Instituto de Matemática
Universidade Federal do Rio de Janeiro

UFRJ

Topological, ergodic and stochastic properties of Iterated Function Systems on the circle

Graccyela Rosybell Salcedo Pirela

Rio de Janeiro, Brasil

20 de agosto de 2022

Topological, ergodic and stochastic properties of Iterated Function Systems on the circle

Graccyela Rosybell Salcedo Pirela

Tese de doutorado apresentada ao Programa de Pós-graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Matemática

Universidade Federal do Rio de Janeiro

Instituto de Matemática

Programa de Pós-Graduação em Matemática

Supervisor: Katrin Grit Gelfert

Rio de Janeiro, Brasil

20 de agosto de 2022

CIP - Catalogação na Publicação

M962t

Mengano, Fulano
Titulo / Fulano Mengano. -- Rio de Janeiro, 2019.
101 f.

Orientador: Fulanaorientadora Menganaorientadora.
Coorientador: Fulanacoorientadora Menganacoorientadora.
Tese (doutorado) - Universidade Federal do Rio de Janeiro, Instituto de Matemática, Programa de Pós Graduação em Matemática, 2019.

1. análise. 2. álgebra. 3. geometria. I. Menganaorientadora, Fulanaorientadora, orient. II. Menganacoorientadora, Fulanacoorientadora, coorient. III. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

Graccyela Rosybell Salcedo Pirela

Topological, ergodic and stochastic properties of Iterated Function Systems on the circle

Tese de doutorado apresentada ao Programa de Pós-graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Matemática

Trabalho aprovado por

Katrin Grit Gelfert
Orientadora

Manuel Stadlbauer

Edgar Matias da Silva

Leandro Martins Cioletti

Bruno Rodrigues Santiago

Rio de Janeiro, Brasil
20 de agosto de 2022

À minha eternamente amada irmã Gabriela.

Acknowledgements

A Deus, por estar sempre comigo, guiando meus passos.

A minha mãe, que cuidou sempre da minha educação. Obrigada Norma Pirela, além de me oferecer a oportunidade de estudar, você sempre esteve presente e estou muito feliz por isso.

A toda a minha família, pelo apoio incondicional. Especialmente, aos meus três sobrinhos: Vicky, Alice e Jesus Eduardo; por me motivarem a dar um bom exemplo para eles.

A minha orientadora, que me inspirou para trabalhar na área de sistemas dinâmicos, sempre ficarei muito grata por isso. Fico também eternamente agradecida pelo apoio e muita paciência que teve Katrin Gelfert, que no meio da pandemia esteve sempre presente fazendo da melhor forma o seu papel.

A todos os professores do IM. Aqueles que direta ou indiretamente deixaram um aprendizado em mim. Em especial à coordenadora do programa Zezé, que me apoiou em diversas ocasiões.

A todos os membros da banca por terem atendido ao convite para desempenhar este papel, dispondo de seu tempo e conhecimento para analisar este trabalho.

Aos meus amigos da salinha da Pós do IM. Sempre ir por um café com eles foi parte essencial para me manter firme, na pandemia teve muita saudade disso.

A meu namorado Miguel Soto, além de me apoiar, ele sempre me fez sorrir e isso foi fundamental para continuar.

A todo o pessoal do IM-UFRJ (limpeza, secretaria, biblioteca, etc.) que fazem um grande trabalho criando um ótimo ambiente de estudo, aconchegante e confortável.

Ao IM-UFRJ, pela oportunidade de fazer parte dele.

A Capes e CNPq pelo suporte financeiro.

Abstract

Investigamos propriedades ergódicas, topológicas e estocásticas de Sistemas de Funções Iteradas (IFS) de mapas contínuos em um espaço métrico que são escolhidos de forma aleatória, idêntica e independente.

Investigamos as condições de contração após mudanças de métricas. Estamos principalmente interessados em mudanças de métricas que tornem um IFS em um que seja contrativo em média. Para um IFS de C^1-difeomorfismos do círculo que é proximal e não tem uma medida de probabilidade simultaneamente invariante por cada mapa, obtemos uma métrica fortemente equivalente que contrai em média.

Para IFSs de C^1-difeomorfismos no círculo estabelecemos uma lei forte dos grandes números e um teorema central do limite. Além disso, descrevemos grandes desvios para a derivada em órbitas aleatórias em relação ao expoente de Lyapunov esperado.

Finalmente, estudamos IFSs no círculo que são induzidos por cociclos de matrices. Mostramos a equivalência entre certas funções de pressão: para o cociclo e para o produto torcido. Provamos a unicidade dos estados de equilíbrio para alguns potenciais.

Palavras-chave: Sistemas de funções iteradas, contração em média, sincronização, sistemas dinamicos aleátorios, difeomorfismos no circulo, exponente de Lyapunov, teorema central do limite, grandes desvios, produtos torcidos, cociclos de matrices.

Abstract

We investigate ergodic, topological and stochastic properties of Iterated Function Systems (IFS) of continuous maps on a metric space which are chosen randomly, identically and independently.

We investigate contraction conditions after metric changes. We are mainly interested in changes of metrics which turn the IFS into one which is contractive on average. For the particular case of a system of C^1-diffeomorphisms of the circle which is proximal and does not have a probability measure simultaneously invariant by every map, we derive a strongly equivalent metric which contracts on average.

For IFSs of C^1-diffeomorphisms on the circle we establish a strong law of large numbers and a central limit theorem. Moreover, we describe large deviations for the derivative on random orbits relative to the expected Lyapunov exponent.

Finally, we study IFSs on the circle which are induced by matrix cocycles. We establish equivalence between pressure functions: for the cocycle and for the associated skew-product. Moreover, we prove the uniqueness of the equilibrium states for some potentials.

Keywords: Iterated function systems, contracting on average, synchronization, random dynamical systems, diffeomorphisms on the circle, Lyapunov exponent, central limit theorem, large deviation, skew-product map, matrix cocycles.

List of Figures

List of abbreviations and acronyms

CA	Contractive on Average (1.1.1)
CLT	Central Limit Theorem (1.2.4)
ECA	Eventually Contractive on Average (2.1.6)
ESCA	Eventually Strongly Contracting on Average (2.1.16)
LCA	Local Contractive on Average (2.1.20)
LCWS	Locally Contractive in the Weak Sense (2.1.22)
LECA	Locally Eventually Contractive on Average (2.1.10)
$\log$-CA	$\log$-Contractive on Average (2.1.9)
NEA	Non-Expansive on Average (2.1.8)
S_{exp}	Exponentially Synchronizing (2.1.2)
SA	Synchronizing on Average (2.1.4)
SA_{exp}	Exponentially Synchronizing on Average (2.1.5)
SLLN	Strong Law of Large Numbers (1.2.3)

Contents

1 Introduction

For the quantitative understanding of deterministic dynamical systems, the use of probabilistic methods has become of increasing importance. In fact, the long-time behavior of "chaotic" systems shows many features of stochastic systems. Such features can rigorously be analyzed and effectively described for so-called hyperbolic dynamical systems. The theory of uniformly hyperbolic dynamical systems has been well studied since the 1960s, based on the fundamental works of Smale. However, such systems fail to contemplate a large number of interesting examples such as the Lorenz flow, as time-continuous system, and the Hénon map, as time-discrete system. So it is necessary to extend their foundations to systems that do not present uniform hyperbolicity. Thus, naturally arises the interest in studying dynamical systems that generalize the uniformly hyperbolic ones, such as non-uniformly hyperbolic or partially hyperbolic ones. The latter still show uniformity, but complete hyperbolicity is replaced by "some hyperbolicity". Understanding its behavior from the topological point of view (structure of the attractors), from the statistical point of view (stationarity of dynamics) and from the ergodic point of view (average behavior along the orbits) is, therefore, of fundamental interest and has guided several works.

One most challenging problem is to study quantifiers of "chaotic" behavior and of objects that remain invariant under the time-evolution of a dynamical system, such as fractal dimensions or entropies, or the scaling and self-similar properties of invariant probability distributions. A second aspect is to investigate statistical properties of a system by means of certain limit laws (for example, a central limit theorem) and to reveal its stochastic-like behavior. Part of this can be achieved using the so-called thermodynamic formalism, originally developed by theoretical physicists. One main object there is the topological pressure, that is, a particular functional on the space of observables, which encodes several important quantities of the underlying dynamical system. The pressure functional ties together, for example, Lyapunov exponents, entropy, fractal dimensions, multifractal spectra, correlation decay rates.

The main focus of this work is on random dynamical systems, in particular on iterated function systems with probabilities. A special focus will be on systems on the circle. Hutchinson [Hut81] was the pioneer in studying iterated function systems. Assuming contraction, he showed various properties of these systems, such as existence and uniqueness of a stationary measure. He also characterized the support of the stationary measure, specified its Hausdorff dimension and showed that this set has only three options: it is either the full space, a finite set, or a fractal (which in 1-dimension this is a Cantor set). In the case of an iterated function system on the circle, instead of contraction other, purely topological, conditions such as minimality guarantee analogous properties. This is for instance studied in [Nav11]. Another topological property, which is in a certain sense an opposing effect, is synchronization. For iterated function systems on the circle the following so-called Invariance Principle [Mal17, Theorem F] holds true:

- either each map of the iterated function system preserves a common probability measure,
- or the iterated function system has the local contraction property: given any point of the circle, typical compositions of the homeomorphisms contract some neighborhood of the point.

The synchronization properties studied in [Mal17] play an important role in this thesis.

1.1 Contracting on average Iterated Function Systems

Given a complete metric space (M, d), an *iterated function system (IFS)* is a finite set $\mathcal{F} = \{f_0, \dots, f_{N-1}\}$, $N \geq 2$, of continuous maps $f_i \colon M \to M$, $i = 0, \dots, N-1$. One important goal is to understand the asymptotic behavior of consecutive concatenations of maps in $\mathcal{F}$ where the choice at each step is according to probabilities p_i, $i = 0, \dots, N-1$, of some probability vector $\mathbf{p} = (p_0, \dots, p_{N-1})$, defining the triple $(\mathcal{F}, \mathbf{p}, d)$. This behavior is very well understood under the hypothesis that every map contracts uniformly. However, this is a rather strong requirement. Several weaker hypotheses imply also good stochastic properties of the associated Markov chain generated by the IFS. For example, an IFS which is contracting on average (the concatenated maps do not necessarily shrink the distance between two points at every step and everywhere, but in expectation they do; see (1.1.1)) has a contracting (hence unique) stationary Borel probability measure.

Though, much less is known for an IFS of maps which either “just do not expand” or have simultaneously some “contracting regions” and some “expanding regions” or even repelling fixed points (compare the examples depicted in Figure 4). In such a general context, to gain any average contraction, one least topological requirement was coined in [Ste01] is that “the orbit of a point wanders sufficiently around the space to pick up an average contraction”. Such property would, for example, call for an IFS which is minimal[1].

Observe that the existence and uniqueness of a Markov chain-stationary measure does not depend on the metric (as long as metrics generate the same topology), while contraction properties do. A natural question is when for $(\mathcal{F}, \mathbf{p}, d)$ there exists some metric D on M equivalent to d such that $(\mathcal{F}, \mathbf{p}, D)$ is, for example, contracting on average. Besides surveys such as [Kai81, DF99, Ios09], we point out [Ste12] which provides an ample discussion of many kinds of contracting conditions and [LSS20] which reviews IFSs from a more topological point of view, both mentioning also the method of metric change. In Section 3.1 we briefly discuss some stochastic properties that remain true for the contracting on average IFS after a metric change.

[1] The IFS $\mathcal{F}$ is *forward minimal* if for every nonempty closed set $A \subset M$ satisfying $f_i(A) \subset A$ for every $i = 0, \dots, N-1$ it holds $A = M$. The IFS $\mathcal{F}$ is *backward minimal* if $\mathcal{F}^{-1} = \{f_i^{-1}\}$ is forward minimal.

One step towards this direction was done in [GS17] where a convenient metric change turns a backward minimal IFS of homeomorphisms on $\mathbb{S}^1$ into a non-expansive on average one (see also [SZ20]). Here, for a non-expansive on average IFS of Lipschitz maps on a compact metric space, we give (sufficient and necessary) conditions to guarantee a metric change which turns an IFS into a contracting on average one. Moreover, we thoroughly discuss several local and global contraction-like properties intimately related with it.

The existence of a metric which makes a *dynamical system* “contracting” or “expanding” has been widely studied. Let us cite some key results. If a map “eventually” contracts (expands) in the sense that its kth iteration has this property, a convenient change of the metric turns it into a contraction (expansion) in its first iteration (see, for example, [PU10, Chapter 4]). As explained, for example, in [Shu87, Chapter 4] a hyperbolic structure of a diffeomorphism is independent of the Riemannian metric on the ambient manifold. In [Fri87, Fat89], using Frink’s metrization theorem, it is shown that for any expansive homeomorphism of a compact topological space there exists an equivalent metric such that the map contracts (expands) on stable (unstable) sets. Analogously, there exists a metric which turns a (positively) expansive continuous map of a compact metric space into an expanding one (see [PU10, Chapter 4]).[2]

Before stating the first main result, let us introduce the main contraction properties which we are going to investigate. A precursor in [DF99] requires *contraction in mean*: f_i are Lipschitz with Lipschitz constants $\mathrm{Lip}(f_i)$, $i = 0, \dots, N-1$, satisfying

$$\sum_{i=0}^{N-1} p_i \,\mathrm{Lip}(f_i) < 1.$$

Weaker concepts are proposed in [BDEG88, Pei93], where $(\mathcal{F}, \mathbf{p}, d)$ is assumed to *contract on average (CA)* in the sense that there exists $\lambda \in (0,1)$ so that

$$\sum_{i=0}^{N-1} p_i d(f_i(x), f_i(y)) \le \lambda d(x,y) \quad \text{for every } x, y \in M. \tag{1.1.1}$$

In fact, [BDEG88] requires even weaker assumptions allowing for place-dependent probabilities and for contraction in L^q norm, $q > 0$, while [Pei93] assumes “contraction on average after some iterations” (that is, it is k-eventually contracting in average, for some $k \in \mathbb{N}$, in the sense defined in (2.1.6) below). Less is known about an IFS if we put only the weaker hypothesis that $(\mathcal{F}, \mathbf{p}, d)$ is *non-expansive on average (NEA)* in the sense that

$$\sum_{i=0}^{N-1} p_i d(f_i(x), f_i(y)) \le d(x,y) \quad \text{for every } x, y \in M.$$

This property implies, for example, that the associated Markov chain is non-expansive (see [Sza03] and references therein). There are variations of these definitions such as being *eventually*

[2] The existence of so-called adapted metrics in partially hyperbolic dynamics and dynamics with a dominated splitting was investigated in [Gou07, HPS77]. The use of adapted norms and metrics is also common in the study of nonuniform hyperbolicity when analyzing the size of local un-/stable manifolds (see, for example, [BP07] in the $C^{1+\varepsilon}$ case and [ABC11, Section 8] for a C^1 dominated setting).

strongly contracting on average (ESCA), *synchronizing (S)*, *synchronizing on average (SA)*, *locally eventually contractive on average (LECA)*, and *proximal*, that we postpone to Section 2.1. To simplify the exposition, we will mainly use these short notations.

Recall that metrics d and D on some common space M are *(topologically) equivalent* if they generate the same topology. They are *strongly equivalent* if there exist positive constants a and b such that $ad(x,y) \le D(x,y) \le bd(x,y)$ for every $x,y \in M$. Clearly, strong equivalence implies topological one, but not vice versa. Given $\alpha \in (0,1)$, note that $d^\alpha \colon M \times M \to [0,\infty)$ defined by $d^\alpha(x,y) \overset{\text{def}}{=} (d(x,y))^\alpha$ is a metric on M, and d and d^α are equivalent.

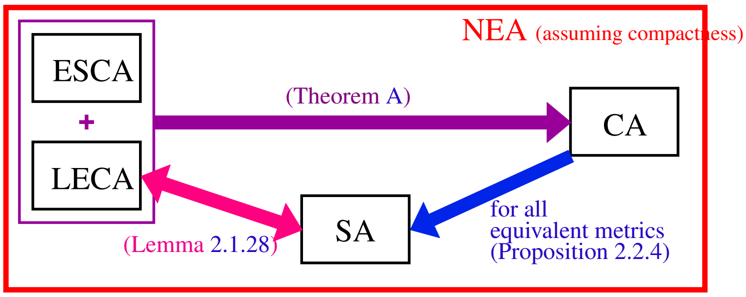

Figure 1 – Some implications, assuming non-expansive on average (NEA) on a compact metric space: eventually strongly contracting on average (ESCA), contracting on average (CA), synchronizing on average (SA), locally eventually contractive on average (LECA).

Let us now state our first main result. Let $\Sigma_N^+ \overset{\text{def}}{=} \{0,\dots,N-1\}^{\mathbb{N}}$ be the space of one-sided sequences and denote by μ the Bernoulli measure on Σ_N^+ determined by $\mathbf{p}$. For any sequence $\xi = (\xi_1\xi_2\dots) \in \Sigma_N^+$, $n \ge 1$, and $x \in M$ let

$$f_\xi^n(x) \overset{\text{def}}{=} f_{\xi_1\dots\xi_n} \overset{\text{def}}{=} f_{\xi_n} \circ f_{\xi_{n-1}} \circ \dots \circ f_{\xi_1}(x), \quad f_\xi^0(x) \overset{\text{def}}{=} x.$$

Given $x,y \in M$ and $n \in \mathbb{N}$, let

$$Z_{n,d}^{x,y}(\xi) \overset{\text{def}}{=} d(f_\xi^n(x), f_\xi^n(y)), \quad Z_{0,d}^{x,y}(\xi) \overset{\text{def}}{=} d(x,y).$$

Given $\lambda > 0$ and $n \in \mathbb{N}$, consider the metric defined by

$$d_{n,\lambda}(x,y) \overset{\text{def}}{=} d(x,y) + \frac{1}{\lambda^{1/n}}\mathbb{E}(Z_{1,d}^{x,y}) + \dots + \frac{1}{\lambda^{(n-1)/n}}\mathbb{E}(Z_{n-1,d}^{x,y}),$$

where $\mathbb{E}(\cdot)$ denotes the expected value according to the probability distribution.

The following result, Theorem A, is proved in Section 2.2.

Theorem A. *Consider a compact metric space (M,d) and a triple $(\mathcal{F},\mathbf{p},d)$, where $\mathbf{p}$ is a non-degenerate probability vector and $\mathcal{F}$ is an IFS of Lipschitz maps that is non-expansive on average. If $(\mathcal{F},\mathbf{p},d)$ satisfies LECA and ESCA, then there exist $\lambda \in (0,1)$ and $n \in \mathbb{N}$ such that $(\mathcal{F},\mathbf{p},d_{n,\lambda})$ is contracting on average. Moreover, LECA is equivalent to SA.*

The second main result concerns the particular case of an IFS of C^1 diffeomorphisms of the circle $\mathbb{S}^1$ equipped with the usual metric $d(x,y) \stackrel{\text{def}}{=} \min\{|x-y|, 1-|x-y|\}$. Theorem B is proved in Section 2.3

Theorem B. *Assume that $\mathcal{F}$ is an IFS of C^1-diffeomorphisms on $\mathbb{S}^1$. Assume that $(\mathcal{F}, d)$ is proximal and there does not exist a probability measure which is invariant by every map in $\mathcal{F}$. Then for every non-degenerate probability vector $\mathbf{p}$ there exist $\alpha \in (0,1]$, $\lambda \in (0,1)$, and $n \in \mathbb{N}$ such that $(\mathcal{F}, \mathbf{p}, D)$, with $D \stackrel{\text{def}}{=} (d^\alpha)_{n,\lambda}$, is contracting on average. Moreover, $d \leq D \leq Cd^\alpha$ for some $C > 0$, and hence d and D are strongly equivalent.*

In Section 2.4 we illustrate and discuss our results in two classes of homeomorphisms on $\mathbb{S}^1$ (compare Figure 4). We summarize their main properties.

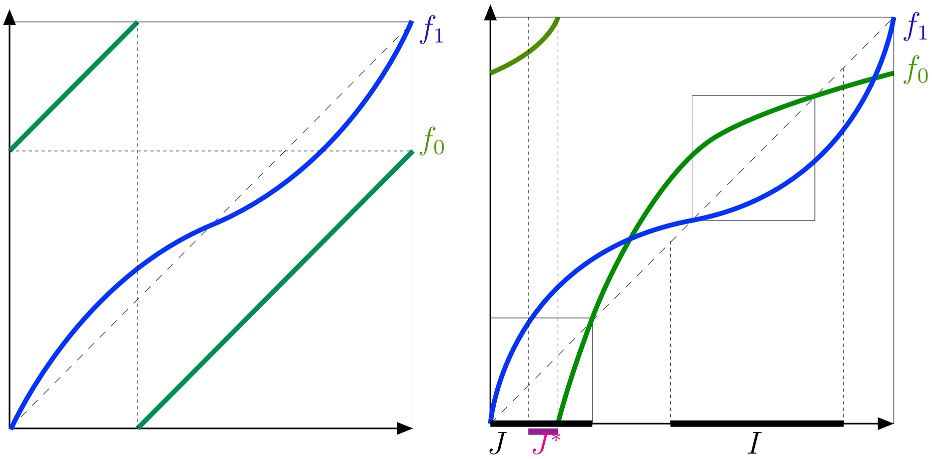

Figure 2 – Examples studied in Section 2.4.2 (left) and Section 2.4.3 (right)

Example 1.1.1 ($(\mathcal{F}, \mathbf{p}, d)$ in Section 2.4.2, see Figure 4 (left))**.** This example of an IFS of C^1-diffeomorphisms is proximal (and hence SA and LECA) but fails to be NEA. The choice of metric ρ in [GS17] forces $(\mathcal{F}, \mathbf{p}, \rho)$ to satisfy NEA, SA, and LECA, but $(\mathcal{F}, \mathbf{p}, \rho)$ fails to be ESCA. As $(\mathcal{F}, \mathbf{p}, \rho)$ verifies the hypotheses of Theorem B, there exist $\alpha \in (0,1]$, $\lambda \in (0,1)$ and $n \in \mathbb{N}$ such that for $D \stackrel{\text{def}}{=} (d^\alpha)_{n,\lambda}$ the triple $(\mathcal{F}, \mathbf{p}, D)$ is CA. In particular, $(\mathcal{F}, \mathbf{p}, D)$ satisfies NEA, SA, LECA, and ESCA. However, for all $\beta \in (0,1]$ the metric D^β fails to be strongly equivalent to d.

Example 1.1.2 ($(\mathcal{F}, \mathbf{p}, d)$ in Section 2.4.3, see Figure 4 (right))**.** In this example the approach in [GS17] does not apply. This example fails to be NEA and ε-LCA, but it is proximal, S, SA, and LECA. For this example, it is shown that for appropriate $\alpha \in (0,1]$ the metric

$$\hat{D}(x,y) \stackrel{\text{def}}{=} \mathbb{E}\Big(\sup_{n\geq 0} Z^{x,y}_{n,d^\alpha}\Big),$$

is strongly equivalent to d^α and $(\mathcal{F}, \mathbf{p}, \hat{D})$ satisfies NEA, SA (and hence LECA), and ESCA. Hence, by Theorem A, there is a metric D which is strongly equivalent to $\hat{D}$ (and hence to d^α)

such that $(\mathcal{F}, \mathbf{p}, D)$ is CA. In particular, $(\mathcal{F}, \mathbf{p}, D)$ satisfies NEA, LECA, and ESCA. Moreover, if f_0 and f_1 are C^1-diffeomorphisms which have no common fixed points, then Theorem B applies.

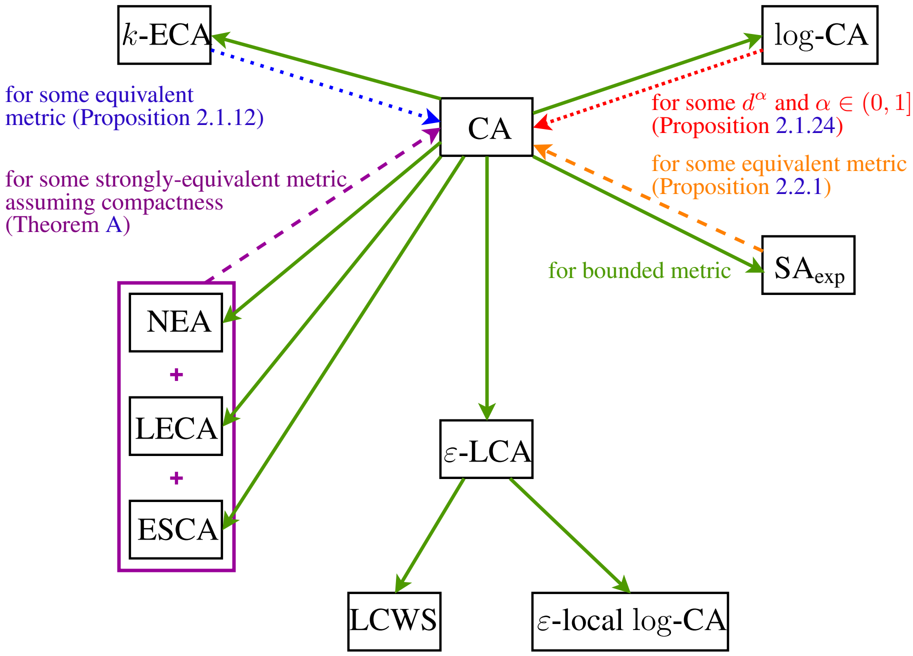

Figure 3 – Implications between: contracting on average (CA), eventually contracting on average (ECA), locally eventually contractive on average (LECA), non-expansive on average (NEA), eventually strongly contracting on average (ESCA), log-contractive on average (log-CA), ε-local contractive on average (ε-LCA), locally contractive in the weak sense (LCWS), and ε-local log-CA

1.2 Stochastic properties for contracting on average IFSs

As discussed above, a change of metrics does not alter the distribution of orbits in the space. Thus, a statistical property for $(\mathcal{F}, \mathbf{p}, d)$ holds also for $(\mathcal{F}, \mathbf{p}, D)$, provided that d and D are equivalent. We describe stochastic properties for contracting on average IFSs and discuss the impact of a metric change. The main results of Chapter 3 are established for IFS on the circle satisfying the hypotheses of Theorem B, because in this case we have satisfied the contraction on average property after metric change. Recall again that contraction on average implies the existence and uniqueness of a stationary measure which in turn implies interesting statistical properties. The first result is immediate consequence of [BDEG88, Theorem 2.1] and Theorem B and establishes a strong law of large numbers and a central limit theorem. The second result states large deviations for the derivative on random orbits with respect to Lyapunov exponent.

Recall that a Borel probability measure ν on $\mathbb{S}^1$ is *stationary* for the IFS $(\mathcal{F}, \mathbf{p})$ if

$$\mathcal{F}_*\nu = \nu, \qquad \text{where} \qquad \mathcal{F}_*\nu \stackrel{\text{def}}{=} \sum_{i=0}^{N-1} p_i (f_i)_*\nu. \tag{1.2.1}$$

We will provide some further details and justification of the term stationary below, see Section 3.1.

For $x \in \mathbb{S}^1$ consider the Dirac measure δ_x at x. Fixed $x \in \mathbb{S}^1$, define the Markov chain $(W_n^x)_{n\in\mathbb{N}}$ on the probability space (Σ_N^+, μ) taking values in $\mathbb{S}^1$ with initial distribution δ_x given by

$$W_n^x(\xi) \stackrel{\text{def}}{=} f_\xi^n(x), \quad W_0^x(\xi) \stackrel{\text{def}}{=} x, \tag{1.2.2}$$

where $\xi = (\xi_1, \xi_2, \dots) \in \Sigma_N^+$ and $f_\xi^n = f_{\xi_n} \circ \dots \circ f_{\xi_1}$. We will study the following asymptotic behavior of $(W_n^x)_{n\in\mathbb{N}}$: given an integrable function $h : \mathbb{S}^1 \to \mathbb{R}$

(SLLN) $(\mathcal{F}, \mathbf{p})$ satisfies the *strong law of large numbers* for h, if for every $x \in \mathbb{S}^1$,

$$\frac{1}{n}\sum_{k=0}^{n-1} h\left(W_k^x\right) \stackrel{a.s.}{\to} \nu(h) \stackrel{\text{def}}{=} \int h\, d\nu \quad \text{as } n \to \infty. \tag{1.2.3}$$

(CLT) $(\mathcal{F}, \mathbf{p})$ satisfies the *central limit theorem* for h, if for every $x \in \mathbb{S}^1$,

$$\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} h\left(W_k^x\right) \to N(\nu(h), \sigma^2(h)) \quad \text{as } n \to \infty \tag{1.2.4}$$

where

$$\sigma^2(h) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \int \left(\sum_{k=0}^{n-1} h\left(W_k^x(\xi)\right) - n\nu(h)\right)^2 d(\mu \otimes \nu)(\xi, x). \tag{1.2.5}$$

The following result is proved in Section 3.1.

Theorem C. *Assume that $\mathcal{F}$ is an IFS of C^1-diffeomorphisms on $\mathbb{S}^1$. Assume that $(\mathcal{F}, d)$ is proximal and there does not exist a probability measure which is invariant by every map in $\mathcal{F}$. Then for every non-degenerate probability vector $\mathbf{p}$ there is a unique stationary probability ν for $(\mathcal{F}, \mathbf{p})$. Moreover, for every $x \in \mathbb{S}^1$, the Markov chain $(W_n^x)_{n\in\mathbb{N}}$ associated to $(\mathcal{F}, \mathbf{p})$ defined in (1.2.2) satisfies (SLLN) and (CLT) for any Lipschitz function $h \colon \mathbb{S}^1 \to \mathbb{R}$, further, for $\sigma^2(h)$ as in (1.2.5) we have $\sigma^2(h) > 0$ whenever there does not exist a function g such that $h = g - g \circ f$, for all $f \in \mathcal{F}$.*

The assertion of uniqueness of the stationary measure is not new, to the best of our knowledge a first reference under the hypotheses of Theorem C is in [DKN07, Proposition 5.5.]. Theorem C provides new sufficient conditions for the CLT to hold. It complements the central limit theorem [SZ21, Theorem 9] which is stated for Hölder observables φ and an IFS of circle homeomorphisms additionally assuming that $(\mathcal{F}, d)$ acts minimally.

When the IFS is formed by diffeomorphisms, to know the behavior of the derivative $(f_\xi^n)'(x)$ helps to predict the behavior of the orbits $f_\xi^n(y)$ with $y \in \mathbb{S}^1$ near x. We are now going to discuss large deviation of finite-time Lyapunov exponents $\frac{1}{n}\log|(f_\xi^n)'(x)|$. The theory of large deviations deals with the probabilities of rare events (or fluctuations) that are exponentially small as a function of some parameter, e.g., the number of random components of a system, the time over which a stochastic system is observed, the amplitude of the noise perturbing a dynamical system or the temperature of a chemical reaction. The theory has applications in many different scientific fields, ranging from queuing theory to statistics and from finance to engineering. In its basic form, the theory of large deviations considers the asymptotic behavior of $\log \mathbb{P}(E_n)$ for a sequence of events $\{E_n\}$ with asymptotically vanishing probability. To be more precise, for $\varepsilon > 0$ and a sequence $\{X_n\}$ of random $\mathbb{R}$-valued variables, consider $S_n \stackrel{\text{def}}{=} X_0 + X_1 + \ldots + X_{n-1}$. If this process is independent, identically distributed and if its first moment is finite, then the average $\frac{1}{n}S_n$ converges almost surely to the mean $\mathbb{E}(X_0)$. In particular it also converges in measure:

$$\mathbb{P}\left[\left|\frac{1}{n}S_n - \mathbb{E}(X_0)\right| > \varepsilon\right] \to 0 \quad \text{as } n \to 0.$$

The event $|\frac{1}{n}S_n - \mathbb{E}(X_0)| > \varepsilon$ is called a *tail event*. The asymptotic behavior of tail events is the object of the theory of large deviations (see [RAS15]). A classical result in this theory is, for example, the Large deviation principle by H. Cramér in [DZ98, Theorem 2.2.3].

Consider $(\mathcal{F}, d)$ satisfying the hypothesis of Theorem C. Let F be the *skew-product map* given by

$$F(\xi, x) \stackrel{\text{def}}{=} (\sigma(\xi), f_{\xi_1}(x)). \tag{1.2.6}$$

Note that, the product measure $\mu \otimes \nu$ is F-invariant (this is immediate consequence of the stationarity of ν). Hence, by Birkhoff's Ergodic Theorem, for $(\mu \otimes \nu)$-almost every $(\xi, x) \in \Sigma_N^+ \times \mathbb{S}^1$ we have

$$\lim_{n\to\infty} \frac{1}{n} \log |(f_\xi^n)'(x)| = \gamma(\mathbf{p}) \stackrel{\text{def}}{=} \int_{\Sigma_N^+ \times \mathbb{S}^1} \log |(f_{\xi_1})'(x)| d\mu \otimes \nu(\xi, x). \tag{1.2.7}$$

The constant $\gamma \stackrel{\text{def}}{=} \gamma(\mathbf{p})$ is called *Lyapunov exponent* and depends on the probability vector $\mathbf{p}$. Lyapunov exponents are natural quantifiers that characterize dynamical systems attractors and their sensitivity on initial conditions. Note that $\gamma < 0$ (see, for example, [Mal17, Theorem F]).

In the following result, we describe the deviation of finite-time Lyapunov exponents from $\gamma(\mathbf{p})$. In the second item, we strengthen the exponential pointwise synchronization obtained in [Mal17] in two aspects (see Remark 2.3.3): we obtain that the "expected" exponential synchronization rate is $\gamma(\mathbf{p})$ and we describe the exponentially small large deviations. We will prove it in Section 3.2.4.

