadruple (i, j, k, l), necessarily $k, l \geq 2r_n$ which implies

$$0 \leq \theta(i,j,k,l) \leq 2, \qquad \forall (i,j,k,l) \in I,$$

which allows us to split (4.2.16) again into three components,

$$\mathbb{E}[\mathcal{S}_n^2] = \left(\sum_{I_0} + \sum_{I_1} + \sum_{I_2} \right) \mu \left(\sigma^{-i} S_k(r_n) \cap \sigma^{-j} S_l(r_n) \right),$$

where $I_t = \{(i,j,k,l) \in I : \theta(i,j,k,l) = t\}$.

Clearly,

$$\#I_0 \leq \#I \leq \frac{1}{4}(n-2r_n)^4.$$

For each $(i,j,k,l) \in I_1$, if we fix any three indices, for example, if i, j, k are fixed, $j+l$ can be r_n-close to either i or $i+k$ as it is automatically $2r_n$-apart from j, hence there are at most $4r_n$ choices for the remaining index l. Hence

$$\#I_1 \leq 2r_n(n-2r_n)^3,$$

and similarly if we fix any two of i, j, k, l in I_2, there are at most $2r_n^2$ choices for the remaining two indices, therefore

$$\#I_2 \leq 2r_n^2(n-2r_n)^2.$$

Contributions of indices in I_0:

We will consider the sum over indices in I_0 first. Since $(i,j,k,l) \in I_0$ implies no overlapping, $x \in \sigma^{-i} S_k(r_n) \cap \sigma^{-j} S_l(r_n)$ implies $x(i, r_n) = x(i+k, r_n)$ and $x(j, r_n) = x(j+l, r_n)$ while the symbols in these two r_n-strings are independent, e.g. when $i+k < j$, x has the following form:

$$\sigma^i(x) = (x(i,r_n), \ldots, x(i,r_n), \ldots, x(j,r_n), \ldots, x(j,r_n), \ldots).$$

Hence by ψ-mixing property

$$\mu(\sigma^{-i} S_k(r_n) \cap \sigma^{-j} S_l(r_n)) \leq (1+\psi(\gamma_{i\,j\,k\,l}))^3 F_{r_n}(1)^2, \tag{4.2.17}$$

where

$$\gamma_{i\,j\,k\,l} = \min\left\{ |a-b| - r_n : a, b \in \{i, j, i+k, j+l\} \right\}.$$

Let $I_0' \subseteq I_0$ be defined as

$$I_0' := \{(i,j,k,l) \in I_0 : \gamma_{ijkl} \geq r_n\},$$

and $I_0'' := I_0 \setminus I_0'$. Notice also that $\#(I_0'') \leq 2r_n(n-2r_n)^3$.

Define the notation for any $G \subseteq I$,

$$\mathbb{E}[\mathcal{S}_n^2|G] := \sum_{i,j,k,l\in G} \mu\left(\sigma^{-i}S_k(r_n)\cap\sigma^{-j}S_l(r_n)\right).$$

Then, using (4.2.17),

$$\mathbb{E}[\mathcal{S}_n^2|I_0'] = \sum_{i,j,k,l\in I_0'} \mu\left(\sigma^{-i}S_k(r_n)\cap\sigma^{-j}S_l(r_n)\right) \leq (n-2r_n)^4(1+\psi(r_n))^3F_{r_n}(1)^2,$$

By Lemma 1.3.7 $\psi(r_n) \leq r_n^{-1}$ for all n large enough, then

$$\begin{aligned}&(1+\psi(r_n))^3-(1-\psi(r_n))^2 \leq (1+\psi(r_n))^3-(1-\psi(r_n))^3\\&=6\psi(r_n)+2\psi(r_n)^3 \leq 8r_n^{-1} \leq \frac{8(\alpha+\varepsilon)}{\log n+\lambda\log\log n} \preceq \frac{1}{\log n}.\end{aligned} \tag{4.2.18}$$

It is easy to see that $\#(I_0') \leq \frac{1}{4}(n-2r_n)^4$. Using (4.2.14) with (4.2.18), as $(1-\psi(r_n))^2 \geq \left(1-r_n^{-1}\right)^2 \geq \frac{1}{4}$, for some constant $K_2>0$,

$$\begin{aligned}\frac{\mathbb{E}[\mathcal{S}_n^2|I_0']-\mathbb{E}[\mathcal{S}_n]^2}{\mathbb{E}[\mathcal{S}_n]^2} &\leq \frac{(n-2r_n)^4F_{r_n}(1)^2\left((1+\psi(r_n))^3-(1-\psi(r_n))^2\right)}{(n-2r_n)^4(1-\psi(r_n))^2F_{r_n}(1)^2}\\&\leq \frac{(1+\psi(r_n))^3-(1-\psi(r_n))^3}{(1-\psi(r_n))^2} \leq K_2\frac{1}{\log n}.\end{aligned} \tag{4.2.19}$$

For the sum over I_0'', the term $1+\psi(\gamma_{ijkl})$ in (4.2.17) is uniformly bounded above by $1+\psi(0)$, and $1-\psi(r_n) \geq \frac{1}{2}$ for all n sufficiently large, therefore

$$\frac{\mathbb{E}[\mathcal{S}_n^2|I_0'']}{\mathbb{E}[\mathcal{S}_n]^2} \approx \frac{r_n(n-2r_n)^3(1+\psi(0))^3F_{r_n}(1)^2}{(1-\psi(r_n))^2(n-2r_n)^4F_{r_n}(1)^2} \preceq \frac{r_n}{n-2r_n}$$

hence for some $K_3>0$ and all n sufficiently large,

$$\frac{\mathbb{E}[\mathcal{S}_n^2|I_0'']}{\mathbb{E}[\mathcal{S}_n]^2} \leq K_3\frac{1}{\log n} \tag{4.2.20}$$

Contributions of indices in I_1:

Next, for $(i,j,k,l)\in I_1$, without loss of generality, suppose only $|i-j|=r<r_n$, $i<j$ and $i+k<j+l$. The other cases are treated exactly the same since the order of the r_n-strings does not have any effects on estimations of the upper bounds for $\mu(\sigma^{-i}S_k(r_n)\cap\sigma^{-j}S_l(r_n))$.

An $x\in\sigma^{-i}S_k(r_n)\cap\sigma^{-j}S_l(r_n)$ means $x_{i+r}=x_j$, $x_{i+r+1}=x_{j+1}$, $\ldots$, $x_{i+r_n}=x_{j+r}$, so $\sigma^i x$ has the following form:

$$\begin{aligned}\sigma^i(x) &= (x_i, \dots, x_{i+r-1}, x_j, \dots, x_{j+r_n-1}, \dots, x_{i+k}, \dots, x_{j+l}, \dots)\\ &= (x(i,r), x(j, r_n - r), x(i + r_n, r), \dots, x(i,r), x(j, r_n - r), \dots, x(j, r_n - r), x(i + r_n, r), \dots).\end{aligned}$$

Again using the quasi-Bernoulli property and (4.2.4),(4.2.5), for $B \geq 1$ the relevant quasi-Bernoulli constant, since there are nine concatenations of words as explained above (we can count the number of commas),

$$\begin{aligned}\mu(\sigma^{-i}S_k(r_n) \cap \sigma^{-j}S_l(r_n)) &\leq B^9 \sum_{\substack{A,B \in \mathcal{C}_r\\ C \in \mathcal{C}_{r_n - r}}} \mu(A)^2 \mu(B)^2 \mu(C)^3 = B^9 F_r(1)^2 F_{r_n - r}(2)\\ &\preceq B^9 e^{-4\alpha r_n} e^{-3\alpha(r_n - r)} \leq B^9 e^{-3\alpha r_n}.\end{aligned}$$

Recall that $r_n = \frac{1}{\alpha+\varepsilon}(\log n + \lambda \log\log n)$ and

$$e^{-\alpha r_n} = \left(n(\log n)^\lambda\right)^{-\frac{\alpha}{\alpha+\varepsilon}} = n^{-\frac{\alpha}{\alpha+\varepsilon}} (\log n)^{-\frac{\alpha\lambda}{\alpha+\varepsilon}}. \tag{4.2.21}$$

For all n large enough such that

$$\frac{n^{\frac{\alpha}{\alpha+\varepsilon}}}{(n - 2r_n)} \leq 1, \tag{4.2.22}$$

by (4.2.14) and (4.2.22) above,

$$\begin{aligned}\frac{\mathbb{E}[\mathcal{S}_n^2 | I_1]}{\mathbb{E}[\mathcal{S}_n]^2} &\preceq \frac{2r_n (n - 2r_n)^3 B^{10} e^{-3\alpha r_n}}{(n - 2r_n)^4 (1 - \psi(r_n))^2 F_{r_n}(1)^2}\\ &\approx \frac{r_n e^{-3\alpha r_n}}{(n - 2r_n) e^{-4\alpha r_n}} = \frac{1}{\alpha+\varepsilon} \frac{\log n + \lambda \log\log n}{(n - 2r_n) e^{-\alpha r_n}}\\ &\preceq \frac{\log n}{(n - 2r_n) e^{-\alpha r_n}} = \frac{n^{\frac{\alpha}{\alpha+\varepsilon}}}{n - 2r_n} (\log n)^{1 + \frac{\lambda\alpha}{\alpha+\varepsilon}}\\ &\leq (\log n)^{\left(1 + \lambda(1 - \frac{\varepsilon}{\alpha+\varepsilon})\right)}.\end{aligned}$$

For all $0 < \varepsilon \leq \alpha$, $1 - \frac{\varepsilon}{\alpha+\varepsilon} \geq \frac{1}{2}$, so we can choose $\lambda = -4$ so that $1 + \lambda\left(1 - \frac{\varepsilon}{\alpha+\varepsilon}\right) \leq -1$. It is then sufficient to conclude that for some constant $K_4 > 0$,

$$\frac{\mathbb{E}[\mathcal{S}_n^2 | I_1]}{\mathbb{E}[\mathcal{S}_n]^2} \leq K_4 \frac{1}{\log n}. \tag{4.2.23}$$

Contributions of indices in I_2:

Finally for indices in I_2, an $x \in \sigma^{-i}S_k(r_n) \cap \sigma^{-j}S_l(r_n)$ has very complicated overlapping behaviour in the subwords $x(i, r_n)$ and $x(j, r_n)$. But this can be trivially and happily reduced to considering that some r_n-subword is repeated twice from the $i-$th entry and the $(i+k)$-th entry without overlapping each other. Using Lemma 4.2.2, we bound the measure of $\sigma^{-i}S_k(r_n) \cap \sigma^{-j}S_l(r_n)$ for each $(i,j,k,l) \in I_2$

by

$$B^4 \sum_{C\in\mathcal{C}_{r_n}} \mu(C)^2 \approx B^4 e^{-2\alpha r_n}.$$

The number of indices in I_2 can be bounded by $2r_n^2(n-2r_n)^2$. Then for $\lambda = -4$ and all n verifying (4.2.22), by (4.2.14), (4.2.21) and the fact that $r_n \preceq \log n$,

$$\begin{aligned}\frac{\mathbb{E}[\mathcal{S}_n^2|I_2]}{\mathbb{E}[\mathcal{S}_n]^2} &\preceq \frac{2r_n^2(n-2r_n)^2B^4e^{-2\alpha r_n}}{(n-2r_n)^4(1-\psi(r_n))^2e^{-4\alpha r_n}}\\ &\approx \frac{r_n^2}{(n-2r_n)^2e^{-2\alpha r_n}} \preceq \frac{n^{\frac{2\alpha}{\alpha+\varepsilon}}}{(n-2r_n)^2}(\log n)^{2+2\frac{\lambda\alpha}{\alpha+\varepsilon}}\\ &\leq (\log n)^{2(1+\lambda(1-\frac{\varepsilon}{\alpha+\varepsilon}))} \leq (\log n)^{-2}.\end{aligned}$$