Theorem D. *Assume that $\mathcal{F}$ is an IFS of $C^{1+\beta}$ diffeomorphisms, for some $\beta > 0$, on $\mathbb{S}^1$ so that $(\mathcal{F}, d)$ is proximal and there does not exist a probability measure which is invariant by every element of $\mathcal{F}$. For every non-degenerate probability vector $\mathbf{p}$ consider the Bernoulli measure μ on Σ_N^+ determined by $\mathbf{p}$, the stationary probability measure ν for $(\mathcal{F}, \mathbf{p}, d)$ and the Lyapunov exponent $\gamma = \gamma(\mathbf{p}) < 0$. Then there exist $h, \varepsilon_0 > 0$ and $c > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ and $n \in \mathbb{N}$*

1. *for every* $x \in \mathbb{S}^1$

$$\mu\left(\xi \in \Sigma_N^+ : \left|\log |(f_\xi^n)'(x)| - n\gamma\right| > n\varepsilon\right) \le ce^{-nh\varepsilon^2}$$

2. *for every* $x, y \in \mathbb{S}^1$, $x \ne y$, *there exists* $C = C(x,y) > 0$ *so that*

$$\mu\left(\xi \in \Sigma_N^+ : \left|\log \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} - n\gamma\right| > n\varepsilon\right) \le Ce^{-n\min\{\varepsilon/2, h\varepsilon^2/4\}/2}.$$

One further interesting property of an IFS is the convergence on average of the orbits. From Lemma 2.1.8 and Proposition 2.2.1, when the metric space is bounded the conditions CA and $\mathrm{SA}_{\mathrm{exp}}$ are equivalent, after a possible metric change. Additionally, we show in Section 3.2.2 that $\mathrm{SA}_{\mathrm{exp}}$ with rate $\lambda \in (0,1)$ implies exponential decay of the amount of functions that do not synchronize any two orbits with the same rate λ. One more time, in the context of C^1-diffeomorphisms on $\mathbb{S}^1$ we show $\mathrm{SA}_{\mathrm{exp}}$ with respect d^α, for d usual metric on $\mathbb{S}^1$ and some $\alpha \in (0,1)$.

1.3 Matrix cocycles

In Chapter 4, we restrict ourselves further and study diffeomorphisms on the projective space which are generated by matrix cocycles. As the projective space is homeomorphic to the circle, this means, that we study a very special class of circle diffeomorphisms. This restriction is due to some technical tools which, up to now, could not yet be put into larger generality. Indeed, matrix cocycles provide much stronger tools. One of them is a type of quasi-multiplicativity. This property was first shown in the case of non-negative matrices in [FL02] and, recently in almost complete generality in [Fen09a]. So far it was very little explored further. It has very strong immediate consequences, for example to investigate thermodynamic aspects.

To formulate the main results of Chapter 4, let us introduce some notation. Let $\mathrm{SL}(2,\mathbb{R})$ be the set of 2×2 matrices with real coefficients and determinant one. One can associate to any matrix $A \in \mathrm{SL}(2,\mathbb{R})$ a circle diffeomorphism as follows. Consider the projective map $f_A \colon \mathbb{P}^1 \to \mathbb{P}^1$ defined by

$$f_A(v) \stackrel{\text{def}}{=} \frac{Av}{\|Av\|},$$

where $\mathbb{P}^1$ denotes the projective line, which topologically is the circle $\mathbb{S}^1$.

Denoting by $\sigma \colon \Sigma_N^+ \to \Sigma_N^+$ the usual left shift (see (3.2.4)).Let us fix a finite family $\mathbb{M} \stackrel{\text{def}}{=} \{M_0, \ldots, M_{N-1}\} \subset \mathrm{SL}(2,\mathbb{R})^N$. Let $\mathcal{F}_{\mathbb{M}}$ be the IFS on the projective space formed by the maps induced by the matrices in $\mathbb{M}$, that is, $\mathcal{F}_{\mathbb{M}} \stackrel{\text{def}}{=} \{f_{M_0}, \ldots, f_{M_{N-1}}\}$. The *linear cocycle* over $(\Sigma_N^+, \mu, \sigma)$ associated to $\mathbb{M}$ is the skew-product map (analogous to (1.2.6)) given by

$$F \colon \Sigma_N^+ \times \mathbb{P}^1 \to \Sigma_N^+ \times \mathbb{P}^1, \quad F(\xi, x) \stackrel{\text{def}}{=} (\sigma(\xi), f_{M_{\xi_1}}(x)).$$

Let $\mathcal{M}(\Sigma_N^+)$ denote the set of all shift-invariant Borel probability measures on Σ_N^+. Let $\mathcal{M}_{\text{erg}}(\Sigma_N^+)$ be the set of the measures in $\mathcal{M}(\Sigma_N^+)$ which is ergodic. Let $\mathcal{M}(\Sigma_N^+ \times \mathbb{P}^1)$ be the set of all F-invariant Borel probability measures on $\Sigma_N^+ \times \mathbb{P}^1$, and let $\mathcal{M}_{\text{erg}}(\Sigma_N^+ \times \mathbb{P}^1)$ be the set of the ergodic measures. Denote by $\pi\colon \Sigma_N^+ \times \mathbb{P}^1 \to \Sigma_N^+$ the projection $\pi(\xi, v) \overset{\text{def}}{=} \xi$. It is not hard to see that for any $\mu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+ \times \mathbb{P}^1)$ the measure $\nu = \pi_*\mu$ is in $\mathcal{M}_{\text{erg}}(\Sigma_N^+)$.

For $\xi \in \Sigma_N^+$, let

$$\mathbb{M}^n(\xi) \overset{\text{def}}{=} M_{\xi_n} \circ \ldots \circ M_{\xi_1} \quad \text{and} \quad f_\xi^n \overset{\text{def}}{=} f_{\mathbb{M}^n(\xi)} = f_{M_{\xi_n}} \circ \ldots \circ f_{M_{\xi_1}}.$$

By Kingman's subadditive ergodic theorem, for any σ-ergodic Borel probability measure ν on Σ_N^+, for almost every ξ the following limit exists

$$\lambda(\mathbb{M}, \xi) \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \log \|\mathbb{M}^n(\xi)\| = \mathbb{M}_*(\nu), \tag{1.3.1}$$

where $\mathbb{M}_*(\nu)$ is called *Lyapunov exponent* relative to ν and is given by

$$\mathbb{M}_*(\nu) \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \int \log \|\mathbb{M}^n(\xi)\|\, d\nu(\xi). \tag{1.3.2}$$

Following [BL85, page 48], a family $\mathbb{M}$ is *irreducible* if there is no non-zero proper linear subspace V of $\mathbb{R}^2$ such that $M_i V \subset V$ for all $i \in \{0, \ldots, N-1\}$. As we will see, this property implies a type of quasi-multiplicativity of the matrix norm over $\mathbb{M}$ (see Proposition 4.3.1).

Remark. *It is not hard to show that $\mathbb{M}$ is irreducible if and only if there is no common fixed point by all maps in the IFS $\mathcal{F}_{\mathbb{M}}$. Indeed, for all $A \in \mathrm{SL}(2, \mathbb{R})$ every proper linear subspace of $\mathbb{R}^2$ is associated to a unique fixed point of f_A, and vice versa. More specifically, if a proper linear subspace V of $\mathbb{R}^2$ holds $AV \subset V$, then there exists $a \in \mathbb{R}$ such that $Av = av$, for all $v \in V$. Thus, for $v \in \mathbb{P}^1$ representing the space V we have $f_A(v) = v$. On the other hand, if there exists $v \in \mathbb{P}^1$ such that $f_A(v) = v$, then for $V = \{tv : t \in \mathbb{R}\}$ we have $Aw = rw$ for all $w \in V$, where $r = \pm\|Av\|$. That is, V is a proper linear subspace of $\mathbb{R}^2$ invariant by A. Consequently, there is a proper linear subspace of $\mathbb{R}^2$ invariant by all matrices in $\mathbb{M}$, if and only if there is a fixed point by all maps in $\mathcal{F}_{\mathbb{M}}$.*

Let us consider a particular case that $\mathbb{M}$ contains (at least) two elements: a hyperbolic matrix and a matrix representing an irrational rotation. Note that this induces an IFS which is a particular case of the example studied in Section 2.4.2. Further, this is a particular example of elliptic cocycles having "some hyperbolicity" as defined in [DGR19, Sect. 11.7]. This set of cocycles, denoted by $\mathfrak{E}_{N,\text{shyp}}$, is defined as follows. The set $\mathfrak{E}_{N,\text{shyp}}$ consists of cocycles $\mathbb{M} \subset \mathrm{SL}(2, \mathbb{R})^N$ having the following properties:

- Some hyperbolicity: There exist $\xi \in \Sigma_N^+$ and $n \in \mathbb{N}$ such that the matrix $\mathbb{M}^n(\xi)$ is hyperbolic, that is, has one eigenvalue with absolute value bigger than one and one smaller than one.

- Transitions in finite time: There exist $\zeta \in \Sigma_N^+$ and $m \in \mathbb{N}$ such that $\mathbb{M}^m(\zeta)$ is an irrational rotation.

In fact, in [DGR19], there is studied a slightly more general class of cocycles.

Remark 1.3.1. It is not hard to check that any $\mathbb{M} \in \mathfrak{E}_{N,\mathrm{shyp}}$ is irreducible.

To state our next result, let us define pressure function and equilibrium states. The *pressure function*, relative to the cocycle, is given by

$$P(q) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \log \sum_{\xi_1,\ldots,\xi_n} \|\mathbb{M}^n(\xi)\|^q, \quad q \geq 0, \tag{1.3.3}$$

where the above limit exists by sub-additivity arguments. Let us also consider the following *variational pressure* given by

$$P_{\text{var}}(q) \stackrel{\text{def}}{=} \sup_{\nu\in\mathcal{M}(\Sigma_N^+)} \left(h_\nu(\sigma) + q\mathbb{M}_*(\nu)\right), \quad q \in \mathbb{R}. \tag{1.3.4}$$

By [CFH08, Corollary 1.2], for every $q \geq 0$ we have the following *variational principle*

$$P(q) = P_{\text{var}}(q). \tag{1.3.5}$$

A measure $\nu \in \mathcal{M}(\Sigma_N^+)$ is an *q-equilibrium state* for P if it realizes the supremum in (1.3.4). Let $\mathcal{I}_q$ be the set of q- equilibrium states. Note that both $\mathbb{M}_*(\cdot)$ and $h_{(\cdot)}(\sigma)$ are upper semi-continuous on $\mathcal{M}(\Sigma_N^+)$. Hence, $\mathcal{I}_q$ is a non-empty closed convex subset of $\mathcal{M}(\Sigma_N^+)$. In particular, $\mathcal{I}_q$ contains ergodic elements (each extreme point of $\mathcal{I}_q$ is an ergodic measure).

Remark. *A priori, the definition of the pressure function in* (1.3.3) *can be extended to the domain $q \in \mathbb{R}$. Indeed, for $q = 0$ it is easy to check that $P(0) = \log N$. By sup-additivity, for $q < 0$ the above limit exists. However, in the present state of the art, further tools such as, a variational principle as in* (1.3.5) *were not established, except for the case when $\mathbb{M}$ is formed by strictly positive matrices, see [Fen04, Theorem 1.1]. Moreover, the uniqueness of equilibrium states fails in general in the case $q < 0$, see [Fen09a, Example 6.6].*

The following theorem answers parts of open questions stated in [DGR19, Remark 2.2] in the context of step skew-products induced by $\mathrm{SL}(2,\mathbb{R})$ cocycles. Denote by $\dim_{\mathrm{H}}$ the Hausdorff dimension, and by h_{top} the topological entropy. As we consider sets that in general are non-compact, we use the concept of entropy defined by Bowen in [Bow73]. It is shown in Section 4.3.3.

Theorem E. *If $\mathbb{M} \in \mathfrak{E}_{N,shyp}$, then for any $q > 0$ there exists a unique q-equilibrium state $\nu_q \in \mathcal{M}(\Sigma_N^+)$ for P, P is differentiable at q, $P'(q) = \mathbb{M}_*(\nu_q)$ and for $\alpha = P'(q)$ we have*

$$\begin{aligned}\dim_H E(\mathbb{M},\alpha) &= \frac{1}{\log N}(P(q) - \alpha q)\\ &= \min_{p\in\mathbb{R}} \frac{1}{\log N}(P(p) - \alpha p)\\ &= \frac{1}{\log N} h_{\nu_q}(\sigma)\\ &= \frac{1}{\log N} h_{top}(E(\mathbb{M},\alpha)),\end{aligned}$$

where

$$E(\mathbb{M},\alpha) \overset{\text{def}}{=} \left\{ \xi \in \Sigma_N^+ \colon \lim_{n\to\infty} \frac{1}{n} \log \|\mathbb{M}^n(\xi)\| = \alpha \right\}.$$

There exists a "translation" between matrix-Lyapunov exponents as defined in (1.3.1) and Lyapunov exponents relative to the circle diffeomorphisms induced by the matrix cocycles. This we will explain in Section 4.1. First note that, every vector in $\mathbb{P}^1$ is the form $(\cos\theta, \sin\theta)$ for some $\theta \in [0,\pi)$, so we can identify $\mathbb{P}^1$ by $[0,\pi)$. Hence, the function f_ξ^n can be considered a real function and so we can calculate its derivative, moreover

$$|(f_\xi^n)'(v)| = \frac{1}{\|\mathbb{M}^n(\xi)v\|^2}, \tag{1.3.6}$$

see Section C.1 for more details. Consider also the *potential* $\varphi\colon \Sigma_N^+ \times \mathbb{P}^1 \to \mathbb{R}$, given by

$$\varphi(\xi, v) \overset{\text{def}}{=} \log |f'_{\xi_1}(v)| \tag{1.3.7}$$

and for every $\mu \in \mathcal{M}(\Sigma_N^+ \times \mathbb{P}^1)$ denote

$$\varphi(\mu) \overset{\text{def}}{=} \int \varphi \, d\mu,$$

Again, by Birkhoff's ergodic theorem, for every $\mu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+ \times \mathbb{P}^1)$ and for μ-almost every (ξ, v) it holds

$$\chi(\mathbb{M},\xi,v) \overset{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n} \log |(f_\xi^n)'(v)| = \int \log |f'_{\xi_1}(v)| \, d\mu(\xi,v) = \varphi(\mu). \tag{1.3.8}$$

Let us also consider the following *variational pressure* function for the skew-product F and $q \in \mathbb{R}$,

$$P_F(q) \overset{\text{def}}{=} \sup_{\mu\in\mathcal{M}(\Sigma_N^+\times\mathbb{P}^1)} \left(h_\mu(F) + \frac{q}{2}\varphi(\mu) \right). \tag{1.3.9}$$

A measure $\mu \in \mathcal{M}(\Sigma_N^+ \times \mathbb{P}^1)$ is called *q-equilibrium state for P_F* if it realizes the supremum in (1.3.9). We will see in Proposition 4.1.4 that, in fact, the pressure functions in (1.3.4) and in (1.3.9) coincide.

Our final main result is a translation of [FK11, Proposition 1.2] (see Proposition 4.3.2) to the skew-product F. It is proved in Section 4.3.2.

Theorem F. *If $\mathbb{M} \in \mathfrak{E}_{N,shyp}$, then for any $q > 0$ there exists a unique q-equilibrium state $\mu_q \in \mathcal{M}(\Sigma_N^+ \times \mathbb{P}^1)$ for P_F, P_F is differentiable at q, $P_F'(q) = \frac{1}{2}\varphi(\mu_q)$ and there exists $C > 1$ such that for every $n \in \mathbb{N}$ and $\xi \in \Sigma_N^+$ there exists $v_n \in \mathbb{P}^1$ such that, with $\alpha = P_F'(q)$*

$$\frac{1}{C} \leq \frac{\mu_q\left(\{(\zeta, v) \in [\xi_1 \ldots \xi_n] \times \mathbb{P}^1 \colon\ \chi(\mathbb{M}, \zeta, v) = 2\alpha\}\right)}{\exp(-nP(q))|(f_\xi^n)'(v_n)|^{-q/2}} \leq C.$$

In fact,

$$\frac{1}{C} \leq \frac{\pi_*\mu_q([\xi_1 \ldots \xi_n])}{\exp(-nP(q))|(f_\xi^n)'(v_n)|^{-q/2}} \leq C.$$

Let us briefly summarize the content of the thesis. Chapter 2 deals with contracting on average IFSs and the effect of metric change. In Section 2.1, we introduce the concepts of synchronization and contraction and we discuss their relation. In Section 2.2, we investigate sufficient and necessary conditions for contraction on average. Section 2.3 discusses the particular case of IFSs on the circle. In Section 2.4 we discuss some examples, in particular those indicated in Figure 4. Chapter 3 explores stochastic properties with special focus on IFSs on the circle. In Section 3.1, we briefly discuss what impact a metric change has on stochastic properties for a contracting on average system. Convergence on average at an exponential rate is stated in Section 3.2 where we also investigate the phenomenon of synchronization in more detail. Chapter 4 investigates IFSs induced by matrix cocycles. In Section 4.2 we provide details of the cocycle. Section 4.1 discusses relations between ergodic measures for the skew-product and for the base map. Finally, in Section 4.3 we prove Theorems E and F.

2 Contracting on average Iterated Function Systems

2.1 Synchronization and contraction (on average)

In this section we discuss several types of synchronization-like and contraction conditions and their relations between each other. Unless stated otherwise, we always assume that (M, d) is a general metric space, $\mathcal{F} = \{f_0, \ldots, f_{N-1}\}$ an IFS of continuous maps, and $\mathbf{p}$ a non-degenerate probability vector. Let μ be the Bernoulli measure on Σ_N^+ determined by $\mathbf{p}$. When X is a random variable on (Σ_N^+, μ), we write

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \int_{\Sigma_N^+} X \, d\mu.$$

2.1.1 Synchronization

The study of synchronization effects goes back to, at least, the 17th century, when Huygens [Huy73] observed the synchronization of linked pendulums. In the theory of dynamical systems, *synchronization* usually refers to the phenomenon that for any two initially fixed distinct points their randomly chosen trajectories converge to each other. Let us now recall related concepts and some properties.

One says that $(\mathcal{F}, \mathbf{p}, d)$ is *synchronizing (S)* if random orbits of different initial points converge to each other with probability 1, that is, for every $x, y \in M$ and almost every $\xi \in \Sigma_N^+$ it holds

$$\lim_{n\to\infty} Z_{n,d}^{x,y}(\xi) = 0. \tag{2.1.1}$$

The triple $(\mathcal{F}, \mathbf{p}, d)$ is *exponentially synchronizing (S_{exp})* if the convergence in (2.1.1) is exponentially fast, that is, if for every $x, y \in M$ and almost every $\xi \in \Sigma_N^+$ there exist $\lambda \in (0, 1)$, $n \in \mathbb{N}$ and $C > 0$ such that

$$Z_{n,d}^{x,y}(\xi) \le C\lambda^n. \tag{2.1.2}$$

The pair $(\mathcal{F}, d)$ is *proximal* if for every $x, y \in M$, there exist $\xi \in \Sigma_N^+$ and an increasing sequence $(n_k)_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} Z_{n_k,d}^{x,y}(\xi) = 0. \tag{2.1.3}$$

The triple $(\mathcal{F}, \mathbf{p}, d)$ is *synchronizing on average (SA)*[1] if for every $x, y \in M$ it holds

$$\lim_{n\to\infty} \mathbb{E}\Big(Z_{n,d}^{x,y}\Big) = 0. \tag{2.1.4}$$

[1] Note that the definition in [MM20, GK16] differs from the one given here: $(\mathcal{F}, \mathbf{p}, d)$ is *synchronizing on average*

The triple $(\mathcal{F}, \mathbf{p}, d)$ is called *exponentially synchronizing on average* ($\mathrm{SA_{exp}}$) if there exist $\lambda \in (0,1)$ and $C > 0$ such that for every $x, y \in M$ and $n \geq 1$ it holds

$$\mathbb{E}\Big(Z_{n,d}^{x,y}\Big) \leq C\lambda^n. \tag{2.1.5}$$

The following general relations hold between the above defined properties.

Remark 2.1.1. The systems studied in [MS21] (see Section 2.4.1) are an example of IFSs $\mathrm{S_{exp}}$ and $\mathrm{SA_{exp}}$.

Lemma 2.1.2. *Assuming that (M, d) is bounded, S implies SA.*

Proof. This is an immediate consequence of the Dominated Convergence Theorem. □

Lemma 2.1.3. *S and proximal are invariant under change of equivalent metrics.*

Proof. It is enough to note that if d and D are equivalent then a sequence in M converges in (M, d) if and only if it converges in (M, D). □

Lemma 2.1.4. *Let d and D be two metrics on M. Assume that there exist $C > 0$ and $\alpha \in (0,1]$ such that $D \leq Cd^\alpha$. If $(\mathcal{F}, \mathbf{p}, d)$ is SA_{exp}, then $(\mathcal{F}, \mathbf{p}, D)$ is SA_{exp}.*

Proof. Assume that $(\mathcal{F}, \mathbf{p}, d)$ is $\mathrm{SA_{exp}}$ and let $c > 0$, $\lambda \in (0,1)$ and $\alpha \in (0,1]$ such that $\mathbb{E}\Big(Z_{n,d}^{x,y}\Big) \leq c\lambda^n$ for all $n \geq 1$ and $x, y \in M$. Hence, using Jensen's inequality we get

$$\mathbb{E}\Big(Z_{n,D}^{x,y}\Big) \leq C\,\mathbb{E}\Big(Z_{n,d^\alpha}^{x,y}\Big) \leq C\left(\mathbb{E}\Big(Z_{n,d}^{x,y}\Big)\right)^\alpha \leq c^\alpha\, C(\lambda^\alpha)^n,$$

that is, $(\mathcal{F}, \mathbf{p}, D)$ is $\mathrm{SA_{exp}}$. □

Lemma 2.1.5. *The following implications hold*

$$S_{exp} \quad \Rightarrow \quad S \quad \Rightarrow \quad SA \quad \Rightarrow \quad proximal.$$

Proof. The first two implications are immediate. Let us assume $(\mathcal{F}, \mathbf{p}, d)$ is SA. Fix $x, y \in M$. Then $Z_{n,d}^{x,y}$ converges to 0 in L^1 as $n \to \infty$. Hence, applying Chebyshev's inequality, for every $\varepsilon > 0$ it holds

$$\mu(Z_{n,d}^{x,y} \geq \varepsilon) \leq \varepsilon^{-1}\mathbb{E}(Z_{n,d}^{x,y}) \to 0$$

as $n \to \infty$, that is, $Z_{n,d}^{x,y}$ converges to 0 in probability. By [Dur19, Theorem 2.3.2], there exists a sub-sequence $(n_k)_k$ such that $Z_{n_k,d}^{x,y}$ converges almost surely to 0 as $k \to \infty$. This implies proximality. □

---

if for every $x, y \in M$, for almost every $\xi \in \Sigma_N^+$ it holds

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} Z_{j,d}^{x,y}(\xi) = 0.$$

Remark 2.1.6. If $\mathcal{F} = \{f_i\}_{i=0}^{N-1}$ is an IFS of homeomorphisms of the circle $\mathbb{S}^1$ which do note have a common fixed point and $\mathbf{p}$ a non-degenerate probability vector, then by [Mal17, Theorem E] the concepts $\mathrm{S}_{\mathrm{exp}}$, S, and proximal (and hence SA) are equivalent.

2.1.2 Global average contraction conditions

The IFS $\mathcal{F}$ is *contracting (C)* if every map in $\mathcal{F}$ is a contraction. Given $k \in \mathbb{N}$, denote $\mathcal{F}^k \stackrel{\text{def}}{=} \{f_{\xi_1 \ldots \xi_k} : (\xi_1, \ldots, \xi_k) \in \{0, \ldots, N-1\}^k\}$. The IFS $\mathcal{F}$ is *k-eventually contracting (k-EC)* if $\mathcal{F}^k$ is contracting and $\mathcal{F}$ is *eventually contracting (EC)* if it is k-eventually contracting for some $k \in \mathbb{N}$.

2.1.2.1 CA

The triple $(\mathcal{F}, \mathbf{p}, d)$ is called *contractive on average (CA)* if there is some *contraction rate* $\lambda \in (0,1)$ such that

$$\mathbb{E}(Z_{1,d}^{x,y}) \leq \lambda d(x,y) \quad \text{for every } x, y \in M.$$

Remark 2.1.7. CA is a particular case of the first hypothesis of [BDEG88, Theorem 2.1], which guarantees the existence and uniqueness of a stationary measure.

Lemma 2.1.8. *If $(\mathcal{F}, \mathbf{p}, d)$ is CA with contraction rate $\lambda \in (0,1)$, then for every $n \in \mathbb{N}$*

$$\mathbb{E}(Z_{n,d}^{x,y}) \leq \lambda^n d(x,y) \quad \textit{for every } x, y \in M.$$

Proof. Let us proceed by induction. For $n = 1$ just apply the definition of CA. Suppose now that for $k \geq 1$, we have

$$\mathbb{E}(Z_{k,d}^{x,y}) \leq \lambda^k d(x,y) \quad \text{for every } x, y \in M.$$

Using that $(\mathcal{F}, \mathbf{p}, d)$ is CA, for every $x, y \in M$ we have

$$\begin{aligned}
\mathbb{E}(Z_{k+1,d}^{x,y}) &= \sum_{i=0}^{N-1} p_i \mathbb{E}(Z_{k,d}^{f_i(x),f_i(y)}) \\
&\leq \lambda^k \sum_{i=0}^{N-1} p_i d(f_i(x), f_i(y)) \\
&\leq \lambda^{k+1} d(x,y).
\end{aligned}$$

This prove the lemma. □

Remark 2.1.9. If (M, d) is bounded, then it follows from Lemma 2.1.8 that CA implies $\mathrm{SA}_{\mathrm{exp}}$.

Given $k \in \mathbb{N}$, the triple $(\mathcal{F}, \mathbf{p}, d)$ is called *k-eventually contractive on average (k-ECA)* if there exist some *contraction rate* $\lambda \in (0,1)$ such that

$$\mathbb{E}(Z_{k,d}^{x,y}) \leq \lambda d(x,y) \quad \text{for every } x, y \in M. \tag{2.1.6}$$

The triple $(\mathcal{F}, \mathbf{p}, d)$ is *eventually contracting on average (ECA)* if it is k-ECA for some $k \in \mathbb{N}$.

The following lemma justifies that a, perhaps obvious, first choice of a metric to establish CA is well defined.

Lemma 2.1.10. *For every $n \in \mathbb{N}$, the function $\psi_n \colon M \times M \to \mathbb{R}$ defined by*

$$\psi_n(x, y) \stackrel{def}{=} \mathbb{E}(Z^{x,y}_{n,d})$$

is continuous and defines a pseudometric on M. For every $x, y \in M$, it holds

$$\sum_{i=0}^{N-1} p_i \psi_n(f_i(x), f_i(y)) = \psi_{n+1}(x, y).$$

Moreover, if $\mathcal{F}$ is an IFS of homeomorphisms then ψ_n is a metric on M which is equivalent to d.

Proof. The continuity, symmetry and triangle inequality are immediate consequences of the fact that d is a metric. Thus, ψ_n is a pseudo metric on M. The second property is immediate.

Finally note that $\psi_n(x, y) = 0$ if and only if, $d(f_{\xi_1 \dots \xi_n}(x), f_{\xi_1 \dots \xi_n}(y)) = 0$ for all $\xi_1, \dots, \xi_n \in \{0, \dots, N-1\}$. Hence, if every f_i is a homeomorphism, then $x = y$. In this case, it also is immediate to see that ψ_n generates the same topology as d. $\square$

Given $k \in \mathbb{N}$ and $\lambda \in (0, 1)$, consider $d_{k,\lambda} \colon M \times M \to [0, \infty)$ defined by

$$d_{k,\lambda}(x, y) \stackrel{\text{def}}{=} d(x, y) + \frac{1}{\lambda^{1/k}} \mathbb{E}(Z^{x,y}_{1,d}) + \cdots + \frac{1}{\lambda^{(k-1)/k}} \mathbb{E}(Z^{x,y}_{k-1,d}). \tag{2.1.7}$$

The following is an immediate consequence of Lemma 2.1.10.

Lemma 2.1.11. *The function $d_{k,\lambda}$ defined in (2.1.7) is a metric which is equivalent to d. Moreover, if all the maps in $\mathcal{F}$ are Lipschitz, then in $d_{k,\lambda}$ is strongly equivalent to d.*

Proof. First note that by Lemma 2.1.10 and definition (2.1.7), $d_{k,\lambda}$ is indeed a metric. It is easy to see that a sequence $(x_n)_{n \geq 1}$ converges is (M, d) if only it converges in $(M, d_{k,\lambda})$. Thus, d and $d_{k,\lambda}$ are equivalent. Moreover, for every $x, y \in M$

$$\begin{aligned} d_{k,\lambda}(x, y) &= d(x, y) + \frac{1}{\lambda^{1/k}} \mathbb{E}(Z^{x,y}_{1,d}) + \cdots + \frac{1}{\lambda^{(k-1)/k}} \mathbb{E}(Z^{x,y}_{k-1,d}) \\ &\leq \sum_{j=0}^{k-1} \frac{L^j}{\lambda^{j/k}} d(x, y), \end{aligned}$$

where L is the maximum of Lipschitz constants of the maps in $\mathcal{F}$. Thus, with

$$c = \sum_{j=0}^{k-1} \frac{L^j}{\lambda^{j/k}},$$

we have

$$d \leq d_{k,\lambda} \leq cd,$$

that is, d and $d_{k,\lambda}$ are strongly equivalent. $\square$

Proposition 2.1.12. *If $(\mathcal{F}, \mathbf{p}, d)$ is k-ECA with contraction rate $\lambda \in (0,1)$, then $(\mathcal{F}, \mathbf{p}, d_{k,\lambda})$ is CA with contraction rate $\lambda^{1/k}$.*

Proof. Assume that $(\mathcal{F}, \mathbf{p}, d)$ is k-ECA with contraction rate $\lambda \in (0,1)$ and $k \in \mathbb{N}$. If $k = 1$, then $d_{k,\lambda} = d$ and $(\mathcal{F}, \mathbf{p}, d)$ is CA. If $k \geq 2$, then it follows from the definition of $d_{k,\lambda}$ that

$$\begin{aligned}
\mathbb{E}\Big(Z^{x,y}_{1,d_{k,\lambda}}\Big) &= \sum_{i=0}^{N-1} p_i d_{k,\lambda}(f_i(x), f_i(y)) \\
&= \sum_{i=0}^{N-1} p_i\Big(d(f_i(x), f_i(y)) + \frac{1}{\lambda^{1/k}}\mathbb{E}\Big(Z^{f_i(x),f_i(y)}_{1,d}\Big) + \cdots + \frac{1}{\lambda^{(k-1)/k}}\mathbb{E}\Big(Z^{f_i(x),f_i(y)}_{k-1,d}\Big)\Big) \\
&= \mathbb{E}\Big(Z^{x,y}_{1,d}\Big) + \frac{1}{\lambda^{1/k}}\mathbb{E}\Big(Z^{x,y}_{2,d}\Big) + \cdots + \frac{1}{\lambda^{(k-2)/k}}\mathbb{E}\Big(Z^{x,y}_{k-1,d}\Big) + \frac{1}{\lambda^{(k-1)/k}}\mathbb{E}\Big(Z^{x,y}_{k,d}\Big) \\
&\leq \mathbb{E}\Big(Z^{x,y}_{1,d}\Big) + \frac{1}{\lambda^{1/k}}\mathbb{E}\Big(Z^{x,y}_{2,d}\Big) + \cdots + \frac{1}{\lambda^{(k-2)/k}}\mathbb{E}\Big(Z^{x,y}_{k-1,d}\Big) + \frac{1}{\lambda^{(k-1)/k}}\lambda d(x,y) \\
&= \lambda^{1/k}\Big(d(x,y) + \frac{1}{\lambda^{1/k}}\mathbb{E}\Big(Z^{x,y}_{1,d}\Big) + \frac{1}{\lambda^{2/k}}\mathbb{E}\Big(Z^{x,y}_{2,d}\Big) + \cdots + \frac{1}{\lambda^{(k-1)/k}}\mathbb{E}\Big(Z^{x,y}_{k-1,d}\Big)\Big) \\
&= \lambda^{1/k} d_{k,\lambda}(x,y).
\end{aligned}$$

Hence, $(\mathcal{F}, \mathbf{p}, d_{k,\lambda})$ is CA with contraction rate $\lambda^{1/k}$. ☐

2.1.2.2 NEA

The triple $(\mathcal{F}, \mathbf{p}, d)$ is said to be *non-expansive on average (NEA)* if

$$\mathbb{E}(Z^{x,y}_{1,d}) \leq d(x,y) \quad \text{for every } x, y \in M. \tag{2.1.8}$$

Remark 2.1.13. The NEA property was introduced in [JT01] as *non-separating on average*. See also [Sza03, Part II] for a study of NEA iterated function systems and associated non-expansive Markov operators.

The following is an immediate consequence of Jensen's inequality.

Lemma 2.1.14. *If $(\mathcal{F}, \mathbf{p}, d)$ is NEA then for any $\alpha \in (0,1)$ the triple $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is NEA.*

Lemma 2.1.15. *Assume (M, d) is bounded. If $(\mathcal{F}, \mathbf{p}, d)$ is SA and D is a metric equivalent to d such that $(\mathcal{F}, \mathbf{p}, D)$ is NEA, then $(\mathcal{F}, \mathbf{p}, D)$ is SA.*

Proof. If $(\mathcal{F}, \mathbf{p}, d)$ is SA, then $\mathbb{E}\Big(Z^{x,y}_{n,d}\Big) \to 0$ as $n \to \infty$ for every $x, y \in M$. Arguing as in the proof of Lemma 2.1.5, there exists a sub-sequence $(n_k)_k$ such that $Z^{x,y}_{n_k,d} \to 0$ almost surely to 0 as $k \to \infty$. Since D and d are assumed to be equivalent, $Z^{x,y}_{n_k,D} \to$ almost surely as $k \to \infty$. Hence, by the Dominated Convergence Theorem, $\mathbb{E}(Z^{x,y}_{n_k,D}) \to 0$ as $k \to \infty$. Since $(\mathcal{F}, \mathbf{p}, D)$ is NEA, $\mathbb{E}\Big(Z^{x,y}_{n,D}\Big) \to 0$ as $n \to \infty$. This proves that $(\mathcal{F}, \mathbf{p}, D)$ is SA. ☐

Lemma 2.1.16. *The following implications hold*

1. *C $\Rightarrow$ CA,*
2. *CA $\Rightarrow$ NEA, and*
3. *k-EC $\Rightarrow$ k-ECA.*

Proof. To prove the implication *(1.)*, assume that $(\mathcal{F}, \mathbf{p}, d)$ is C and take $\lambda_i \in (0,1)$ the contraction constant of $f_i \in \mathcal{F}$. Then, for every $x, y \in M$ we have

$$\mathbb{E}(Z^{x,y}_{1,d}) = \sum_{i=0}^{N-1} p_i d(f_i(x), f_i(y)) \leq \left(\sum_{i=0}^{N-1} p_i \lambda_i\right) d(x,y),$$

and note that,

$$\sum_{i=0}^{N-1} p_i \lambda_i \in (0,1)$$

to conclude that $(\mathcal{F}, \mathbf{p}, d)$ is CA.

Item *(2.)* is an immediate consequence of the definition.

Applying *(1.)* to the IFS $\mathcal{F}^k$, we obtain *(3.)*. □

Remark 2.1.17. If $(\mathcal{F}, \mathbf{p}, d)$ is NEA, then for all $k \in \mathbb{N}$ and $\lambda \in (0,1)$ the metric $d_{k,\lambda}$ defined in (2.1.7) is strongly equivalent to d.

Lemma 2.1.18. *For $n \geq 0$, let $\psi_n(x,y) \overset{def}{=} \mathbb{E}(Z^{x,y}_{n,d})$ be as in Lemma 2.1.10. If $(\mathcal{F}, \mathbf{p}, d)$ is NEA, then for every $x, y \in M$ it holds*

1. *$\psi_1(x,y) \leq d(x,y)$,*
2. *$\psi_n(x,y)$ is non-increasing in n and hence the limit $\lim_{n\to\infty} \psi_n(x,y)$ exists,*
3. *Assuming that M is compact, if for every $x, y \in M$ we have $\lim_{n\to\infty} \psi_n(x,y) = 0$, then $\psi_n \to 0$ uniformly.*

In the following proof and below we use the following simplifying notation

$$p_{\xi_1 \ldots \xi_n} \overset{\text{def}}{=} p_{\xi_1} \cdots p_{\xi_n}.$$

Proof. Property *(1.)* just restates the definition of NEA. To show *(2.)*, check that

$$\begin{aligned}
\psi_{n+1}(x,y) = \mathbb{E}(Z^{x,y}_{n+1,d}) &= \sum_{\xi_1,\ldots,\xi_{n+1}=0}^{N-1} p_{\xi_1,\ldots,\xi_{n+1}} d(f_{\xi_1,\ldots,\xi_{n+1}}(x), f_{\xi_1,\ldots,\xi_{n+1}}(y)) \\
&= \sum_{\xi_1,\ldots,\xi_n=0}^{N-1} p_{\xi_1,\ldots,\xi_n} \sum_{\ell=1}^{N} p_\ell \, d\Big(f_\ell(f_{\xi_1,\ldots,\xi_n}(x)), f_\ell(f_{\xi_1,\ldots,\xi_n}(y))\Big) \\
&= \sum_{\xi_1,\ldots,\xi_n=0}^{N-1} p_{\xi_1,\ldots,\xi_n} \mathbb{E}\Big(Z_{1,d}^{f_{\xi_1,\ldots,\xi_n}(x), f_{\xi_1,\ldots,\xi_n}(y)}\Big) \\
\text{(by NEA)} \quad &\leq \sum_{\xi_1,\ldots,\xi_n=0}^{N-1} p_{\xi_1,\ldots,\xi_n} d(f_{\xi_1,\ldots,\xi_n}(x), f_{\xi_1,\ldots,\xi_n}(y)) = \mathbb{E}(Z^{x,y}_{n,d}) = \psi_n(x,y).
\end{aligned}$$

This, together with $\psi_n \geq 0$ implies item *(2.)*.

To prove *(3.)*, assume that M is compact and $\psi_n \to 0$ point-wise. As the limit function is continuous, by Dini's theorem, convergence is uniform. □

Remark 2.1.19. By [GS17, Proposition 1], for any IFS $\mathcal{F} = \{f_i\}_{i=0}^{N-1}$ of homeomorphisms of $(\mathbb{S}^1, d)$ which is backward minimal there exists a metric ρ on $\mathbb{S}^1$ equivalent to d on $\mathbb{S}^1$ such that $(\mathcal{F}, \mathbf{p}, \rho)$ is NEA.

2.1.2.3 log-CA

The triple $(\mathcal{F}, \mathbf{p}, d)$ is called log-*contractive on average (*log-*CA)* if there exists $\lambda < 1$ such that

$$\prod_{j=0}^{N-1} d(f_j(x), f_j(y))^{p_j} \leq \lambda d(x,y) \quad \text{for every } x, y \in M,$$

or, equivalently,

$$\mathbb{E}\left(\ln \frac{Z_{1,d}^{x,y}}{d(x,y)}\right) \leq \ln \lambda < 0 \quad \text{for every } x, y \in M, x \neq y. \tag{2.1.9}$$

Remark 2.1.20. Let $\mathcal{F}$ be an IFS of Lipschitz maps. In [DF99], the condition

$$\mathbb{E}\left(\ln \sup_{x \neq y} \frac{Z_{1,d}^{x,y}}{d(x,y)}\right) < 0,$$

was called *contracting on average* which is stronger than log-CA. The condition log-CA was introduced in [Elt87, Page 84].

Lemma 2.1.21. *CA implies* log-*CA*.

Proof. Assuming CA with contraction rate $\lambda \in (0,1)$, note that $\mathbb{E}(Z_{1,d}^{x,y}/d(x,y)) \leq \lambda$. Hence, by Jensen's inequality

$$\mathbb{E}\left(\ln \frac{Z_{1,d}^{x,y}}{d(x,y)}\right) \leq \ln \mathbb{E}\left(\frac{Z_{1,d}^{x,y}}{d(x,y)}\right) \leq \ln \lambda < 0,$$

proving the lemma. □

Remark 2.1.22. The concept of log-CA was introduced in [BE88]. Assuming (M, d) to be a complete metric space, $\mathcal{F}$ to be an IFS of Lipschitz maps, and $(\mathcal{F}, \mathbf{p}, d)$ to be log-CA, they prove the existence of an attractive (hence unique) stationary measure (extending previous results obtained in the case when M is compact, see references in [BE88]).

The following example is presented in [Eda96] to illustrate that log-CA is weaker than C. Indeed, it also shows that log-CA is weaker than CA.

Example 2.1.23 (log-CA, but not NEA and not CA)**.** Let $M = [0,1]$ and $d(x,y) \overset{\text{def}}{=} |x-y|$. Let $\mathbf{p} = (\frac{1}{2}, \frac{1}{2})$. Consider the IFS $\mathcal{F} = \{f_0, f_1\}$ given by

$$f_0, f_1 \colon M \to M, \quad f_0(x) \overset{\text{def}}{=} \frac{x}{3}, \ f_1(x) \overset{\text{def}}{=} \min\{1, 2x\}.$$

Note that for all $x, y \in [0,1]$ it holds

$$[d(f_0(x), f_0(y))]^{1/2} [d(f_1(x), f_1(y))]^{1/2} \leq \frac{2}{3}|x-y|$$

and hence $(\mathcal{F}, \mathbf{p}, d)$ is log-CA. On the other hand, for $x, y \in [0, 1/2]$ it holds $f_1(x) = 2x$ and $f_1(y) = 2y$, so that

$$\mathbb{E}(Z_{1,d}^{x,y}) = \frac{1}{2}\left(d(f_0(x), f_0(y)) + d(f_1(x), f_1(y))\right) = \frac{1}{2}\left(\frac{1}{3}d(x,y) + 2d(x,y)\right),$$

which implies that $(\mathcal{F}, \mathbf{p}, d)$ is not NEA and thus not CA.

The proof of the following proposition is a bit technical. For completeness, we provided all the details.

Proposition 2.1.24. *Assume that $\mathcal{F}$ is an IFS of Lipschitz maps and $(\mathcal{F}, \mathbf{p}, d)$ is* log*-CA. Then, there exists $\alpha \in (0,1]$ such that $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is CA.*

Proof. Following the proof of [BDEG88, Lemma 2.6], we apply auxiliary Lemma B.0.1 to the function $h_{x,y} : \Sigma_N^+ \to (0, +\infty)$, $x \neq y$, given by

$$h_{x,y}(\xi) \overset{\text{def}}{=} \max\left\{\frac{Z_{1,d}^{x,y}(\xi)}{d(x,y)}, \left(\frac{\lambda}{s}\right)^{1/\delta}\right\}.$$

Here

- $s \geq 1$ is a constant greater than all Lipschitz constants of maps in $\mathcal{F}$. Hence for every $x \neq y$ and $\xi \in \Sigma_N^+$
$$\frac{Z_{1,d}^{x,y}(\xi)}{d(x,y)} \leq s.$$
- $\delta \in (0,1)$ is the minimum of the entries of $\mathbf{p}$, and
- $\lambda \in (0,1)$ is so that (2.1.9) holds.

Note that for every $x, y \in M$, $x \neq y$,

$$\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}} \leq h_{x,y} \leq s.$$

Therefore, for $r>0$, using that $x-1\geq \ln x$ and Jensen's inequality as in proof of Lemma B.0.1 (see (B.0.3), for more details), we get

$$\begin{aligned}0&\leq\left(\int(h_{x,y})^r\,d\mu\right)^{\frac{1}{r}}-\exp\left(\int\ln h_{x,y}d\mu\right)\\&\leq\exp\left(\frac{1}{r}\left(\int(h_{x,y})^r d\mu-1\right)\right)-\exp\left(\int\ln h_{x,y}d\mu\right)\\&\leq\exp\left(\int\ln h_{x,y}d\mu\right)\left[\exp\left(\int\left(\frac{1}{r}\left((h_{x,y})^r-1\right)-\ln h_{x,y}\right)d\mu\right)-1\right].\end{aligned}$$

Note that the function

$$x\mapsto\frac{1}{r}(x^r-1)-\ln x$$

is increasing for $x>1$ and decreasing for $0<x<1$, so that

$$\begin{aligned}&\int\left(\frac{1}{r}\left((h_{x,y})^r-1\right)-\ln h_{x,y}\right)d\mu\\&\leq\int_{\{h_{x,y}>1\}}\left(\frac{s^r-1}{r}-\ln s\right)d\mu+\int_{\{h_{x,y}\leq 1\}}\left(\frac{\left(\frac{\lambda}{s}\right)^{\frac{r}{\delta}}-1}{r}-\ln\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}}\right)d\mu\\&\leq\left(\frac{s^r-1}{r}-\ln s\right)+\left(\frac{\left(\frac{\lambda}{s}\right)^{\frac{r}{\delta}}-1}{r}-\ln\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}}\right),\end{aligned}$$

so that

$$\begin{aligned}0&\leq\left(\int(h_{x,y})^r\,d\mu\right)^{\frac{1}{r}}-\exp\left(\int\ln h_{x,y}d\mu\right)\\&\leq\left(\frac{1}{r}(s^r-1)-\ln s\right)+\left(\frac{1}{r}\left(\left(\frac{\lambda}{s}\right)^{\frac{r}{\delta}}-1\right)-\ln\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}}\right).\end{aligned}$$

Consequently, the following convergence is uniform on $x,y\in M$, $x\neq y$,

$$\lim_{r\downarrow 0}\left(\int(h_{x,y}(\xi))^r\,d\mu(\xi)\right)^{\frac{1}{r}}=\exp\left(\int\ln h_{x,y}(\xi)d\mu(\xi)\right).$$

Therefore,

$$\lim_{r\downarrow 0}\sup_{x\neq y}\left(\int(h_{x,y}(\xi))^r\,d\mu(\xi)\right)^{\frac{1}{r}}=\sup_{x\neq y}\exp\left(\int\ln h_{x,y}(\xi)d\mu(\xi)\right)$$

On the other hand, for $x\neq y$ we have

$$\begin{aligned}&\int\ln h_{x,y}(\xi)d\mu(\xi)\\&=\int_{\left\{\frac{Z^{x,y}_{1,d}(\xi)}{d(x,y)}\leq\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}}\right\}}\ln\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}}d\mu(\xi)+\int_{\left\{\frac{Z^{x,y}_{1,d}(\xi)}{d(x,y)}>\left(\frac{\lambda}{s}\right)^{\frac{1}{\delta}}\right\}}\ln\frac{Z^{x,y}_{1,d}(\xi)}{d(x,y)}d\mu(\xi),\end{aligned}$$

using the hypothesis we get

$$\int\ln h_{x,y}(\xi)d\mu(\xi)\leq\ln\lambda.$$

which implies that

$$\lim_{r\downarrow 0}\sup_{x,y}\left(\int (h_{x,y}(\xi))^r\, d\mu(\xi)\right)^{\frac{1}{r}} \leq \lambda,$$

and so the existence of $\alpha \in (0,1]$ such that

$$\sup_{x\neq y}\int\left(\frac{Z_{1,d}^{x,y}(\xi)}{d(x,y)}\right)^{\alpha} d\mu(\xi) \leq \sup_{x,y}\int (h_{x,y}(\xi))^{\alpha}\, d\mu(\xi) < 1,$$

thus we conclude this proof. □

Remark 2.1.25. It is common in the literature to assume a log-CA condition instead of CA, see for example [Ste12]. In the present setting, by Lemma 2.1.21 and Proposition 2.1.24, these conditions are equivalent (changing d by d^{α} if necessary, for some $\alpha \in (0,1]$).

2.1.3 Local average contraction conditions: LECA and ESCA

In this section, we discuss several types of local average contraction conditions for IFSs. In particular, we introduce LECA and ESCA, which are, besides NEA, key properties towards Theorem A. Such conditions have been studied, for example, in [Kai78] and later, independently, in [Ste99, Ste01] and [Car02]. See also [JT01, LsS05]. Here our focus is on conditions which are sufficient or necessary for CA, possibly after some change of metric.

The first property is a generalization of CA. We say that $(\mathcal{F},\mathbf{p},d)$ is *locally eventually contractive on average (LECA)* if

$$\text{for every } x,y\in M,\ x\neq y, \text{ there exists } \ell\geq 1:\quad \mathbb{E}(Z_{\ell,d}^{x,y}) < d(x,y). \tag{2.1.10}$$

Remark 2.1.26. If $(\mathcal{F},\mathbf{p},d)$ is LECA then for any $\alpha\in(0,1)$ the triple $(\mathcal{F},\mathbf{p},d^{\alpha})$ is LECA.

Lemma 2.1.27. *The following implications hold*

1. *CA ⇒ LECA, and*
2. *SA ⇒ LECA.*

Proof. To prove *(1.)* assume that $(\mathcal{F},\mathbf{p},d)$ is CA with contraction rate $\lambda\in(0,1)$, Then, take for every $x,y\in M$, $\ell = 1$ to get

$$\mathbb{E}(Z_{1,d}^{x,y}) \leq \lambda d(x,y),$$

therefore, $(\mathcal{F},\mathbf{p},d)$ is LECA.

Now, let us prove *(2.)*. If $(\mathcal{F},\mathbf{p},d)$ is SA, then for every x,y there exists $\ell\geq 1$ large enough such that

$$\mathbb{E}(Z_{\ell,d}^{x,y}) \leq \frac{1}{2}d(x,y).$$

The last statement is clear when $x=y$. In the case $x\neq y$, use that $\lim_{n\to\infty}\mathbb{E}(Z_{n,d}^{x,y}) = 0$ and that $d(x,y)>0$. □

Lemma 2.1.27 shows that synchronization (on average) is intimately related with contraction (on average). The next result indeed proves that, assuming NEA, those properties are equivalent. For its proof we need to recall some more concepts. The set Σ_N^+ is naturally equipped with the product topology on $\{0,\dots,N-1\}^{\mathbb{N}}$, where $\{0,\dots,N-1\}$ is given the discrete topology. A basis is given by the family of *cylinders*

$$[i;\xi_1,\dots,\xi_n] \stackrel{\text{def}}{=} \{\eta\in\Sigma_N^+ \colon \eta_{i+1}=\xi_1,\dots,\eta_{i+n}=\xi_n\}.$$

We simply write $[\xi_1,\dots,\xi_n] \stackrel{\text{def}}{=} [1;\xi_1,\dots,\xi_n]$. Every cylinder is clopen. Every open set in Σ_N^+ is a countable union of cylinders.

Lemma 2.1.28. *Suppose that $(\mathcal{F},\mathbf{p},d)$ is NEA on some compact metric space (M,d). Then, $(\mathcal{F},\mathbf{p},d)$ is SA if and only if it is LECA.*

Proof. Assume that $(\mathcal{F},\mathbf{p},d)$ is NEA. By Lemma 2.1.27, SA implies LECA. To prove the reverse implication, let us assume that $(\mathcal{F},\mathbf{p},d)$ is LECA. By Lemma 2.1.18 (2), for every $x,y\in M$ the limit

$$\delta(x,y) \stackrel{\text{def}}{=} \lim_{n\to\infty}\psi_n(x,y)\ge 0, \qquad \text{where} \qquad \psi_n(x,y)\stackrel{\text{def}}{=}\mathbb{E}(Z_{n,d}^{x,y}), \tag{2.1.11}$$

exists. Arguing by contradiction, let us suppose that $(\mathcal{F},\mathbf{p},d)$ is not SA and hence there exist $x,y\in M$ such that $\delta=\delta(x,y)>0$. Define

$$B_\delta \stackrel{\text{def}}{=} \left\{(z,w)\colon d(z,w)\ge\frac{\delta}{2}\right\},$$

which is a closed subset of $M\times M$ and hence compact. Note that d and ψ_n are continuous on $M\times M$. Thus, for $n\ge 1$ and $t\in(0,1)$ the set

$$C_{n,t} \stackrel{\text{def}}{=} \{(z,w)\colon \mathbb{E}(Z_{n,d}^{z,w})<td(z,w)\}$$

is open in $M\times M$ and so

$$A_{n,t}\stackrel{\text{def}}{=} B_\delta\cap C_{n,t}$$

is open in B_δ. As we assume LECA, for each $(z,w)\in B_\delta$ there exist $\ell\ge 1$ and $\lambda'\in(0,1)$ such that $\mathbb{E}(Z_{\ell,d}^{z,w})\le\lambda' d(z,w)$ and hence, $(z,w)\in A_{\ell,t}$ for every $t\in(\lambda',1)$. In particular, it holds

$$B_\delta=\bigcup_{t\in(0,1)}\bigcup_{n\in\mathbb{N}}A_{n,t}.$$

By compactness of B_δ, there exist $n_1,\dots,n_k\in\mathbb{N}$ and $t_1,\dots,t_k\in(0,1)$ such that

$$B_\delta=A_{n_1,t_1}\cup\dots\cup A_{n_k,t_k}.$$

Define

$$N\stackrel{\text{def}}{=}\max\{n_1,\dots,n_k\} \qquad \text{and} \qquad \lambda\stackrel{\text{def}}{=}\max\{t_1,\dots,t_k\}.$$

If (z,w) in B_δ, then $(z,w) \in A_{n_i,t_i}$ for some $i \in \{1,\dots,k\}$. This together with Lemma 2.1.18 (2) implies that

$$\mathbb{E}(Z^{z,w}_{N,d}) \le \mathbb{E}(Z^{z,w}_{n_i,d}) < t_i d(z,w) \le \lambda d(z,w). \tag{2.1.12}$$

Let

$$\Gamma_n \stackrel{\text{def}}{=} \left\{\xi \colon (f^n_\xi(x), f^n_\xi(y)) \in B_\delta\right\}.$$

Note that Γ_n is the union of cylinder sets. Indeed, if $\xi \in \Gamma_n$ then every $\eta \in [\xi_1,\dots,\xi_n] \in \Gamma_n$. For $n \ge 0$, we have

$$\mathbb{E}(Z^{x,y}_{N+n,d}) = \mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n}\right) + \mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n^c}\right). \tag{2.1.13}$$

For the first term, we observe

$$\begin{aligned}
&\mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n}\right)\\
&= \sum_{\xi_{n+1},\dots,\xi_{N+n}} \sum_{[\xi_1\dots\xi_n]\subset\Gamma_n} p_{\xi_{n+1}\dots\xi_{N+n}} p_{\xi_1\dots\xi_n} d(f_{\xi_1\dots\xi_{N+n}}(x), f_{\xi_1\dots\xi_{N+n}}(y))\\
&= \sum_{[\xi_1,\dots,\xi_n]\subset\Gamma_n} p_{\xi_1\dots\xi_n} \sum_{\xi_{n+1},\dots,\xi_{N+n}} p_{\xi_{n+1}\dots\xi_{N+n}} d(f_{\xi_1\dots\xi_{N+n}}(x), f_{\xi_{n+1}\dots\xi_{N+n}}(y))\\
&= \sum_{[\xi_1,\dots,\xi_n]\subset\Gamma_n} p_{\xi_1\dots\xi_n}\mathbb{E}\left(Z^{f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)}_{N,d}\right).
\end{aligned}$$

By definition of Γ_n it holds $(f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)) \in B_\delta$, it follows from (2.1.12) that

$$\begin{aligned}
\mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n}\right) &\le \sum_{[\xi_1,\dots,\xi_n]\subset\Gamma_n} p_{\xi_1\dots\xi_n}\lambda\, d\left(f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)\right)\\
&= \lambda\,\mathbb{E}(Z^{x,y}_{n,d}\mathbb{1}_{\Gamma_n}).
\end{aligned} \tag{2.1.14}$$

For the second term, we have

$$\mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n^c}\right) = \sum_{[\xi_1,\dots,\xi_n]\subset\Gamma_n^c} p_{\xi_1\dots\xi_n}\mathbb{E}\left(Z^{f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)}_{N,d}\right).$$

Since $(\mathcal{F},\mathbf{p},d)$ is NEA, again using Lemma 2.1.18 (2), we get

$$\mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n^c}\right) \le \sum_{[\xi_1,\dots,\xi_n]\subset\Gamma_n^c} p_{\xi_1\dots\xi_n} d\left(f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)\right).$$

By definition of Γ_n we have that $(f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)) \notin B_\delta$ so that

$$\mathbb{E}\left(Z^{x,y}_{N+n,d}\mathbb{1}_{\Gamma_n^c}\right) < \frac{\delta}{2}\mu(\Gamma_n^c). \tag{2.1.15}$$

From (2.1.13), (2.1.14) and (2.1.15) we get

$$\mathbb{E}(Z^{x,y}_{N+n,d}) < \lambda\,\mathbb{E}\left(Z^{x,y}_{n,d}\mathbb{1}_{\Gamma_n}\right) + \frac{\delta}{2}\mu\left(\Gamma_n^c\right).$$

Moreover, by induction on $k \in \mathbb{N}$, it follows

$$\begin{aligned}
\mathbb{E}(Z^{x,y}_{kN,d}) &= \mathbb{E}\left(Z^{x,y}_{kN,d}\mathbb{1}_{\Gamma_{(k-1)N}}\right) + \mathbb{E}\left(Z^{x,y}_{kN,d}\mathbb{1}_{\Gamma^c_{(k-1)N}}\right) \\
&\leq \lambda\,\mathbb{E}\left(Z^{x,y}_{(k-1)N,d}\mathbb{1}_{\Gamma_{(k-1)N}}\right) + \frac{\delta}{2}\mu\left(\Gamma^c_{(k-1)N}\right) \\
&\leq \lambda^2\,\mathbb{E}\left(Z^{x,y}_{(k-2)N,d}\mathbb{1}_{\Gamma_{(k-1)N}\cap\Gamma_{(k-2)N}}\right) + \frac{\delta}{2}\mu\left(\Gamma^c_{(k-1)N}\cup\Gamma^c_{(k-2)N}\right) \\
\leq \ldots &\leq \lambda^k\,\mathbb{E}\left(Z^{x,y}_{0,d}\mathbb{1}_{\bigcap_{j=0}^{k-1}\Gamma_{jN}}\right) + \frac{\delta}{2}\mu\Big(\bigcup_{j=0}^{k-1}\Gamma^c_{jN}\Big) \\
&= \lambda^k d(x,y)\mu\Big(\bigcap_{j=0}^{k-1}\Gamma_{jN}\Big) + \frac{\delta}{2}\mu\Big(\bigcup_{j=0}^{k-1}\Gamma^c_{jN}\Big).
\end{aligned}$$

Hence, recalling that $\lambda \in (0,1)$ and using $\mu(\cdot) \leq 1$, we get

$$\lim_{k\to\infty}\mathbb{E}(Z^{x,y}_{kN,d}) \leq \frac{\delta}{2},$$

which is a contradiction. This implies SA. □

The following is a consequence of Lemmas 2.1.28 and 2.1.15.

Corollary 2.1.29. *Assume (M,d) is compact. Assume that d and D are equivalent metrics such that $(\mathcal{F},\mathbf{p},d)$ and $(\mathcal{F},\mathbf{p},D)$ are NEA. Then $(\mathcal{F},\mathbf{p},d)$ is LECA if and only if $(\mathcal{F},\mathbf{p},D)$ is LECA.*

The following is a generalization of the definition of locally contractive with respect to the reverse system [Ste99, Definition 5] and of the definition ε-local (average) contractive [Ste12, Definition 1]. We say that $(\mathcal{F},\mathbf{p},d)$ is *eventually strongly contracting on average (ESCA)* if for every $x \in M$ there exist $\ell \geq 1$ and an open neighborhood $V_{(x,x)} \subset M \times M$ of (x,x) such that

$$\sup_{(y,z)\in V_{(x,x)},y\neq z}\frac{\mathbb{E}(Z^{z,y}_{\ell,d})}{d(z,y)} < 1. \tag{2.1.16}$$

Lemma 2.1.30. *If $(\mathcal{F},\mathbf{p},d)$ is ESCA then for any $\alpha \in (0,1)$ the triple $(\mathcal{F},\mathbf{p},d^\alpha)$ is ESCA.*

Proof. Let $\alpha \in (0,1)$. Using Jensen's inequality, we get

$$\left(\sup_{(y,z)\in V_{(x,x)},y\neq z}\frac{\mathbb{E}(Z^{z,y}_{\ell,d})}{d(z,y)}\right)^\alpha \geq \sup_{(y,z)\in V_{(x,x)},y\neq z}\frac{\mathbb{E}(Z^{z,y}_{\ell,d^\alpha})}{d^\alpha(z,y)},$$

which implies that, if $(\mathcal{F},\mathbf{p},d)$ is ESCA then $(\mathcal{F},\mathbf{p},d^\alpha)$ is ESCA. □

Remark 2.1.31. The example $(\mathcal{F},\mathbf{p},d^\alpha)$ given in Section 2.4.3 is ESCA with parameter $\ell \neq 1$. Moreover, $(\mathcal{F},\mathbf{p},d^\alpha)$ is LECA, but fails to be NEA.

Lemma 2.1.32. *If $(\mathcal{F},\mathbf{p},d)$ is CA, then for every $\alpha \in (0,1]$ and a metric D which is strongly equivalent to d^α, for every $k \in \mathbb{N}$ large enough the triple $(\mathcal{F},\mathbf{p},D)$ is k-ECA.*

Proof. Suppose $(\mathcal{F}, \mathbf{p}, d)$ be CA with contracting rate $\lambda \in (0,1)$. Given $\alpha \in (0,1]$ and a metric D on M strongly equivalent to d^α, take $b > a > 0$ and such that $aD \leq d^\alpha \leq bD$. Note that $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is CA with contracting rate $\lambda^\alpha \in (0,1)$. Take $k \in \mathbb{N}$ so that

$$\frac{b}{a}\lambda^{\alpha k} < 1.$$

First, strong equivalence implies that

$$a\mathbb{E}(Z^{x,y}_{k,D}) \leq \mathbb{E}(Z^{x,y}_{k,d^\alpha}) \quad \text{for every } x, y \in M.$$

For every $x, y \in M$, $x \neq y$, CA with contraction rate λ^α together with the above implies

$$\lambda^{\alpha k} \geq \frac{\mathbb{E}(Z^{x,y}_{k,d^\alpha})}{d^\alpha(x,y)} = \frac{\mathbb{E}(Z^{x,y}_{k,d^\alpha})}{\mathbb{E}(Z^{x,y}_{k,D})} \cdot \frac{\mathbb{E}(Z^{x,y}_{k,D})}{D(x,y)} \cdot \frac{D(x,y)}{d^\alpha(x,y)} \geq a \cdot \frac{\mathbb{E}(Z^{x,y}_{k,D})}{D(x,y)} \cdot \frac{1}{b}.$$

Hence, it follows

$$\mathbb{E}(Z^{x,y}_{k,D}) \leq \frac{b}{a}\lambda^{\alpha k} D(z,y),$$

which proves that $(\mathcal{F}, \mathbf{p}, D)$ is k-ECA with contraction rate $\frac{b}{a}\lambda^{\alpha k} \in (0,1)$. □

2.1.4 Further contraction conditions

In this section, we continue our discussion of contraction conditions and put them into the context of the ones defined above. Although, none of the concepts defined in this section will be implemented in the remainder of this paper. Let us define for every $x \in M$ the sequence of random variables $(X^x_n)_{n\geq 0}$ on (Σ^+_N, μ) by

$$X^x_n(\xi) := f_{\xi_n} \circ f_{\xi_{n-1}} \circ \cdots \circ f_{\xi_1}(x); \quad X^x_0(\xi) := x. \tag{2.1.17}$$

For $A \subset M$ measurable, denote by $\tau_A(x) \overset{\text{def}}{=} \inf\{n \geq 1 \colon X^x_n \in A\}$ the first time the process X^x_n hits the set A. Following [JT01], $(\mathcal{F}, \mathbf{p}, d)$ satisfies the *local contraction property relative to* $A \subset M$ if there is $\lambda \in (0,1)$ such that

$$\mathbb{E}\Big(d\Big(X^x_{\tau_A(x)\vee\tau_A(y))}, X^y_{\tau_A(x)\vee\tau_A(y))}\Big)\Big) \leq \lambda d(x,y) \quad \text{for every } x, y \in M, \tag{2.1.18}$$

that is, "there is some contraction after the set $A \subset M$ is reached". In other terms, this condition states that if we start two chains, at x and y, respectively, and run them simultaneously using the same maps, then at the time both of them have visited A, in average they will be closer to each other by a factor λ.

Remark 2.1.33. In [JT01, Section 3], it is shown that $\tau_A(x) \vee \tau_A(y) < \infty$ almost surely, assuming that there exists a function $V \colon M \to [1, \infty)$, satisfying $\sup_{x\in A} V(x) < \infty$, and constants $r \in (0,1)$ and $b < \infty$ such that for every $x \in M$

$$\mathbb{E}V\left(X^x_1\right) \leq rV(x) + b\mathbb{1}_A(x). \tag{2.1.19}$$

Furthermore, in [JT01] the function V is used to control the behavior outside of A. Moreover, assuming that $(\mathcal{F}, \mathbf{p}, d)$ is NEA and M is complete separable metric space with bounded metric d, thus guaranteeing the existence of a unique stationary probability measure (see [JT01, Theorem 2.1]).

Lemma 2.1.34. *Assume that $(\mathcal{F}, \mathbf{p}, d)$ satisfies the local contraction property relative to $A = M$, then $(\mathcal{F}, \mathbf{p}, d)$ is CA and (2.1.19) is satisfied for $V \equiv 1$.*

Proof. If (2.1.18) holds for $A = M$, then $\tau_A(x) = 1$, for every $x \in M$, and

$$\mathbb{E}(Z^{x,y}_{1,d}) = \mathbb{E}\Big(d\Big(X^x_{\tau_A(x)\vee\tau_A(y))}, X^y_{\tau_A(x)\vee\tau_A(y))}\Big)\Big) \le \lambda d(x, y) \quad \text{for every } x, y \in M,$$

that is, $(\mathcal{F}, \mathbf{p}, d)$ is CA.

Moreover, in the case $A = M$ consider $V \equiv 1$, any $r \in (0, 1)$ and $b = 1 - r$, to get (2.1.19) holds. □

The following example $(\mathcal{F}, \mathbf{p}, d)$ satisfies the local contraction property (2.1.18), but is not CA. Furthermore, it is NEA and satisfies (2.1.19).

Example 2.1.35 (NEA, but not CA)**.** Adapting an example in [JT01, Section 6] to our context, let $M = [0, 2]$ and $d(x, y) \overset{\text{def}}{=} |x - y|$. Let $\mathbf{p} = (p, 1 - p)$ for some $p \in (1/2, 1)$. Consider the IFS $\mathcal{F} = \{f_0, f_1\}$ given by

$$f_0, f_1 \colon M \to M, \quad f_0(x) \overset{\text{def}}{=} \begin{cases} x - \frac{2}{3} & x \ge 1, \\ \frac{x}{3} & x \le 1, \end{cases} \text{ and } f_1(x) \overset{\text{def}}{=} \min\left\{x + \frac{2}{3}, 2\right\}.$$

It is easy to see that $(\mathcal{F}, \mathbf{p}, d)$ is NEA. On the other hand, for $x, y \in [1, 4/3]$ it holds

$$\mathbb{E}(Z^{x,y}_{1,d})) = \mathbb{E}(d(X^x_1, X^y_1)) = d(x, y) = |x - y|$$

and hence $(\mathcal{F}, \mathbf{p}, d)$ is not CA.

Now let us show that $(\mathcal{F}, \mathbf{p}, d)$ satisfies the local contraction property (2.1.18) relative to $A = [0, 1]$. Since f_0 and f_1 are non-decreasing functions we have that for every $x, y \in M$, such that $x < y$,

$$\tau_A(x) \le \tau_A(y)$$

and if $y \in A$, then $x \in A$. Also, note that for $x, y \in A$ we have that

$$\mathbb{E}(Z^{x,y}_{1,d}) = \left(\frac{p}{3} + (1 - p)\right) d(x, y).$$

Therefore, (2.1.18) holds with $\lambda \overset{\text{def}}{=} \frac{p}{3} + (1 - p) < 1$.

On the other hand, consider $f(t) = pe^{-\frac{2}{3}t} + (1 - p)e^{\frac{2}{3}t}$. Note that $f(0) = 1$ and $f'(0) < 0$. Thus, to see that (2.1.19) holds fix $t > 0$ such that $f(t) < 1$. Let $r \overset{\text{def}}{=} f(t) \in (0, 1)$, $b = e^{\frac{5}{3}t}$ and $V(x) \overset{\text{def}}{=} e^{tx}$. If $x \in A$, then

$$\mathbb{E}V(X^x_1) = pe^{t\frac{x}{3}} + (1 - p)e^{t\left(x+\frac{2}{3}\right)},$$

and hence

$$\mathbb{E}V(X_1^x) \le pe^{\frac{t}{3}} + (1-p)e^{\frac{5}{3}t} \le e^{\frac{5}{3}t} \le r + b \le rV(x) + b\mathbb{1}_A(x).$$

If $x \in (1,2] = M \setminus A$, then

$$\mathbb{E}V(X_1^x) \le pe^{t\left(x-\frac{2}{3}\right)} + (1-p)e^{t\left(x+\frac{2}{3}\right)} = V(x)\left(pe^{-t\frac{2}{3}} + (1-p)e^{\frac{2}{3}t}\right) = rV(x) + \mathbb{1}_A(x).$$

This proves (2.1.19).