It follows that for some constant $K_5 > 0$,

$$\frac{\mathbb{E}[\mathcal{S}_n^2|I_2]}{\mathbb{E}[\mathcal{S}_n]^2} \leq K_5\frac{1}{(\log n)^2}. \tag{4.2.24}$$

Then, combining (4.2.13) (4.2.16) (4.2.19)-(4.2.24), there is some constant $K_6 > 0$ such that

$$\mu(\{M_n < r_n\}) \leq K_6\frac{1}{\log n}.$$

We can repeat the trick of picking a subsequence $n_k = \lceil e^{k^2}\rceil$, and apply the Borel-Cantelli Lemma to the sum $\sum_{k=1}^{\infty}\mu(\{M_{n_k} < r_{n_k}\}) < +\infty$, which means for all k large enough,

$$\frac{M_{n_k}(x)}{\log n_k} \geq \frac{1}{\alpha+\varepsilon}\left(1-\frac{4\log\log n_k}{\log n_k}\right).$$

Taking the liminf on both sides and applying the subsequence argument (4.2.11),

$$\liminf_{n\to+\infty}\frac{M_n(x)}{\log n} = \liminf_{n\to+\infty}\frac{M_{n_k}(x)}{\log n_k} \geq \frac{1}{\alpha+\varepsilon}.$$

Thus (4.2.3) holds as $\varepsilon > 0$ is arbitrarily small. □

Therefore, for (Σ,ϕ,σ) a topologically mixing countable Markov subshift where ϕ is a locally Hölder potential and μ the corresponding unique Gibbs measure,

$$\lim_{n\to\infty}\frac{M_n(x)}{\log n} = \frac{2}{H_2(\mu)} \text{ for } \mu\text{-almost every } x\in\Sigma.$$

4.3 Shortest distance problem for Gibbs-Markov interval maps

We will say goodbye to our great friend symbolic dynamics and consider a general compact metric space (X,d) with $f : X \to X$. One can study the *shortest distance* between iterates of typical points. For two orbits generated by typical points, we wish to prove the limiting behaviours for the following quantity:

$$m_n(x,y) := \min_{0\le i,j\le n-1} d\left(f^i x, f^j y\right).$$

For symbolic systems with the natural symbolic metric d_s, $M_n(x,y) = -\log m_n(x,y)/\log 2$. The quantity $m_n(x,y)$ is related to various objects, *e.g.* the correlation integral, extremal value theory with clustering phenomena, shrinking target problems; and as cover times in Chapter 3, this problem is linked to hitting times:

$$\text{writing } W_r(x,y) := \inf\left\{j\ge 1 : f^j x \in B(y,r)\right\},\ W_r(x,y)\le n \text{ implies } m_n(x,y) < r.$$

Also note that if $m_n^{LY}(x,y) := \min_{0\le i<n} d\left(f^i x, f^i y\right)$, then one can see this is closely linked to Li-Yorke pairs and one might expect $\lim_{n\to\infty} \frac{m_n^{LY}(x,y)}{m_n(x,y)} = \frac{1}{2}$ almost surely.

Coming back to m_n, the first almost sure result concerning its asymptotic growth was given in [BLR], the same paper dealing with almost sure growth of substring matching, the authors proved the following theorem.

Theorem 4.3.1. *[BLR, Theorem 1,Theorem 3] Let (X,d,f,μ) be a probability preserving system such that $\underline{D}_2(\mu) > 0$. Then for $\mu\times\mu$ almost every (x,y),*

$$\limsup_{n\to\infty} \frac{\log m_n(x,y)}{-\log n} \le \frac{2}{\underline{D}_2(\mu)}.$$

Furthermore, if (X,d) is tight[2] and (f,μ) has polynomial decay of correlations (see Definition 4.3.3 below),

$$\liminf_{n\to\infty} \frac{\log m_n(x,y)}{-\log n} \ge \frac{2}{\overline{D}_2(\mu)}.$$

Subsequent research, just as in the previous section for substring matching in symbolic dynamics, proves analogous statements for multiple orbits and random dynamics [BR21][GRS]. The notion of decay of correlations will be defined in Definition 4.3.3. In the theorem above, the quantities $\underline{D}_2(\mu)$ and $\overline{D}_2(\mu)$ are called *correlation dimensions*, which are generalisations of the Rényi entropies.

Definition 4.3.2. *Let (X,d) be a metric space and μ a probability measure on the Borel sets in X. The* upper and lower correlation dimensions *of μ are*

$$\overline{D}_2(\mu) = \limsup_{r\to 0} \frac{\log\int \mu(B(x,r))\,d\mu(x)}{\log r},\quad \underline{D}_2(\mu) = \liminf_{r\to 0} \frac{\log\int \mu(B(x,r))\,d\mu(x)}{\log r}$$

respectively, and we write $D_2(\mu)$ when the two limits coincide.

[2] The terminology is inherited from [BLR], which is similar to the notion of doubling for metric spaces: a metric space is *tight* if there exists $r_0 > 0,\ N_0 \in \mathbb{N}$ such that for every $0 < r < r_0$ and all $x \in X$, the ball $B(x,2r)$ can be covered by at most N_0 number of r-balls. This is similar to the notion of *bounded local complexity* discussed before Lemma 4.3.5.

Clearly, if $X \subset \mathbb{R}$ and Leb is the Lebesgue measure then $D_2(Leb) = 1$. If μ is an acip with a bounded density function then $D_2(\mu) = 1$ as well. The aim of this section is to prove the following theorem, which is a twinned statement of Theorem 4.2.1 for Gibbs-Markov maps defined in Example 3.3.1, but we allow that the alphabet of the corresponding symbolic shift to be countably infinite. Let X be a closed interval in $\mathbb{R}$ and $f : X \to X$ a Gibbs-Markov map as in Example 3.3.1, with a countable partition. A Gibbs-Markov f admits an invariant probability measure μ with the Gibbs property, and by [You] μ has exponential decay of correlations and verifies the CLT.

Definition 4.3.3. *Say $f : X \to X$ has* exponential (or polynomial) decay of correlation *for $\mathcal{BV}$ against L^1 observables, where $\mathcal{BV} := \{f \in L^1(\mu) : f \text{ has bounded variation.}\}$, if there is $\rho : \mathbb{N} \to \mathbb{R}$ with $\rho(n) = C_1 e^{-c_1 n}$ or ($\rho(n) = C_1 n^{-c_1}$) for some C_1, $c_1 > 0$, and for all $\phi, \varphi : X \to \mathbb{R}$, $f \in \mathcal{BV}$ and $g \in L^1$,*

$$\left| \int \phi \cdot \varphi \circ f^n \, d\mu - \int \phi \, d\mu \int \varphi \, d\mu \right| \leq \|\phi\|_{\mathcal{BV}} \|\varphi\|_1 \rho(n)$$

where the norm $\|\phi\|_{\mathcal{BV}} := \|\phi\|_1 + \mathrm{TV}(\phi)$, and $\mathrm{TV}(\phi)$ is the total variation of ϕ. For $\mathbf{1}_E$ an indicator function of some measurable $E \subseteq X$, $\|\mathbf{1}_E\|_{\mathcal{BV}} = 2$ and $\|\mathbf{1}_E\|_1 = \mu(E)$.

A messier decay of correlations for multiple functions will be proved later in Lemma 4.3.12. The main theorem of this section is stated below.

Theorem 4.3.4. *Let X be a closed interval of $\mathbb{R}$, (X, f) a Gibbs-Markov system and μ a Gibbs probability measure admitting exponential decay of correlations for L^1 against $\mathcal{BV}$ observables. Then if its upper correlation dimension $\overline{D}_2(\mu)$ is bounded away from 0,*

$$\liminf_{n\to\infty} \frac{\log m_n(x)}{-\log n} \geq \frac{2}{\overline{D}_2(\mu)}$$

for μ–almost every x in the repeller Λ. If μ is absolutely continuous with respect to the Lebesgue measure, then

$$\lim_{n\to\infty} \frac{\log m_n(x)}{-\log n} = \frac{2}{D_2(\mu)} \ \mu\text{-a.e.}$$

In this case, as the invariant density (with respect to Lebesgue) is bounded, $\underline{D}_2(\mu) = \overline{D}_2(\mu) = 1$.

This theorem is applicable to the following systems.

Example 4.3.1 (k-multiplying maps). $f : [0, 1] \to [0, 1]$, $f(x) = kx \pmod 1$ for $k = 2, 3, \ldots$, and $\mu = Leb$. In these cases the uniform k-Bernoulli measure $\tilde{\mu}$ on $\{1, \ldots, k\}^{\mathbb{N}_0}$ satisfies $\mu = \pi_* \tilde{\mu}$ for all cylinder sets, and Bernoulli measures are clearly Gibbs and ψ-mixing.

Example 4.3.2 (Piecewise affine interval maps). Let $\{a_k\}_k$ be a monotone decreasing sequence with

$a_1 = 1$ and $\lim_k a_k = 0$. Then $f : [0,1] \to [0,1]$ with

$$f|_{[a_{k+1},a_k)} = \frac{1}{a_k - a_{k+1}} (x - a_{k+1})$$

satisfies the assumptions of Theorem 4.3.4.

Example 4.3.3 (Gauss Map). Define the Gauss map $G : [0,1] \to [0,1]$ by

$$G(x) = \begin{cases} \frac{1}{x} \text{ (mod 1)} & x \in (0,1], \\ 0 & x = 0. \end{cases}$$

It is a full-branched map. Let μ_G be the Gauss measure, which is the Gibbs measure for the potential $-\log DF$ with density $\frac{d\mu_G}{dLeb} = \frac{1}{(1+x)\log 2}$ which is bounded for all $x \in (0,1]$. Then Theorem 4.3.4 holds for (G, μ_G).

Example 4.3.4 (An induced map). Let F be the first return function to $[0, \frac{1}{2})$ of a Manneville–Pomeau map $f : [0,1] \to [0,1]$:

$$f(x) = \begin{cases} x(1 + 2^a x^a) & x \in [0, 1/2), \\ 2x - 1 & x \in [1/2, 1). \end{cases}$$

for $a \in (0,1)$. There exists μ_F a Gibbs probability measure with respect to the potential $-\log |DF|$ (see [LSV]) and is an acip.

As in the symbolic setting, the one-point orbit case involves short return behaviour which complicates things slightly: approximating short return to balls is crucial for obtaining the upper bound of $\frac{\log m_n(x)}{-\log n}$, that is also generally harder than the recurrence analysis of cylinders. Another challenge appearing here but not for the symbolic setting is that the open balls defined by the Euclidean metric and the cylinders generated by the natural partitions disagree, and for points located on the boundary of adjacent cylinder sets, their symbolic representation may not be unique.