The idea of locally contractive Markov chains can be expressed in several ways. Globally contracting on average-type CA and $\log$-CA are convenient because they can be analysed by many different methods. In [Ste12], the following local average contraction conditions is considered. Given $\varepsilon > 0$, the triple $(\mathcal{F}, \mathbf{p}, d)$ is called *ε-local contractive on average (ε-LCA)* if there exists $\lambda \in (0,1)$ such that

$$\sup_{0<d(x,y)<\varepsilon} \frac{\mathbb{E}(Z_{1,d}^{x,y})}{d(x,y)} \le \lambda. \tag{2.1.20}$$

The triple $(\mathcal{F}, \mathbf{p}, d)$ is called *ε-local* $\log$*-CA* if for some $\lambda \in (0,1)$

$$\sup_{0<d(x,y)<\varepsilon} \mathbb{E}\left(\ln \frac{Z_{1,d}^{x,y}}{d(x,y)}\right) \le \ln\lambda < 0. \tag{2.1.21}$$

Lemma 2.1.36. *ε-LCA implies ε-local* $\log$*-LCA.*

Proof. By Jensen's inequality,

$$\sup_{0<d(x,y)<\varepsilon} \mathbb{E}\left(\ln \frac{Z_{1,d}^{x,y}}{d(x,y)}\right) \le \sup_{0<d(x,y)<\varepsilon} \ln \mathbb{E}\left(\frac{Z_{1,d}^{x,y}}{d(x,y)}\right).$$

Since logarithm is an increasing function, we get

$$\sup_{0<d(x,y)<\varepsilon} \mathbb{E}\left(\ln \frac{Z_{1,d}^{x,y}}{d(x,y)}\right) \le \ln\left(\sup_{0<d(x,y)<\varepsilon} \frac{\mathbb{E}\left(Z_{1,d}^{x,y}\right)}{d(x,y)}\right),$$

which implies the lemma. ☐

The triple $(\mathcal{F}, \mathbf{p}, d)$ is *locally contractive in the weak sense (LCWS)*, if for some $\lambda \in (0,1)$ it holds

$$\sup_{x\in M}\left[\limsup_{y\to x} \frac{\mathbb{E}(Z_{1,d}^{x,y})}{d(x,y)}\right] \le \lambda. \tag{2.1.22}$$

It is clear that (2.1.20) implies (2.1.22). In [Ste12, Remark 9] an example of a triple $(\mathcal{F}, \mathbf{p}, d)$ that is ε-local $\log$-CA but not LCWS is shown.

Lemma 2.1.37. *ε-LCA implies LCWS and ESCA.*

Proof. The first implication is immediate. To check the second one, take $k = 1$ and $V_{(x,x)} = \{(y,z)\colon d(y,z) < \varepsilon\}$ for every $x \in M$. ☐

2.2 Conditions to guarantee CA

Let us begin this section by presenting a sufficient condition for the existence of a metric D equivalent to d for which the system is CA. We always consider an IFS $\mathcal{F}$ of a metric space (M, d).

Proposition 2.2.1. *Assume that $(\mathcal{F}, \mathbf{p}, d)$ is SA$_{exp}$, that is, there exist constants $C > 0$ and $\lambda \in (0, 1)$ such that*

$$\mathbb{E}(Z_{n,d}^{x,y}) \leq C\lambda^n \quad \textit{for every } x, y \in M \textit{ and } n \in \mathbb{N}. \tag{2.2.1}$$

For every $q \in (\lambda, 1)$

$$D(x,y) \stackrel{\text{def}}{=} \sum_{n\geq 0} \frac{q^n}{\lambda^n} \mathbb{E}(Z_{n,d}^{x,y}). \tag{2.2.2}$$

defines a metric on M which is equivalent to d such that $(\mathcal{F}, \mathbf{p}, D)$ is CA.

Proof. Pick $q \in (\lambda, 1)$ and define $D\colon M \times M \to \mathbb{R}$ by (2.2.2). Note that it follows from our hypothesis that D is well-defined. It is easy to check that D is a metric. It remains to show that D has the claimed properties. By Lemma 2.1.10, for every $x, y \in M$

$$\begin{aligned}\sum_{i=0}^{N-1} p_i D(f_i(x), f_i(y)) &= \sum_{n\geq 0} \frac{q^n}{\lambda^n} \sum_{i=0}^{N-1} p_i \mathbb{E}\Big(Z_{n,d}^{f_i(x),f_i(y)}\Big) = \sum_{n\geq 0} \frac{q^n}{\lambda^n} \mathbb{E}(Z_{n+1,d}^{x,y}) \\ &= \frac{\lambda}{q} \sum_{n\geq 1} \frac{q^n}{\lambda^n} \mathbb{E}(Z_{n,d}^{x,y}) \leq \frac{\lambda}{q} \sum_{n\geq 0} \frac{q^n}{\lambda^n} \mathbb{E}(Z_{n,d}^{x,y}) = \frac{\lambda}{q} D(x,y).\end{aligned}$$

Hence, $(\mathcal{F}, \mathbf{p}, D)$ is CA with contraction rate $\lambda/q \in (0, 1)$.

It remains to see that D and d are equivalent. First note that $d \leq D$, which implies that the topology of (M, d) is a subset of the topology of (M, D). Now, let us prove that the topology of (M, D) is a subset of the topology of (M, d). Given V an open set of (M, D) and $x \in V$, there exists $r > 0$ such that $B_D(x, r) \subset V$, where $B_D(x, r)$ is the open ball relative to the metric D with center x and radius r. Let $\varepsilon \stackrel{\text{def}}{=} (1 - q)r/(2 - q) > 0$. Take $L \in \mathbb{N}$ such that

$$C \sum_{n\geq L} q^n < \varepsilon.$$

It follows from Lemma 2.1.10, that

$$U \stackrel{\text{def}}{=} \bigcap_{n=0}^{L-1} U_n, \qquad \text{where} \qquad U_n = \{y \in M : \mathbb{E}(Z_{n,d}^{x,y}) < \lambda^n \varepsilon\}$$

is an open set of (M, d). Furthermore, for every $y \in U$

$$D(x,y) \leq \sum_{n=0}^{L-1} \frac{q^n}{\lambda^n} \mathbb{E}(Z_{n,d}^{x,y}) + \sum_{n=L}^{\infty} \frac{q^n}{\lambda^n} \mathbb{E}(Z_{n,d}^{x,y}) < \sum_{n=0}^{L-1} \varepsilon q^n + \varepsilon < \varepsilon \frac{1}{1-q} + \varepsilon = r.$$

Thus, $x \in U \subset B_D(x, r)$ and hence $x \in U \subset V$, which proves the desired. Therefore, the topologies of (M, d) and (M, D) coincide. $\square$

Remark 2.2.2. Proposition 2.2.1 can be applied to IFS $\mathcal{F}$ on $\mathbb{S}^1$ induced by the projective action of $GL(2,\mathbb{R})$ matrix cocycles and implies the existence of a metric D that makes $(\mathcal{F}, \mathbf{p}, D)$ CA. We refrain from providing the details. In this context, the existence of a unique stationary measure is well known (see, for example, [BL85, Chapter II]) and no further immediate application of CA is given.

However, in this respect, it is reasonable to ask if the existence of a unique stationary measure implies the existence of some metric that preserves the topology and makes the system CA?

Remark 2.2.3. If $(\mathcal{F}, \mathbf{p}, d)$ is not ESCA, then D provided by Proposition 2.2.1 is in general not strongly equivalent. Indeed, if ESCA fails, then there exists $x \in M$ such that for every $n \geq 1$ there exists a sequence $\{x_k^n\}_{k\in\mathbb{N}}$ in M such that $\lim_{k\to\infty} x_k^n = x$ in (M, d) and

$$\lim_{k\to\infty} \frac{\mathbb{E}(Z_{n,d}^{x,x_k^n})}{d(x, x_k^n)} \geq 1.$$

Therefore, we can find a sequence $\{y_n\}_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} \frac{D(x, y_n)}{d(x, y_n)} = \infty,$$

We now invoke the results obtained in Sections 2.1.2 and 2.1.3 to prove Theorem A.

Proof of Theorem A. We will prove Theorem A as follows. Using the LECA and ESCA conditions we will cover $M \times M$ by open sets. The compactness of M will allow us to find a finite subcover. On each set of coverage we will have the condition CA satisfied. Finally, since the NEA condition controls the iterates on average, we can conclude the condition CA globally.

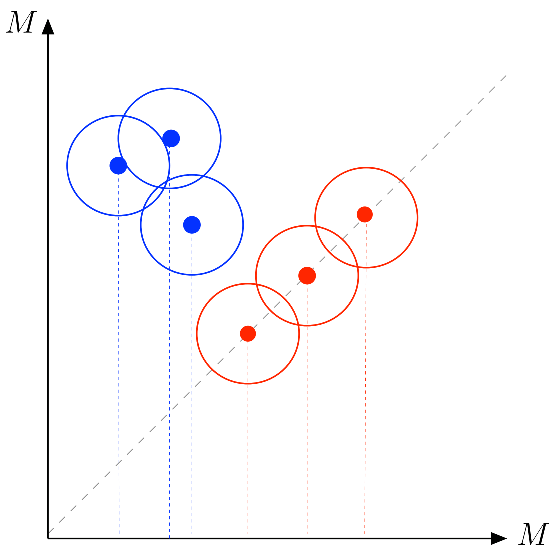

Figure 4 – Using conditions LECA (blue colour) and ESCA (red colour), respectively, to find coverings of the product space $M \times M$.

Assume that $(\mathcal{F}, \mathbf{p}, d)$ is NEA, LECA, and ESCA.

First, let us construct for every $(x, y) \in M \times M$ an open neighborhood $V_{(x,y)}$ as follows. For $(x, y) \in M \times M$, $x \neq y$, take $\ell = \ell(x, y) \in \mathbb{N}$ and $\lambda(x, y) \in (0, 1)$ as in the definition of (see (2.1.10)) satisfying

$$\mathbb{E}(Z^{x,y}_{\ell,d}) \leq \lambda(x, y)d(x, y)$$

and let

$$V_{(x,y)} \overset{\text{def}}{=} \left\{(z, w) \in M \times M : z \neq w \text{ and } \frac{\mathbb{E}(Z^{z,w}_{\ell,d})}{d(z, w)} < \sqrt{\lambda(x, y)}\right\}.$$

As the function

$$f : M \times M \setminus \{(z, z) : z \in M\} \to \mathbb{R}, \quad f(z, w) \overset{\text{def}}{=} \frac{\mathbb{E}(Z^{z,w}_{\ell,d})}{d(z, w)},$$

is continuous and the set $\{(z, z) : z \in M\}$ is closed, $V_{(x,y)}$ is an open subset of $M \times M$ containing (x, y). For $x = y \in M$, take $\ell = \ell(x, x) \in \mathbb{N}$ and an open neighborhood $V_{(x,x)}$ of (x, x) as in the definition of ESCA (see (2.1.16)) satisfying

$$\lambda(x, x) \overset{\text{def}}{=} \sup_{(z,w)\in V_{(x,y)} : y\neq z} \frac{\mathbb{E}(Z^{z,w}_{\ell,d})}{d(z, w)} < 1$$

to get

$$\mathbb{E}(Z^{z,w}_{\ell,d}) \leq \lambda(x, x)d(z, w),$$

for all $(z, w) \in V_{(x,y)}$.

As $M \times M$ is compact, it has a finite sub-cover $\{V_{(x_1,y_1)}, \ldots, V_{(x_m,y_m)}\}$. Hence, for every $i \in \{1, \ldots, m\}$ there are $\ell_i = \ell(x_i, y_i) \in \mathbb{N}$ and $\lambda_i = \lambda(x_i, y_i) \in (0, 1)$ such that for all $(z, w) \in V_{(x_i,y_i)}$ it holds

$$\mathbb{E}(Z^{z,w}_{\ell_i,d}) \leq \lambda_i d(z, w).$$

Take $k \overset{\text{def}}{=} \max_{1\leq i\leq m} \ell_i$ and $\lambda \overset{\text{def}}{=} \max_{1\leq i\leq m} \lambda_i$. Hence, together with Lemma 2.1.18 (2), for every $(x, y) \in M \times M$ there exists $i \in \{1, \ldots, m\}$ such that

$$\mathbb{E}(Z^{x,y}_{k,d}) \leq \mathbb{E}(Z^{x,y}_{\ell_i,d}) \leq \lambda_i d(x, y) \leq \lambda d(x, y).$$

But this implies that $(\mathcal{F}, \mathbf{p}, d)$ is k-ECA with contraction rate $\lambda \in (0, 1)$. By Proposition 2.1.12, $(\mathcal{F}, \mathbf{p}, D)$, where $D \overset{\text{def}}{=} d_{k,\lambda}$ is defined in (2.1.7), is CA with contraction rate $\lambda^{1/k}$. Hence, invoking Lemma 2.1.11, d and D are strongly equivalent.

Together with Lemma 2.1.28, this proves the theorem. □

The following provides a necessary condition for the existence of a metric D equivalent to d for which the system is CA.

Proposition 2.2.4. *Suppose that $(\mathcal{F}, \mathbf{p}, d)$ is NEA on some compact metric space (M, d). If there exists a metric D on K equivalent to d such that $(\mathcal{F}, \mathbf{p}, D)$ is CA, then $(\mathcal{F}, \mathbf{p}, d)$ is SA.*

Proof. Assuming that $(\mathcal{F}, \mathbf{p}, D)$ is CA, by Lemma 2.1.8 for every $n \in \mathbb{N}$ it holds

$$\mathbb{E}(Z^{x,y}_{n,D}) \le \lambda^n D(x,y) \quad \text{for every } x, y \in M.$$

Fix $x, y \in M$. By the above, it holds

$$\lim_{n\to\infty} \mathbb{E}(Z^{x,y}_{n,D}) = 0, \tag{2.2.3}$$

that is, $Z^{x,y}_{n,D}$ converges to 0 in L^1. By Chebyshev's inequality, for every $\varepsilon > 0$ it holds

$$\mu(Z^{x,y}_{n,D} \ge \varepsilon) \le \varepsilon^{-1}\mathbb{E}(Z^{x,y}_{n,D}).$$

and hence it follows that $Z^{x,y}_{n,D}$ converges to 0 in probability. By [Dur19, Theorem 2.3.2], there exists a sub-sequence $(n_k)_k$ such that $Z^{x,y}_{n_k,D}$ converges almost surely to 0 as $k \to \infty$.

The fact that D and d are equivalent implies $Z^{x,y}_{n_k,d}$ converges almost surely to 0 as $k \to \infty$. By dominated convergence theorem, we conclude that

$$\lim_{k\to\infty} \mathbb{E}(Z^{x,y}_{n_k,d}) = 0.$$

As we assume that $(\mathcal{F}, \mathbf{p}, d)$ is NEA on a compact space and x, y were arbitrary, by Lemma 2.1.18 (2)–(3) it follows

$$\lim_{n\to\infty} \mathbb{E}(Z^{(\cdot),(\cdot)}_{n,d}) = 0$$

uniformly. This implies SA. $\square$

2.3 CA for IFSs on $\mathbb{S}^1$

In this section, we will study the particular case of an IFS $\mathcal{F}$ of homeomorphisms on $M = \mathbb{S}^1$ (equipped with the usual metric $d(x,y) \stackrel{\text{def}}{=} \min\{|x-y|, 1-|x-y|\}$. In particular, we prove Theorem B.

We first recall the following results which are an immediate consequence of [Mal17, Theorem A and Proposition 4.2], respectively.

Proposition 2.3.1. *Let $\mathcal{F}$ be an IFS of homeomorphisms of $\mathbb{S}^1$ and assume that there does not exist a probability measure which is invariant by every element of $\mathcal{F}$. Then for every non-degenerate probability vector $\mathbf{p}$ there is a constant $\lambda \in (0,1)$ such that for every $x \in \mathbb{S}^1$ and almost every $\xi \in \Sigma_N^+$ there exists an open neighborhood $I_x(\xi) \subset \mathbb{S}^1$ of x such that for all $n \in \mathbb{N}$ it holds*

$$Z^{w,z}_{n,d}(\xi) = d(f^n_\xi(w), f^n_\xi(z)) \le \lambda^n \quad \textit{for every } w, z \in I_x(\xi).$$

For the statement of the next proposition, consider the *shift* $\sigma\colon \Sigma_N^+ \to \Sigma_N^+$ defined by

$$(\sigma(\xi))_j = \xi_{j+1}, \quad j \ge 1.$$

Recall that it is continuous.

Proposition 2.3.2. *Under the hypotheses of Proposition 2.3.1, consider the map*

$$G\colon \Sigma_N^+ \times \mathbb{S}^1 \times \mathbb{S}^1 \to \Sigma_N^+ \times \mathbb{S}^1 \times \mathbb{S}^1, \quad G(\xi, x, y) \stackrel{\text{def}}{=} (\sigma(\xi), f_{\xi_1}(x), f_{\xi_1}(y)).$$

Let $\mathcal{E} = \bigcup_{\xi \in \Sigma_N^+} \{\xi\} \times U(\xi) \subset \Sigma_N^+ \times \mathbb{S}^1 \times \mathbb{S}^1$ such that $G^{-1}(\mathcal{E}) \subset \mathcal{E}$ and $U(\xi)$ is open in $\mathbb{S}^1 \times \mathbb{S}^1$ for every $\xi \in \Sigma_N^+$. Let $\mathbf{p}$ be a non-degenerate probability vector and μ its its associate Bernoulli measure μ on Σ_N^+ and assume that

$$(\mu \otimes \nu)(\mathcal{E}) > 0$$

for every stationary[2] probability measure ν on $\mathbb{S}^1 \times \mathbb{S}^1$. Then actually,

$$(\mu \otimes \nu)(\mathcal{E}) = 1$$

for every probability measure ν on $\mathbb{S}^1 \times \mathbb{S}^1$ (not necessarily stationary).

Remark 2.3.3. Under the hypothesis of Proposition 2.3.1, the local contraction property holds. Now, let us show that if $(\mathcal{F}, d)$ is proximal then we have $\mathrm{S}_{\mathrm{exp}}$ (as defined in (2.1.2)) holds with the same rate.

Now, let us prove the following lemma.

Lemma 2.3.4. *Assume that $(\mathcal{F}, \mathbf{p}, d)$ and $\lambda \in (0,1)$ are as in Proposition 2.3.1. If $(\mathcal{F}, d)$ is proximal, then*

$$\mu(\Omega^{x,y}) = 1 \quad \textit{for every } x, y \in \mathbb{S}^1,$$

where

$$\Omega^{x,y} \stackrel{\text{def}}{=} \{\xi \in \Sigma_N^+ \colon \textit{ there exists } C > 0 \textit{ such that } Z_{n,d}^{x,y}(\xi) \le C\lambda^n \textit{ for all } n \in \mathbb{N}\}. \tag{2.3.1}$$

Proof. For every $z \in \mathbb{S}^1$ and $k \in \mathbb{N}$ let

$$\Gamma_k(z) \stackrel{\text{def}}{=} \Big\{\eta \in \Sigma_N^+ \colon d(f_\eta^n(z_1), f_\eta^n(z_2)) \le \lambda^n \text{ for all } n \in \mathbb{N}, z_1, z_2 \in \Big(z - \frac{1}{k}, z + \frac{1}{k}\Big)\Big\}.$$

Clearly, $\Gamma_k(z) \subset \Gamma_{k+1}(z)$. By Proposition 2.3.1, for every $z \in \mathbb{S}^1$

$$\mu\Big(\bigcup_{k \in \mathbb{N}} \Gamma_k(z)\Big) = 1.$$

Hence, there is $k_0 = k_0(z) \in \mathbb{N}$ such that

$$\mu(\Gamma_{k_0}(z)) > 0. \tag{2.3.2}$$

Let $\mathcal{E}$ be the set of points $(\xi, x, y) \in \Sigma_N^+ \times \mathbb{S}^1 \times \mathbb{S}^1$ such that there exist $z \in \mathbb{S}^1$, $k_0 \in \mathbb{N}$, and $k_1 \in \mathbb{N}$ satisfying

$$f_\xi^{k_1}(x), f_\xi^{k_1}(y) \in \Big(z - \frac{1}{k_0}, z + \frac{1}{k_0}\Big), \quad \xi \in \sigma^{-k_1}(\Gamma_{k_0}(z)) \quad \text{and} \quad \mu(\Gamma_{k_0}(z)) > 0.$$

By the following claim, $\mathcal{E}$ is nonempty. More precisely, for every $(x, y) \in \mathbb{S}^1 \times \mathbb{S}^1$ the set $\mathcal{E} \cap \Big(\Sigma_N^+ \times \{(x,y)\}\Big)$ is nonempty.

[2] Recall that here ν is *stationary* if and only if $\mu \otimes \nu$ is invariant by the skew product G.

Claim 2.3.5. *For every $x, y \in \mathbb{S}^1$ it holds $\left(\mu \otimes \delta_{(x,y)}\right)(\mathcal{E}) > 0$, where $\delta_{(x,y)}$ is the Dirac measure at (x, y).*

Proof. Fix $x, y \in \mathbb{S}^1$. By proximality, there exist $\xi \in \Sigma_N^+$ and an increasing sequence $(n_k)_{k\in\mathbb{N}}$ such that we have

$$\lim_{k\to\infty} Z^{x,y}_{n_k,d}(\xi) = 0.$$

By compactness of $\mathbb{S}^1$, there are $z \in \mathbb{S}^1$ and a subsequence $(n_{k_j})_{j\geq 1}$ such that

$$f_\xi^{n_{k_j}}(x) \to z \quad \text{and} \quad f_\xi^{n_{k_j}}(y) \to z,$$

as $j \to \infty$. Hence, taking $k_0 = k_0(z)$ as in (2.3.2), there exists $k_1 \in \mathbb{N}$ large enough such that $f_\xi^{k_1}(x)$ and $f_\xi^{k_1}(y)$ are both in $\left(z - \frac{1}{k_0}, z + \frac{1}{k_0}\right)$.

As μ is σ-invariant, it holds $\mu(\sigma^{-k_1}(\Gamma_{k_0}(z))) = \mu(\Gamma_{k_0}(z)) > 0$. Furthermore,

$$\sigma^{-k_1}\left(\Gamma_{k_0}(z)\right) = \underbrace{\Sigma_N^+ \times \cdots \times \Sigma_N^+}_{k_1\text{-times}} \times \Gamma_{k_0}(z),$$

which implies that $\mu\left([\xi_1, \ldots, \xi_{k_1}] \cap \sigma^{-k_1}\left(\Gamma_{k_0}(z)\right)\right) > 0$. Since

$$\left([\xi_1, \ldots, \xi_{k_1}] \cap \sigma^{-k_1}\left(\Gamma_{k_0}(z)\right)\right) \times \{(x, y)\} \subset \mathcal{E},$$

this implies the claim. □

Integrating over $(x, y) \in \mathbb{S}^1 \times \mathbb{S}^1$ with respect to any stationary probability measure ν on $\mathbb{S}^1 \times \mathbb{S}^1$, it follows that

$$(\mu \otimes \nu)(\mathcal{E}) > 0.$$

Claim 2.3.6. *The set $\mathcal{E}$ is G-invariant, that is, $G^{-1}(\mathcal{E}) \subset \mathcal{E}$.*

Proof. If $(\xi, x, y) \in G^{-1}(\mathcal{E})$ then $(\sigma(\xi), f_{\xi_1}(x), f_{\xi_1}(y)) \in \mathcal{E}$. Hence there are $k_1, k_0 \in \mathbb{N}$ and $z \in \mathbb{S}^1$ satisfying

$$\begin{aligned}
&f_{\sigma(\xi)}^{k_1}(f_{\xi_1}(x)), f_{\sigma(\xi)}^{k_1}(f_{\xi_1}(y)) \in \left(z - \frac{1}{k_0}, z + \frac{1}{k_0}\right),\\
&\sigma(\xi) \in \sigma^{-k_1}(\Gamma_{k_0}(z)), \quad \mu\left(\Gamma_{k_0}(z)\right) > 0.
\end{aligned}$$

As $f_{\sigma(\xi)}^{k_1} \circ f_{\xi_1} = f_\xi^{k_1+1}$, this implies

$$\begin{aligned}
&f_\xi^{k_1+1}(x), f_\xi^{k_1+1}(y) \in \left(z - \frac{1}{k_0}, z + \frac{1}{k_0}\right),\\
&\xi \in \sigma^{-(k_1+1)}(\Gamma_{k_0}(z)), \quad \mu\left(\Gamma_{k_0}(z)\right) > 0.
\end{aligned}$$

But this implies $(\xi, x, y) \in \mathcal{E}$. □

Claim 2.3.7. *For every $\xi \in \Sigma_N^+$ the set $U(\xi) \overset{\text{def}}{=} \{(x,y) \in \mathbb{S}^1 \times \mathbb{S}^1 \colon (\xi, x, y) \in \mathcal{E}\}$ is open in $\mathbb{S}^1 \times \mathbb{S}^1$.*

Proof. Fix $\xi \in \Sigma_N^+$. Given $x, y \in \mathbb{S}^1$ such that $(\xi, x, y) \in \mathcal{E}$, there exist $z \in \mathbb{S}^1$ and $k_0, k_1 \in \mathbb{N}$ satisfying

$$f_\xi^{k_1}(x), f_\xi^{k_1}(y) \in \Big(z - \frac{1}{k_0}, z + \frac{1}{k_0}\Big),$$
$$\xi \in \sigma^{-k_1}\Big(\Gamma_{k_0}(z)\Big), \quad \mu\Big(\Gamma_{k_0}(z)\Big) > 0.$$

The continuity of $f_\xi^{k_1}$ implies that the set

$$V_{(x,y)} \overset{\text{def}}{=} \Big(f_\xi^{k_1}\Big)^{-1}\Big(z - \frac{1}{k_0}, z + \frac{1}{k_0}\Big) \times \Big(f_\xi^{k_1}\Big)^{-1}\Big(z - \frac{1}{k_0}, z + \frac{1}{k_0}\Big)$$

is an open neighborhood of (x,y) in $\mathbb{S}^1 \times \mathbb{S}^1$. For every $(w_1, w_2) \in V_{(x,y)}$, it holds $f_\xi^{k_1}(w_1)$, $f_\xi^{k_1}(w_2) \in (z - 1/k_0, z + 1/k_0)$. Thus, $V_{(x,y)} \subset U(\xi)$. This proves the claim. □

By Proposition 2.3.2, for every $x, y \in \mathbb{S}^1$ we have

$$\Big(\mu \otimes \delta_{(x,y)}\Big)(\mathcal{E}) = 1.$$

Therefore, by definition of the sets $\Gamma_{k_0}(z)$ for every $x, y \in \mathbb{S}^1$ and almost every $\xi \in \Sigma_N^+$ there exists $k_1 \in \mathbb{N}$ such that

$$Z_{n+k_1,d}^{x,y}(\xi) = d(f_\xi^{k_1+n}(x), f_\xi^{k_1+n}(y)) \leq \lambda^n.$$

This proves the lemma. □

The following is an immediate consequence of Lemma 2.3.4 by the dominated convergence theorem.

Corollary 2.3.8. *Assume that $(\mathcal{F}, \mathbf{p}, d)$ and $\lambda \in (0,1)$ are as in Proposition 2.3.1. Then $(\mathcal{F}, \mathbf{p}, d)$ is SA (and hence LECA).*

In the rest of this section, $\mathcal{F}$ is a finite family of C^1-diffeomorphisms. Hence, there exist $L > 1$ such that for all $i \in \{0, \dots, N-1\}$

$$L^{-1} d(x,y) \leq d(f_i(x), f_i(y)) \leq L d(x,y). \tag{2.3.3}$$

Lemma 2.3.9. *Assume that $(\mathcal{F}, \mathbf{p}, d)$ and $\lambda \in (0,1)$ are as in Proposition 2.3.1. Assume also that each map in $\mathcal{F}$ is a C^1-diffeomorphism. Then, for every $t \in (\lambda, 1)$ and $x \in \mathbb{S}^1$ and almost every ξ there exists an open neighborhood $J_x(\xi) \subset \mathbb{S}^1$ of x such that there exists $C > 0$ satisfying for all $n \in \mathbb{N}$ we have*

$$\max_{z \in J_x(\xi)} |(f_\xi^n)'(z)| \leq C t^n.$$

Proof. Fix $t \in (\lambda, 1)$ and $x \in \mathbb{S}^1$. By Proposition 2.3.1, for almost every $\xi \in \Sigma_N^+$ there exists an open arc $I_x(\xi) \subset \mathbb{S}^1$ of x such that for every $n \in \mathbb{N}$ and $y, z \in I_x(\xi)$

$$Z_{n,d}^{y,z}(\xi) \leq \lambda^n. \tag{2.3.4}$$

Denote by $\omega_\xi(\cdot)$ the modulus of continuity of $\ln|f'_{\xi_1}|$. Since $\omega_\xi(\varepsilon)$ tends to 0 as $\varepsilon \to 0$ and is uniformly bounded, by dominated convergence it follows

$$\lim_{\varepsilon \to 0} \int \omega_\xi(\varepsilon) d\mu(\xi) = 0.$$

Fix $\varepsilon > 0$ such that

$$\int \omega_\xi(\varepsilon) d\mu(\xi) \leq \ln \frac{\lambda + t}{2} - \ln \lambda.$$

By Birkhoff ergodic theorem, almost every $\xi \in \Sigma_N^+$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \omega_{\sigma^j(\xi)}(\varepsilon) = \int \omega_\eta(\varepsilon) d\mu(\eta). \tag{2.3.5}$$

Now, fix $\xi \in \Sigma_N^+$ such that (2.3.4) and (2.3.5) hold. Take $k \geq 1$ so that $\lambda^k < \varepsilon$. Define

$$I_x(\xi, \varepsilon) \stackrel{\text{def}}{=} I_x(\xi) \cap \bigcap_{j=0}^{k} (f_\xi^j)^{-1} \left(f_\xi^j(x) - \frac{\varepsilon}{2}, f_\xi^j(x) + \frac{\varepsilon}{2} \right),$$

and note that $I_x(\xi, \varepsilon)$ is an open arc containing x. For every $y, z \in I_x(\xi, \varepsilon)$ it holds

$$\ln \frac{|(f_\xi^n)'(y)|}{|(f_\xi^n)'(z)|} = \sum_{j=0}^{n-1} \ln|f'_{\sigma^j(\xi)}(f_\xi^j(y))| - \ln|f'_{\sigma^j(\xi)}(f_\xi^j(z))| \leq \sum_{j=0}^{n-1} \omega_{\sigma^j(\xi)}(\varepsilon).$$

Let x_1 and x_2 be the extreme points of $I_x(\xi, \varepsilon)$. Note that for every $z \in I_x(\xi, \varepsilon)$, it holds

$$\frac{\ln|(f_\xi^n)'(z)|}{n} \leq \frac{1}{n} \ln \left(\frac{Z_{n,d}^{x_1,x_2}}{d(x_1, x_2)} \right) + \frac{1}{n} \sum_{j=0}^{n-1} \omega_{\sigma^j(\xi)}(\varepsilon).$$

Hence, using (2.3.4) for all $n \geq 1$

$$\frac{1}{n} \ln \left(\max_{z \in I_x(\xi,\varepsilon)} |(f_\xi^n)'(z)| \right) \leq \ln \lambda - \frac{1}{n} \ln d(x_1, x_2) + \frac{1}{n} \sum_{j=0}^{n-1} \omega_{\sigma^j(\xi)}(\varepsilon),$$

so that

$$\limsup_{n \to \infty} \frac{1}{n} \ln \left(\max_{z \in I_x(\xi,\varepsilon)} |(f_\xi^n)'(z)| \right) \leq \ln \frac{\lambda + t}{2}.$$

Then, there exists $C > 0$ such that for all $n \geq 1$

$$\max_{z \in I_x(\xi,\varepsilon)} |(f_\xi^n)'(z)| \leq C t^n.$$

This proves the lemma. □

The following result together with Proposition 2.1.12 immediately implies Theorem B.

Proposition 2.3.10. *Assume the hypotheses of Theorem B. Then there exist $k \in \mathbb{N}$ and $\alpha_0 \in (0,1)$ such that for all $\alpha \in (0,\alpha_0]$ the triple $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is k-ECA with rate λ^α.*

Proof. Fix points $x, y \in \mathbb{S}^1$. Let

$$A_n \stackrel{\text{def}}{=} \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{n,d}(\xi)}{d(x,y)}\right) d\mu(\xi).$$

The sequence $(A_n)_{n\geq 0}$ is a subadditive. Indeed,

$$\begin{aligned} A_{n+m} &= \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{n+m,d}(\xi)}{d(x,y)}\right) d\mu(\xi) \\ &= \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{n+m,d}(\xi)}{Z^{x,y}_{n,d}(\xi)}\,\frac{Z^{x,y}_{n,d}(\xi)}{d(x,y)}\right) d\mu(\xi) \\ &\leq \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{n+m,d}(\xi)}{Z^{x,y}_{n,d}(\xi)}\right) d\mu(\xi) + \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{n,d}(\xi)}{d(x,y)}\right) d\mu(\xi), \end{aligned}$$

using that μ is a Bernoulli measure, we get

$$A_{n+m} \leq \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{m,d}(\xi)}{d(x,y)}\right) d\mu(\xi) + \sup_{x\neq y} \int_{\Sigma_N^+} \ln\left(\frac{Z^{x,y}_{n,d}(\xi)}{d(x,y)}\right) d\mu(\xi) = A_m + A_n.$$

Hence, by Fekete's Lemma, the limit $A \stackrel{\text{def}}{=} \lim_{n\to\infty} A_n/n = \inf_{n\geq 1} A_n/n \in [-\infty, \infty)$ exists.

All hypotheses of Proposition 2.3.1 are satisfied and we can consider $\lambda \in (0,1)$ as provided by this proposition.

Claim 2.3.11. $A \leq \ln\lambda$.

Proof. Arguing by contradiction, suppose that $\ln\lambda < A$. Then, for all $n \in \mathbb{N}$

$$\ln\lambda < A \leq \sup_{x\neq y} \int_{\Sigma_N^+} F_n(x,y,\xi)\, d\mu(\xi), \qquad \text{where} \qquad F_n(x,y,\xi) \stackrel{\text{def}}{=} \frac{1}{n} \ln\left(\frac{Z^{x,y}_{n,d}(\xi)}{d(x,y)}\right).$$

Thus, for all $n \in \mathbb{N}$ there exist x_n, y_n in $\mathbb{S}^1$, $x_n \neq y_n$, such that

$$\ln\lambda < \int_{\Sigma_N^+} F_n(x_n, y_n, \xi)\, d\mu(\xi). \tag{2.3.6}$$

By compactness, there exist a subsequence $(n_k)_{k\geq 1}$ and points $x, y \in \mathbb{S}^1$ such that

$$\lim_{k\to\infty} x_{n_k} = x, \qquad \lim_{k\to\infty} y_{n_k} = y. \tag{2.3.7}$$

In the following two cases we consider ξ in a appropriate set of measure 1 to obtain that the limit superior of $F_{n_k}(x_{n_k}, y_{n_k}, \xi)$ as $k \to \infty$ is less than or equal to $\ln\lambda$. We then will apply Fatou's Lemma to contradict (2.3.6). Note that all hypotheses of Lemmas 2.3.4 and 2.3.9 are satisfied. Given x, y as above, let $\Omega^{x,y}$ be as in (2.3.1).

Case $x \neq y$. Fix any $t \in (\lambda, 1)$. Denote by Γ_t the set of sequences $\xi \in \Sigma_N^+$ such that there exist $C > 0$ and open arcs $J_x(\xi)$ and $J_y(\xi)$ containing x and y, respectively, and satisfying

$$\max_{z \in J_x(\xi) \cup J_y(\xi)} |(f_\xi^n)'(z)| \leq C t^n, \quad Z_{n,d}^{x,y}(\xi) \leq C \lambda^n \text{ for all } n \in \mathbb{N}$$

By Lemma 2.3.4 and Lemma 2.3.9, the set Γ_t has measure 1. Given $\xi \in \Gamma_t$, by the triangle inequality and the mean value inequality it follows

$$\begin{aligned} Z_{n_k,d}^{x_{n_k},y_{n_k}}(\xi) &\leq Z_{n_k,d}^{x_{n_k},x}(\xi) + Z_{n_k,d}^{x,y}(\xi) + Z_{n_k,d}^{y,y_{n_k}}(\xi) \\ &= d(f_\xi^{n_k}(x_{n_k}), f_\xi^{n_k}(x)) + d(f_\xi^{n_k}(x), f_\xi^{n_k}(y)) + d(f_\xi^{n_k}(y), f_\xi^{n_k}(y_{n_k})) \\ &\leq |(f_\xi^{n_k})'(\hat{x}_k)| d(x_{n_k}, x) + Z_{n,d}^{x,y}(\xi) + |(f_\xi^{n_k})'(\hat{y}_k)| d(y_{n_k}, y), \end{aligned}$$

for some points $\hat{x}_k$ and $\hat{y}_k$ between x_{n_k} and x and between y_{n_k} and y, respectively. By (2.3.7), $\hat{x}_k \to x$ and $\hat{y}_k \to y$ as $k \to \infty$. Then, for k large enough $\hat{x}_k \in J_x(\xi)$ and $\hat{y}_k \in J_y(\xi)$ and it follows

$$Z_{n_k,d}^{x_{n_k},y_{n_k}}(\xi) \leq C t^{n_k} d(x_{n_k}, x) + C \lambda^{n_k} + C t^{n_k} d(y_{n_k}, y).$$

This implies

$$\limsup_{k \to \infty} F_{n_k}(x_{n_k}, y_{n_k}, \xi) \leq \ln t.$$

Given $L > 1$ satisfying (2.3.3), as

$$-\ln L \leq F_{n_k}(x_{n_k}, y_{n_k}, \xi) \leq \ln t < 0,$$

Fatou's Lemma implies

$$\limsup_{k \to \infty} \int F_{n_k}(x_{n_k}, y_{n_k}, \xi)\, d\mu(\xi) \leq \int \limsup_{k \to \infty} F_{n_k}(x_{n_k}, y_{n_k}, \xi)\, d\mu(\xi) \leq \ln t.$$

As $t \in (\lambda, 1)$ was arbitrary, it follows

$$\limsup_{k \to \infty} \int F_{n_k}(x_{n_k}, y_{n_k}, \xi) d\mu(\xi) \leq \ln \lambda,$$

which contradicts (2.3.6).

Case $x = y$. By Lemma 2.3.9, for every $t \in (\lambda, 1)$ and $x \in \mathbb{S}^1$ and for almost every ξ there exist an open neighborhood $J_x(\xi) \subset \mathbb{S}^1$ of x and a constant $C > 0$ such that for all $n \in \mathbb{N}$ it holds

$$\max_{z \in J_x(\xi)} |(f_\xi^n)'(z)| \leq C t^n. \tag{2.3.8}$$

Fix $\xi \in \Sigma_N^+$ such that (2.3.8) holds. Take $k_1 \in \mathbb{N}$ large enough so that x_{n_k} and y_{n_k} are both $J_x(\xi)$, for all $k \geq k_1$. By the mean value inequality,

$$F_{n_k}(x_{n_k}, y_{n_k}, \xi) \leq \frac{1}{n_k} \ln C + \ln t,$$

for all $k \geq k_1$. Thus,

$$\limsup_{k \to \infty} F_{n_k}(x_{n_k}, y_{n_k}, \xi) \leq \ln t.$$

By Fatou's lemma and using again that $t \in (\lambda, 1)$ was arbitrary, it follows

$$\limsup_{k\to\infty} \int F_{n_k}(x_{n_k}, y_{n_k}, \xi)\, d\mu(\xi) \le \int \limsup_{k\to\infty} F_{n_k}(x_{n_k}, y_{n_k}, \xi)\, d\mu(\xi) \le \ln \lambda,$$

which contradicts (2.3.6).

This proves the claim. □

By Claim 2.3.11, there exists $k \in \mathbb{N}$ sufficiently large so that $k > 4$ and

$$\frac{1}{k} A_k < \frac{1}{2} \ln \lambda < 0.$$

Consider again $L > 1$ as in (2.3.3).

Hence, for $\xi \in \Sigma_N^+$ and $x \neq y$

$$-k \ln L \le \left| \ln \frac{d(f_\xi^k(x), f_\xi^k(y))}{d(x,y)} \right| \le k \ln L$$

Using the above and that $e^x \le 1 + x + x^2 e^{|x|}/2$, for every $\alpha \in (0,1)$ it follows that

$$\begin{aligned}
\int_{\Sigma_N^+} \frac{d^\alpha(f_\xi^k(x), f_\xi^k(y))}{d^\alpha(x,y)}\, d\mu(\xi) &= \int_{\Sigma_N^+} e^{\alpha \ln(d(f_\xi^k(x), f_\xi^k(y))/d(x,y))}\, d\mu(\xi) \\
&\le \int_{\Sigma_N^+} \Big(1 + \alpha \ln \frac{d(f_\xi^k(x), f_\xi^k(y))}{d(x,y)} + \\
&\qquad + \frac{\alpha^2}{2} \ln^2 \left(\frac{d(f_\xi^k(x), f_\xi^k(y))}{d(x,y)} \right) e^{|\ln(d(f_\xi^k(x), f_\xi^k(y))/d(x,y))|} \Big)\, d\mu(\xi) \\
&\le 1 + \alpha \int_{\Sigma_N^+} \ln \left(\frac{Z_{k,d}^{x,y}(\xi)}{d(x,y)} \right) d\mu(\xi) + \frac{\alpha^2}{2} (k \ln L)^2 L^k \\
&\le 1 + \alpha A_k + \frac{\alpha^2}{2} (k \ln L)^2 L^k < 1 + \frac{\alpha}{2} k \ln \lambda + \frac{\alpha^2}{2} (k \ln L)^2 L^k.
\end{aligned}$$

To finish this proof, note that for α small enough such that

$$\alpha \le -\frac{\ln \lambda}{2k(\ln L)^2 L^k}$$

we have

$$\int_{\Sigma_N^+} \frac{d^\alpha(f_\xi^k(x), f_\xi^k(y))}{d^\alpha(x,y)}\, d\mu(\xi) \le 1 + \frac{\alpha}{4} k \ln \lambda \le \lambda^\alpha < 1.$$

This implies the assertion. □

Proof of Theorem B. By Proposition 2.3.10, there exist $\alpha \in (0,1)$, $\lambda \in (0,1)$, and $k \in \mathbb{N}$ so that $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is k-ECA with rate λ. By Proposition 2.1.12, $(\mathcal{F}, \mathbf{p}, D)$ is CA with contraction rate $\lambda^{1/k}$, where

$$D(x,y) \overset{\text{def}}{=} d^\alpha(x,y) + \frac{1}{\lambda^{1/k}} \mathbb{E}(Z_{1,d^\alpha}^{x,y}) + \cdots + \frac{1}{\lambda^{(k-1)/k}} \mathbb{E}(Z_{k-1,d^\alpha}^{x,y})$$

is as in (2.1.7) for d^α instead of d. By Lemma 2.1.11, it holds $d \le D \le C d^\alpha$ for some $C > 0$ and hence D is strongly equivalent to d^α. □

2.4 Examples

The first example illustrates an application of Proposition 2.2.1. The last two examples illustrate that the hypotheses in Theorems A and B are sharp. For the examples in Sections 2.4.2 and 2.4.3, we consider $\mathbb{S}^1$ equipped with the usual metric $d(x,y) \overset{\text{def}}{=} \min\{|x-y|, 1-|x-y|\}$.

2.4.1 IFSs on $\mathbb{R}^k$

Let us introduce the class of IFSs studied in [MS21]. First, let's define a partial order on $\mathbb{R}^k$. Given $J \subset \{1,\dots,k\}$, define a partial order as follows: for $x, y \in \mathbb{R}^k$, we write $x <_J y$ if and only if

$$x_i < y_i \quad \text{for} \quad i \in J \quad \text{and} \quad x_i > y_i \quad \text{for} \quad i \notin J.$$

Let $S \subset \mathbb{R}^k$. A function $f : S \to S$ is called *J-increasing* (or, *J-decreasing*) if

$$x <_J y \Rightarrow f(x) <_J f(y) \quad (\text{or,} \quad x <_J y \Rightarrow f(y) <_J f(x))\,.$$

And, f is called *J-monotone* if f is either J-increasing or J-decreasing.

Now, let us consider $\mathcal{F}$ a IFS of functions from S itself, $\mathbf{p}$ a non-degenerate probability vector and d the taxicab distance on $\mathbb{R}^k$. In this context, the triple $(\mathcal{F}, \mathbf{p}, d)$ satisfies the *J-splitting condition* if all maps in $\mathcal{F}$ are J-monotone and there exist $m \in \mathbb{N}$, $\xi, \eta \in \Sigma_N^+$ such that

$$f_\xi^m(S) <_J f_\eta(S).$$

By [MS21, Theorem 2], if S is bounded in $\mathbb{R}^k$ and $(\mathcal{F}, \mathbf{p}, d)$ satisfies the J-splitting condition, then there exists $\lambda \in (0,1)$ and an integrable map $c : \Sigma_N^+ \to [0,\infty)$ such that for almost every $\xi \in \Sigma_N^+$ and all $n \in \mathbb{N}$

$$\operatorname{diam} f_\xi^n(S) \le c(\xi)\lambda^n.$$

That is, $(\mathcal{F}, \mathbf{p}, d)$ is $\mathrm{S}_{\mathrm{exp}}$ and by integrability of c the triple $(\mathcal{F}, \mathbf{p}, d)$ is also $\mathrm{SA}_{\mathrm{exp}}$. Therefore, Proposition 2.2.1 applies to get existence of a metric D given by (2.2.2) equivalent to d such that $(\mathcal{F}, \mathbf{p}, D)$ is CA.

2.4.2 LECA, but not ESCA

Let $\mathcal{F} = \{f_0, f_1\}$ be the family of two diffeomorphisms of $\mathbb{S}^1$ such that f_0 has two fixed points, one attracting and one repelling, and f_1 is an irrational rotation. Note that $(\mathcal{F}, d)$ is proximal. Let $\mathbf{p} = (p, 1-p)$ be a non-degenerate probability vector. Note that $(\mathcal{F}, \mathbf{p}, d)$ fails to be NEA. The triple $(\mathcal{F}, \mathbf{p}, d)$ satisfies the hypotheses of Theorem B (and hence of Proposition 2.3.1). Hence, by Corollary 2.3.8, $(\mathcal{F}, \mathbf{p}, d)$ is SA and LECA. It is easy to check that for $(\mathcal{F}, \mathbf{p}, d)$ condition (2.1.8) for NEA and condition (2.1.6) for k-ECA ($k \in \mathbb{N}$ arbitrary) are violated at the repelling fixed point of f_0.

Note that $(\mathcal{F}, \mathbf{p}, d)$ verifies the hypotheses of Theorem B. Hence, there exist $\alpha \in (0,1]$, $\lambda \in (0,1)$ and $n \in \mathbb{N}$ such that for $D \stackrel{\text{def}}{=} (d^\alpha)_{n,\lambda}$ the triple $(\mathcal{F}, \mathbf{p}, D)$ is CA. In particular, $(\mathcal{F}, \mathbf{p}, D)$ satisfies NEA, SA, LECA, and ESCA. However, by the latter together with Lemma 2.1.32, for all $\beta \in (0,1]$ the metric D^β cannot be strongly equivalent to d.

Let μ be a stationary probability for IFS with probabilities $(\mathcal{F}, \mathbf{p})$. By stationarity, it holds

$$\mu = p\,(f_0)_*\mu + (1-p)\,(f_1)_*\mu.$$

By [GS17, Lemma 2.6], μ is non-atomic and has full support. Consider the metric ρ on $\mathbb{S}^1$, given by $\rho(x,y) \stackrel{\text{def}}{=} \min\{\mu([x,y]), \mu([y,x])\}$. By [GS17, Proposition 1.2], $(\mathcal{F}, \mathbf{p}, \rho)$ is NEA.

Lemma 2.4.1. *$(\mathcal{F}, \mathbf{p}, \rho)$ is SA and LECA.*

Proof. Since $(\mathcal{F}, \mathbf{p}, d)$ is SA and $(\mathcal{F}, \mathbf{p}, \rho)$ is NEA, Lemma 2.1.15 implies that $(\mathcal{F}, \mathbf{p}, \rho)$ is SA. By Lemma 2.1.27, $(\mathcal{F}, \mathbf{p}, \rho)$ is LECA. □

The following result checking that the IFS is "ρ-isometric in average" if and only if it is "ρ-isometric" is straightforward.

Claim 2.4.2. *Assume $\rho(x,y) = \mu([x,y])$. Then, it holds $\mathbb{E}(Z^{x,y}_{n,\rho}) = \rho(x,y)$ if only if for all $\xi_1, \dots, \xi_n$ we have*

$$\rho(f_{\xi_1\dots\xi_n}(x), f_{\xi_1\dots\xi_n}(y)) = \mu(f_{\xi_1\dots\xi_n}([x,y])).$$

Lemma 2.4.3. *The triple $(\mathcal{F}, \mathbf{p}, \rho)$ is not ESCA.*

Proof. Given $x \in \mathbb{S}^1$ and $\varepsilon \in (0,1)$, denote by $\mathrm{Arc}(x,\varepsilon)$ the open arc centered at x and with μ-measure equal to ε (recall that μ is nonatomic and has full support, hence $\mathrm{Arc}(x,\varepsilon)$ is a nontrivial interval). Given x and $\ell \in \mathbb{N}$, consider the set

$$V_\ell(x) \stackrel{\text{def}}{=} \mathrm{Arc}(x, 4^{-1}) \cap \bigcap_{\xi_1,\dots,\xi_\ell} (f_{\xi_1\dots\xi_\ell})^{-1}\, \mathrm{Arc}(f_{\xi_1\dots\xi_\ell}(x), 4^{-1}) \subset \mathrm{Arc}(x, 4^{-1}),$$

which is also a nontrivial open interval. For every $y, z \in V_\ell(x)$ such that $[y,z] \subset V_\ell(x)$ for every $\xi_1, \dots, \xi_\ell \in \{0,1\}$, it hence holds

$$f_{\xi_1\dots\xi_\ell}([y,z]) \subset \mathrm{Arc}(f_{\xi_1\dots\xi_\ell}(x), 4^{-1})$$

and therefore

$$4^{-1} \geq \mu(f_{\xi_1\dots\xi_\ell}([y,z])) \; = \rho(f_{\xi_1\dots\xi_\ell}(y), f_{\xi_1\dots\xi_\ell}(z)).$$

Hence, from Claim 2.4.2, it follows

$$\rho(y,z) = \mu([y,z]) = \mathbb{E}(Z^{y,z}_{\ell,\rho}). \tag{2.4.1}$$

Any other neighborhood V of x, contains an open arc W containing x, so that for all $\ell \geq 1$, $W \cap V_\ell(x) \subset V$ is and open arc containing x. Now it is enough to consider $y, z \in W \cap V_\ell(x)$ to get (2.4.1). This completes the proof that $(\mathcal{F}, \mathbf{p}, \rho)$ is not ESCA. □

2.4.3 LECA, but not NEA

Let $\mathbf{p} = (p, 1-p)$ be a non-degenerate probability vector and μ its associated Bernoulli measure. Without loss of generality, we can assume $p = \max\{p, 1-p\}$. Let $f_0, f_1 \colon \mathbb{S}^1 \to \mathbb{S}^1$ be orientation preserving homeomorphisms, such that there exist two open arcs $I, J \subset \mathbb{S}^1$ with disjoint closures having the following properties (compare also Figure 4):

1. The extreme points of $\overline{J}$ are fixed points y_0 and y_1 of f_0 and f_1, respectively. Here we are assuming that $y_0 \neq y_1$.
2. There exists an open arc $J^* \subset J$ such that $f_0(J^*), f_1(J^*) \subset \mathbb{S}^1 \setminus \overline{J}$.
3. The arc I is (forward) invariant, that is, $f_0(I), f_1(I) \subset I$.
4. For every $x \in \mathbb{S}^1 \setminus \overline{J}$ there exists $n \geq 0$ such that $f_\xi^n(x) \in I$ for all $\xi \in \Sigma_2^+$.
5. There is $r \in (0,1)$ so that
$$d\left(f_\xi^n(x), f_\xi^n(y)\right) \leq r^n d(x,y) \quad \text{for every } n \in \mathbb{N} \text{ and } x, y \in I.$$
6. $d(f_i(x), f_i(y)) \geq d(x,y)$ for every $x, y \in J \cap f_0^{-1}(J) \cap f_1^{-1}(J)$ and $i = 0, 1$.
7. Every f_i is Lipschitz: there is $c > 1$ so that
$$d(f_i(x), f_i(y)) \leq c d(x,y) \quad \text{for every } x, y \in \mathbb{S}^1 \text{ and } i = 0, 1.$$

By (6), for the IFS $\mathcal{F} = \{f_0, f_1\}$ the triple $(\mathcal{F}, \mathbf{p}, d)$ is not ε-LCA. An appropriately chosen example also fails to be NEA (just choose f_0, f_1 being expanding in J). Since $\overline{J} \subsetneq \mathbb{S}^1$ and $f_0^{-1} \cup f_1^{-1}\left(\overline{J}\right) \subset \overline{J}$ we have that $\mathcal{F}$ fails to be backward minimal, hence methods from [GS17] do not apply immediately. Below we prove the following.

Lemma 2.4.4. *The triple $(\mathcal{F}, \mathbf{p}, d)$ is proximal, S, SA, and LECA.*

We will construct a metric $\hat{D}$ that will be equivalent to d, for which $(\mathcal{F}, \mathbf{p}, \hat{D})$ is NEA, LECA, and ESCA. Then we will construct a metric D equivalent to $\hat{D}$ (and hence d), such that $(\mathcal{F}, \mathbf{p}, D)$ is CA.

By (3) and (5), $\mathcal{F}$ induces a contracting IFS on I. Together with (4), every $x \notin \overline{J}$ eventually enters and remains in I. On the other hand, it follows from (2) that

$$f_0^{-1}(\overline{J}), f_1^{-1}(\overline{J}) \subset \overline{J}$$

and there is a set of points in $X \subset \overline{J}$ and for every $x \in X$ some sequence $\xi = \xi(x)$ such that $f_\xi^n(x) \in X$ for every $n \in \mathbb{N}$, though other forward iterates under the IFS $\mathcal{F}$ eventually leave $\overline{J}$. Though, as counterpart and first preliminary result we show that for every x

$$\{\xi \in \Sigma_2^+ \colon f_\xi^n(x) \in J \quad \text{for all} \quad n \in \mathbb{N}\}$$

has measure zero. For $k \geq 1$ and $x \in \mathbb{S}^1$, define

$$\Gamma_x^0 \stackrel{\text{def}}{=} \begin{cases} \Sigma_2^+ & \text{if } x \in I, \\ \emptyset & \text{otherwise,} \end{cases}$$
$$\Gamma_x^k \stackrel{\text{def}}{=} \{\xi \in \Sigma_2^+ \colon f_\xi^k(x) \in I, f_\xi^{k-1}(x) \notin I\}, \quad \Gamma_x \stackrel{\text{def}}{=} \bigcup_{k\geq 0} \Gamma_x^k.$$

Lemma 2.4.5. *There exists $N \in \mathbb{N}$ such that*

$$\mu\Big(\bigcup_{k=0}^{N+1} \Gamma_x^k\Big) > 0 \quad \textit{for every } x \in \mathbb{S}^1. \tag{2.4.2}$$

and for all $m \geq 1$

$$\mu\Big(\Gamma_x^{N+m+1}\Big) \leq \mu\Big(\Sigma_2^+ \setminus \bigcup_{k=0}^{N+m} \Gamma_x^k\Big) \leq p^m. \tag{2.4.3}$$

Moreover, for every $x \in \mathbb{S}^1$ it holds $\mu(\Gamma_x) = 1$.

Proof. Fix $y \in J^*$. By (4), it holds $f_0(y), f_1(y) \notin \overline{J}$. Let $K, L \subset \mathbb{S}^1 \setminus J$ be the open arcs with extremes $f_0(y)$ and y_1 and $f_1(y)$ and y_0, respectively. By (4), there are $k = k(K) \in \mathbb{N}$ and $\ell = \ell(L) \in \mathbb{N}$ so that

$$f_\xi^k(K), f_\xi^\ell(L) \in I \quad \text{for every } \xi \in \Sigma_2^+.$$

Let $N \stackrel{\text{def}}{=} k + \ell$. By (3), for all $\xi \in \Sigma_2^+$

$$f_\xi^{k+\ell}(K), f_\xi^{k+\ell}(L) \in I \quad \text{for every } \xi \in \Sigma_2^+.$$

Let $W \subset \mathbb{S}^1 \setminus J$ be the closed arc with extremes $f_0(y)$ and $f_1(y)$. As f_0 and f_1 preserve orientation, f_ξ^n preserves orientation. Hence, $f_\xi^n(W) \subset I$ for all $\xi \in \Sigma_2^+$.

Now, let us prove (2.4.2) and (2.4.3). Fix $x \in \mathbb{S}^1$.

Case $x \in W$. As $f_\xi^{k+\ell}(x) \in I$ for every $\xi \in \Sigma_2^+$, (2.4.2) and (2.4.3) are immediate.

Case $x \notin W$. Let us construct a sequence $\xi \in \Sigma_2^+$ such that for all $m \geq 1$

$$[\xi_1] \cup [\eta_1, \xi_2] \cup \cdots \cup [\eta_1, \ldots, \eta_{m-1}, \xi_m] \subset \bigcup_{k=0}^{m+N} \Gamma_x^k, \tag{2.4.4}$$

where $\eta_i \in \{0,1\} \setminus \{\xi_i\}$. Since x is either between y and $f_0(y)$ or between y and $f_1(y)$, there exists $\xi_1 \in \{0,1\}$ such that x is between y and $f_{\xi_1}(y)$. As f_{ξ_1} preserves orientation, $f_{\xi_1}(x) \in W$. So that

$$[\xi_1] \subset \bigcup_{k=0}^{N+1} \Gamma_x^k,$$

that is, (2.4.4) holds for $m = 1$. Let $\eta_1 \in \{0,1\} \setminus \{\xi_1\}$. If $f_{\eta_1}(x) \in W$ then $\bigcup_{k=0}^{N+1} \Gamma_x^k = \Sigma_2^+$ which implies (2.4.4) for all $m \geq 2$ and $\xi_m \in \{0,1\}$. If $f_{\eta_1}(x) \notin W$ then $f_{\eta_1}(x)$ is either between y

and $f_0(y)$ or between y and $f_1(y)$ so there exists $\xi_2 \in \{0,1\}$ such that $f_{\eta_1}(x)$ is between y and $f_{\xi_2}(y)$. As f_{ξ_2} preserves orientation, $f_{\xi_2}(f_{\eta_1}(x)) \in W$. So that

$$[\xi_1] \cup [\eta_1, \xi_2] \subset \bigcup_{k=0}^{N+2} \Gamma_x^k,$$

that is, (2.4.4) holds for $m = 2$. Let $\eta_2 \in \{0,1\} \setminus \{\xi_2\}$. If $f_{\eta_2}(f_{\eta_1}(x)) \in W$ then $\bigcup_{k=0}^{N+2} \Gamma_x^k = \Sigma_2^+$ which implies (2.4.4) for all $m \geq 3$ and $\xi_m \in \{0,1\}$. If $f_{\eta_2}(f_{\eta_1}(x)) \notin W$ then $f_{\eta_2}(f_{\eta_1}(x))$ is either between y and $f_0(y)$ or between y and $f_1(y)$ so that there exists $\xi_3 \in \{0,1\}$ such that $f_{\eta_2}(f_{\eta_1}(x))$ is between y and $f_{\xi_3}(y)$. Continuing this process inductively on m we conclude (2.4.4).

Therefore

$$[\xi_1] \subset \bigcup_{k=0}^{N+1} \Gamma_x^k \quad \text{and hence} \quad \mu\Big(\bigcup_{k=0}^{N+1} \Gamma_x^k\Big) \geq 1 - p > 0,$$

and so (2.4.2) holds. Moreover

$$\Gamma_x^{N+m+1} \subset \Sigma_2^+ \setminus \bigcup_{k=0}^{m+N} \Gamma_x^k \subset [\eta_1, \dots, \eta_m]$$

so that $\mu(\Gamma_x^{N+m+1}) \leq p^m$. This proves the lemma. □

Proof of Lemma 2.4.4. Given any $x, y \in \mathbb{S}^1$, let $\xi \in \Gamma_x \cap \Gamma_y$ and choose $k \in \mathbb{N}$ such that $f_\xi^k(x), f_\xi^k(y) \in I$. By (4), it holds

$$\lim_{n\to\infty} d(f_\xi^n(x), f_\xi^n(y)) = \lim_{n\to\infty} d(f_\xi^{k+n}(x), f_\xi^{k+n}(y)) = 0.$$

As by Lemma 2.4.5, $\mu(\Gamma_x \cap \Gamma_y) = 1$ holds, it follows that $(\mathcal{F}, \mathbf{p}, d)$ is S. By Lemma 2.1.5, it is SA and proximal. By Lemma 2.1.27, it is LECA. □

Fix $\ell \geq 1$ and pick $\alpha \in (0,1)$ such that

$$r^\ell c < 1 \quad \text{and} \quad c^\alpha < c^{(\ell+1)\alpha} < \frac{1}{p}. \tag{2.4.5}$$

The choice of these numbers will be apparent in the proof of Lemma 2.4.7. Consider the metric d^α on $\mathbb{S}^1$ and define the metric $\hat{D}$ on $\mathbb{S}^1$ by

$$\hat{D}(x,y) \stackrel{\text{def}}{=} \mathbb{E}\left(\sup_{n\geq 0} Z_{n,d^\alpha}^{x,y}\right).$$

Since d and d^α are equivalent, $\hat{D}$ and d are equivalent by Lemma 2.1.10.

Lemma 2.4.6. *It holds*

$$d^\alpha \leq \hat{D} \leq C d^\alpha \quad \textit{where} \quad C \stackrel{\text{def}}{=} c^{(N+1)\alpha}\Big(1 + 2\sum_{k\geq 1} (c^\alpha p)^k\Big).$$

Proof. Clearly $d^\alpha \le \hat{D}$. Let us show the other inequality. For $k \ge 1$ and $x, y \in \mathbb{S}^1$ define

$$\Gamma^k_{x,y} \overset{\text{def}}{=} \Big\{\xi \in \Sigma^+_2 \colon f^k_\xi(x) \in I, f^k_\xi(y) \in I\Big\} \cap \Big\{\xi \in \Sigma^+_2 \colon f^{k-1}_\xi(x) \notin I \text{ or } f^{k-1}_\xi(y) \notin I\Big\}.$$

If $x, y \in I$, then let $\Gamma^0_{x,y} \overset{\text{def}}{=} \Sigma^+_2$. If $x \notin I$ or $y \notin I$, then let $\Gamma^0_x \overset{\text{def}}{=} \emptyset$. Note that $\{\Gamma^k_{x,y}\}_k$ is a family of pairwise disjoint sets and $\Gamma^k_{x,y} \subset \Gamma^k_x \cup \Gamma^k_y$. Moreover, it is immediate from the definition that

$$\Gamma_x \cap \Gamma_y \subset \Gamma_{x,y} \overset{\text{def}}{=} \bigcup_{n \ge 0} \Gamma^n_{x,y}.$$

Hence, together with Lemma 2.4.5, it follows

$$1 = \mu(\Gamma_x \cap \Gamma_y) = \mu(\Gamma_{x,y}) = \sum_{n \ge 0} \mu\left(\Gamma^n_{x,y}\right).$$

Fix $x, y \in \mathbb{S}^1$. For every $\xi \in \Gamma^k_{x,y}$, by (7) and (5), we have that

$$\sup_{n \ge 0} Z^{x,y}_{n,d^\alpha}(\xi) = \sup_{n \ge 0} d^\alpha(f^n_\xi(x), f^n_\xi(y)) < c^{k\alpha} d^\alpha(x, y).$$

It follows that

$$\begin{aligned}
\hat{D}(x, y) &= \mathbb{E}\Big(\sup_{n \ge 0} Z^{x,y}_{n,d^\alpha}\Big) = \sum_{k \ge 0} \mathbb{E}\Big(\sup_{n \ge 0} Z^{x,y}_{n,d^\alpha} \mathbb{1}_{\Gamma^k_{x,y}}\Big) \\
&\le c^{(N+1)\alpha} d^\alpha(x, y) \mu\Big(\Gamma^0_{x,y} \cup \ldots \cup \Gamma^{N+1}_{x,y}\Big) + d^\alpha(x, y) \sum_{k \ge N+2} c^{k\alpha} \mu\Big(\Gamma^k_{x,y}\Big) \\
&\le c^{(N+1)\alpha} d^\alpha(x, y) \Big(1 + \sum_{k \ge N+2} c^{(k-N-1)\alpha} \Big(\mu(\Gamma^k_x) + \mu(\Gamma^k_y)\Big)\Big) \\
&\le c^{(N+1)\alpha} d^\alpha(x, y) \Big(1 + 2 \sum_{k \ge N+2} c^{(k-N-1)\alpha} p^{k-N-1}\Big) \\
&\le C d^\alpha(x, y).
\end{aligned}$$

As x, y were arbitrary, this finishes the proof. □

Lemma 2.4.7. *The triple $(\mathcal{F}, \mathbf{p}, \hat{D})$ is SA, NEA, LECA and ESCA.*

Proof. By Lemmas 2.4.6 and 2.1.3, $(\mathcal{F}, \mathbf{p}, \hat{D})$ is S. Hence, by which Lemma 2.1.2, $(\mathcal{F}, \mathbf{p}, \hat{D})$ is SA, proving the first assertion.

To show NEA and LECA, check that

$$\begin{aligned}
\mathbb{E}\left(Z^{x,y}_{1,\hat{D}}\right) &= p\, \hat{D}(f_0(x), f_0(y)) + (1-p)\, \hat{D}(f_1(x), f_1(y)) \\
&= p\, \mathbb{E}\left(\sup_{n \ge 0} Z^{f_0(x), f_0(y)}_{n,d^\alpha}\right) + (1-p)\, \mathbb{E}\left(\sup_{n \ge 0} Z^{f_1(x), f_1(y)}_{n,d^\alpha}\right) \\
&= p\, \mathbb{E}\left(\mathbb{1}_{[0]} \sup_{n \ge 0} Z^{x,y}_{n+1,d^\alpha}\right) + (1-p)\, \mathbb{E}\left(\mathbb{1}_{[1]} \sup_{n \ge 0} Z^{x,y}_{n+1,d^\alpha}\right) = \mathbb{E}\left(\sup_{n \ge 1} Z^{x,y}_{n,d^\alpha}\right).
\end{aligned}$$

This implies $\mathbb{E}(Z^{x,y}_{1,\hat{D}}) \le \hat{D}(x, y)$, that is, $(\mathcal{F}, \mathbf{p}, \hat{D})$ is NEA. Hence, as $(\mathcal{F}, \mathbf{p}, \hat{D})$ is NEA. As $(\mathcal{F}, \mathbf{p}, \hat{D})$ is NEA and SA, Lemma 2.1.28 property LECA follows.