For Λ the repeller (see (3.2.1) for definition) of a Gibbs-Markov map f and $d(x,y) = |x-y|$, (Λ, d) trivially satisfies the *bounded local complexity* condition: (X,d) has bounded local complexity if there exists $C_0 \in \mathbb{N}$ such that for each $r > 0$, there is $k(r) < \infty$, and $\{x_1^r, x_2^r, \ldots, x_{k(r)}^r\} \subseteq X$ such that

$$X \subset \bigcup_{p=1}^{k(r)} B\left(x_p^r, r\right)$$

and each $x \in X$ belongs to at most C_0 elements of $\{B(x_p^r, 2r)\}_{p=1}^{k(r)}$. Any compact subset of $\mathbb{R}$ has bounded local complexity: compact implies totally bounded which gives $k(r) < \infty$ for all r and C_0 can be chosen to be 4 because one can choose an r-net such that $d(x_i^r, x_j^r) \geq r$ for $i \neq j \in \{1, \ldots, k(r)\}$. The property verifies an alternative way to compute $D_2(\mu)$. The following lemmas are analogous to

[GRS, Lemma 12, Lemma 13].

Lemma 4.3.5. *For all $x, y \in X$, let $\mathbf{1}_{p,r} := \mathbf{1}_{B\left(x_p^r, 2r\right)}$. If X has bounded local complexity,*

$$\mathbf{1}_{B(x,r)}(y) \leq \sum_{p=1}^{k(r)} \mathbf{1}_{p,r}(x)\mathbf{1}_{p,r}(y) \leq C_0 \mathbf{1}_{B(x,4r)}(y)$$

Proof. Given $r > 0$ and $x \in X$, as $\cup_p B\left(x_p^r, r\right)$ is a cover of X, there is at least one $p \in \{1, \ldots, k(r)\}$ such that $d(x, x_p^r) < r$. Then if $d(x, y) < r$, $d(y, x_p^r) < 2r$ hence $1 = \mathbf{1}_{p,r}(x)\mathbf{1}_{p,r}(y)$ and the left inequality is proved.

On the other hand, by bounded local complexity, there are at most C_0 elements of $p \in \{1, \ldots, k(r)\}$ such that $\mathbf{1}_{p,r}(x) \neq 0$, and for each such p, $\mathbf{1}_{p,r}(y) \neq 0$ implies $d(x, y) < 4r$, which proves the right inequality. □

Lemma 4.3.6. *The following identities hold.*

$$\limsup_{r\to 0} \frac{\log \sum_{p=1}^{k(r)} \left(\int \mathbf{1}_{p,r} d\mu\right)^2}{\log r} = \overline{D}_2(\mu), \quad \liminf_{r\to 0} \frac{\log \sum_{p=1}^{k(r)} \left(\int \mathbf{1}_{p,r} d\mu\right)^2}{\log r} = \underline{D}_2(\mu), \tag{4.3.1}$$

which means for any $\varepsilon > 0$, there is $r_0 > 0$ such that for all $0 < r < r_0$,

$$r^{\overline{D}_2(\mu)+\varepsilon} \leq \sum_{p=1}^{k(r)} \left(\int \mathbf{1}_{p,r}\, d\mu\right)^2 \leq r^{\underline{D}_2(\mu)-\varepsilon}. \tag{4.3.2}$$

Proof. Proof for this lemma can be readily adapted from the proof of [GRS, Lemma 13]. □

4.3.1 Proof of Theorem 4.3.4

Define the following quantities inspired by [GRS],

$$\epsilon(n) = (\log n)^2$$

$$m_n^{\leq}(x) := \min_{\substack{0\leq i<j<n \\ |i-j|\leq\epsilon(n)}} d\left(f^i x, f^j x\right),$$

$$m_n^{>}(x) := \min_{\substack{0\leq i<j<n \\ |i-j|>\epsilon(n)}} d\left(f^i x, f^j x\right),$$

$$m_n^{\gg}(x) := \min_{\substack{0\leq i\leq n/3 \\ 2n/3\leq j<n}} d\left(f^i x, f^j x\right),$$

Obviously, $m_n(x) = \min\left\{m_n^{\leq}(x), m_n^{>}(x)\right\} \leq m_n^{\gg}(x)$ for all $x \in \Lambda$ and $n \in \mathbb{N}$. Then under the conditions of Theorem 4.3.4, we prove the following.

Proposition 4.3.7. *Let $T : X \to X$ be a Gibbs-Markov map defined above, and μ its invariant Gibbs measure admitting exponential decay of correlations for $\mathcal{BV}$ against L^1 observables, then one has for*

μ-every $x \in \Lambda$,

$$\limsup_{n\to\infty} \frac{\log m_n^{>}(x)}{-\log n} \le \frac{2}{\underline{D}_2}. \tag{4.3.3}$$

If μ is absolutely continuous with respect to Lebesgue measure, $\underline{D}_2 = \overline{D}_2 = 1$, and

$$\limsup_{n\to\infty} \frac{\log m_n^{\le}(x)}{-\log n} \le \frac{2}{D_2}, \tag{4.3.4}$$

for μ-almost every $x \in \Lambda$.

Proposition 4.3.8. *For all Gibbs measures μ, for μ-almost every $x \in \Lambda$,*

$$\liminf_{n\to\infty} \frac{\log m_n^{\gg}(x)}{-\log n} \ge \frac{2}{\overline{D}_2}. \tag{4.3.5}$$

Putting these two propositions together imply Theorem 4.3.4, *i.e.*, if μ is an acip,

$$\lim_{n\to\infty} \frac{\log m_n(x)}{\log n} = 2$$

μ-almost every $x \in \Lambda$.

The proof of (4.3.3) is basically a practice of applying decay of correlations, whereas the proof of (4.3.4) requires estimating the measures of sets of short return points. During the proof we will see also that (4.3.3) and (4.3.5) hold for all Gibbs invariant measures with exponential decay of correlations and $D_2(\mu) > 0$. Also for Gibbs *acip* μ, the correlation dimension is well-defined in the sense that $\overline{D}_2(\mu) = \underline{D}_2(\mu) = 1$, because the invariant density with respect to Lebesgue measure is uniformly bounded; hence $D_2(\mu) = D_2(Leb)$.

Also, for simplicity of calculation, the following definition is introduced in [GRS].

Definition 4.3.9. *A term is said to be* admissible *if it has the form $r^{-k}g(n)$, for some $k \ge 0$ and a function g which decays in n faster than any polynomial of n, hence for any $k \in \mathbb{N}$, by* (4.3.2) *and choosing the scale of r as in* (4.3.6) *below we can bound any admissible error by $\mathcal{O}(n^{-k})$ for all n large.*

Proof of (4.3.3). Let $\varepsilon, r > 0$ be given, in particular, r should be small enough that it satisfies (4.3.2). Define the random variable $\mathcal{S}_n^{>}$,

$$\mathcal{S}_n^{>}(x) := \sum_{p=1}^{k(r)} \sum_{\substack{0\le i<j<n \\ |i-j|>\epsilon(n)}} \mathbf{1}_{p,r}\left(f^i x\right) \mathbf{1}_{p,r}\left(f^j x\right).$$

By Lemma 4.3.5, $\{m_n^{>}(x) \le r\} \subseteq \{\mathcal{S}_n^{>} \ge 1\}$. Therefore, by Markov's inequality and decay of correlation,

$$\begin{aligned}
\mu(x : \mathcal{S}_n^{>}(x) \geq 1) &\leq \mathbb{E}[S_n^{>}] = \sum_{\substack{0\leq i<j<n \\ j-i>\epsilon(n)}} \sum_{p=1}^{k(r)} \int \mathbf{1}_{p,r}\left(f^i x\right) \mathbf{1}_{p,r}\left(f^j x\right)\, d\mu(x) \\
&\leq \sum_{\substack{0\leq i<j<n \\ j-i>\epsilon(n)}} \sum_{p=1}^{k(r)} \left(\left(\int \mathbf{1}_{p,r}\, d\mu \right)^2 + \rho(\epsilon(n)) \|\mathbf{1}_{p,r}\|_{\mathcal{BV}} \|\mathbf{1}_{p,r}\|_1 \right) \\
&\leq \sum_{0\leq i<j<n} r^{\underline{D}_2-\varepsilon} + \sum_{0\leq i<j<n} \sum_{p=1}^{k(r)} \|\mathbf{1}_{p,r}\|_{\mathcal{BV}} \|\mathbf{1}_{p,r}\|_1 \rho(\epsilon(n)) \\
&\leq n^2 r^{\underline{D}_2-\varepsilon} + C_1 e^{-c_1(\log n)^2} n^2 \sum_{p=1}^{k(r)} 2\mu(B(x_p, 2r)) \\
&\leq n^2 r^{\underline{D}_2-\varepsilon} + 2C_0C_1 e^{-c_1(\log n)^2} n^2,
\end{aligned}$$

where the penultimate line holds due to (4.3.2), and the last inequality follows from the definition of bounded local complexity. As $e^{-c_1(\log n)^2}$ decays faster than any polynomial of n, replace all r terms above with

$$r_n = \exp\left(-\frac{2+2\varepsilon}{\underline{D}_2 - \varepsilon} (\log n + \log\log n) \right) \leq n^{-\frac{2+2\varepsilon}{\underline{D}_2-\varepsilon}}, \tag{4.3.6}$$

for all n sufficiently large,

$$n^2 r_n^{\underline{D}_2-\varepsilon} + 2C_0C_1 n^2 e^{-c_1(\log n)^2} \leq n^{-2\varepsilon} + C_1 e^{-c_1(\log n)^2} r_n^{-C_2}.$$

The second term on the right is admissible by definition, thus there is some constant $C_2 > 0$ such that for all n large enough that $n^{-\varepsilon} \leq \frac{1}{\log n}$:

$$\mu(m_n^{>}(x) \leq r) \leq \mathbb{E}[S_n^{>}] \leq C_2 n^{-\varepsilon} \preceq \frac{1}{\log n}.$$

Applying Borel-Cantelli to a subsequence n_k as in Section 4.2.1, eventually for μ-a.e. x,

$$\frac{\log m_{n_k}^{>}(x)}{-\log n_k} \leq \frac{2+\varepsilon}{\underline{D}_2 - \varepsilon} \left(1 + \frac{\log\log n_k}{\log n_k} \right).$$

Although $m_n^{>}$ is not monotonically increasing, for each $n \in [n_s, n_{s+1}]$, we define

$$-\log m'_{n_{k+1}}(x) := -\log \min_{\substack{0\leq i<j<n_{k+1} \\ j-i>\epsilon(n_k)}} \geq -\log m_n^{>}(x), -\log m''_{n_k}(x) := -\log \min_{\substack{0\leq i<j<n_k \\ j-i>\epsilon(n_{k+1})}} \leq -\log m_n^{>}(x). \tag{4.3.7}$$

By modifying the arguments we have done so far, one can show exactly that, for μ-almost every x, for all k large,

$$\frac{-\log m'_{n_k}(x)}{\log n_k} \leq \frac{2+\varepsilon}{\underline{D}_2 - \varepsilon} \left(1 + \frac{\log\log n_k}{\log n_k} \right),$$

and

$$\frac{-\log m''_{n_k}(x)}{\log n_k} \leq \frac{2+\varepsilon}{\underline{D}_2 - \varepsilon}\left(1 + \frac{\log\log n_k}{\log n_k}\right).$$

By (4.3.7), these limits can be passed to the whole tail of $m_n^{>}$. As $\varepsilon > 0$ is arbitrarily small, (4.3.3) is proved. $\square$

Note also that for $|i-j|$ relatively large, the measure of the sets $\left\{x : d\left(f^i x, f^j x\right) < r\right\}$ scales like r^{D_2} which is similar to the sequence matching problem in the symbolic setting, and this matches our intuition because $D_2(\mu)$ is analogous to H_2 in some ways.