Now, let us prove that $(\mathcal{F}, \mathbf{p}, \hat{D})$ is ESCA. Let N be as in Lemma 2.4.5, ℓ as in (2.4.5), and C is as in Lemma 2.4.6. Recalling that $r, p \in (0,1)$, we can fix $n \in \mathbb{N}$ sufficiently large such that

$$(r^\ell c)^{n\alpha} + c^{n(\ell+1)\alpha} p^{n-N} < \frac{1}{2C}. \tag{2.4.6}$$

Recall that $\bigcup_{k=0}^{N+1} \Gamma_x^k \neq \emptyset$ for every $x \in \mathbb{S}^1$. Note that $\bigcup_{k=0}^{N+1} \Gamma_x^k$ is covered by cylinders of length N and, in particular, $\{f_\xi^N : \xi \in \bigcup_{k=0}^{N+1} \Gamma_x^k\}$ is a collection of at most 2^N homeomorphisms. Hence, given $x \in \mathbb{S}^1$, the set

$$V_x \stackrel{\text{def}}{=} \bigcap_{\xi \in \bigcup_{k=0}^{N} \Gamma_x^k} \left(f_\xi^N\right)^{-1}(I),$$

is, as an intersection of finitely many open intervals containing x, an open neighborhood of x. Moreover, for every $y, z \in V_x$ it holds

$$\begin{aligned}
\mathbb{E}\left(Z^{y,z}_{(\ell+1)n,d^\alpha}\right) &= \mathbb{E}\left(\mathbb{1}_{\bigcup_{k=0}^{n}\Gamma_x^k} Z^{y,z}_{(\ell+1)n,d^\alpha}\right) + \mathbb{E}\left(\mathbb{1}_{\Sigma_2^+ \setminus \bigcup_{k=0}^{n}\Gamma_x^k} Z^{y,z}_{(\ell+1)n,d^\alpha}\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\bigcup_{k=0}^{n}\Gamma_x^k} d^\alpha(f^{\ell n}_{\sigma^n(\xi)}(f^n_\xi(y)), f^{\ell n}_{\sigma^n(\xi)}(f^n_\xi(z))\right) + \mathbb{E}\left(\mathbb{1}_{\Sigma_2^+ \setminus \bigcup_{k=0}^{n}\Gamma_x^k} Z^{y,z}_{(\ell+1)n,d^\alpha}\right) \\
\text{(by (7))} &\leq c^{n\alpha} r^{\ell n\alpha} d^\alpha(y,z)\mu\left(\bigcup_{k=0}^{n}\Gamma_x^k\right) + c^{(\ell+1)n\alpha} d^\alpha(y,z)\mu\left(\Sigma_2^+ \setminus \bigcup_{k=0}^{n}\Gamma_x^k\right) \\
&\leq \left((r^\ell c)^{n\alpha} + c^{(\ell+1)n\alpha} p^{n-N}\right) d^\alpha(y,z) \\
&\leq \frac{1}{2C} d^\alpha(y,z).
\end{aligned}$$

By Lemma 2.4.6, for every $y, z \in V_x$

$$\frac{\mathbb{E}\left(Z^{y,z}_{(m+1)N,\hat{D}}\right)}{\hat{D}(y,z)} \leq C \frac{\mathbb{E}\left(Z^{y,z}_{(m+1)N,d^\alpha}\right)}{d^\alpha(y,z)} \leq \frac{1}{2},$$

which implies that

$$\sup_{(y,z) \in V_x \times V_x} \frac{\mathbb{E}\left(Z^{y,z}_{(m+1)N,\hat{D}}\right)}{\hat{D}(y,z)} \leq \frac{1}{2} < 1.$$

Since x is arbitrary, this proves ESCA, and completes the proof. □

Remark 2.4.8. The constant $\frac{1}{2}$ in equation (2.4.6) is insignificant, in fact, we can change $\frac{1}{2}$ for any $t \in (0,1)$ and get the same result.

By Theorem A, Lemma 2.4.7 allows us to conclude that there exists a metric D strongly equivalent to $\hat{D}$ such that $(\mathcal{F}, \mathbf{p}, D)$ is CA. Hence, D is strongly equivalent to d^α.

Remark 2.4.9. Note that in the proof of Lemma 2.4.7 we show that $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is ESCA. The triple $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is an example of a system that is ESCA such that the ℓ required in the definition is not the constant 1.

3 Stochastic properties for contracting on average Iterated Function Systems

3.1 Stochastic properties after metric change

Let us provide some details on how a change of metric impacts (or not) statistical properties of an IFS which is contracting on average (CA, recall this definition in (1.1.1)). As before, consider $\mathcal{F}$ an IFS on $\mathbb{S}^1$, $\mathbf{p}$ a probability vector and d the usual metric on $\mathbb{S}$.

Let us first comment on the concept of stationarity. This term is justified by the fact that if $(I_n)_{n\in\mathbb{N}}$ is a stochastic sequence taking values in $\{0,\dots,N-1\}$ and being independently and equally distributed according to the probability vector $\mathbf{p}$ and X is a ν-distributed random variable, independent of $(I_n)_{n\in\mathbb{N}}$, then

$$W_n^X \overset{\text{def}}{=} (f_{I_n}\circ\cdots\circ f_{I_1})(X), \quad W_0^X \overset{\text{def}}{=} X, \tag{3.1.1}$$

defines a stationary stochastic sequence. Even more, $(W_n^X)_{n\in\mathbb{N}}$ is a Markov chain with initial distribution ν. The continuity of the maps in $\mathcal{F}$ and the compactness of $\mathbb{S}^1$ imply that the chain in (3.1.1) has the *weak Feller property*[1] and hence there exists at least one stationary measure ν. Note that, as μ is a Bernoulli measure, the Markov chains (3.1.1) and (1.2.2) coincide, but have different initial distributions. Under the hypothesis that there is no measure simultaneously invariant by all maps in $\mathcal{F}$, it follows from [Mal17, Theorem F] that the stationary measure can not be Dirac. Although, if a statistical property holds for the Markov chain in (1.2.2) for every $x\in\mathbb{S}^1$, then it holds also for the Markov chain in (3.1.1).

If there exists a metric D such that $(\mathcal{F},\mathbf{p},D)$ is CA, then [BDEG88, Theorem 2.1] implies that there exists a unique stationary Borel probability measure ν on $\mathbb{S}^1$.

Remark 3.1.1. In the case of a unique stationary measure, its support is a closed set invariant by all maps in the IFS, which is either a finite set, a Cantor set, or the full circle (for a proof in the case orientation preserving homeomorphisms on the circle see [Nav11, Theorem 2.1.1]). For an IFS of homeomorphisms acting minimally on the circle there exists a unique stationary measure which is fully supported (see for example [GS17, Corollary 2 and Remark 7]).

[1] The *transfer operator* T associated to the pair $(\mathcal{F},\mathbf{p})$ acts on the space of bounded measurable functions $\varphi\colon M\to\mathbb{R}$ by

$$T\varphi(x) \overset{\text{def}}{=} \sum_{i=0}^{N-1} p_i\varphi(f_i(x)).$$

It has the *weak Feller property* if it maps the space of real valued continuous functions on $\mathbb{S}^1$ to itself.

It is relevant to know under what conditions such stationary measure ν is unique and, moreover, under what hypotheses it is true that for every Borel set $B \subset \mathbb{S}^1$

$$P^n(x,B) \stackrel{\text{def}}{=} \mu\{W_n^x \in B\} \to \nu(B)$$

as $n \to \infty$, and what is the speed of such a convergence. By [JT01, Corollary 2.1], for any initial conditions x, the distribution of W_n^x converges exponentially fast to ν in the Prokhorov metric, that is, for every measurable set B and $n \in \mathbb{N}$ it holds

$$P^n(x,B) \le \nu(B_n) + A_x r^n, \quad \nu(B) \le P^n(x,B_n) + A_x r^n,$$

where $B_n \stackrel{\text{def}}{=} \{y \in M \colon D(y,B) < A_x r^n\}$. Here the rate of convergence $r \in (0,1)$ does neither depend on n nor on x. Furthermore, the constant A_x does not depend on n and is uniformly bounded. Note that all previous facts do not depend on the metric on $\mathbb{S}^1$ (within the class of metrics which generate the same topology). Only the explicit convergence in the Prokhorov metric was given in terms of D. Observe that if D and another metric d are such that $C^{-1}d^\alpha \le D \le Cd^\alpha$ for some constants $C^{-1}, \alpha \in (0,1]$ (that is, D and d^α are strongly equivalent), then we still obtain exponential contraction taking

$$B_n \stackrel{\text{def}}{=} \{y \in M \colon d(y,B) < C_x r^{n/\alpha}\}, \quad \text{where} \quad C_x \stackrel{\text{def}}{=} (CA_x)^{1/\alpha}.$$

Assuming CA, in [Elt87, page 484] an Ergodic Theorem was shown, whose assertion is unaltered under any metric change. Assuming the slightly more general property k-ECA for any $k \in \mathbb{N}$, in [Pei93, Theorem 5.1] a strong law of large numbers and a central limit theorem are stated; again these assertions remain the same under metric change.

We are now ready to prove Theorem C.

Proof of Theorem C. We first invoke the idea of metric change to prove Theorem C. By Theorem B, there exist $\alpha, \lambda \in (0,1)$ such that the metric D on $\mathbb{S}^1$ defined by

$$D(x,y) \stackrel{\text{def}}{=} d^\alpha(x,y) + \frac{1}{\lambda^{1/n}} \int d^\alpha(W_1^x, W_1^y) d\mu + \cdots + \frac{1}{\lambda^{(n-1)/n}} \int d^\alpha(W_{n-1}^x, W_{n-1}^y) d\mu \tag{3.1.2}$$

is equivalent to the usual metric d on $\mathbb{S}^1$ and has the property that $(\mathcal{F}, \mathbf{p})$ is contracting on average on $(\mathbb{S}^1, D)$.

Note that none of the assertions of Theorem C depends on the chosen metric on $\mathbb{S}^1$ (within the class of metrics which generate the same topology). By Theorem B, the system is contracting on average with respect to the metric D in (3.1.2). In particular, [Pei93, condition (H3)] is satisfied for $k_0 = 1$. By Theorem B and [BDEG88, Theorem 2.1], there is a unique stationary probability and by [Pei93, Theorem 5.1], the strong law of large numbers SLLN and the central limit theorem CLT hold true. $\square$

3.2 Synchronization on average at an exponential rate

In this section, we study in more detail synchronization on average of orbits with respect to an IFS. In very rough terms, given two points we are interested in the amount of functions that at time n have not yet sufficiently synchronized the orbit of those points. When an IFS is $\mathrm{SA}_{\mathrm{exp}}$, in Proposition 3.2.3 we prove that this amount decays exponentially. Here, "the amount" refers to the ν-measure of the set of sequences ξ whose associated functions f_ξ^n have the desired property.

In the context of IFSs of circle homeomorphism, by [Mal17, Theorems A], the *local contraction* property holds (see Proposition 2.3.1 for more details). Recall the concept of proximality given in (2.1.3). Moreover, [Mal17, Theorem E], proximality is equivalent to exponential synchronization, where *exponential synchronization* means that for every $x, y \in \mathbb{S}^1$ and almost every $\xi \in \Sigma_N^+$ the sequence $\{Z_{n,d}^{x,y}(\xi)\}_{n\in\mathbb{N}}$ converges to 0 exponentially fast as $n \to \infty$. Note that here a priori the rate of convergence of $\{Z_{n,d}^{x,y}(\xi)\}_{n\in\mathbb{N}}$ depends on x, y and ξ. In this section, we show that for an IFS of C^1-diffeomorphisms on $\mathbb{S}^1$ proximality implies synchronization on average with uniform exponential decay rate.

3.2.1 Some auxiliary CA results

Let us state the following result which complements Theorem C. We will use it below to prove our large deviation results.

Proposition 3.2.1. *Assume that $\mathcal{F}$ satisfies hypotheses in Theorem C. Let $\mathbf{p}$ a non-degenerate probability vector. There exist $\lambda \in (0,1)$, $c > 0$, $k \in \mathbb{N}$, and $\alpha_0 \in (0,1)$ such that for all $\alpha \in (0, \alpha_0)$ the following hold*

1. *for every $x, y \in \mathbb{S}^1$, $x \neq y$,*
$$\int_{\Sigma_N^+} \frac{Z_{n,d^\alpha}^{x,y}(\xi)}{d^\alpha(x,y)}\, d\mu(\xi) \leq c\left(\lambda^{\frac{\alpha}{k}}\right)^n,$$
2. *for every $x \in \mathbb{S}^1$,*
$$\int_{\Sigma_N^+} \left[(f_\xi^n)'(x)\right]^\alpha\, d\mu(\xi) \leq c\left(\lambda^{\frac{\alpha}{k}}\right)^n.$$

Remark 3.2.2. For $n \in \mathbb{N}$ large enough, $c\left(\lambda^{\frac{\alpha}{k}}\right)^n < 1$, therefore item *(1.)* implies that $(\mathcal{F}, \mathbf{p}, d^\alpha)$ is ECA.

Proof of Proposition 3.2.1. Let us start this proof by remembering that for all $\alpha \in (0,1)$, $n \in \mathbb{N}$, $\xi \in \Sigma_N^+$ and $x, y \in \mathbb{S}^1$
$$Z_{n,d^\alpha}^{x,y}(\xi) = d^\alpha(f_\xi^n(x), f_\xi^n(y))$$
By Proposition 2.3.10, there exist $k \in \mathbb{N}$ and $\alpha_0 \in (0,1)$ such that for $\alpha \in (0, \alpha_0]$ we have
$$\sup_{x\neq y} \int_{\Sigma_N^+} \frac{Z_{k,d^\alpha}^{x,y}(\xi)}{d^\alpha(x,y)}\, d\mu(\xi) \leq \lambda^\alpha. \tag{3.2.1}$$

On the other hand, define the sequence $(b_n)_{n\geq 1}$ by

$$b_n \stackrel{\text{def}}{=} \sup_{x\neq y} \int_{\Sigma_N^+} \frac{Z^{x,y}_{n,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi).$$

The sequence $(b_n)_{n\geq 1}$ is submultiplicative, that is, $b_{n+m} \leq b_n b_m$ for all $n, m \geq 0$. Indeed, using that

$$Z^{f^m_\xi(x), f^m_\xi(y)}_{n,d^\alpha}(\sigma^m(\xi)) = Z^{x,y}_{n+m,d^\alpha}(\xi),$$

we get

$$\begin{aligned}
b_n b_m &= \left(\sup_{z\neq w} \int_{\Sigma_N^+} \frac{Z^{z,w}_{n,d^\alpha}(\eta)}{d^\alpha(z,w)} d\mu(\eta)\right) \left(\sup_{x\neq y} \int_{\Sigma_N^+} \frac{Z^{x,y}_{m,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi)\right) \\
&= \sup_{x\neq y} \int_{\Sigma_N^+} \left(\sup_{z\neq w} \int_{\Sigma_N^+} \frac{Z^{z,w}_{n,d^\alpha}(\eta)}{d^\alpha(z,w)} d\mu(\eta)\right) \frac{Z^{x,y}_{m,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi) \\
&\geq \sup_{x\neq y} \int_{\Sigma_N^+} \left(\int_{\Sigma_N^+} \frac{Z^{f^m_\xi(x), f^m_\xi(y)}_{n,d^\alpha}(\eta)}{d^\alpha(f^m_\xi(x), f^m_\xi(y))} d\mu(\eta)\right) \frac{Z^{x,y}_{m,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi).
\end{aligned}$$

Using that μ is a Bernoulli measure and that $Z^{x,y}_{m,d^\alpha}(\xi)$ depends only on the first m entries of the sequence ξ, we conclude that

$$\begin{aligned}
b_n b_m &\geq \sup_{x\neq y} \int_{\Sigma_N^+} \int_{\Sigma_N^+} \frac{Z^{f^m_\xi(x), f^m_\xi(y)}_{n,d^\alpha}(\eta)}{Z^{x,y}_{m,d^\alpha}(\xi)} \frac{Z^{x,y}_{m,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\eta) d\mu(\xi) \\
&= \sup_{x\neq y} \int_{\Sigma_N^+} \frac{Z^{f^m_\xi(x), f^m_\xi(y)}_{n,d^\alpha}(\sigma^m(\xi))}{d^\alpha(x,y)} d\mu(\xi) \\
&= \sup_{x\neq y} \int_{\Sigma_N^+} \frac{Z^{x,y}_{n+m,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi) = b_{n+m}.
\end{aligned} \tag{3.2.2}$$

Therefore, for k satisfying (3.2.1) and $n = mk + r \in \mathbb{N}$ with $r \in \{0, \dots, k-1\}$ we have

$$b_n \leq (b_k)^m b_r$$

that is, for all $n \in \mathbb{N}$

$$\sup_{x\neq y} \int_{\Sigma_N^+} \frac{Z^{x,y}_{n,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi) \leq c \left(\lambda^{\frac{\alpha}{k}}\right)^n,$$

where $c = \max_{r\in\{0,\dots,k-1\}} b_r \left(\lambda^\alpha\right)^{-\frac{r}{k}}$. Hence, for $x, y \in \mathbb{S}^1$, $x \neq y$, since the integral

$$\int_{\Sigma_N^+} \frac{Z^{x,y}_{n,d^\alpha}(\xi)}{d^\alpha(x,y)} d\mu(\xi)$$

is a finite sum, we make y tend to x to get

$$\int_{\Sigma_N^+} \left[(f^n_\xi)'(x)\right]^\alpha d\mu(\xi) \leq c \left(\lambda^{\frac{\alpha}{k}}\right)^n.$$

The proof of Proposition 3.2.1 is finished. $\square$

3.2.2 Some large deviation results

In this section, we first state a general large deviation results, which holds for a general metric space (in particular we do not require that it is bounded). We also establish $\mathrm{SA_{exp}}$ for IFSs on the circle.

Let $\mathcal{F}$ be an IFS on a metric space (M,d). Consider $\mathbf{p}$ a probability vector and its associated Bernoulli measure μ.

Proposition 3.2.3. *Assume that $(\mathcal{F},\mathbf{p},d)$ is SA_{exp}, that is, there exist constants $C>0$ and $\lambda\in(0,1)$ such that*

$$\mathbb{E}(Z^{x,y}_{n,d}) \le C\lambda^n \quad \textit{for every } x,y\in M \textit{ and } n\in\mathbb{N}.$$

For every $\varepsilon>0$ and $x,y\in M$ we have

$$\mu\left(\left\{\xi\in\Sigma_N^+\colon \frac{1}{n}\ln\, Z^{x,y}_{n,d}(\xi) > \ln\,\lambda+\varepsilon\right\}\right)\le Ce^{-\varepsilon n}.$$

Proof. Given $\varepsilon>0$, $x,y\in M$. Note that,

$$\begin{aligned}\mathbb{E}(Z^{x,y}_{n,d}) &\ge \int_{\left\{\xi\in\Sigma_N^+\colon Z^{x,y}_{n,d}(\xi)>(e^{\varepsilon}\lambda)^n\right\}} Z^{x,y}_{n,d}(\xi)d\mu(\xi)\\ &\ge (e^{\varepsilon}\lambda)^n\,\mu\left(\left\{\xi\in\Sigma_N^+\colon Z^{x,y}_{n,d}(\xi)>(e^{\varepsilon}\lambda)^n\right\}\right),\end{aligned}$$

so that, by hypothesis

$$\mu\left(\left\{\xi\in\Sigma_N^+\colon Z^{x,y}_{n,d}(\xi)>(e^{\varepsilon}\lambda)^n\right\}\right)\le Ce^{-\varepsilon n}.$$

This proposition is proved. □

Remark 3.2.4. Proposition 3.2.3 does not require (M,d) to be bounded and applies to the example studied in Section 2.4.1.

Corollary 3.2.5. *If $(\mathcal{F},\mathbf{p},d)$ is CA with contraction rate $\lambda\in(0,1)$ and (M,d) is bounded, then there exists $C>0$ such that for every $\varepsilon>0$ and $x,y\in M$ we have*

$$\mu\left(\left\{\xi\in\Sigma_N^+\colon \frac{1}{n}\ln\, Z^{x,y}_{n,d}(\xi) > \ln\,\lambda+\varepsilon\right\}\right)\le Ce^{-\varepsilon n}.$$

Proof. By Lemma 2.1.8, for every x,y

$$\mathbb{E}(Z^{x,y}_{n,d})\le\lambda^n d(x,y).$$

Now, using that d is bounded we conclude that $(\mathcal{F},\mathbf{p},d)$ is $\mathrm{SA_{exp}}$. Taking $C\overset{\text{def}}{=}\operatorname{diam}M$, Proposition 3.2.3 applies. □

Remark 3.2.6. The Corollary 3.2.5 applies to examples of Section 2.4 with the respective metric change.

In the context of IFSs on the circle, we have the following result, which slightly extends Proposition 3.2.1). It implies that the IFS is $\mathrm{SA}_{\mathrm{exp}}$ with respect to the original metric d on $\mathbb{S}^1$.

Proposition 3.2.7. *Assume that $\mathcal{F}$ satisfies hypotheses in Theorem C. For every probability vector $\mathbf{p}$, $(\mathcal{F}, \mathbf{p}, d)$ is SA_{exp}.*

Remark 3.2.8. Assuming the hypotheses in Proposition 3.2.7, by Proposition 2.2.1, for every $q \in (\lambda, 1)$

$$D(x,y) \stackrel{\text{def}}{=} \sum_{n\geq 0} \frac{q^n}{\lambda^n} \int_{\Sigma_N^+} Z_{n,d}^{x,y}(\xi) d\mu(\xi).$$

defines a metric on $\mathbb{S}^1$ which is equivalent to d such that $(\mathcal{F}, \mathbf{p}, D)$ is contracting on average with contraction rate $\lambda/q \in (0,1)$.

Proof of Proposition 3.2.7. By Theorem B and Lemma 2.1.8 there exist $\alpha \in (0,1)$ and $\lambda \in (0,1)$ such that for the metric D as in Theorem B, for $C = \sup_{x,y\in\mathbb{S}^1} D(x,y)$, for all $n \geq 1$ and for every $x, y \in \mathbb{S}^1$

$$\int_{\Sigma_N^+} Z_{n,d}^{x,y}(\xi) d\mu(\xi) \leq \int_{\Sigma_N^+} Z_{n,D}^{x,y}(\xi) d\mu(\xi) \leq \lambda^n D(x,y) \leq C\lambda^n,$$

therefore the desired result follows. $\square$

The following is a consequence of Proposition 3.2.3 and Proposition 3.2.7.

Corollary 3.2.9. *Assume that $\mathcal{F}$ is an IFS of C^1-diffeomorphisms on $\mathbb{S}^1$. If $(\mathcal{F}, \mathbf{p}, d)$ is proximal and there does not exist a probability measure which is invariant by every element of $\mathcal{F}$, then there exist $\lambda \in (0,1)$ and $c > 0$ such that for every $\varepsilon > 0$ and $x, y \in \mathbb{S}^1$ we have*

$$\mu\left(\left\{\xi \in \Sigma_N^+ \colon \frac{1}{n} \ln\, Z_{n,d}^{x,y}(\xi) > \ln\, \lambda + \varepsilon\right\}\right) \leq Ce^{-\varepsilon n}.$$

3.2.3 A Markov system for IFSs

In this section, we collect some preliminary results about Markov systems. They will be used to prove Theorem D. For more details about Markov systems see Appendix D.2.

Let d be the usual metric on $\mathbb{S}^1$. Fix $N \geq 2$ and define $\Lambda = \{0, \dots, N-1\}$. For $k \in \Lambda$, let $f_k : \mathbb{S}^1 \to \mathbb{S}^1$ be a C^1-diffeomorphism. Let $\mathcal{F} = \{f_0, \dots, f_{N-1}\}$. Let $L \geq 1$ be such that

$$L^{-1} d(x,y) \leq d(f_j(x), f_j(y)) \leq L d(x,y), \tag{3.2.3}$$

for all $j \in \Lambda$ and every $x, y \in \mathbb{S}^1$.

Let $\mathbf{p} = (p_0, \dots, p_{N-1})$ be a non-degenerate probability vector and let μ be the Bernoulli measure on $\Sigma_N^+ = \Lambda^{\mathbb{N}}$ determined by $\mathbf{p}$. Define the probability θ on Λ by

$$\theta \stackrel{\text{def}}{=} \sum_{i=0}^{N-1} p_i \delta_i,$$

where δ_i denotes the Dirac at i. Let $\sigma : \Sigma_N^+ \to \Sigma_N^+$ be the left shift map defined by

$$(\sigma(\xi))_j = \xi_{j+1}, \text{ for } j \geq 1. \tag{3.2.4}$$

Consider the Markov kernel by

$$K : \Lambda \times \mathbb{S}^1 \to \text{Prob}(\Lambda \times \mathbb{S}^1), \quad K_{j,x} \stackrel{\text{def}}{=} \int_\Lambda \delta_{(\xi_1, f_j(x))} d\theta(i) = \sum_{i=1}^{N-1} p_i \delta_{(i, f_j(x))}.$$

For a set X denote by $\mathscr{B}(x)$ the σ-algebra of Borel sets in X.

Lemma 3.2.10. *$(K, \theta \otimes \nu)$ is a Markov System on $\Lambda \times \mathbb{S}^1$.*

Proof. Let $A \times B \in \mathscr{B}(\Lambda) \times \mathscr{B}(\mathbb{S}^1)$, then

$$\begin{aligned}\int_{\Lambda \times \mathbb{S}^1} K_{j,x}(A \times B) d(\theta \otimes \nu)(j,x) &= \int_{\mathbb{S}^1} \int_\Lambda K_{j,x}(A \times B) d\theta(j) d\nu(x) \\ &= \int_{\mathbb{S}^1} \sum_{j=0}^{N-1} p_j K_{j,x}(A \times B) d\nu(x).\end{aligned}$$

Applying the definition of $K_{j,x}$ we get

$$\begin{aligned}\int_{\Lambda \times \mathbb{S}^1} K_{j,x}(A \times B) d\theta \otimes \nu(j,x) &= \int_{\mathbb{S}^1} \sum_{j=0}^{N-1} p_j \sum_{i=1}^{N-1} p_i \delta_i(A) \delta_{f_j(x)}(B) d\nu(x) \\ &= \sum_{j=0}^{N-1} p_j \sum_{i=1}^{N-1} p_i \delta_i(A) \int_{\mathbb{S}^1} \delta_{f_j(x)}(B) d\nu(x) \\ &= \sum_{i=1}^{N-1} p_i \delta_i(A) \sum_{j=0}^{N-1} p_j (f_j)_* \nu(B).\end{aligned}$$

Hence, the stationarity of ν implies the following

$$\int_{\Lambda \times \mathbb{S}^1} K_{j,x}(A \times B) d(\theta \otimes \nu)(j,x) = \sum_{i=1}^{N-1} p_i \delta_i(A) \nu(B) = \theta \otimes \nu(A \times B).$$

Since $\mathscr{B}(\Lambda) \times \mathscr{B}(\mathbb{S}^1)$ generates the product σ-algebra $\mathscr{B}(\Lambda \times \mathbb{S}^1)$, we conclude that for all $E \in \mathscr{B}(\Lambda \times \mathbb{S}^1)$

$$\int_{\Lambda \times \mathbb{S}^1} K_{j,x}(E) d(\theta \otimes \nu)(j,x) = (\theta \otimes \nu)(E),$$

which implies that $\theta \otimes \nu$ is K-stationary and hence $(K, \theta \otimes \nu)$ is a Markov System on $\Lambda \times \mathbb{S}^1$. $\square$

From now on consider $\alpha_0 \in (0,1)$ and $\lambda \in (0,1)$ be as in Proposition 3.2.1. Consider $\alpha \stackrel{\text{def}}{=} \min\{\beta, \alpha_0\}$. Then, the compactness of $\mathbb{S}^1$ implies that each map in $\mathcal{F}$ is $\mathcal{C}^{1+\alpha}$. Consider the space $L^\infty(\Lambda \times \mathbb{S}^1)$ of bounded measurable functions $\phi : \Lambda \times \mathbb{S}^1 \to \mathbb{C}$. Given a function $\phi \in L^\infty(\Lambda \times \mathbb{S}^1)$, define

$$\|\phi\|_\alpha \stackrel{\text{def}}{=} |\phi|_\alpha + \|\phi\|_\infty, \text{ where } |\phi|_\alpha \stackrel{\text{def}}{=} \sup_{j \in \Lambda, x \neq y} \frac{|\phi(j,x) - \phi(j,y)|}{d^\alpha(x,y)}, \tag{3.2.5}$$

and set

$$\mathcal{H}_\alpha(\Lambda \times \mathbb{S}^1) \stackrel{\text{def}}{=} \{\phi \in L^\infty(\Lambda \times \mathbb{S}^1) : \|\phi\|_\alpha < \infty\}.$$

It is not hard to show that $(\mathcal{H}_\alpha(\Lambda \times \mathbb{S}^1), \|\phi\|_\alpha)$ is a Banach algebra with unity.

Note that K determines the Markov operator $\mathcal{Q} : L^\infty(\Lambda \times \mathbb{S}^1) \to L^\infty(\Lambda \times \mathbb{S}^1)$ given by

$$\mathcal{Q}(\phi)(j,x) \stackrel{\text{def}}{=} \int_\Lambda \phi(i, f_j(x))d\theta(i) = \int_{\Sigma_N^+} \phi(\xi_1, f_j(x))d\theta(\xi_1) = \sum_{i=1}^{N-1} p_i\phi(i, f_j(x)).$$

Let us show a relation between the operator $\mathcal{Q}$ and the map F defined in (1.2.6). Define $\pi : \Sigma_N^+ \times \mathbb{S}^1 \to \Lambda \times \mathbb{S}^1$ as the projection given by $\pi(\xi, x) \stackrel{\text{def}}{=} (\xi_1, x)$.

Lemma 3.2.11. *For any function $\phi \in L^\infty(\Lambda \times \mathbb{S}^1)$ and for all $n \geq 1$ we have*

$$(\mathcal{Q}^n(\phi))(j,x) = \int_{\Sigma_N^+} \phi \circ \pi \circ F^{n-1}(\xi, f_j(x))d\mu(\xi). \tag{3.2.6}$$

Proof. The case $n = 1$ is clear. By induction on n, suppose that (3.2.6) holds for $n = k$ and let us prove that (3.2.6) holds for $n = k + 1$. Indeed, using the inductive hypothesis we get

$$\begin{aligned}
(\mathcal{Q}^{k+1}(\phi))(j,x) &= \int_{\Sigma_N^+} \mathcal{Q}^k(\phi)(\xi_1, f_j(x))d\mu(\xi) \\
&= \int_{\Sigma_N^+}\int_{\Sigma_N^+} \phi \circ \pi \circ F^{k-1}(\eta, f_{\xi_1}(f_j(x)))d\mu(\eta)d\mu(\xi) \\
&= \sum_{i=0}^{N-1} p_i \int_{\Sigma_N^+} \phi \circ \pi \circ F^{k-1}(\eta, f_i(f_j(x)))d\mu(\eta).
\end{aligned}$$

Since μ is σ-invariant, we get

$$(\mathcal{Q}^{k+1}(\phi))(j,x) = \sum_{i=0}^{N-1} p_i \int_{\Sigma_N^+} \phi \circ \pi \circ F^{k-1}(\sigma(\eta), f_i(f_j(x)))d\mu(\sigma(\eta)).$$

Note that $\sigma(\eta) = \sigma(i, \eta_2, \eta_3, \ldots, \eta_n, \ldots)$ for all $i \in \Lambda$. Hence,

$$\begin{aligned}
(\mathcal{Q}^{k+1}(\phi))(j,x) &= \sum_{\eta_1=0}^{N-1} p_{\eta_1} \int_{\Sigma_N^+} \phi \circ \pi \circ F^{k-1}(\sigma(\eta), f_{\eta_1}(f_j(x)))d\mu(\sigma(\eta)) \\
&= \int_{\Sigma_N^+} \phi \circ \pi \circ F^{k-1}(\sigma(\eta), f_{\eta_1}(f_j(x)))d\mu(\eta) \\
&= \int_{\Sigma_N^+} \phi \circ \pi \circ F^k(\eta, f_j(x))d\mu(\eta),
\end{aligned}$$

which concludes the proof. $\square$

Proposition 3.2.12. *The Markov operator $\mathcal{Q}$ acts simply and quasi-compactly on $\mathcal{H}_\alpha(\Lambda \times \mathbb{S}^1)$ with stationary measure $\theta \otimes \nu$.*

Proof. To prove this proposition, we need to show that there are constants $C > 0$ and $\gamma \in (0,1)$ such that for all $\phi \in \mathcal{H}_\alpha(\Lambda \times \mathbb{S}^1)$ and all $n \geq 0$

$$\left\| \mathcal{Q}_K^n \phi - \left(\int_M \phi d\theta \right) \mathbf{1} \right\|_\alpha \leq C\sigma^n \|\phi\|_\alpha.$$

We split the proof into two parts. First, we show that

$$|\mathcal{Q}^n \phi|_\alpha = \left| \mathcal{Q}^n \phi - \left(\int \phi d(\theta \otimes \nu) \right) \mathbf{1} \right|_\alpha$$

is bounded above by $C'\lambda^n|\phi|_\alpha$. Second, we show that the norm of $\mathcal{Q}^n\phi - (\int \phi d(\theta \otimes \nu))\,\mathbf{1}$ is also bounded above by $C''\lambda^n|\phi|_\alpha$, which allows us to conclude that is norm in $\mathcal{H}_\alpha(\Lambda, \mathbb{S}^1)$ is also bounded above by $C\lambda^n|\phi|_\alpha$, for some constant $C > 0$. Therefore, the Markov operator $\mathcal{Q}$ acts simply and quasi-compactly on $\mathcal{H}_\alpha(\Lambda \times \mathbb{S}^1)$.

Claim 3.2.13. *There exist $C > 0$ and $r \in (0,1)$ such that for all $\phi \in L^\infty(\Lambda \times \mathbb{S}^1)$ and $n \in \mathbb{N}$*

$$|\mathcal{Q}^n \phi|_\alpha \leq Cr^n |\phi|_\alpha.$$

Proof. Recall that $\pi(\xi, x) = (\xi_1, x)$. By Lemma 3.2.11, together with (3.2.3) we get

$$\begin{aligned}
|\mathcal{Q}^n\phi|_\alpha &= \sup_{j\in\Lambda, x\neq y} \left| \int_{\Sigma_N^+} \frac{(\phi\circ\pi\circ F^{n-1}(\xi, f_j(x)) - \phi\circ\pi\circ F^{n-1}(\xi, f_j(y)))}{d^\alpha(x,y)} d\mu(\xi) \right| \\
&\leq \sup_{j\in\Lambda, x\neq y} \int_{\Sigma_N^+} \frac{\left|\phi\circ\pi(\sigma^{n-1}(\xi), f_\xi^{n-1}\circ f_j(x)) - \phi\circ\pi(\sigma^{n-1}(\xi), f_\xi^{n-1}\circ f_j(y))\right|}{d^\alpha(x,y)} d\mu(\xi),
\end{aligned}$$

Multiplying and dividing the integrand by $d^\alpha\left(f_\xi^{n-1}\circ f_j(x), f_\xi^{n-1}\circ f_j(y)\right)$ we obtain

$$\begin{aligned}
|\mathcal{Q}^n\phi|_\alpha &\leq |\phi|_\alpha \left(\sup_{j\in\Lambda, x\neq y} \int_{\Sigma_N^+} \frac{d^\alpha\left(f_\xi^{n-1}\circ f_j(x), f_\xi^{n-1}\circ f_j(y)\right)}{d^\alpha\left(f_j(x), f_j(y)\right)} \frac{d^\alpha\left(f_j(x), f_j(y)\right)}{d^\alpha(x,y)} d\mu(\xi) \right) \\
&\leq L^\alpha |\phi|_\alpha \left(\sup_{j\in\Lambda, x\neq y} \int_{\Sigma_N^+} \frac{d^\alpha\left(f_\xi^{n-1}\circ f_j(x), f_\xi^{n-1}\circ f_j(y)\right)}{d^\alpha\left(f_j(x), f_j(y)\right)} d\mu(\xi) \right),
\end{aligned}$$

where L is as in (3.2.3). Using Proposition 3.2.1, we get

$$|\mathcal{Q}^n\phi|_\alpha \leq L^\alpha c \lambda^{\alpha(n-1)/k} |\phi|_\alpha.$$

Letting $C \stackrel{\text{def}}{=} L^\alpha c \lambda^{-\alpha/k}$ and $r \stackrel{\text{def}}{=} \lambda^{\alpha/k}$, this proves the claim. □

Claim 3.2.14. *There exist $C > 0$ and $r \in (0,1)$ such that for all $\phi \in L^\infty(\Lambda \times \mathbb{S}^1)$ and $n \in \mathbb{N}$*

$$\left\| \mathcal{Q}^n \phi - \left(\int_{\Lambda\times\mathbb{S}^1} \phi(j,x) d(\theta\otimes\nu)(j,x) \right) \mathbf{1} \right\|_\infty \leq Cr^n |\phi|_\alpha.$$

Proof. Let $\phi \in L^\infty(\Lambda \times \mathbb{S}^1)$. As ν is stationary, $\mu \otimes \nu$ is F-invariant. Thus,

$$\int_{\Sigma_N^+\times\mathbb{S}^1} \phi\circ\pi(\xi,x) d(\mu\otimes\nu)(\xi,x) = \int_{\Sigma_N^+\times\mathbb{S}^1} \phi\circ\pi\circ F^{n-1}(\xi,x) d(\mu\otimes\nu)(\xi,x)$$

so that, using Lemma 3.2.11

$$\begin{aligned}
&\left\| \mathcal{Q}^n\phi - \Big(\int_{\Lambda\times\mathbb{S}^1} \phi(j,x)d(\theta\otimes\nu)(j,x)\Big)\mathbf{1}\right\|_\infty \\
&= \sup_{i\in\Lambda, y\in\mathbb{S}^1} \left| \mathcal{Q}^n\phi(i,y) - \Big(\int_{\Sigma_N^+\times\mathbb{S}^1} \phi\circ\pi(\xi,x)d(\mu\otimes\nu)(\xi,x)\Big)\right| \\
&= \sup_{i\in\Lambda, y\in\mathbb{S}^1} \left| \int_{\Sigma_N^+} \phi\circ\pi\circ F^{n-1}(\xi, f_i(y))d\mu(\xi) - \Big(\int_{\Sigma_N^+\times\mathbb{S}^1} \phi\circ\pi\circ F^{n-1}(\xi,x)d(\mu\otimes\nu)(\xi,x)\Big)\right|.
\end{aligned}$$

By definition of $|\phi|_\alpha$ and Proposition 3.2.1, we get

$$\begin{aligned}
&\left\| \mathcal{Q}^n\phi - \Big(\int_{\Lambda\times\mathbb{S}^1} \phi(j,x)d(\theta\otimes\nu)(j,x)\Big)\mathbf{1}\right\|_\infty \\
&\leq \sup_{j\in\Lambda, y\in\mathbb{S}^1} \int_{\mathbb{S}^1}\int_{\Sigma_N^+} \left|\phi\circ\pi\circ F^{n-1}(\xi, f_j(y)) - \phi\circ\pi\circ F^{n-1}(\xi,x)\right| d\mu(\xi)d\nu(x) \\
&\leq |\phi|_\alpha \left(\sup_{j\in\Lambda, y\in\mathbb{S}^1} \int_{\mathbb{S}^1}\int_{\Sigma_N^+} d^\alpha\left(f_\xi^{n-1}(f_j(y)), f_\xi^{n-1}(x)\right) d\mu(\xi)d\nu(x)\right) \\
&\leq c\lambda^{\alpha(n-1)/k}|\phi|_\alpha \left(\sup_{j\in\Lambda, y\in\mathbb{S}^1} \int_{\mathbb{S}^1} d^\alpha\left(f_j(y), x\right) d\nu(x)\right).
\end{aligned}$$

Letting $C = c\lambda^{-\alpha/k}|\phi|_\alpha$ and $r = \lambda^{\alpha/k}$ proves the claim. □

Recalling the definition of $\|\cdot\|_\alpha$ in (3.2.5), it follows from Claim 3.2.13 and Claim 3.2.14 that there exist $C > 0$ and $r \in (0,1)$ so that

$$\|Q^n\phi - \Big(\int \phi\, d\theta\Big)\mathbf{1}\|_\alpha \leq 2Cr^n|\phi|_\alpha \leq 2Cr^n\|\phi\|_\alpha.$$

This proves the proposition. □

3.2.4 Proof of Theorem D

Proof of Theorem D (1.) We will apply the Proposition D.3.1 to the function $\phi : \Lambda\times\mathbb{S}^1 \to \mathbb{R}$ given by

$$\phi(j,x) \stackrel{\text{def}}{=} \ln|(f_j)'(x)|. \tag{3.2.7}$$

Let us start by establishing a appropriate space of functions where ϕ belongs. Let ϕ be as in (3.2.7). Consider $k \in \mathbb{N}$ and $\alpha_0 \in (0,1)$ as in Proposition 2.3.10, such that (3.2.1) satisfies for $\alpha \in (0, \alpha_0]$. Fix $\alpha \in (0, \min\{\beta, \alpha_0\}]$. Since every f_{ξ_1} is $\mathcal{C}^{1+\beta}$-diffeomorphism is also $\mathcal{C}^{1+\alpha}$-diffeomorphism. Hence, $\phi \in \mathcal{H}_\alpha(\Lambda\times\mathbb{S}^1)$.

We are now in the position to apply Proposition D.3.1 to the Markov system $(K, \theta\otimes\nu)$ and the function ϕ. It guarantees that there exist constants $\varepsilon_0, c, h > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$, $(j,x) \in \Lambda\times\mathbb{S}^1$ and $n \in \mathbb{N}$,

$$\mu\left(\left\{\xi\in\Sigma_N^+ : \xi_1 = j, \quad \left|\ln|(f_\xi^n)'(x)| - n\gamma\right| > n\varepsilon\right\}\right) \leq ce^{-nh\varepsilon^2}.$$

Averaging in $j \in \Lambda$ with respect to θ, for every $x \in \mathbb{S}^1$ we get

$$\mu\left(\left\{\xi \in \Sigma_N^+ : \left|\ln|(f_\xi^n)'(x)| - n\gamma\right| > n\varepsilon\right\}\right) \leq ce^{-nh\varepsilon^2}.$$

This proves the first assertion of Theorem D. $\square$

Proof of Theorem D (2.) Let $x, y \in \mathbb{S}^1$. For $\xi \in \Omega^{x,y}$, by mean value theorem and the triangle inequality

$$\left|\ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} - n\gamma\right| \leq \ln\left(\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right|\right) + |\ln|(f_\xi^n)'(x)| - n\gamma|.$$

Therefore,

$$\mu\left(\xi : \left|\ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} - n\gamma\right| > n\varepsilon\right) \tag{3.2.8}$$
$$\leq \mu\left(\xi : \ln\left(\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right|\right) > \frac{n\varepsilon}{2}\right) + \mu\left(\xi : |\ln|(f_\xi^n)'(x)| - n\gamma| > \frac{n\varepsilon}{2}\right).$$

Let us limit the first term of the right-hand sum in (3.2.8). Note that,

$$\mu\left(\xi : \ln\left(\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right|\right) > \frac{n\varepsilon}{2}\right) = \mu\left(\xi : \max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right| > e^{\frac{n\varepsilon}{2}}\right).$$

It follows from Chebyshev's inequality, that

$$\mu\left(\xi : \ln\left(\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right|\right) > \frac{n\varepsilon}{2}\right) \leq e^{-n\varepsilon/2}\int_{\Sigma_N^+}\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right| d\mu(x). \tag{3.2.9}$$

From (3.2.12) we have

$$\lim_{n\to\infty}\frac{1}{n}\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right| = 1,$$

so that dominated convergence theorem implies

$$\lim_{n\to\infty}\frac{1}{n}\int_{\Sigma_N^+}\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right| d\mu(\xi) = 1.$$

Consequently, there exists a constant $\hat{c} > 0$ (depending on x and y) such that for all $n \geq 1$

$$\int_{\Sigma_N^+}\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right| d\mu(\xi) \leq \hat{c}n,$$

and so from (3.2.9) we obtain

$$\mu\left(\xi : \ln\left(\max_{z,w\in I_\xi}\left|\frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)}\right|\right) > \frac{n\varepsilon}{2}\right) \leq \hat{c}ne^{-n\varepsilon/2}. \tag{3.2.10}$$

Now, for the second term of the sum on the right-hand sum in (3.2.8) apply Theorem D to obtain that there exists $c, h, \varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$

$$\mu\left(\xi : |\ln|(f_\xi^n)'(x)| - n\gamma| > \frac{n\varepsilon}{2}\right) \leq ce^{-nh\varepsilon^2/4}. \tag{3.2.11}$$

By (3.2.8), (3.2.10) and (3.2.11), for all $n \geq 1$

$$\mu\left(\xi : \left|\ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} - n\gamma\right| > n\varepsilon\right) \leq \hat{c}ne^{-n\varepsilon/2} + ce^{-nh\varepsilon^2/4},$$

with which we conclude this proof. $\square$

3.2.5 Application of Theorem D

In this section, we discuss some consequences of Theorem D. Let us start with the following result:

Theorem 3.2.15. *Assume that $\mathcal{F}$ is an IFS of $C^{1+\beta}$ diffeomorphisms, for some $\beta > 0$, on $\mathbb{S}^1$ so that $(\mathcal{F}, d)$ is proximal and there does not exist a probability measure which is invariant by every element of $\mathcal{F}$. Then, for every non-degenerate probability vector $\mathbf{p}$ and the Bernoulli measure μ determined by $\mathbf{p}$, the following hold*

1. *for every $x \in \mathbb{S}^1$,*
$$\lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_N^+} \ln |(f_\xi^n)'(x)| d\mu(\xi) = \gamma < 0,$$
2. *and for every $x, y \in \mathbb{S}^1$,*
$$\lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_N^+} \ln Z_{n,d}^{x,y}(\xi) d\mu(\xi) = \lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_N^+} \ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} d\mu(\xi) = \gamma.$$

Remark 3.2.16. As we discussed in Chapter 1, we have that
$$\gamma = \lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_N^+ \times \mathbb{S}^1} \ln |(f_\xi^n)'(x)| d(\mu \otimes \nu)(\xi, x).$$
Note that item *(1.)* in Theorem 3.2.15 improves this result slightly. Analogously, we have that the process $(\ln Z_{n,d}^{x,y}(\xi))_{n\in\mathbb{N}}$ is additive, in the following sense
$$\ln \frac{Z_{n+m,d}^{x,y}(\xi)}{d(x,y)} = \ln \frac{Z_{m,d}^{f_\xi^n(x), f_\xi^n(y)}(\sigma^n(\xi))}{d(f_\xi^n(x), f_\xi^n(y))} + \ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)}.$$
Therefore, by Birkhoff's Ergodic Theorem, we get
$$\lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_N^+} \ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} d(\mu \otimes \nu \otimes \nu)(\xi, x, y) = \gamma.$$
As before, item *(2.)* in Theorem 3.2.15 slightly improves this result.

Proof Theorem 3.2.15 (1). Let $\varepsilon_0, h, c > 0$ be as in Theorem D. For $\varepsilon \in (0, \varepsilon_0)$ define
$$\Lambda_\varepsilon \stackrel{\text{def}}{=} \left\{\xi \in \Sigma_N^+ : \gamma - \varepsilon \le \liminf_{n\to\infty} \frac{1}{n} \ln |(f_\xi^n)'(x)| \le \limsup_{n\to\infty} \frac{1}{n} \ln |(f_\xi^n)'(x)| \le \gamma + \varepsilon\right\}.$$
For all $n \in \mathbb{N}$, by Theorem D we have
$$\mu\left(\xi \in \Sigma_N^+ : \left|\ln |(f_\xi^n)'(x)| - n\gamma\right| > n\varepsilon\right) \le ce^{-nh\varepsilon^2}.$$
The Borel–Cantelli lemma implies $\mu(\Lambda_\varepsilon) = 1$, for all $\varepsilon \in (0, \varepsilon_0)$. Then, for $k \ge 1$ and $\varepsilon_k \in (0, \varepsilon_0)$ such that $\varepsilon_k \to 0$ as $k \to \infty$, we have that
$$\mu\left(\cap_{k=1}^{\infty} \Lambda_{\varepsilon_k}\right) = 1.$$

Since

$$\cap_{k=1}^{\infty} \Lambda_{\varepsilon_k} = \{\xi \in \Sigma_N^+ : \lim_{n\to\infty} \frac{1}{n} \ln |(f_\xi^n)'(x)| = \gamma\}$$

we conclude that for almost every $\xi \in \Sigma_N^+$

$$\lim_{n\to\infty} \frac{1}{n} \ln |(f_\xi^n)'(x)| = \gamma.$$

Note that for every $n \in \mathbb{N}$ and $\xi \in \Sigma_N^+$ we have

$$-\ln L \leq \frac{1}{n} \ln |(f_\xi^n)'(x)| \leq \ln L,$$

where $L \geq 1$ are as in (3.2.3). This proof ends by applying the Dominated Convergence Theorem. $\square$

Now, before proving *(2)* of the Theorem 3.2.15 let us establish some general properties of the system. For every $x, y \in \mathbb{S}^1$ let $\Omega^{x,y}$ as in (2.3.1). Since $(\mathcal{F}, \mathbf{p}, d)$ is proximal, by Proposition 2.3.1 we get that for every $x, y \in \mathbb{S}^1$

$$\mu(\Omega^{x,y}) = 1.$$

If $x \neq y$, for $\xi \in \Omega^{x,y}$ define

$$I_\xi \stackrel{\text{def}}{=} \begin{cases} [x,y], & \text{if} \quad \lim_{n\to\infty} \left| f_\xi^n([x,y] \right| = 0 \\ [y,x], & \text{if} \quad \lim_{n\to\infty} \left| f_\xi^n([y,x] \right| = 0 \end{cases} \qquad \text{and} \qquad \delta_n(\xi) \stackrel{\text{def}}{=} |f_\xi^n(I_\xi)|.$$

Consider the following modulus of continuity

$$\omega(\delta) \stackrel{\text{def}}{=} \max_{i\in\Lambda} \max_{d(z,w)\leq\delta} |\ln |f_i'(z)| - \ln |f_i'(w)||$$

and note that $\lim_{\delta\to 0^+} \omega(\delta) = 0$.

In this context let us prove the following lemma:

Lemma 3.2.17.

$$\lim_{n\to\infty} \frac{1}{n} \int_{\Sigma_N^+} \ln \left(\max_{z,w\in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| \right) d\mu(\xi) = 0.$$

Proof. For $\xi \in \Omega^{x,y}$ we have that

$$\ln \left(\max_{z,w\in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| \right) \leq \sum_{k=0}^{n-1} \omega(\delta_k(\xi)).$$

Recall that $\xi \in \Omega^{x,y}$ implies that $\delta_k(\xi) \to 0$, as $k \to \infty$, and hence

$$\lim_{k\to\infty} \omega(\delta_k(\xi)) = 0.$$

It follows that

$$\lim_{n\to\infty} \frac{1}{n} \ln \left(\max_{z,w\in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| \right) = 0. \qquad (3.2.12)$$

Since $\mu(\Omega^{x,y}) = 1$ and

$$0 \le \frac{1}{n} \ln \left(\max_{z,w \in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| \right) \le 2 \ln L,$$

the dominated convergence theorem implies the claim. □

Proof of Theorem 3.2.15 (2). Let $x, y \in \mathbb{S}^1$, $x \neq y$. By the mean value inequality, for every $x, y \in \mathbb{S}^1$, $x \neq y$, and $\xi \in \Omega^{x,y}$

$$\min_{z,w \in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| |(f_\xi^n)'(x)| \le \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} \le \max_{z,w \in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| |(f_\xi^n)'(x)|.$$

Notice that

$$\min_{z,w \in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| = \left(\max_{z,w \in I_\xi} \left| \frac{(f_\xi^n)'(z)}{(f_\xi^n)'(w)} \right| \right)^{-1}.$$

Hence, by Lemma 3.2.17, for almost every $\xi \in \Sigma_N^+$

$$\lim_{n \to \infty} \frac{1}{n} \ln Z_{n,d}^{x,y}(\xi) = \lim_{n \to \infty} \frac{1}{n} \ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} = \lim_{n \to \infty} \frac{1}{n} \ln |(f_\xi^n)'(x)| = \gamma.$$

Applying dominated convergence theorem, we conclude that

$$\lim_{n \to \infty} \frac{1}{n} \int_{\Sigma_N^+} \ln Z_{n,d}^{x,y}(\xi) d\mu(\xi) = \lim_{n \to \infty} \frac{1}{n} \int_{\Sigma_N^+} \ln \frac{Z_{n,d}^{x,y}(\xi)}{d(x,y)} d\mu(\xi) = \gamma.$$

This proves the theorem. □

4 Matrix cocycles

4.1 The twin measures

As there is a relation between Lyapunov exponents of the cocycles and the associated skew-product, analogously for invariant measures for cocycles and the skew-product and their averages. In this section, we explore this relation and, in particular, we prove the equality of the pressure functions (1.3.4) and (1.3.9).

Let us first establish a relation between the Lyapunov exponents. By [DGR19, Theorem 11.1] for every $\alpha > 0$ and for ξ it holds $\lambda(\mathbb{M}, \xi) = \alpha$ if only if there are $v, w \in \mathbb{P}^1$ such that

$$2\alpha = \chi(\mathbb{M}, \xi, v) = -\chi(\mathbb{M}, \xi, w). \tag{4.1.1}$$

By Oseledets' Theorem (for example, for context of cocycles see [Via14, Theorem 3.14]), for $\xi \in \Sigma_N^+$ such that $\lambda(\mathbb{M}, \xi) = \alpha$ there exist two unique non-collinear vectors v_ξ^+ and v_ξ^- in $\mathbb{P}^1$ satisfying

1. $f_{M_{\xi_1}}(v_\xi^\pm) = v_{\sigma(\xi)}^\pm$
2. $\chi(\mathbb{M}, \xi, v_\xi^+) = \chi(\mathbb{M}^{-1}, \xi, v_\xi^-) = 2\alpha$ and $\chi(\mathbb{M}, \xi, v_\xi^-) = -2\alpha$,

where $\mathbb{M}^{-1} = \{M_0^{-1}, \dots, M_{N-1}^{-1}\}$. Let $L_\xi^\pm$ be the sub-space generated by $v^\pm$, respectively. The decomposition $\mathbb{R}^2 = L_\xi^+ \oplus L_\xi^-$ is known as *Oseledets decomposition* [Ose68].

The following result is proved in [Led84, Proposition 5.1], below we give an alternative proof. Given $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$, we denote by $\mathcal{M}(\nu)$ the set of the F-ergodic measures μ such that $\pi_*\mu = \nu$.

Lemma 4.1.1. *For every $\nu \in \mathcal{M}_{erg}(\Sigma_N^+)$ with $\mathbb{M}_*(\nu) > 0$, there exist two unique F-ergodic measures $\mu^\pm$ satisfying $\pi_*\mu^\pm = \nu$ and*

$$2\mathbb{M}_*(\nu) = \varphi(\mu^+) = -\varphi(\mu^-) = \sup_{\mu \in \mathcal{M}(\nu)} \varphi(\mu).$$

Moreover, μ^+ and μ^- are the only elements in $\mathcal{M}(\nu)$.

Proof. Let us first prove the existence of μ^+, the existence of μ^- is demonstrated analogously. Let ξ be ν-generic. By (4.1.1), there exists $v \in \mathbb{P}^1$ such that

$$\alpha \stackrel{\text{def}}{=} \mathbb{M}_*(\nu) = \lambda(\mathbb{M}, \xi) = \frac{1}{2}\chi(\mathbb{M}, \xi, v).$$

Consider the probability measures

$$\mu_n \stackrel{\text{def}}{=} \frac{1}{n}\Big(\delta_{\xi,v} + F_*\delta_{\xi,v} + \ldots + (F^{n-1})_*\delta_{\xi,v}\Big)$$

and let μ be some weak* accumulation point of the sequence $(\mu_n)_n$. Note that

$$2\alpha = \chi(\mathbb{M}, \xi, v) = \lim_{n\to\infty} \varphi(\mu_n) = \varphi(\mu),$$

where the last equality follows from weak* convergence and the continuity of φ. It is immediate to check that μ is F-invariant Borel probability measure. Moreover, $\pi_*\mu = \nu$. Consider its ergodic decomposition $\mu = \int \mu_\theta \, d\tau(\theta)$ into F-ergodic measures μ_θ. Note that τ-almost every ergodic component satisfies $\pi_*\mu_\theta = \nu$. Let us check that additionally, it holds $\varphi(\mu_\theta) = 2\alpha$.

Indeed, consider an ergodic component μ_θ. Then μ_θ-almost every point (η, w) is μ_θ-generic, and in particular it holds $\chi(\mathbb{M}, \eta, w) = \varphi(\mu_\theta)$. Moreover, it holds that η is ν-generic and $\lambda(\mathbb{M}, \eta)$ is well defined. Moreover, by (4.1.1),

$$\alpha = \lambda(\mathbb{M}, \eta) = \frac{1}{2} \left|\chi(\mathbb{M}, \eta, w)\right| = \frac{1}{2} \left|\varphi(\mu_\theta)\right|.$$

As

$$2\alpha = \varphi(\mu) = \int \varphi(\mu_\theta) \, d\tau(\theta)$$

it follows that for τ-almost every θ, $\varphi(\mu_\theta) = 2\alpha$.

This proves the existence of μ^+ and the inequality $2\mathbb{M}_*(\nu) \le \sup_\mu \varphi(\mu)$. To prove the opposite inequality, take $\mu \in \mathcal{M}(\nu)$. Let (ξ, v) be μ-generic. Without loss of generality, we can assume that ξ is ν-generic and $\lambda(\mathbb{M}, \xi)$ is well defined and equal to $\mathbb{M}_*(\nu)$. Check that

$$\varphi(\mu) = \chi(\mathbb{M}, \xi, v) = \lim_{n\to\infty} \frac{1}{n} \ln |(f_\xi^n)'(v)| \le \lim_{n\to\infty} \frac{2}{n} \ln \|\mathbb{M}^n(\xi)\| \le 2\mathbb{M}_*(\nu). \tag{4.1.2}$$

This finishes the proof of the three equalities.

Let us finally prove the uniqueness of μ^+ (for μ^- it is analogous). The following was shown in [Led84, Ch. I, sect. 5], see also [Via14, Theorem 6.1]. If $\mu \in \mathcal{M}(\nu)$ is a F-ergodic measure, then

$$\mu\left(\{(\xi, v_\xi^+) \colon \lambda(\mathbb{M}, \xi) = \alpha\}\right) = 1, \quad \text{and} \quad \varphi(\mu) = 2\alpha,$$

or

$$\mu\left(\{(\xi, v_\xi^-) \colon \lambda(\mathbb{M}, \xi) = \alpha\}\right) = 1, \quad \text{and} \quad \varphi(\mu) = -2\alpha,$$

where $v_\xi^\pm$ are as in Section 4.2.1.

Let μ_1^+ and μ_2^+ be F-ergodic measures such that $\pi_*\mu_i^+ = \nu$ and $2\alpha = \varphi(\mu_i^+)$, for $i = 1, 2$. Let $G_1, G_2 \subset \{(\xi, v_\xi^+) \colon \lambda(\mathbb{M}, \xi) = \alpha\}$ be F-invariant sets such that $\mu_i^+(G_i) = 1$, for $i = 1, 2$. As was discussed in Section 4.2.1, we have that $f_{M_{\xi_1}}(v_\xi^+) = v_{\sigma(\xi)}^+$, so that the projections $\pi(G_1)$ and $\pi(G_2)$ are σ-invariant sets. Then, $\nu(\pi(G_1)) = \nu(\pi(G_2)) = 1$. Hence, for ν-almost every ξ we have that $(\xi, v_\xi^+) \in G_1 \cap G_2$ and so (by ergodicity) $\mu_1^+ = \mu_2^+$. This ends the proof of the Lemma. □

Remark 4.1.2. Assume that ν is a Bernoulli measure on Σ_N^+. By [BL85, Part A, Ch. II, sec. 4, Theorem 4.4], if there is no distribution m on $\mathbb{P}^1$ simultaneously invariant for all matrices in $\mathbb{M}$, then $\mathbb{M}_*(\nu) > 0$ and there exists a unique ν-stationary[1] distribution for $\mathbb{M}$.

Remark 4.1.3. Let ν be a Bernoulli measure on Σ_N^+. By [Via14, Lemma 5.25], if $\mathbb{M}_*(\nu) > 0$ then μ^+ and μ^- in Lemma 4.1.1 are unique. When $\mathbb{M}_*(\nu) = 0$, [Via14, Linear invariance principle] implies that every $\mu \in \mathcal{M}(\nu)$ there exists a distribution m on $\mathbb{P}^1$ which is both ν-stationary distribution for $\mathbb{M}$ and ν-stationary distribution for $\mathbb{M}^{-1} = \{M_0^{-1}, \dots, M_{N-1}^{-1}\}$ such that $\mu = \nu \otimes m$.

Now let us show the relationship between the pressure functions for the cocycle in (1.3.4) and for the skew-product in (1.3.9).

Proposition 4.1.4. *For every $q \in \mathbb{R}$, we have*

$$P_{var}(|q|) = P_F(q).$$

Proof. Given $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$, we have $h_\mu(F) = h_\nu(\sigma)$ for all $\mu \in \mathcal{M}(\nu)$ (see Appendix C.2). For $\mu^\pm$ as in Lemma 4.1.1 and $q \geq 0$ we have

$$\sup_{\mu \in \mathcal{M}(\nu)} \Big(h_\mu(F) + \frac{q}{2}\varphi(\mu) \Big) = h_\nu(\sigma) + \frac{q}{2}\varphi(\mu^+) = h_\nu(\sigma) + q\mathbb{M}_*(\nu),$$

and for $q < 0$

$$\begin{aligned}\sup_{\mu \in \mathcal{M}(\nu)} \Big(h_\mu(F) + \frac{q}{2}\varphi(\mu) \Big) &= h_\nu(\sigma) + \frac{q}{2}\varphi(\mu^-) \\ &= h_\nu(\sigma) - q\mathbb{M}_*(\nu) \\ &= h_\nu(\sigma) + |q|\mathbb{M}_*(\nu).\end{aligned}$$

Now, taking supremum over $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$ we conclude this result. □

4.2 Properties of the cocycle

4.2.1 The level sets

Given $\alpha \geq 0$, consider the following level set of Lyapunov exponents of the cocycle

$$E(\mathbb{M}, \alpha) \stackrel{\text{def}}{=} \{\xi \in \Sigma_N^+ \colon \lambda(\mathbb{M}, \xi) = \alpha\}.$$

We also study in parallel the exponents for the induced skew-product. Analogously, we consider the following level set

$$\mathcal{E}(F, 2\alpha) \stackrel{\text{def}}{=} \{(\xi, v) \colon \chi(\mathbb{M}, \xi, v) = 2\alpha\}. \tag{4.2.1}$$

[1] Analogous to the context of IFSs on the circle defined in (1.2.1). In the context of cocycles, a probability measure m on $\mathbb{P}^1$ is ν-stationary for $\mathbb{M}$ if $m(B) = \int m(M_{\xi_1}^{-1}B)d\nu(\xi)$ for every measurable set $B \subset \mathbb{P}^1$.

The following variational principle was proved in [DGR19, Lemma 5.2]. Let us rewrite it in terms of the cocycle.

Lemma 4.2.1. *Let $\mathbb{M} \in \mathfrak{E}_{N,shyp}$. For $\alpha \geq 0$ such that $E(\mathbb{M},\alpha) \neq \emptyset$ we have*

$$h_{top}(\sigma, E(\mathbb{M},\alpha)) = \sup\{h_\nu(\sigma)\colon \nu \in \mathcal{M}_{erg}(\Sigma_N^+), \mathbb{M}_*(\nu) = \alpha\}.$$

Proof. Given $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$, note that $\mathbb{M}_*(\nu) = \alpha$ if and only if $\nu(E(\mathbb{M},\alpha)) = 1$. Then, by [Bow73, Theorem 1] we get

$$\sup\{h_\nu(\sigma)\colon \nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+), \mathbb{M}_*(\nu) = \alpha\} \leq h_{\text{top}}(\sigma, E(\mathbb{M},\alpha)).$$

On the other hand, by [DGR19, Theorem A] (in the case $\alpha > 0$) and [DGR, Theorem A] (in the case $\alpha = 0$), the following restricted variational principle holds for every $\alpha > 0$:

$$h_{\text{top}}(F, \mathcal{E}(F, 2\alpha)) = \sup\Big\{h_\mu(F)\colon \mu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+ \times \mathbb{P}^1), \int \ln|(f'_{\xi_1})(v)|\, d\mu(\xi, v) = 2\alpha\Big\}.$$

where $\mathcal{E}(F, 2\alpha)$ is as in (4.2.1). From (4.1.1) we have that $(\xi, v) \in \mathcal{E}(F, 2\alpha)$ for some $v \in \mathbb{P}^1$, if and only if $\xi \in E(\mathbb{M},\alpha)$. Therefore,

$$E(\mathbb{M},\alpha) = \pi(\mathcal{E}(F, 2\alpha)).$$

Hence,

$$h_{\text{top}}(\sigma, E(\mathbb{M},\alpha)) = h_{\text{top}}(\sigma, \pi(\mathcal{E}(F, 2\alpha))) \leq h_{\text{top}}(F, \mathcal{E}(F, 2\alpha)).$$

Given $\mu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+ \times \mathbb{P}^1)$ satisfying $\int \ln|(f'_{\xi_1})(v)|\, d\mu(\xi, v) = 2\alpha$, $\nu \overset{\text{def}}{=} \pi_*\mu$ is ergodic and, again by Remark 4.2.1, $\mathbb{M}_*\nu = \alpha$. Moreover, from the fact that the fiber entropy is zero we get $h_\mu(F) = h_\nu(\sigma)$, see Section C.2. This implies

$$\begin{aligned}\sup\Big\{h_\mu(F)\colon \mu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+ \times \mathbb{P}^1), \int \ln|(f'_{\xi_1})(v)|\, d\mu(\xi, v) = 2\alpha\Big\}\\ \leq \sup\{h_\nu(\sigma)\colon \nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+), \mathbb{M}_*(\nu) = \alpha\}\end{aligned}$$

and hence equality of both expressions, and equality with

$$h_{\text{top}}(F, \mathcal{E}(F, 2\alpha)) = h_{\text{top}}(\sigma, E(\mathbb{M},\alpha)),$$

proving the assertion. □

4.3 Proof of results

4.3.1 Gibbs property for equilibrium states

We know that the standard matrix norm $\|\cdot\|$ is sub-multiplicative, that is, $\|AB\| \leq \|A\|\|B\|$. The following result establishes a kind of quasi-super-multiplicativity.

Proposition 4.3.1 ([Fen09b, Proposition 2.8])**.** *If $\mathbb{M}$ is irreducible, then there exist $D > 0$ and $k \in \mathbb{N}$ such that for any $n, m \in \mathbb{N}$ and $\xi, \eta \in \Sigma_N^+$ there exist $r \leq k$, $\zeta \in \Sigma_N^+$ so that*

$$\|\mathbb{M}^n(\xi) \cdot \mathbb{M}^r(\zeta) \cdot \mathbb{M}^m(\eta)\| \geq D\|\mathbb{M}^n(\xi)\|.\|\mathbb{M}^m(\eta)\|.$$

The above result is the key in the proof of the following proposition.

Proposition 4.3.2 ([FK11, Proposition 1.2])**.** *If $\mathbb{M}$ is irreducible, then for every $q > 0$, $P(q)$ has a unique q-equilibrium state ν_q, P is differentiable at q, and $P'(q) = \mathbb{M}_*(\nu_q)$. Moreover, ν_q has the following Gibbs property: there exists $C > 1$ such that for every $n \in \mathbb{N}$ and $\xi \in \Sigma_N^+$ it holds*

$$\frac{1}{C} \leq \frac{\nu_q([\xi_1 \dots \xi_n])}{\exp(-nP(q))\|\mathbb{M}^n(\xi)\|^q} \leq C.$$

Remark 4.3.3. Under the hypothesis of Proposition 4.3.2 we have that for every $q > 0$

$$\nu_q\left(E(\mathbb{M}, P'(q))\right) = 1,$$

where ν_q is the q-equilibrium state of P. Indeed, since ν_q is the unique element in $\mathcal{I}_q$, ν_q is σ-ergodic. By Kingman's subadditive ergodic theorem, for ν_q-almost every ξ

$$\lambda(\mathbb{M}, \xi) = \mathbb{M}_*(\nu_q) = P'(q),$$

which proves the claim.