To prove (4.3.4), we need to deal with iterates of x which return to an r-neighbourhood of itself within $\epsilon(n)$ iterations, so again we need to approximate the measure of some short return sets as those $S_m(k)$ sets defined in Section 4.2.1 for the symbolic case. But we cannot expect a similar upper bound for short returns as in (4.2.8) or (4.2.9). This is because for symbolic structures, an r_n-cylinder is itself an r_n open ball with respect to the symbolic metric, so analysing the returns is equivalent to analysing the repetition of letters in cylinders and one does not need to consider the case that two iterates $\sigma^i(x)$, $\sigma^j(x)$ are close to the boundaries of two open balls with a common boundary but belong to different cylinders.

For interval maps, although the Gibbs-Markov structure prescribes a natural partition hence a way to define cylinders, the metric balls and symbolic cylinders are different objects so one needs to take more caution and include the case that two points belong to different cylinders $U, V \in \mathcal{P}_n$ but accumulate on a common boundary of U, V with distance smaller than the contraction scale of n-cylinders. Luckily, for Gibbs-Markov maps there is the following lemma.

Lemma 4.3.10. *[HNL, Lemma 3.4] Define the sets*

$$\mathcal{E}_n(r) := \{x \in X : |x - f^n x| \leq r\}.$$

Then for f satisfying Gibbs-Markov property with the invariant Gibbs measure μ absolutely continuous with respect to Lebesgue measure and with exponential decay of correlation for $\mathcal{BV}$ against L^1 observables, there is some constant C_3 such that for all $n \in \mathbb{N}$ and r small enough,

$$\mu\left(\mathcal{E}_n(r)\right) \leq C_3 r.$$

Proof. The original lemma states that $m\left(\mathcal{E}_n(r)\right) \preceq r$, and since m is equivalent to μ on Λ with $d\mu/dm$ bounded away from 0 and $+\infty$, there is a uniform constant C_μ such that

$$C_\mu^{-1} m(A) \leq \mu(A) \leq C_\mu m(A)$$

for each measurable A; therefore there is some $C_3 > 0$ such that

$$\mu\left(\mathcal{E}_n(\epsilon)\right) \le C_3\epsilon. \qquad \square$$

Proof of (4.3.4). Define the random variable $\mathcal{S}_n^{\le}$ by

$$\mathcal{S}_n^{\le}(x) := \sum_{i=0}^{n-1} \sum_{k=1}^{\epsilon(n)\wedge(n-i-1)} \mathbf{1}_{B(f^i x, r)}(f^{i+k}x),$$

where $a \wedge b = \min\{a, b\}$.

As $\left\{x : \mathbf{1}_{B(f^i x, r)}\left(f^{i+k}(x)\right) = 1\right\} \subseteq \left\{x : f^i x \in \mathcal{E}_k(r)\right\} = f^{-i}\mathcal{E}_k(r)$, using the Markov inequality and Lemma 4.3.10 we obtain the following bound:

$$\mu(\mathcal{S}_n^{\le} \ge 1) \le \mathbb{E}_m[\mathcal{S}_n^{\le}] \le \sum_{i=0}^{n-1} \sum_{k=1}^{\epsilon(n)\wedge(n-i-1)} \mu(f^{-i}\mathcal{E}_k(r)) \le n\epsilon(n)C_3 r$$

Pick $r = r_n$ as in (4.3.6), for all n large enough such that

$$\epsilon(n) = (\log n)^2 \le n^{\frac{\varepsilon}{2-\varepsilon}}, \qquad n^{-\frac{\varepsilon}{2-\varepsilon}} \le \frac{1}{\log n}.$$

As the invariant density $d\mu/dm$ is uniformly bounded, $D_2(\mu) = 1 < 2$, one has

$$\mu\left(x : m_n(x) \le r_n\right) \le C_3 n^{1+\frac{\varepsilon}{2-\varepsilon}} r \le n^{\frac{2}{2-\varepsilon}} C_3 n^{-\frac{2+2\varepsilon}{D_2-\varepsilon}} \le C_3 n^{-\frac{\varepsilon}{2-2\varepsilon}} \preceq \frac{1}{\log n}.$$

Therefore, by picking a subsequence $n_k = \lceil e^{k^2} \rceil$, by Borel-Cantelli Lemma we have that $\mu\{x : m_{n_k} \le r_{n_k}$ for infinitely many $k\} = 0$, so for μ-almost every x, for all k large enough,

$$m_{n_k}^{\le}(x) \ge r_{n_k}.$$

We then repeat the subsequence trick to obtain (4.3.4) for μ-almost every x. $\square$

Remark 4.3.11. *The condition that μ is an acip may be not sharp; if ν is another invariant probability measure with exponential decay of correlation and satisfies $\nu(\mathcal{E}_k(r)) \approx r$, then Proposition 4.3.7 remains valid.*

As in the symbolic case the proof for the lower bound of $\frac{\log m_n(x)}{-\log n}$ is slightly more complicated and also requires a second-moment computation which exploits the following notion of mixing.

Lemma 4.3.12. *The unique Gibbs measure μ with respect to the geometric potential $-\log|Df|$ for a Gibbs-Markov interval map f has exponential* 4-mixing*, that is, for $a < b \le c$ in $\mathbb{N}$, there are $C_1', c_1' > 0$*

such that for all $u_1, u_2 \in \mathcal{BV}$*,* $u_3, u_4 \in L^\infty$*, such that*

$$\left| \int u_1 \left(u_2 \circ f^a\right) \left(u_3 \circ f^b\right) \left(u_4 \circ f^c\right) \, d\mu - \int u_1 \left(u_2 \circ f^a\right) \, d\mu \int u_3 \left(u_4 \circ f^{c-b}\right) \, d\mu \right| \leq C_1' e^{-c_1'(b-a)}.$$

The constant C_1' *depends on the functions* u_i*. In particular, for any given* $r > 0$*,* $0 \leq p, q \leq k(r)$*,* $u_1 = u_2 = \mathbf{1}_{p,r}$*,* $u_3 = u_4 = \mathbf{1}_{q,r}$*, the constant* C_1' *does not depend on* r*.*

Proof. Consider the transfer operator $\mathcal{L}$ associated with the geometric potential $-\log|Df|$, that acts on the space of functions of bounded variation, $\mathcal{BV} = \mathcal{BV}(X)$,

$$\mathcal{L} = \mathcal{L} : \mathcal{BV} \to \mathcal{BV}, \quad \mathcal{L}u(x) = \sum_{Ty=x} e^{-\log|Df(y)|} u(y).$$

Let ν be the eigenmeasure of $\mathcal{L}$ and h the invariant density, $\frac{d\mu}{d\nu} = h$. By the following well-known fact, (see for example [Kel1, (3)]) for topologically mixing Gibbs-Markov maps, there are $C_{\mathcal{BV}} > 0$, $\kappa \in (0,1)$ such that for any $u \in \mathcal{BV}$,

$$\left\| \mathcal{L}^n u - h \int u \, d\nu \right\|_{\mathcal{BV}} \leq C_{\mathcal{BV}} \cdot \kappa^n \|u\|_{\mathcal{BV}}.$$

Then,

$$\begin{aligned}
&\left| \int u_1 \, \left(u_2 \circ f^a\right) \, \left(u_3 \circ f^b\right) \, \left(u_4 \circ f^c\right) \, d\mu - \int u_1 \left(u_2 \circ f^a\right) \, d\mu \int u_3 \left(u_4 \circ f^{c-b}\right) \, d\mu \right| \\
=& \left| \int \mathcal{L}^{b-a}(hu_1 u_2 \circ f^a)(u_3 \circ f^a) \left(u_4 \circ f^{c-b+a}\right) d\nu - \int hu_1(u_2 \circ f^a) \, d\nu \int h \left(u_3 \circ f^a\right) \left(u_4 \circ f^{c-b+a}\right) d\nu \right| \\
=& \left| \int \left(\mathcal{L}^{b-a} \left(hu_1 u_2 \circ f^a\right) - h \int hu_1 u_2 \circ f^a \, d\nu \right) u_3 \circ f^a \, u_4 \circ f^{c-b+a} \, d\nu \right| \\
\leq& \left\| \mathcal{L}^{b-a}(hu_1 u_2 \circ f^a) - h \int hu_1 u_2 \circ f^a \, d\nu \right\|_1 \|u_3\|_\infty \|u_4\|_\infty \\
\leq& C_{\mathcal{BV}} \cdot \kappa^{b-a} \|hu_1 u_2 \circ f^a \, \|_{\mathcal{BV}} \|u_3\|_\infty \|u_4\|_\infty, \qquad (\star)
\end{aligned}$$

where $\|\cdot\|_1$ denotes the L^1 norm with respect to ν, and the first equality holds by the duality of ν. $hu_1 u_2 \circ f^a$ is of bounded variation because h is of bounded variation, and the product of functions in $\mathcal{BV}$ has bounded variation, and the first part of the lemma is proved.

Now we deal with the case where u_i's are indicator functions $\mathbf{1}_{p,r}$ or $\mathbf{1}_{q,r}$, and find a suitable upper bound for

$$\|h\mathbf{1}_{p,r}\mathbf{1}_{q,r} \circ f^a\|_{\mathcal{BV}} = \|h\mathbf{1}_{p,r}\mathbf{1}_{q,r} \circ f^a\|_1 + \mathrm{TV}(h\mathbf{1}_{p,r}\mathbf{1}_{q,r} \circ f^a).$$

For the 1-norm,

$$\int |h\mathbf{1}_{p,r}\mathbf{1}q, r \circ f^a| \, d\nu \leq \|h\|_\infty.$$

For the total variation, first recall that for any functions $u, v \in \mathcal{BV}$,

$$\mathrm{TV}(uv) \le \|u\|_\infty \mathrm{TV}(v) + \|v\|_\infty \mathrm{TV}(u).$$

Then, as for any indicator function $\mathbf{1}_{p,r}$, $\mathrm{TV}(\mathbf{1}_{p,r}) \le 2$, we have

$$\begin{aligned} \mathrm{TV}(h\mathbf{1}_{p,r}\mathbf{1}_{q,r} \circ f^a) &\le \|h\|_\infty \mathrm{TV}(\mathbf{1}_{p,r}\mathbf{1}_{q,r} \circ f^a) + \|\mathbf{1}_{p,r}\mathbf{1}_{q,r} \circ f^a\|_\infty \\ &\le \|h\|_\infty (\|\mathbf{1}_{p,r}\|_\infty \mathrm{TV}(\mathbf{1}_{q,r}) + \|\mathbf{1}_{q,r} \circ f^a\|_\infty \mathrm{TV}(\mathbf{1}_{p,r})) + \mathrm{TV}(h) \\ &\le 4\|h\|_\infty + \|h\|_{\mathcal{BV}}. \end{aligned}$$

Therefore, for any p, q and $r > 0$, if $u_1, u_3 = \mathbf{1}_{p,r}$ and $u_2, u_4 = \mathbf{1}_{q,r}$, ($\star$) is bounded from above by $C_{\mathcal{BV}}\ 5\|h\|_\infty\ \kappa^{b-a}$. □

Proof of lower bound in Theorem 4.3.4

Now we can finish the remaining proof of Theorem 4.3.4 which, similar to that of Theorem 4.2.1, involves a second moment argument where the 4-mixing property becomes useful.

Proof of (4.3.5). Let $\varepsilon > 0$ small be given. Consider the quantity $m_n^{\gg}$ and the random variable $\mathcal{S}_n^{\gg}$:

$$m_n^{\gg}(x) := \min_{\substack{0 \le i \le n/3 \\ 2n/3 \le j < n}} d\left(f^i x, f^j x\right), \quad \mathcal{S}_n^{\gg}(x) := \sum_{\substack{0 \le i \le n/3 \\ 2n/3 \le j < n}} \sum_{p=1}^{k(r)} \mathbf{1}_{p,r}(f^i x)\mathbf{1}_{p,r}(f^j x).$$

By Lemma 4.3.5, $m_n^{\gg}(x) > 4r$ implies for all pairs of $0 \le i \le \frac{n}{3}$, $\frac{2n}{3} \le j < n$, if for some p, $d(f^i x, x_p^r) < 2r$, then $d(f^j x, x_p^r) \ge 2r$ hence $\mathcal{S}_n^{\gg}(x) = 0$. By the Paley-Zygmund inequality,

$$\mu\left(m_n > 4r\right) \le \mu\left(x : \mathcal{S}_n^{\gg}(x) = 0\right) \le \frac{\mathbb{E}[(\mathcal{S}_n^{\gg})^2] - \mathbb{E}[\mathcal{S}_n^{\gg}]^2}{\mathbb{E}[\mathcal{S}_n^{\gg}]^2}.$$

Using decay of correlations and invariance of μ,

$$\mathbb{E}[\mathcal{S}_n^{\gg}(x)] = \sum_{\substack{0 \le i \le n/3 \\ 2n/3 \le j < n}} \sum_p \int \mathbf{1}_{p,r}(f^i x)\mathbf{1}_{p,r}(f^j x)\, d\mu(x) \le \left(\frac{n}{3}\right)^2 \sum_p \left(\left(\int \mathbf{1}_{p,r}\, d\mu(x)\right)^2 + 2\rho\left(n/3\right)\right). \tag{4.3.8}$$

Consider

$$\left(\mathcal{S}_n^{\gg}(x)\right)^2 = \sum_{i,j}\sum_{s,t}\sum_{p,q} \mathbf{1}_{p,r}(f^i x)\mathbf{1}_{p,r}(f^j x)\mathbf{1}_{q,r}(f^s x)\mathbf{1}_{q,r}(f^t x).$$

As in the proof of symbolic case, we will split this sum in terms of the distance between the indices i, j, s, t. Recall that

$$\epsilon(n) = (\log n)^2.$$

Let Q be the collection of all possible quadruples of indices (i,j,s,t), and define the counting function

$$\tau : Q \to \mathbb{N} \cup \{0\}, \quad \tau(i,j,s,t) = \sum_{\substack{a \in \{i,s\} \\ b \in \{j,t\}}} \mathbf{1}_{[a-\epsilon(n), a+\epsilon(n)]}(b).$$

Then $\tau \leq 2$ since i,j and s,t are at both at least $\frac{n}{3}$ iterates apart. This allows us to split Q into $Q_m := \{(i,j,s,t) \in F : \tau = m\}$ for $m = 0,1,2$. Obviously, the following upper bounds hold for the cardinality of each Q_m,

$$\# Q_m \leq (2\epsilon(n))^m \left(\frac{n}{3}\right)^{4-m}. \tag{4.3.9}$$

Recall the notation

$$\mathbb{E}[\left(\mathcal{S}_n^{\gg}\right)^2 | Q_m] = \sum_{(i,j,s,t) \in Q_m} \sum_{p,q} \int \mathbf{1}_{p,r}(f^i x) \mathbf{1}_{p,r}(f^j x) \mathbf{1}_{q,r}(f^s x) \mathbf{1}_{q,r}(f^t x) \, d\mu(x),$$

also for simplicity, let

$$R_p = \int \mathbf{1}_{p,r} \, d\mu = \mu(B(x_p, 2r)).$$

Contribution of indices in Q_0:

For each $(i,j,s,t) \in Q_0$, without loss of generality, suppose $i + \epsilon(n) < s$ and $j + \epsilon(n) < t$, as the alternative cases can be treated equally by exchanging the roles of i,s or j,t which makes no difference to the calculation. As $\min\{j,t\} - \max\{i,s\} \geq \frac{n}{3}$, by Lemma 4.3.12 and invariance, one obtains the following upper bound for each such quadruple (i,j,s,t):

$$\begin{aligned}
&\sum_{p,q} \int \mathbf{1}_{p,r}(f^i x) \mathbf{1}_{p,r}(f^j x) \mathbf{1}_{q,r}(f^s x) \mathbf{1}_{q,r}(f^t x) \, d\mu \\
&= \sum_{p,q} \int \mathbf{1}_{p,r} \mathbf{1}_{q,r} \circ f^{s-i} \mathbf{1}_{p,r} \circ f^{j-i} \mathbf{1}_{q,r} \circ f^{t-i} \, d\mu \\
&\leq C_1' e^{-c_1' \frac{n}{3}} k(r)^2 + \sum_{p,q} \int \mathbf{1}_{p,r} \mathbf{1}_{q,r} \circ f^{s-i} \, d\mu \int \mathbf{1}_{p,r} \mathbf{1}_{q,r} \circ f^{t-j} \, d\mu \\
&\leq C_1' e^{-c_1' \frac{n}{3}} k(r)^2 + \sum_{p} \sum_{q} \left(R_p R_q + 2\rho(\epsilon(n))\right)^2 \\
&\leq C_1' e^{-c_1' \frac{n}{3}} r^{-2C_0'} + 8\rho(\epsilon(n)) r^{-2C_0'} + \sum_{p,q} \left(R_p R_q\right)^2.
\end{aligned}$$

The last inequality holds as $R_p, R_q \leq 1$ for any p,q, and by [GRS, Lemma 3.3] $k(r) \leq r^{-C_0'}$ for some $C_0' = 4 \log C_0$. Any term in the inequality above involving $\rho(\epsilon(n))$ or $C_1' e^{-c_1' \frac{n}{3}}$ is admissible, hence for each $k \in \mathbb{R}$ it is bounded by $\mathcal{O}(n^{-k})$ for all n sufficiently large, and now we pick

$$r = r_n = n^{-\frac{2-4\varepsilon}{D_2+\varepsilon}}.$$

Then by (4.3.2)

$$r^{\overline{D}_2+\varepsilon} \leq \sum_p \left(\int \mathbf{1}_{p,r} \, d\mu \right)^2 = \sum_p R_p^2. \tag{4.3.10}$$

Therefore, the contribution of indices in Q_0 is bounded from above up to an admissible error by

$$\left(\frac{n}{3}\right)^4 \sum_p \sum_q R_p^2 R_q^2 \leq \left(\frac{n}{3}\right)^4 \sum_p R_p^2 \sum_q R_q^2 = \left(\frac{n}{3}\right)^4 \left(\sum_p R_p^2 \right)^2,$$

combining with (4.3.8), up to an admissible error term,

$$\mathbb{E}[(\mathcal{S}_n^{\gg})^2 | Q_0] - \mathbb{E}[\mathcal{S}_n^{\gg}]^2 \preceq (n^{-\varepsilon}).$$

Also by (4.3.8), as $\rho(\frac{n}{3})$ is admissible we can bound it by n^{-3},

$$\mathbb{E}[\mathcal{S}_n^{\gg}]^2 \geq \left(\frac{n}{3}\right)^4 \left(r_n^{\overline{D}_2+\varepsilon} - 2\rho(\frac{n}{3}) \right) \geq \left(\frac{n}{3}\right)^4 (n^{-2-4\varepsilon} - n^{-3})^2 \approx n^{8\varepsilon}$$

allowing us to conclude that there is some constant $C_4 > 0$:

$$\frac{\mathbb{E}[(\mathcal{S}_n^{\gg})^2 | Q_0] - \mathbb{E}[\mathcal{S}_n^{\gg}]^2}{\mathbb{E}[\mathcal{S}_n^{\gg}]^2} \leq \frac{C_4}{n^{\varepsilon}}. \tag{4.3.11}$$

Contributions of indices in Q_1:

Now we will deal with the indices in Q_1. Without loss of generality, suppose $|i - s| \leq \epsilon(n)$, $i < s$ and $j < t$, the other cases can be treated by exchanging the roles of i, s or j, t. By invariance, Lemma 4.3.12 and decay of correlations, for $i, j, s, t \in Q_1$,

$$\begin{aligned}
&\sum_{p,q} \int \mathbf{1}_{p,r}(f^i x) \mathbf{1}_{p,r}(f^j x) \mathbf{1}_{q,r}(f^s x) \mathbf{1}_{q,r}(f^t x) \, d\mu(x) \\
&= \sum_{p,q} \int \mathbf{1}_{p,r} \mathbf{1}_{q,r} \circ f^{s-i} \mathbf{1}_{p,r} \circ f^{j-i} \mathbf{1}_{q,r} \circ f^{t-i} \, d\mu \\
&\leq C_1' e^{-c_1' \frac{n}{3}} k(r)^2 + \sum_{p,q} \left(\int \mathbf{1}_{p,r} \mathbf{1}_{q,r} \circ f^{t-j} \, d\mu \right) \int \mathbf{1}_{p,r}(x) \mathbf{1}_{q,r}(f^{s-i} x) \, d\mu(x) \\
&\leq C_1' e^{-c_1' \frac{n}{3}} r^{-2C_0'} + 2\rho(\epsilon(n)) + \sum_{p,q} (R_p R_q) \int \mathbf{1}_{p,r}(x) \mathbf{1}_{q,r}(f^{s-i} x) \, d\mu(x).
\end{aligned}$$

Using Cauchy-Schwarz inequality, the last line can be bounded by the following up to an admissible error (recall Definition 4.3.9)

$$\begin{aligned}
&\sum_{p,q}\int R_pR_q\mathbf{1}_{p,r}(x)\mathbf{1}_{q,r}(f^{s-i}x)\,d\mu(x)\\
&=\int\sum_p R_p\mathbf{1}_{p,r}(x)\sum_q R_q\mathbf{1}_{q,r}(f^{s-i})\,d\mu(x)\\
&\leq\left(\int\left(\sum_p R_p\mathbf{1}_{p,r}(x)\right)^2 d\mu\right)^{\frac{1}{2}}\left(\int\left(\sum_q R_q\mathbf{1}_{q,r}\circ f^{s-i}\right)^2 d\mu\right)^{\frac{1}{2}}\\
&=\int\left(\sum_p R_p\mathbf{1}_{p,r}\right)^2 d\mu,
\end{aligned}$$

where the last line is by symmetry and invariance. Notice that for all real numbers $a_1,\dots,a_m\geq 0$,

$$(a_1+a_2+\cdots+a_m)^2\leq m\left(a_1^2+a_2^2+\dots,a_m^2\right).$$

By bounded local complexity assumption, there are at most C_0 non-zero terms in $\{\mathbf{1}_{p,r}(x)\}_{p=1}^{k(r)}$ for any $x\in X$, and $(\mathbf{1}_{p,r})^2\leq\mathbf{1}_{p,r}$,

$$\left(\sum_p R_p\mathbf{1}_{p,r}\right)^2 d\mu\leq\int C_0\sum_p R_p^2\mathbf{1}_{p,r}\,d\mu=\sum_p C_0R_p^2\int\mathbf{1}_{p,r}\,d\mu=C_0\sum_p R_p^3.$$