4.3.2 Proof of Theorem F

Proof of Theorem F. Let ν_q as in Proposition 4.3.2. By Claim 4.1.4, we get

$$h_{\nu_q}(\sigma) + q\mathbb{M}_*(\nu_q) = P_{\text{var}}(q) = \sup_{\mu \in \mathcal{M}_{\text{erg}}(F)} \left(h_\mu(F) + \frac{q}{2}\varphi(\mu)\right).$$

Now, applying Lemma 4.1.1 there exists a unique F-ergodic measure $\mu_q \in \mathcal{M}(\nu_q)$ such that $2\mathbb{M}_*(\nu_q) = \varphi(\mu_q)$ and so $P'(q) = \frac{1}{2}\varphi(\mu_q)$. Hence,

$$h_{\nu_q}(\sigma) + q\mathbb{M}_*(\nu_q) = h_{\mu_q}(F) + \frac{q}{2}\varphi(\mu_q).$$

To see that μ_q is the only measure in $\mathcal{M}_{\text{erg}}(F)$ that realizes the supremum, take $\mu \in \mathcal{M}_{\text{erg}}(F)$ such that

$$h_{\nu_q}(\sigma) + q\mathbb{M}_*(\nu_q) = h_\mu(F) + \frac{q}{2}\varphi(\mu).$$

Then, there exists $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$ such that $\mu \in \mathcal{M}(\nu)$. Applying Lemma 4.1.1 for ν, we get

$$h_\mu(F) + \frac{q}{2}\varphi(\mu) \leq h_\nu(\sigma) + q\mathbb{M}_*(\nu).$$

The uniqueness of the equilibrium state ν_q implies that $\nu = \nu_q$ and so $\mu = \mu_q$. With which we conclude the first part of the proposition.

Now, let us prove the last conclusion. Since $\mu_q\left(\{(\zeta, v_\zeta^+)\colon \lambda(\mathbb{M},\zeta)=\alpha\}\right)=1$ we have for every $n\in\mathbb{N}$ and $\xi\in\Sigma_N^+$

$$\mu_q\left(\left\{(\zeta, v_\zeta^+)\colon \zeta\in[\xi_1\dots\xi_n],\ \lambda(\mathbb{M},\zeta)=\alpha\right\}\right)=\pi_*\mu_q([\xi_1\dots\xi_n])=\nu_q([\xi_1\dots\xi_n]),$$

and, there exists $v_n\in\mathbb{P}^1$ such that

$$\|\mathbb{M}^n(\xi)\|=\|\mathbb{M}^n(\xi)v_n\|.$$

Therefore, it only remains to apply the equality (1.3.6) to conclude. □

Remark 4.3.4. Given $\xi\in\Sigma_N^+$, there exists a unique vector $v_\xi^+\in\mathbb{P}^1$ such that $\chi(\mathbb{M},\zeta,v_\xi^+)=2\alpha$, see Section 4.2.1 for more details. Moreover, for $n\in\mathbb{N}$ take $v_n\in\mathbb{P}^1$ as in Theorem F. By [Via14, Lemma 3.16], the sequence (v_n) is Cauchy in projective space and its limit is v_ξ^+. Moreover, the angle $\measuredangle(v_n,v_{n+1})$ decreases exponentially:

$$\limsup_{n\to\infty}\frac{1}{n}\ln|\sin\measuredangle(v_n,v_{n+1})|\le-2\alpha.$$

4.3.3 Proof of Theorem E

Proposition 4.3.2 states already the uniqueness of q-equilibrium state. Hence, it only remains we prove statements about the level sets. The proof will be based on the study of the L^q-spectrum $t\mapsto\tau_\nu(t)$ of a probability measure ν on Σ_N^+, defined by

$$\tau_\nu(t)\stackrel{\text{def}}{=}\liminf_{n\to\infty}\frac{-1}{n\ln N}\ln\sum_{[\xi_1\dots\xi_n]}\nu([\xi_1\dots\xi_n])^t.$$

We follow some arguments standard in multifractal analysis (see, for example [FL02, Proof of Theorem 1.3]).

Proof of Theorem E. Given $\alpha=P'(q)$ for some q, let ν_q be the corresponding Gibbs measure in Proposition 4.3.2. Note that the Gibbs property of ν_q implies

$$\tau_{\nu_q}(t)=\liminf_{n\to\infty}\frac{-1}{n\ln N}\ln\left(e^{-ntP(q)}\sum_{[\xi_1\dots\xi_n]}\|\mathbb{M}^n(\xi)\|^{qt}\right)=\frac{tP(q)-P(tq)}{\ln N},$$

and

$$E(\mathbb{M},\alpha)=\left\{\xi\in\Sigma_N^+\colon \lim_{n\to\infty}\frac{-1}{n\ln N}\ln\nu_1([\xi_1\dots\xi_n])=\frac{P(1)-\alpha}{\ln N}\right\}.$$

It follows

$$\tau_{\nu_1}(q)=\frac{qP(1)-P(q)}{\ln N}$$

Hence, by [LN99, Theorem 4.1]

$$\dim_{\mathrm{H}}E(\mathbb{M},\alpha)\le\inf_{q\in\mathbb{R}}\left(q\frac{P(1)-\alpha}{\ln N}-\tau_{\nu_1}(q)\right)=\inf_{q\in\mathbb{R}}\frac{1}{\ln N}(P(q)-\alpha q)$$

On the other hand, the Gibbs property implies that for $\xi \in E(\mathbb{M}, \alpha)$ (hence, ν_q-almost every ξ)

$$\lim_{n\to\infty} -\frac{1}{n \ln N} \ln \nu_q([\xi_1 \dots \xi_n]) = \frac{1}{\ln N}(P(q) - \alpha q).$$

Note that for $r \in \left(\frac{1}{N^{n+1}}, \frac{1}{N^n}\right]$, we have $B(\xi, r) = [\xi_1, \dots, \xi_n]$. Hence, for ν_q-almost every ξ

$$\liminf_{r\to 0^+} \frac{\nu_q(B(\xi, r))}{\ln(r)} = \lim -\frac{1}{n \ln N} \ln \nu_q([\xi_1 \dots \xi_n]) = \frac{1}{\ln N}(P(q) - \alpha q).$$

it follows from the mass distribution principle that

$$\dim_{\mathrm{H}} E(\mathbb{M}, \alpha) \geq \frac{1}{\ln N}(P(q) - \alpha q),$$

which shows the first two equalities.

To prove the last two equalities, note that ν_q is a q-equilibrium state and so

$$P(q) = h_{\nu_q}(\sigma) + q\mathbb{M}_*(\nu_q).$$

By Proposition 4.3.2

$$P(q) = h_{\nu_q}(\sigma) + \alpha q, \text{ and } \dim_{\mathrm{H}} E(\mathbb{M}, \alpha) = \frac{1}{\ln N}(P(q) - \alpha q).$$

Therefore,

$$\dim_{\mathrm{H}} E(\mathbb{M}, \alpha) = \frac{1}{\ln N} h_{\nu_q}(\sigma).$$

On the other hand, since ν_q is the q-equilibrium state of P. For all $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$ with $\mathbb{M}_*(\nu) = \alpha$ we have

$$h_\nu(\sigma) \leq P(q) - q\alpha = h_{\nu_q}(\sigma).$$

Thus, applying Lemma 4.2.1, we get $h_{\text{top}}(E(\mathbb{M}, \alpha)) = h_{\nu_q}(\sigma)$, which finishes the proof. $\square$

Bibliography

[ABC11] Flavio Abdenur, Christian Bonatti, and Sylvain Crovisier. Nonuniform hyperbolicity for C^1-generic diffeomorphisms. *Israel J. Math.*, 183:1–60, 2011.

[AR62] L. M. Abramov and V. A. Rohlin. Entropy of a skew product of mappings with invariant measure. *Vestnik Leningrad. Univ.*, 17(7):5–13, 1962.

[BDEG88] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo. Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(3):367–394, 1988.

[BE88] Michael F. Barnsley and John H. Elton. A new class of Markov processes for image encoding. *Adv. in Appl. Probab.*, 20(1):14–32, 1988.

[BL85] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1985.

[Bow73] Rufus Bowen. Topological entropy for noncompact sets. *Trans. Amer. Math. Soc.*, 184:125–136, 1973.

[BP07] Luis Barreira and Yakov Pesin. *Nonuniform hyperbolicity*, volume 115 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. Dynamics of systems with nonzero Lyapunov exponents.

[Car02] Niclas Carlsson. A contractivity condition for iterated function systems. *J. Theoret. Probab.*, 15(3):613–630, 2002.

[CFH08] Yong-Luo Cao, De-Jun Feng, and Wen Huang. The thermodynamic formalism for sub-additive potentials. *Discrete Contin. Dyn. Syst.*, 20(3):639–657, 2008.

[DF99] Persi Diaconis and David Freedman. Iterated random functions. *SIAM Rev.*, 41(1):45–76, 1999.

[DGR] L. J. Díaz, K. Gelfert, and M. Rams. Variational principle for nonhyperbolic ergodic measures: Skew products and elliptic cocycles. arXiv:2104.11638.

[DGR19] L. J. Díaz, K. Gelfert, and M. Rams. Entropy spectrum of Lyapunov exponents for nonhyperbolic step skew-products and elliptic cocycles. *Comm. Math. Phys.*, 367(2):351–416, 2019.

[DK17] Pedro Duarte and Silvius Klein. *Continuity of the Lyapunov exponents of linear cocycles*. Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2017. 31^o Colóquio Brasileiro de Matemática.

[DKN07] Bertrand Deroin, Victor Kleptsyn, and Andrés Navas. Sur la dynamique unidimensionnelle en régularité intermédiaire. *Acta Math.*, 199(2):199–262, 2007.

[Dur19] Rick Durrett. *Probability: Theory and Examples*. Cambridge university press, 2019.

[DZ98] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.

[Eda96] Abbas Edalat. Power domains and iterated function systems. *Inform. and Comput.*, 124(2):182–197, 1996.

[Elt87] John H. Elton. An ergodic theorem for iterated maps. *Ergodic Theory Dynam. Systems*, 7(4):481–488, 1987.

[Fat89] Albert Fathi. Expansiveness, hyperbolicity and Hausdorff dimension. *Comm. Math. Phys.*, 126(2):249–262, 1989.

[Fen04] De-Jun Feng. The variational principle for products of non-negative matrices. *Nonlinearity*, 17(2):447–457, 2004.

[Fen09a] De-Jun Feng. Lyapunov exponents for products of matrices and multifractal analysis. II. General matrices. *Israel J. Math.*, 170:355–394, 2009.

[Fen09b] De-Jun Feng. Lyapunov exponents for products of matrices and multifractal analysis. II. General matrices. *Israel J. Math.*, 170:355–394, 2009.

[FK11] De-Jun Feng and Antti Käenmäki. Equilibrium states of the pressure function for products of matrices. *Discrete Contin. Dyn. Syst.*, 30(3):699–708, 2011.

[FL02] De-Jun Feng and Ka-Sing Lau. The pressure function for products of non-negative matrices. *Math. Res. Lett.*, 9(2-3):363–378, 2002.

[Fri87] David Fried. Finitely presented dynamical systems. *Ergodic Theory Dynam. Systems*, 7(4):489–507, 1987.

[GK16] Anton Gorodetski and Victor Kleptsyn. Synchronization properties of random piecewise isometries. *Comm. Math. Phys.*, 345(3):781–796, 2016.

[Gou07] Nikolaz Gourmelon. Adapted metrics for dominated splittings. *Ergodic Theory Dynam. Systems*, 27(6):1839–1849, 2007.

[GS17] Katrin Gelfert and Örjan Stenflo. Random iterations of homeomorphisms on the circle. *Mod. Stoch. Theory Appl.*, 4(3):253–271, 2017.

[HPS77] M. W. Hirsch, C. C. Pugh, and M. Shub. *Invariant manifolds*. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, 1977.

[HS69] Edwin Hewitt and Karl Stromberg. *Real and abstract analysis. A modern treatment of the theory of functions of a real variable*. Springer-Verlag, New York-Berlin, 1969. Second printing corrected.

[Hut81] John E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.

[Huy73] Christiaan Huygens. *Horologium oscillatorium*. Apud F. Muguet, 1673.

[Ios09] Marius Iosifescu. Iterated function systems. A critical survey. *Math. Rep. (Bucur.)*, 11(61)(3):181–229, 2009.

[JT01] S. F. Jarner and R. L. Tweedie. Locally contracting iterated functions and stability of Markov chains. *J. Appl. Probab.*, 38(2):494–507, 2001.

[Kai78] Thomas Kaijser. A limit theorem for Markov chains in compact metric spaces with applications to products of random matrices. *Duke Math. J.*, 45(2):311–349, 1978.

[Kai81] Thomas Kaijser. On a new contraction condition for random systems with complete connections. *Rev. Roumaine Math. Pures Appl.*, 26(8):1075–1117, 1981.

[Led84] F. Ledrappier. Quelques propriétés des exposants caractéristiques. In *École d'été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 305–396. Springer, Berlin, 1984.

[LN99] Ka-Sing Lau and Sze-Man Ngai. Multifractal measures and a weak separation condition. *Adv. Math.*, 141(1):45–96, 1999.

[LsS05] Andreas Nordvall Lagerå s and Örjan Stenflo. Central limit theorems for contractive Markov chains. *Nonlinearity*, 18(5):1955–1965, 2005.

[LSS20] Krzysztof Leśniak, Nina Snigireva, and Filip Strobin. Weakly contractive iterated function systems and beyond: a manual. *J. Difference Equ. Appl.*, 26(8):1114–1173, 2020.

[Mal17] Dominique Malicet. Random walks on $\mathbf{Homeo}(S^1)$. *Comm. Math. Phys.*, 356(3):1083–1116, 2017.

[MM20] Edgar Matias and Ítalo Melo. Random products of maps synchronizing on average. *Commun. Contemp. Math.*, 22(8):1950086, 18, 2020.

[MS21] Edgar Matias and Eduardo Silva. Random iterations of maps on $\mathbb{R}^k$: asymptotic stability, synchronisation and functional central limit theorem. *Nonlinearity*, 34(3):1577–1597, 2021.

[Nav11] Andrés Navas. *Groups of circle diffeomorphisms*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, spanish edition, 2011.

[Ose68] V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968.

[Pei93] Marc Peigné. Iterated function systems and spectral decomposition of the associated Markov operator. In *Fascicule de probabilités*, volume 1993 of *Publ. Inst. Rech. Math. Rennes*, page 28. Univ. Rennes I, Rennes, 1993.

[PU10] Feliks Przytycki and Mariusz Urbański. *Distance-expanding maps*, page 102–130. London Mathematical Society Lecture Note Series. Cambridge University Press, 2010.

[RAS15] Firas Rassoul-Agha and Timo Seppäläinen. *A course on large deviations with an introduction to Gibbs measures*, volume 162 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.

[Sea07] Mícheál Ó. Searcóid. *Metric spaces*. Springer Undergraduate Mathematics Series. Springer-Verlag London, Ltd., London, 2007.

[Shu87] Michael Shub. *Global stability of dynamical systems*. Springer-Verlag, New York, 1987. With the collaboration of Albert Fathi and Rémi Langevin, Translated from the French by Joseph Christy.

[Ste99] David Steinsaltz. Locally contractive iterated function systems. *Ann. Probab.*, 27(4):1952–1979, 1999.

[Ste01] David Steinsaltz. Random logistic maps and Lyapunov exponents. *Indag. Math. (N.S.)*, 12(4):557–584, 2001.

[Ste12] Örjan Stenflo. A survey of average contractive iterated function systems. *J. Difference Equ. Appl.*, 18(8):1355–1380, 2012.

[SZ20] Tomasz Szarek and Anna Zdunik. The rate of convergence for function systems on the circle. *Colloq. Math.*, 159(1):77–89, 2020.

[SZ21] Tomasz Szarek and Anna Zdunik. The central limit theorem for iterated function systems on the circle. *Mosc. Math. J.*, 21(1):175–190, 2021.

[Sza03] Tomasz Szarek. Invariant measures for nonexpensive Markov operators on Polish spaces. *Dissertationes Math. (Rozprawy Mat.)*, 415:62, 2003. Dissertation, Polish Academy of Science, Warsaw, 2003.

[Via14] Marcelo Viana. *Lectures on Lyapunov exponents*, volume 145 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2014.

APPENDIX A – Equivalences between metrics

Let us recall some standard definitions and facts, see [Sea07]. A metric on a set M is a map $d : M \times M \to [0, \infty)$, required to satisfy the following axioms for all $x, y, z \in M$:

1. $d(x, y) = 0$ if and only if $x = y$;
2. $d(x, y) = d(y, y)$;
3. the triangle inequality $d(x, y) \leq d(x, z) + d(z, y)$.

Two metrics d_1 and d_2 on the same space M are said to be *equivalent* if they generate the same topology on M, or equivalent, every convergent sequence of (X, d_1) is convergent in (X, d_2) with the same limit, and vice versa.

Two metrics d_1 and d_2 on the same space M are *strongly equivalent* if there exist positive constants a and b such that

$$ad_1(x, y) \leq d_2(x, y) \leq bd_1(x, y)$$

for every $x, y \in M$. Of course, strong equivalence of two metrics implies equivalence, but not vice versa.

Lemma A.0.1. *For $\alpha \in (0, 1]$ the function $d^\alpha : M \times M \to [0, \infty)$ given by*

$$d^\alpha(x, y) \stackrel{def}{=} (d(x, y))^\alpha$$

is a metric on M.

Proof. The identity and symmetry properties of d^α are an immediate consequence of the fact that d is a metric. For the triangular inequality we use the following inequality. For $x \in [0, 1]$ we have

$$(1 + x)^\alpha \leq 1 + x \leq 1 + x^\alpha. \tag{A.0.1}$$

Let $x, y, z \in M$. Since d is a metric,

$$d^\alpha(x, y) \leq (d(x, z) + d(z, y))^\alpha. \tag{A.0.2}$$

We assume, without loss of generality, that $d(x, z) \leq d(z, y)$. If $d(z, y) = 0$, then it is clear that (A.0.2) implies

$$d^\alpha(x, y) \leq d^\alpha(x, z) + d^\alpha(z, y).$$

If $d(z, y) = 0$, then (A.0.2) and (A.0.1) imply

$$\begin{aligned} d^\alpha(x, y) &\leq d^\alpha(z, y)\left(\frac{d(x, z)}{d(z, y)} + 1\right)^\alpha \\ &\leq d^\alpha(z, y)\left(\frac{d^\alpha(x, z)}{d^\alpha(z, y)} + 1\right) \\ &= d^\alpha(x, z) + d^\alpha(z, y), \end{aligned}$$

so that the triangular inequality holds for d^α. Therefore, d^α is a metric on M. $\square$

Remark A.0.2. For $\alpha \in (0, 1)$ and d a metric, we have that d and d^α are equivalent, but not necessarily strongly equivalent.

APPENDIX B – Properties of the Integral

In this section, let us enunciate some classic results in probability theory for integrals, see for example [Dur19] for more details.

Jensen's inequality. *Given a probability space $(\Omega, \mathcal{B}, \mu)$ and a convex function $\varphi : \mathbb{R} \to \mathbb{R}$. If $f : \Omega \to \mathbb{R}$ is a measurable function such that $f, \varphi(f) \in L^1(\mu)$, then*

$$\varphi\left(\int f d\mu\right) \leq \int \varphi(f) d\mu$$

Fatou's lemma. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. Assume $(f_n)_{n\in\mathbb{N}}$ be a sequence of μ-measurable non-negative functions $f_n : \Omega \to [0, +\infty]$. Define the function $f : \Omega \to [0, +\infty]$ by setting $f(x) = \liminf_{n\to\infty} f_n(x)$, for every $x \in \Omega$. Then*

$$\liminf_{n\to\infty} \int f_n d\mu \geq \int \left(\liminf_{n\to\infty} f_n\right) d\mu.$$

Monotone convergence theorem. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. Consider a pointwise non-decreasing sequence $(f_n)_{n\in\mathbb{N}}$ of μ-measurable non-negative functions $f_n : \Omega \to [0, +\infty]$. Define the function $f : \Omega \to [0, +\infty]$ by setting $f(x) = \lim_{n\to\infty} f_n(x)$, for every $x \in \Omega$. Then*

$$\int f_n d\mu \uparrow \int f d\mu.$$

Dominated convergence theorem. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. Consider a sequence $(f_n)_{n\in\mathbb{N}}$ of μ-measurable functions $f_n : \Omega \to \mathbb{R}$ converging pointwise to a μ-measurable function $f : \Omega \to \mathbb{R}$. If $g : \Omega \to [0, +\infty]$ is a μ-measurable function such that $|f_n| \leq g$ a.e., for all n, and $g \in L^1(\mu)$, then*

$$\int f_n d\mu \to \int f d\mu.$$

The following result is the unique proved here, we are guided by [HS69, p. 201].

Lemma B.0.1. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. Consider a μ-measurable function $f : \Omega \to (0, +\infty)$. If $f \in L^1(\mu)$ then*

$$\lim_{r\downarrow 0} \left(\int f^r(\xi) d\mu(\xi)\right)^{\frac{1}{r}} = \exp \int \ln f(\xi) d\mu(\xi). \tag{B.0.1}$$

Proof. It is known that for every ξ

$$\lim_{r\to 0} \frac{1}{r}\left(f^r(\xi) - 1\right) = \ln f(\xi).$$

Since, $r \mapsto \frac{1}{r}\left(f^r(\xi) - 1\right)$ is increasing on $r \geq 0$,

$$\lim_{r \downarrow 0} \frac{1}{r}\left(f^r(\xi) - 1\right) = \ln f(\xi).$$

Applying Monotone Convergence Theorem, we get

$$\lim_{r \downarrow 0} \frac{1}{r}\left(\int f^r(\xi) d\mu(\xi) - 1\right) = \int \ln f(\xi) d\mu(\xi). \tag{B.0.2}$$

Now, using the inequality $x - 1 \geq \ln x$ and Jensen's Inequality, we get

$$\begin{aligned} \frac{1}{r}\left(\int f^r(\xi) d\mu(\xi) - 1\right) &\geq \ln \left(\int f^r(\xi) d\mu(\xi)\right)^{\frac{1}{r}} \\ &\geq \frac{1}{r}\int \ln f^r(\xi) d\mu(\xi) \\ &= \int \ln f(\xi) d\mu(\xi), \end{aligned} \tag{B.0.3}$$

which implies (B.0.1). □

APPENDIX C – Some properties of Matrix cocycles

C.1 Derivative of the projective map

Consider the group $GL^+(2,\mathbb{R})$ of all 2×2 matrices with real coefficients and positive determinant. For

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL^+(2,\mathbb{R})$$

define the map f_A by

$$f_A(v) \stackrel{\text{def}}{=} \frac{Av}{\|Av\|}, \quad \text{for} \quad v \in \mathbb{R}^2. \tag{C.1.1}$$

The map $f_\xi^n : \mathbb{P}^1 \to \mathbb{P}^1$ induces a diffeomorphism on the circle $g_\xi^n : \mathbb{S}^1 \to \mathbb{S}^1$. Let us identify $\mathbb{P}^1$ by $[0,\pi)$ and $\mathbb{S}^1$ by $[0,2\pi)$. Define $g : \mathbb{S}^1 \to \mathbb{P}^1$ by

$$g(\theta) \stackrel{\text{def}}{=} \theta/2, \quad \theta \in [0,2\pi),$$

which is differentiable and invertible, its inverse $g^{-1} : \mathbb{P}^1 \to \mathbb{S}^1$ is also differentiable. Therefore, $f_\xi^n : \mathbb{P}^1 \to \mathbb{P}^1$ induces the differentiable map $g_\xi^n : \mathbb{S}^1 \to \mathbb{S}^1$ given by

$$g_\xi^n \stackrel{\text{def}}{=} g^{-1} \circ f_\xi^n \circ g,$$

with

$$(g_\xi^n)'(\theta) = (f_\xi^n)'(\theta/2).$$

For vectors $v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{R}^2$ we define

$$\delta(v,w) := \sin(\angle(v,w)) = \frac{|v_1 w_2 - v_2 w_1|}{\|v\| \, \|w\|},$$

it is easily seen that this defines a distance on $\mathbb{P}^1$.

Lemma C.1.1. *If* $v = (v_1, v_2)$, $w = (w_1, w_2)$ *and* $A \in \mathrm{SL}(2,\mathbb{R})$, *then*

$$\delta(f_A(v), f_A(w)) = |\det A| \frac{\|v\| \, \|w\|}{\|Av\| \, \|Aw\|} \delta(v,w)$$

Proof.

$$\begin{aligned}\delta(f_A(v), f_A(w)) &= \frac{|adv_1w_2 + bcv_2w_1 - bcv_1w_2 - adv_2w_1|}{\|Av\|\,\|Aw\|}\\ &= \frac{|(ad-bc)v_1w_2 - (ad-bc)v_2w_1|}{\|Av\|\,\|Aw\|}\\ &= |\det A|\frac{|v_1w_2 - v_2w_1|}{\|Av\|\,\|Aw\|}\\ &= |\det A|\frac{\|v\|\,\|w\|}{\|Av\|\,\|Aw\|}\delta(v,w).\end{aligned}$$

□

We now consider the subgroup $\mathrm{SL}(2,\mathbb{R}) \subset GL^+(2,\mathbb{R})$ of 2×2 matrices with real coefficients and determinant one. Then for $A \in \mathrm{SL}(2,\mathbb{R})$, the map $f_A : \mathbb{P}^1 \to \mathbb{P}^1$ definided in C.1.1 is called projective map. By Lemma C.1.1, for every $v \in \mathbb{P}^1$ we have

$$(f_A)'(v) = \lim_{w\to v}\frac{\delta(f_A(v), f_A(w))}{\delta(v,w)} = \frac{1}{\|Av\|^2}.$$

C.2 The fiber entropy

Fixed $\nu \in \mathcal{M}_{\text{erg}}(\Sigma_N^+)$. Let μ be a F-invariant measure such that $\pi_*\mu = \nu$. By Rokhlin disintegration theorem, for ν-almost every $\xi \in \Sigma_N^+$ there exists a probability measure μ_ξ on $\mathbb{P}^1$ such that

$$\mu = \int \mu_\xi d\nu(\xi).$$

Let $\mathcal{P}$ be an at most countable partition of the fiber $\mathbb{P}^1$ in measurable sets with finite entropy $H_\xi(\mathcal{P}) < \infty$ for almost every $\xi \in \Sigma_N^+$, where $H_\xi(\mathcal{P}) = -\sum_{C\in\mathcal{P}}\mu_\xi(C)\ln\mu_\xi(C)$. Let us put

$$\mathcal{P}_\xi^n = \bigvee_{k=0}^{n-1}(f_\xi^k)^{-1}(\mathcal{P}).$$

By [AR62], the following limit exists and it is finite

$$h^\sigma(\mathbb{M},\mathcal{P}) = \lim_{n\to\infty}\frac{1}{n}\int H_\xi(\mathcal{P}_\xi^n)d\nu(\xi).$$

The supremum of $h^\sigma(\mathbb{M},\mathcal{P})$ over all measurable partitions

$$h^\sigma(\mathbb{M}) = \sup_{\mathcal{P}} h^\sigma(\mathbb{M},\mathcal{P})$$

is called the *fiber entropy*. Again by [AR62], we have

$$h_\mu(F) = h_\nu(\sigma) + h^\sigma(\mathbb{M})$$

Since $f_0,\dots,f_{N-1}$ are difeomorphisms (so homeomorphisms) on $\mathbb{P}^1$, $h^\sigma(\mathbb{M}) = 0$. Therefore, $h_\mu(F) = h_\nu(\sigma)$. Indeed, since $f_0,\dots,f_{N-1}$ are homeomorphisms on $\mathbb{P}^1$, for every sequence $\xi \in \Sigma_N^+$ and for all $n \in \mathbb{N}$ the function f_ξ^n is also a homeomorphism on $\mathbb{P}^1$. And so for every partition $\mathcal{P}$ the cardinality of $\mathcal{P}_\xi^n$ is at most n times the cardinality of $\mathcal{P}$. Since the number $H_\xi(\mathcal{P}_\xi^n)$ is bounded by the logarithm of the cardinality of $\mathcal{P}_\xi^n$, we conclude that $h^\sigma(\mathbb{M}) = 0$.

APPENDIX D – Large deviations for Markov system

D.1 Markov kernel

Let M be a compact metric space and $\mathscr{B}(M)$ be its Borel σ-field. Let $\mathrm{Prob}(M)$ denote the space of Borel probability measures on M. We denote by $L^\infty(M)$ the Banach space of bounded measurable functions $f : M \to \mathbb{C}$, endowed with the norm $\|\cdot\|_\infty$ given by

$$\|f\|_\infty \stackrel{\text{def}}{=} \sup_{\omega \in M} |f(\omega)|.$$

A *Markov kernel* is a function $K : M \to \mathrm{Prob}(M)$, $\omega \mapsto K_\omega$, such that for any Borel set $E \in \mathscr{B}(M)$, the function $\omega \mapsto K_\omega(E)$ is $\mathscr{B}(M)$-measurable. A Markov kernel K determines the following linear operator $\mathcal{Q}_K : L^\infty(M) \to L^\infty(M)$,

$$(\mathcal{Q}_K f)(\omega) \stackrel{\text{def}}{=} \int_M f(\vartheta) dK_\omega(\vartheta).$$

Following [DK17], we refer to K as the kernel of $\mathcal{Q}_K$ and to $\mathcal{Q}_K$ as the Markov operator of K.

The topological product space $M^{\mathbb{N}}$ is compact and metrizable. Its Borel σ-field $\mathscr{B}(M^{\mathbb{N}})$ is generated by the cylinders, i.e., sets of the form

$$C(E_1, \ldots, E_n) \stackrel{\text{def}}{=} \{(\omega_j)_{j\in\mathbb{N}} : \omega_j \in E_j \text{ for } j \in \{1, \ldots, n\}\}$$

with $E_1, \ldots, E_n \in \mathscr{B}(M)$. Given $\theta \in \mathrm{Prob}(M)$ and a Markov kernel K, the following expression determines a pre-measure over the cylinder semi-algebra on $M^{\mathbb{N}}$

$$\mathbb{P}_\mu[C(E_1, \ldots, E_n)] \stackrel{\text{def}}{=} \int_{E_n} \cdots \int_{E_1} d\theta(\omega_0) \prod_{j=1}^{n-1} dK_{\omega_j}(\omega_{j+1}).$$

By Carathéodory's extension theorem this pre-measure extends to a unique probability measure $\mathbb{P}_\theta$ on $(M^{\mathbb{N}}, \mathscr{B}(M^{\mathbb{N}}))$. Following Kolmogorov, we define the process $X_n : M^{\mathbb{N}} \to M$ by

$$X_n(\varpi) \stackrel{\text{def}}{=} \omega_n, \text{ where } \varpi = (\omega_j)_{j\in\mathbb{N}}.$$

Recall that it satisfies for all $E \in \mathscr{B}(M)$,

1. $\mathbb{P}_\theta[X_1 \in E] = \theta(E)$,
2. $\mathbb{P}_\theta[X_n \in E | X_{n-1} = \omega] = K_\omega(E)$ for all $\omega \in M$ and $n \geq 2$.

By construction $\{X_n\}_{n\in\mathbb{N}}$ is a time-homogeneous[1] Markov process with initial distribution θ and transition kernel K on the probability space $(M^{\mathbb{N}}, \mathscr{B}(M^{\mathbb{N}}), \mathbb{P}_\theta)$.

To simplify notation, the probability $\mathbb{P}_\theta$ for a Dirac mass $\theta = \delta_\omega$ with $\omega \in M$, will be denoted by $\mathbb{P}_\omega$. Notice that $\mathbb{P}_\omega(X_1 = \omega) = 1$, that is, the Markov process $\{X_n\}_{n\in\mathbb{N}}$ starts in state ω.

D.2 Markov system

Given a Markov kernel K on $(M, \mathscr{B}(M))$, a measure $\theta \in \text{Prob}(M)$ is said to be K-*stationary* when for all $E \in \mathscr{B}(M)$,

$$\theta(E) = \int_M K_x(E) d\theta(x).$$

If K is a Markov kernel K on $(M, \mathscr{B}(M))$ and $\theta \in \text{Prob}(X)$ is a K-stationary probability measure, then the pair (K, θ) is called *Markov system*.

Let $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ be a complex Banach algebra *with unity*, that is, $\mathbf{1} \in \mathcal{B}$. A Markov system (K, θ) is said to *act simply and quasi-compactly* on $\mathcal{B}$ if there are constants $C > 0$ and $\sigma \in (0, 1)$ such that for all $\varphi \in \mathcal{B}$ and all $n \geq 0$

$$\left\| \mathcal{Q}_K^n \varphi - \left(\int_M \varphi d\theta \right) \mathbf{1} \right\|_{\mathcal{B}} \leq C \sigma^n \|\varphi\|_{\mathcal{B}}. \tag{D.2.1}$$

Let $\mathcal{L}(\mathcal{B})$ be the Banach algebra of bounded linear operators on $\mathcal{B}$ and denote by $|||T|||_{\mathcal{B}}$ the operator norm of $T \in \mathcal{L}(\mathcal{B})$.

Remark D.2.1. If a Markov system (K, θ) acts simply and quasi-compactly on $\mathcal{B}$ then $\mathcal{Q}_K \in \mathcal{L}(\mathcal{B})$.

D.3 Large deviations

For $\varphi \in L^\infty(M)$, define the *sum process* $S_n(\varphi) : M^{\mathbb{N}} \to \mathbb{C}$ by

$$S_n(\varphi)(\varpi) \stackrel{\text{def}}{=} \varphi(X_1(\varpi)) + \varphi(X_2(\varpi)) + \cdots + \varphi(X_n(\varpi))$$

where $\varpi = (\omega_j)_{j\in\mathbb{N}}$.

Let $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ be a complex Banach algebra with unity which is also a *lattice*[2].Assume also $\mathcal{B} \subset L^\infty(M)$ and that the inclusion $\mathcal{B} \hookrightarrow L^\infty(M)$ is continuous, that is, $\|\varphi\|_\infty \leq \|\varphi\|_{\mathcal{B}}$ for all $\varphi \in \mathcal{B}$.

The following result is proved in [DK17, Theorem 4.4].

[1] The transition probability is independent of n.

[2] A complex Banach algebra $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a *lattice* if $|\varphi|, \overline{\varphi} \in \mathcal{B}$ for all $\varphi \in \mathcal{B}$.

Proposition D.3.1. *Let* (K, θ) *be a Markov system which acts simply and quasi-compactly on a Banach sub-algebra* $\mathcal{B} \subset L^\infty(M)$ *satisfying the above assumptions. Then given* $\varphi \in \mathcal{B}$ *there exist constants* $h, \varepsilon_0 > 0$ *and* $C > 0$ *such that for all* $\omega \in M$*,* $\varepsilon \in (0, \varepsilon_0)$ *and* $n \in \mathbb{N}$

$$\mathbb{P}_\omega \left[\left| \frac{1}{n} S_n(\varphi) - \int_M \varphi d\mu \right| > \varepsilon \right] \leq C e^{-nh\varepsilon^2}.$$