As $(a_1+\cdots+a_m)^{\frac{2}{3}}\leq\sum_{k=1}^m a_k^{\frac{2}{3}}$, clearly $\sum_k a_k\leq\left(\sum_k a_k^{2/3}\right)^{3/2}$, there exists some constant C_5 such that for all n large enough with $\epsilon(n)\leq n^{\varepsilon}$,

$$\begin{aligned}
\frac{\mathbb{E}[(\mathcal{S}_n^{\gg})^2|Q_1]}{\mathbb{E}[\mathcal{S}_n^{\gg}]^2}\leq&\frac{2\epsilon(n)(\frac{n}{3})^3C_0\left(\sum_p R_p^2\right)^{\frac{3}{2}}}{(\frac{n}{3})^4(\sum_p R_p^2-2\rho(\epsilon(n)))^2}\\
=&\frac{6\epsilon(n)C_0\left(\sum_p R_p^2\right)^{\frac{3}{2}}}{n\left(\left(\sum_p R_p^2\right)^{1/2}-4\rho(\epsilon(n))\left(\sum_p R_p^2\right)^{-1/2}+4\rho(\epsilon(n))^2\left(\sum_p R_p^2\right)^{-3/2}\right)}\\
\leq&\frac{6C_0\epsilon(n)}{n\left((r^{\overline{D}_2+\varepsilon})^{1/2}-\mathcal{O}(n^{-1})\right)}=\frac{6C_0\epsilon(n)}{n\left((n^{-1+2\varepsilon}-\mathcal{O}(n^{-1})\right)}\\
\leq&\frac{C_5n^{\varepsilon}}{n\cdot n^{-1+2\varepsilon}}=\frac{C_5}{n^{\varepsilon}},
\end{aligned}\tag{4.3.12}$$

because $(\sum_p R_p^2)^{-\frac{3}{2}}4\rho(\epsilon(n))$ and $\left(\sum_p R_p^2\right)^{-\frac{1}{2}}\rho(\epsilon(n))$ are both admissible errors.

Contribution of indices in Q_2:

Finally, let us consider indices (i,j,s,t) such that $|i-s|,|j-t| \le \epsilon(n)$. By Lemma 4.3.5, $\sum_q \mathbf{1}_{q,r}(f^s x)\mathbf{1}_{q,r}(f^t x) \le C_0$ for any x, therefore for each i,j,s,t in Q_2,

$$\begin{aligned}
&\sum_{p,q} \int \mathbf{1}_{p,r}(f^i x)\mathbf{1}_{p,r}(f^j x)\mathbf{1}_{p,r}(f^s x)\mathbf{1}_{p,r}(f^t x)\, d\mu(x) \\
&\le C_0 \sum_p \int \mathbf{1}_{p,r}(f^i x)\mathbf{1}_{p,r}(f^j x)\, d\mu(x) \\
&\le C_0 \sum_p R_p^2 + C_0 \rho\left(\frac{n}{3}\right) k(r).
\end{aligned}$$

Therefore, as $\# Q_2 \le \frac{4}{9}\epsilon(n)^2 n^2$, by our choice of r_n in (4.3.10), up to an admissible error there is some constant C_6 such that,

$$\begin{aligned}
\frac{\mathbb{E}[(\mathcal{S}_n^{\gg})^2 | Q_2]}{\mathbb{E}[S_n^{\gg}]^2} &\le \frac{4\epsilon(n)^2(\frac{n}{3})^2 C_0 \sum_p R_p^2}{(\frac{n}{3})^4(\sum_p R_p^2 - 2\rho(\epsilon(n)))^2} \\
&= \frac{36 C_0 \epsilon(n)^2}{n^2\left(\sum_p R_p^2 - 4\rho(\epsilon(n)) + 4\rho(\epsilon(n))^2\left(\sum_p R_p^2\right)^{-1}\right)} \\
&\le \frac{36 C_0 n^{2\varepsilon}}{n^2\left(r^{\overline{D}_2+\varepsilon} - \mathcal{O}(n^{-2})\right)} \\
&\le \frac{C_6 n^{2\varepsilon}}{n^2 n^{-2+4\varepsilon}} = \frac{C_6}{n^{2\varepsilon}}.
\end{aligned} \tag{4.3.13}$$

Hence, putting (4.3.11), (4.3.12) and (4.3.13) together, we can conclude that for all n large enough and $r = r_n = n^{-\frac{2-4\varepsilon}{\overline{D}_2+\varepsilon}}$, there is some constant $C_7 > 0$ such that

$$\mu(m_n^{\gg} > 8r_n) \le \frac{\mathrm{Var}[\mathcal{S}_n^{\gg}]}{\mathbb{E}[\mathcal{S}_n^{\gg}]^2} \le \frac{C_7}{n^{\varepsilon}}.$$

Picking a subsequence $n_k = \lceil k^{2/\varepsilon} \rceil$, the probability is summable along the subsequence which means that by the Borel-Cantelli Lemma, for μ-almost every x, for k large

$$-\log m_n^{\gg}(x) \ge -\log 4r_{n_k}.$$

The proof of (4.3.5) is yet complete because $m_n^{\gg}$ is not a monotone sequence so we need to repeat the trick at the end of the Proof of (4.3.3). For each $n \in [n_k, n_{k+1}]$, define

$$-\log m_n^{\gg}(x) \ge -\log \min_{\substack{0 \le i \le n_k \\ 2n_{k+1}/3 \le j < n_k}} d\left(f^i x, f^j x\right) =: -\log m_{n_k}^*(x).$$

As for all k, $m_{n_{k+1}}^{\gg}(x) \le m_{n_k}^*(x)$, repeating the same proof we have done for $m_n^{\gg}$ one can also show that $\liminf_{k\to\infty} \frac{\log m_{n_k}^*}{\log n_k} \ge \frac{2-4\varepsilon}{\overline{D}_2+\varepsilon}$, such a lower bound can be passed to the entire tail of $-\log m_{n_k}^{\gg}(x)$, and then $-\log m_n(x)$. We can conclude

$$\liminf_{n\to\infty} \frac{\log m_n(x)}{-\log n} \geq \frac{2}{\overline{D}_2}$$

for μ-almost every x since $\varepsilon > 0$ was arbitrarily small. $\square$

Remark 4.3.13. (4.3.3) *and* (4.3.5) *still hold if decay of correlations is exponential with respect to other Banach function spaces $\mathcal{B}$, $\mathcal{B}'$ – for example, both observables are in $\mathcal{BV}$ or* Lip*, where* $\mathrm{Lip} := \{f \in C(X) :$ *f is Lipschitz$\}$ – as long as $\rho(\epsilon(n))\|\mathbf{1}_{p,r}\|_{\mathcal{B}}\|\mathbf{1}_{q,r}\|_{\mathcal{B}'}$ remains an admissible term. For example, if the system has decay of correlations for Lipschitz observables, one can replace the $\mathbf{1}_{p,r}$ functions with $\{\rho_p^r\}_{p=1}^{k(r)}$ (a set of discretisation functions defined in [GRS]) although it requires heavier machinery to adjust the proof for* (4.3.5) *and Lemma 4.3.12. Lastly, instead of exponential decay of correlations, the proof remains valid under stretched exponential decay by manipulating the scale of k in $\epsilon(n) = (\log n)^k$.*

4.3.2 Irrational rotations

Just as in Section 3.2, mixing is important in the proofs in last section, so we look at irrational rotations again to see what happens if mixing properties are absent. Recall that given an irrational $\theta \in (0,1)$, $T_\theta : [0,1) \to [0,1)$ is defined by $T_\theta(x) = x + \theta \pmod 1$. For all $x \in [0,1)$, the shortest distance quantity $M_n(x)$ is independent of x:

$$m_n(x) = \min\{\|(i-j)\theta\| : 0 \leq i < j < n\} = \min\{\|k\theta\| : 1 \leq |k| < n\} = m_n(0), \tag{4.3.14}$$

where the norm $\|\cdot\|$ was defined in Definition3.5.1. Then the single-orbit shortest distance problem for circle rotations is simply determining the limiting behaviour of $\log m_n(0)/-\log n$. It should be noted that the waiting time results in [KS] and the proof for [BLR, Theorem 10] are not directly applicable here, since this is essentially a recurrence problem.

Theorem 4.3.14. *Let $\theta \in (0,1)$ be an irrational number with $\eta(\theta) > 1$, then for every $x \in [0,1)$*

$$\liminf_{n\to\infty} \frac{\log m_n(x)}{-\log n} = \frac{1}{\eta}, \text{ and } \limsup_{n\to\infty} \frac{\log m_n(x)}{-\log n} \geq 1.$$

If θ is an algebraic[3] *number, then $\eta(\theta) = 1$ and $\limsup_{n\to\infty} \frac{\log m_n(x)}{-\log r} = 1$.*

Proof for $\limsup$. By (4.3.14) and Lemma 3.5.6 $\{\|j\theta\|\}_{j\in\mathbb{N}}$ is a decreasing sequence, then it suffices to check $\limsup_{n\to\infty} \frac{\log \|n\theta\|}{\log n}$. By Hurwitz's theorem [Hur], for all irrational number $\theta \in (0,1)$ there are infinitely many pairs $p, q \in \mathbb{Z}$ such that

$$\left|\theta - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}, \tag{4.3.15}$$

[3] A number is *algebraic* if it is a root of some polynomial with integer coefficients.

which implies

$$\limsup_{n\to\infty} \frac{\log m_n(0)}{-\log n} = \limsup_{n\to\infty} \frac{\log \|n\theta\|}{-\log n} \geq 1. \qquad \square$$

Now suppose θ is an algebraic number and let $\varepsilon > 0$ be arbitrary, notice that in this case $\eta(\theta) = 1$. By the famous Thue-Siegel-Roth theorem [Rot], for all irrational algebraic number θ, there exists $c(\theta, \varepsilon)$ such that there are only finitely many pairs of $p, q \in \mathbb{Z}$ with

$$\left|\theta - \frac{p}{q}\right| \leq \frac{c(\theta,\varepsilon)}{q^{2+\varepsilon}}.$$

Since $\varepsilon > 0$ was arbitrarily small, we conclude with

$$\limsup_{n\to\infty} \frac{\log \|n\theta\|}{-\log n} \leq 1.$$

Proof for $\liminf$. Now let $n_k = q_{k+1} - 1$ where q_k is the k-th convergent's denominator as in Definition 3.5.3. By Lemma 3.5.6(c) and (4.3.15) above, by sharpness of the constant $\sqrt{5}$, for all $x \in S^1$,

$$\lim_{k\to\infty} \frac{\log m_{n_k}(x)}{\log n_k} = \lim_{k\to\infty} \frac{\log \|q_k\theta\|}{\log q_{k+1}} > \lim_{k\to\infty} \frac{\log q_k}{\log q_{k+1}} = \frac{1}{\eta},$$

therefore

$$\liminf_{n\to\infty} \frac{\log m_n(x)}{-\log n} \leq \frac{1}{\eta}.$$

To prove the lower bound for $\liminf$, recall that $\tau_r(x)$ is a cover time if for all $y \in S^1$, there exists $j \leq \tau_r(x)$ such that $d\left(f^j x, y\right) < r$. So if $\tau_r(x) = k$, there exists $i, j \leq k$ such that $d\left(f^i x, y\right)$ and $d\left(f^j x, y + 2r\right) < r$, which implies $d\left(f^i x, f^j x\right) < 4r$. Let $\varepsilon > 0$ and set $r_n = n^{-\frac{1}{\eta+\varepsilon}}$, by Theorem 3.5.4, for all n large enough there is $\tau_{r_n}(x) \leq r_n^{-(\eta+\varepsilon)} = n$, that is $m_n(x) \leq \frac{4}{n^{1/(\eta+\varepsilon)}}$. As this holds for all n large,

$$\liminf_{n\to\infty} \frac{\log m_n(x)}{-\log n} \geq \liminf_{n\to\infty} \frac{-\frac{1}{\eta+\varepsilon}\log n}{-\log n} = \frac{1}{\eta+\varepsilon}.$$

This concludes the proof for $\liminf$ and the theorem. $\square$

Again, we have shown that just like the two-point orbit case [BLR, Theorem 10], the asymptotic shortest distance in one-point orbit under irrational rotations may not converge.

4.4 Shortest distance for suspension flows

Another immediate application of Theorem 4.2.1 is the shortest distance problem on the suspension flows defined in Section 3.6. Recall the suspension flow setting: Σ is a two-sided Markov subshift of finite type, $\phi : \Sigma \to \mathbb{R}$ a Hölder potential, μ the Gibbs measure for ϕ, and $\varphi : \Sigma \to \mathbb{R}_{\geq 0}$ an $L^1(\mu)$ roof function with $\inf \varphi > 0$. The flow space is

$$Y_\varphi = \{(x,s) \in \Sigma \times \mathbb{R}_{\geq 0} : 0 \leq s \leq \varphi(x)\} / \sim, \text{ equipped with the Bowen-Walters distance } d_Y.$$

Let $\{\Psi_t\}_t$ denote the suspension flow on Y_φ and $\nu = \frac{\mu \times Leb|_{Y_\varphi}}{\mu \times Leb(Y_\varphi)}$ which is a flow-invariant probability measure on Y_φ (see *e.g.* [AK]). For each $x \in \Sigma$ and $T > 0$, let

$$k(x,T) := \max \left\{ n \in \mathbb{N}_0 : \sum_{j=0}^{n-1} \varphi\left(\sigma^j x\right) \leq T \right\}.$$

Then as φ is $L^1(\mu)$, for μ-a.e. $x \in \Sigma$, the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(\sigma^j x\right) = \lim_{T\to\infty} \frac{\sum_{j=0}^{k(x,T)-1} \varphi\left(\sigma^j x\right)/T}{k(x,T)/T} = \int \varphi \, d\mu$$

is finite. By the definition of $k(x,T)$, $\sum_{j=0}^{k(x,T)-1} \varphi\left(\sigma^j x\right)/T \to 1$ so the above implies that for μ-a.e. $x \in \Sigma$,

$$\lim_{T\to\infty} \frac{T}{k(x,T)} = \int \varphi d\mu. \tag{4.4.1}$$

Define $m_T^Y(x,s) = \min\left\{d_Y\left(\Psi_{t_1}(x,s), \Psi_{t_2}(x,s)\right) : t_1 < t_2 \leq T\right\}$ and $m_T^Y((x,s),(y,t))$ the analogous two-point version. By [RT, Theorem 4.1, Theorem 4.2], for $\nu \times \nu$ almost every $((x,s),(y,t))$,

$$\lim_{T\to\infty} \frac{\log m_T^Y((x,s),(y,t))}{-\log T} = \frac{2\log 2}{H_2(\mu)},$$

where $H_2(\mu)$ is the Rényi entropy of μ. Analogously, we will prove the following for $m_T^Y(x,s)$.

Theorem 4.4.1. *For ν-a.e. $(x,s) \in Y_\varphi$,*

$$\lim_{T\to\infty} \frac{\log m_T^Y(x,s)}{-\log T} = \frac{2\log 2}{H_2(\mu)}, \tag{4.4.2}$$

Remark 4.4.2. *The constant $\log 2$ comes from the symbolic metric d_s defined in* (3.6.5) *on the shift; if $d_s(x,y) = c^{-x\wedge y}$ for some $c > 1$ instead, then $\log 2$ in the theorem above is replaced by $\log c$. Anyhow, $\frac{H_2(\mu)}{\log 2}$ (or $\frac{H_2(\mu)}{\log c}$) should be seen as the correlation dimension $D_2(\mu)$ for the symbolic system.*

Proof. We first prove the upper bound. Let $T > 0$, by (3.6.7) and Proposition 3.6.4, for all $0 \leq t_1, t_2 \leq T$,

$$d_\pi(\Psi_{t_1}(x,s), \Psi_{t_2}(x,s)) \leq c_\pi d_Y(\Psi_{t_1}(x,s), \Psi_{t_2}(x,s)).$$

By definition of d_π, for all $(x,s),(y,t) \in Y_\varphi$,

$$d_Y((x,s),(y,t)) \leq \min\left\{d_s(x,y), d_s(\sigma x, y), d_s(\sigma y, x)\right\},$$

which implies that $m_{k(x,T+s)+1}(x) < m_T^Y(x,s)$, where $m_k(x)$ takes the shortest symbolic distance with respect to d_s between iterates of x up to time k. As $d_s(x,y) = 2^{-x\wedge y}$, $-M_k(x)\log 2 = \log m_k(x)$ so

$$\frac{M_{k(x,T+s)+1}(x)\log 2}{\log k(x,s+T)}\frac{\log k(x,s+T)}{\log T} \geq \frac{\log m_T^Y(x,s)}{-\log T}.$$

By (4.4.1) and Theorem 4.2.1 there exists an intersection S of conull sets[4], such that every $x \in S$ in the intersection has

$$\lim_{n\to\infty}\frac{M_{k(x,T+s)+1}(x)}{\log k(x,s+T)} = \frac{2}{H_2(\mu)} \text{ and } \lim_{T\to\infty}\frac{\log k(x,s+T)}{\log T} = 1. \tag{4.4.3}$$

Hence for ν-a.e. (x,s),

$$\limsup_{T\to\infty}\frac{\log m_T^Y(x,s)}{-\log T} \leq \frac{2\log 2}{H_2(\mu)}.$$

Now for the lower bound, for each $(x,s) \in Y_\varphi$ and $T > 0$, there is

$$\begin{aligned}
&m_T^Y(x,s) = \min\left\{d_Y(\Psi_{t_1}(x,s),\Psi_{t_2}(x,s)) : t_1 < t_2 \leq T\right\}\\
&\leq \min\left\{d_Y\left(\left(\sigma^{k(x,s+t_1)}x,0\right),\left(\sigma^{k(x,t_2+s)}x,0\right)\right) : t_1 < t_2 \leq T,\ k(x,t_1+s) \neq k(x,t_2+s)\right\}\\
&\leq c_\pi \min\left\{d_s\left(\sigma^{k(x,s+t_1)}x,\sigma^{k(x,t_2+s)}x\right) : t_1 < t_2 \leq T,\ k(x,t_1+s) \neq k(x,t_2+s)\right\} = 2^{-M_{k(x,s+T)}(x)}.
\end{aligned}$$

Again, taking x from the intersection of the conull sets as in (4.4.3), for ν-a.e $(x,s) \in Y_\varphi$,

$$\liminf_{T\to\infty}\frac{\log m_T^Y(x,s)}{-\log T} \geq \frac{2\log 2}{H_2(\mu)}. \qquad \square$$

Thus, we have shown that for ν-almost every (x,t) in the suspension flow Y_φ, $\lim_{T\to\infty}\frac{\log m_T^Y(x,s)}{-\log T} = \frac{2\log 2}{H_2(\mu)}$. This is just an example to demonstrate that the almost sure result for interval maps as in Theorem 4.2.1 and Theorem 4.3.4 may be proved for flows with some asymptotic independence properties, for example, the class of flows discussed in Section 3.6.

[4]A set is *conull* if its complement has zero measure

Bibliography

[AD] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn. **1** (2001) 193-237.

[ADU] J. Aaronson, M. Denker and M. Urbanski, *Ergodic theory for Markov fibred systems and parabolic rational maps*. Trans. Amer. Math. Soc. **337**(2) (1993) 495-548.

[AK] W. Ambrose and S. Kakutani, *Structure and continuity of measurable flows*, Duke Mathematical Journal, Duke Math. J. **9**(1) (1942) 25-42.

[AW] R. Arratia and M. S. Waterman, *An Erdős-Rényi law with shifts*, Adv. Math. **55** (1985) 13–23.

[Bal] V. Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, **16**, World Scientific, Singapore (2000).

[BJK] B. Bárány, N. Jurga and I. Kolossváry, *On the convergence rate of the chaos game*, Int. Math. Res. Not. IMRN **5** (2023) 4456–4500.

[Bar] M. Barnsley, *Fractals everywhere*, Academic Press Professional, Inc., USA, 1988.

[BS] L. Barreira and B. Saussol, *Multifractal analysis of hyperbolic flows*, Comm. Math. Phys. **214** (2000) 339-371.

[BLR] V. Barros, L. Liao and J. Rousseau, *On the shortest distance between orbits and the longest common substring problem*, Adv. Math. **344** (2019) 311-339.

[BR21] V. Barros and J. Rousseau, *Shortest distance between multiple orbits and generalized fractal dimensions*, Ann. Henri Poincaré, **22** (2021) 1853–1885.

[Bol] L. Boltzmann. *Le cons sur la théorie des gaz*, Gauthier-Villars, Paris, tome I 1902, tome II 1905. Re-édition Jean Gabay, Paris (1987).

[Bow] R. Bowen, *Equilibrium States and The Ergodic Theory of Anosov Diffeomorphisms*, Lect. Notes in Math. **470** Springer (1975).

[BW] R. Bowen and P. Walters. *Expansive one-parameter flows*, J. Differential Equations **12** (1972) 180-193.

[BBG] M. Boyle, J. Buzzi and R. Gómez, *Almost isomorphism for countable state Markov shifts*, J. Reine Angew. Math., **2006**592 (2006) 23-47.

[Bra] R. Bradley, *Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions*, Probab. Surveys **2** (2005) 107-144.

[BDT] H. Bruin, M. Demers and M. Todd, *Hitting and escaping statistics: mixing, targets and holes*, Adv. Math. **328** (2018) 1263–1298.

[BT1] H. Bruin and M. Todd, *Equilibrium states for interval maps: potentials with* $\sup\phi - \inf\phi < h_{top}(f)$, Comm. Math. Phys. **283** (2008) 579–611.

[BT2] H. Bruin and M. Todd, *Equilibrium states for interval maps: the potential* $-t\log|Df|$, Ann. Sci. Éc. Norm., Supér. (4) **42** (2009) 559–600.

[BLS] H. Bruin, S. Luzzatto and S. van Strien, *Decay of correlations in one-dimensional dynamics,* Ann. Sci. École Norm. Sup. (4)**36** (2003) 621– 646.

[Buz] J. Buzzi, *Puzzles of quasi-finite type, zeta functions and symbolic dynamics for multi-dimensional maps*, Ann. Inst. Fourier (Grenoble) **60** (2010) 801–852.

[BCS1] J. Buzzi, S. Crovisier and O. Sarig, *Measures of maximal entropy for surface diffeomorphisms,* Ann. of Math. (2) **195** (2022) 421–508.

[BCS2] J. Buzzi, S. Crovisier and O. Sarig, *Strong positive recurrence and exponential mixing for diffeomorphisms,* arxiv: 2501.07455 (2025)

[BS] J. Buzzi and O. Sarig, *Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps*, Ergodic Theory Dynam. Systems **23**(5) (2003) 1383-1400.

[Cli] V. Climenhaga, *Specification and towers in shift spaces*, Commun. Math. Phys. **364** (2018) 441–504.

[CGR] P. Collet, C. Giardina and F. Redig, *Matching with shift for one-dimensional Gibbs measures*, Ann. Appl. Probab. **19** (2009) 1581–1602.

[Cyr] V. Cyr, *Countable Markov shifts with transient potentials*, Proc. London Math. Soc. **103**(6) (2011) 923-949 .

[CS] V. Cyr and O, Sarig, *Spectral gap and transience for Ruelle Operators on countable Markov shifts*. Commun. Math. Phys. **292** (2009) 637–666.

[DKZa] A. Dembo, S. Karlin and O. Zeitouni, *Critical phenomena for sequence matching with scoring*, Ann.Probab. **22** (1994) 1993 –2021.

[DKZb] A. Dembo, S. Karlin and O. Zeitouni, *Limit distribution of maximal non-aligned two-sequence segmental score* Ann. Probab. **22** (1994) 2022–2039.

[DKU] M. Denker, G. Keller and M. Urbański, *On the uniqueness of equilibrium states for piecewise monotone mappings,* Studia Math. **97** (1990) 27-36.

[DT] N. Dobbs and M. Todd, *Free energy and equilibrium states for families of interval maps*, Mem. Amer. Math. Soc. **286** (2023) no.1417.

[Dob68] R.L. Dobrushin, *The problem of uniqueness of a Gibbsian random field and the problem of phase transitions*, Functional Anal. & Appl. **2** (1968) 302–312 .

[Dol] D. Dolgopyat, *On decay of correlations in Anosov flows*, Ann. Math. **147**(2) (1998) 357–390.

[FFK] K. Falconer, J. Fraser and A. Käenmäki. *Minkowski dimension for measures*, Proc. Amer. Math. Soc.,**151** (2023) 779-794.

[GGL] C. Gabriela, V. Girardin and L. Loïck, *Computation and estimation of generalized entropy rates for denumerable Markov chains*, IEEE Trans. Inform. Theory **57** (2011), 4026–4034.

[GalSch] A. Galves and B. Schmitt, *Inequalities for hitting times in mixing dynamical systems*, Random Comput. Dynam., **5** (1997) 337-347.

[GST] S. Gouëzel, B. Schapira and S. Tapie, *Pressure at infinity and strong positive recurrence in negative curvature,* Comment. Math. Helv. **98** (2023) 431–508.

[GRS] S. Gouëzel, J. Rousseau and M. Stadlbauer, *Minimal distance between random orbits*, Probab. Theory Relat. Fields **189** (2024) 811–847.

[Gur1] B. Gurevich. *Topological Entropy for denumerable Markov chains*, Dokl. Akad. Nauk. SSSR 187 (1969); Engl. transl. in Soviet Math. Dokl. **10** (1969) 911–915

[Gur2] B. Gurevich, *Shift entropy and Markov measures in the path space of a denumerable graph*, Dokl. Akad. Nauk. SSSR 192 (1970); Engl. transl. in Soviet Math. Dokl. 11 (1970) 744–747.

[GurSav] B. Gurevich and S. V. Savchenko, *Thermodynamic formalism for symbolic Markov chains with a countable number of states*, Uspekhi Mat. Nauk, **53**(2(320)) (1998) 3-106.

[HV] N. Haydn and S. Vaienti, *The Rényi entropy function and the large deviation of short return times*, Ergodic Theory Dynam. Systems **30** (2010) 159–179.

[HK] F. Hofbauer and G. Keller, *Equilibrium states for piecewise monotonic transformations,* Ergodic Theory Dynam. Systems **2** (1982) 23–43.

[HNL] M. Holland, M. Nicol and A. Török, *Extreme value theory for non-uniformly expanding dynamical systems*. Trans. Amer. Math. Soc. **364** (2012) 661-688.

[Hur] A. Hurwitz, *Ueber die angenäherte Darstellung der Irrationalzahlen durch rationale Brüche* [On the approximate representation of irrational numbers by rational fractions], Mathematische Annalen (in German)**39** (2) (1891) 279–284.

[IJT] G. Iommi, T. Jordan and M. Todd, *Recurrence and transience for suspension flows*, Isr. J. Math. **209** (2015) 547–592.

[ITV] G. Iommi, M. Todd and A. Velozo, *Escape of entropy for countable Markov shifts,* Adv. Math. **405** (2022), Paper No. 108507.

[IT] G. Iommi and M. Todd, *Transience in dynamical systems,* Ergod. Theory Dynam. Syst. **33** (2013) 1450–1476.

[IV] G. Iommi and A. Velozo, *The space of invariant measures for countable Markov shifts*, J. Anal. Math. **143**(2) (2021) 461–501.

[IR] I. Inoquio-Renteria and J. Rivera-Letelier, *A characterization of hyperbolic potentials of rational maps,* Bull. Braz. Math. Soc. (N.S.) **43** (2012) 99–127.

[JT] N. Jurga, M. Todd, *Cover times in dynamical systems*, In Press: Isr. J. Math. (2024).

[Kel1] G. Keller, *On the rate of convergence to equilibrium in one-dimensional systems*, Comm. Math. Phys. **96** (1984) 181 –193.

[Kel2] G. Keller, *Equilibrium States in Ergodic Theory*, Cambridge University Press, (1998).

[Khi] A. Khintchine, *Continued Fractions*, Univ. Chicago Press, Chicago, (1964).

[KS] D. Kim and B. Seo, *The waiting time for irrational rotations*, Nonlinearity **16** 1861 (2003).

[Kit] B. Kitchens. *Symbolic Dynamics: One-sided, Two-sided and Countable State Markov Shifts*, Springer (Universitext) (1998).

[LanRue] O. Lanford and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. **13** (1969) 194–215.

[LiRiv1] H. Li and J. Rivera-Leterlier, *Equilibrium states of interval maps for hyperbolic potentials*, Nonlinearity **27** (2014) 1779–1904.

[LiRiv2] H. Li and J. Rivera-Leterlier, *Equilibrium states of weakly hyperbolic one-dimensional maps for Hölder potentials,* Comm. Math. Phys. **328** (2014) 397–419.

[LSV] C. Liverani, B. Saussol and S. Vaienti, *Conformal measure and decay of correlation for covering weighted systems,* Ergodic Theory Dynam. Systems **18** (1998) 1399–1420.

[Mat] P. Matthews, *Covering problems for Brownian motion on spheres,* Ann. Probab. **1**(16) (1988) 189–199.

[MU] R. Mauldin and M. Urbański, *Gibbs states on the symbolic space over an infinite alphabet*, Israel J. Maths. **125** (2001) 93-130.

[NS] T. Nowicki and D. Sands, *Non-uniform hyperbolicity and universal bounds for S-unimodal maps,* Invent. Math. **132** (1998) 633–680.

[PZ] Y. Pesin and K. Zhang, *Phase transitions for uniformly expanding maps*, J. Stat. Phys. **122** (2006) 1095–1110.

[Rat] M. Ratner, *The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature*, Israel J. Math. **16** (1973) 181–197.

[Ren] A. Rényi, *Foundations of Probability*. Holden-Day (1970) San Francisco.

[RV] F. Riquelme and A. Velozo, *Ergodic optimization and zero temperature limits in negative curvature*, Annal. Henri Poincaré **23** (2022) 2949–2977.

[Rot] K. Roth, *Rational approximations to algebraic numbers,* Mathematika **2**(1) (1955) 1-20.

[RT] J. Rousseau and M. Todd, *Orbits closeness for slowly mixing dynamical systems*, Ergodic Theory Dynam. Systems **44**(4) (2024) 1192–1208.

[Rue1] D. Ruelle, *Statistical mechanics of a one-dimensional lattice gas,* Comm. Math. Phys. **9** (1968) 267–278.

[Rue2] D. Ruelle, *A measure associated with axiom-A attractors*, Amer. J. Math. **98**(3) (1976) 619–654.

[Rue3] D. Ruelle, *Thermodynamic formalism*. Encyclopedia of Mathematics and its Applications **5**, Addison-Wesley (1978).

[Rue4] D. Ruelle, *Thermodynamic formalism. The mathematical structures of equilibrium statistical mechanics*. Second Edition, Cambridge Univ. Press (2004).

[Ru1] S. Ruette, *On the Vere-Jones classification and existence of maximal measures for countable topological Markov chains,* Pacific J. Math. **209** (2003) 366–380.

[Ru2] S. Ruette, *Transitive topological Markov chains of given entropy and period with or without measure of maximal entropy,* Pacific J. Math. **303** (2019) 317–323.

[RS] R. Rühr and O. Sarig, *Effective intrinsic ergodicity for countable state Markov shifts,* Isr. J. Math. **251** (2022) 679–735.

[Sar1] O. Sarig, *Thermodynamic formalism for countable Markov shifts*, Ergodic Theory Dynam. Systems **19** (1999) 1565-1593.

[Sar2] O. Sarig, Thermodynamic formalism for null recurrent potentials. Isr. J. Math. **121** (2001) 285–311.

[Sar3] O. Sarig, *Phase transitions for countable Markov shifts*, Commun. Math. Phys. **217** (2001) 555–577.

[Sar4] O. Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002) 629–653.

[Sar5] O. Sarig, *Existence of Gibbs measures for countable Markov shifts*, Proc. Amer. Math. Soc. **131** (2003) 1751-1758.

[Sar6] O. Sarig, *Thermodynamic formalism for countable Markov shifts.* Proc. Sympos. Pure Math. **89** (2015) 81-117.

[Sin] Y. Sinai, *Gibbs measures in ergodic theory*, Russian Math. Surveys **27**(4) (1972) 21–69.

[TZ] M. Todd and B. Zhao, *Countable Markov shifts with exponential mixing*, arxiv: 2403.02092 2024.

[VJ] D. Vere–Jones, *Geometric ergodicity in denumerable Markov chains*, Quart. J. Math. Oxford **13**(2) (1962) 7–28.

[You] L. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. **147** (1998) 585–650.

[Zha1] B. Zhao, *Closest distance between iterates of typical points*, Discrete Contin. Dyn. Syst. **44**(8) (2024) 2252–2279.

[Zha2] B. Zhao, *Almost sure cover times for dynamical systems with ψ-mixing property*, arxiv: 2412.17425 (2024).