Statistical Properties of Compact Group Extensions of Non-Uniformly Expanding Dynamical Systems

by

Barinder Singh Matharu

Thesis

Submitted to the University of Warwick
for the degree of

Doctor of Philosophy

Warwick Mathematics Institute

October 2019

Contents

List of Figures

Acknowledgements

First and foremost, I would like to thank Prof. Ian Melbourne for not only suggesting the project, but also for his expert guidance and support throughout the PhD. Ian has always been generous with his time and knowledge, offering valuable advice and feedback throughout the years. For these reasons, I am deeply indebted to him. I am also grateful to Prof. Mark Pollicott and Dr Alexey Korepanov for valuable discussions earlier on in the PhD. I have thoroughly enjoyed my time as part of a friendly and dynamic environment at the Warwick Mathematics Institute. The ETDS and PaW research groups have provided stimulating and vibrant communities to learn. I am grateful to have been financially supported by the EPSRC (Grant No. EP/HO23364/1) as part of the MASDOC DTC. I also extend my gratitude to Prof. Richard Sharp and Prof. Mark Holland for examining this thesis.

Outside of research, I would like to thank WUPC and WUSBC for providing me with much needed downtime and giving me the opportunity to represent Warwick nationally. Special thanks also go to my friends in Cambridge for hosting me during my many visits, and my friends in Birmingham for easing the transition into life as a PhD student. Last but not least, I would like to thank my family, without whom this thesis would not be possible; my siblings Harminder and Gurpreet and their partners Mandeep and Rinesh, for the continuous messages of support and providing an enjoyable environment during my visits home; and my parents Gurdeep and Balwinder, who encouraged me to pursue my interests and have always been supportive of my choices in life. This is for you.

Declarations

This thesis has been submitted to the University of Warwick for the degree of Doctor of Philosophy. I declare that the material in this thesis is my own work except where otherwise indicated in the text. This material has not been submitted for any other degree or qualification.

Abstract

We prove statistical limit theorems for Birkhoff sums of the form $\sum_{k=0}^{n-1} \phi_n \circ T_{h^{(n)}}$, where $T_{h^{(n)}}$ are a sequence of compact group extensions with non-uniformly expanding base and ϕ_n are a sequence of equivariant Hölder observables. This is done by extending the methods of Korepanov, Kosloff, and Melbourne to construct two new martingale-coboundary decompositions.

Even in the case of a fixed observable and compact group extension, these decompositions enable us not only to reprove existing results in the literature, but also to obtain far reaching consequences. Using our primary martingale-coboundary decomposition, we give a new proof of a central limit theorem and weak invariance principle under very general conditions, and obtain moment estimates which are optimal given our setup. Still in the case of a fixed observable and compact group extension, we use our secondary martingale-coboundary decomposition to prove an almost sure invariance principle with excellent error rates.

As an application, we prove a homogenisation result for discrete fast-slow dynamical systems with additive noise, where the fast dynamics are generated by a family of compact group extensions with non-uniformly expanding base.

Chapter 1

Introduction

Broadly speaking, ergodic theory is concerned with studying the statistical properties of deterministic dynamical systems. Given such a system, one would like to describe the behaviour of orbits in time. However, if only approximate information regarding the starting point is available, difficulties may arise due to sensitive dependence on initial conditions. This chaotic nature restricts our ability to make deterministic predictions for large times into the future, and so it makes sense to study such systems from a probabilistic viewpoint. A starting point is Birkhoff's ergodic theorem [11], which says that for typical orbits, the time average coincides with the space average. This is a natural generalisation of the strong law of large numbers [44] from probability theory. Further classical results such as the Lindeberg-Lévy central limit theorem [19] and Donsker's invariance principle [29] have been widely studied for random variables exhibiting weak forms of dependence [8, 15, 32, 48, 69]. In order to appeal to these results, it is necessary to utilise the properties of the dynamics.

The statistical properties of uniformly expanding dynamical systems [1, 12, 80, 88, 90] and non-uniformly expanding dynamical systems [40, 72, 96, 101, 102] have been comprehensively studied in the literature. We are interested in dynamical systems of a product structure, where the dynamics in the base are driven by a chaotic system and the dynamics in the fibre by compact group translations.

More precisely, let $T\colon X \to X$ be a dynamical system on a metric space X and let G be a compact connected Lie group with a fixed representation into $\mathrm{O}(d)$. Let $h\colon X \to G$ be Hölder and define the compact group extension $T_h\colon X\times G \to X\times G$ by $T_h(x,g) = (Tx, gh(x))$. We consider equivariant observables $\phi\colon X \times G \to \mathbb{R}^d$ of the form $\phi(x,g) = g \cdot v(x)$ where v is Hölder, and are interested in the long term behaviour of

$$\sum_{k=0}^{n-1} \phi \circ T_h^k(x,g) = \sum_{k=0}^{n-1} \phi\big(T^k y, gh_k(y)\big) = \sum_{k=0}^{n-1} gh_k(y) \cdot v(T^k y),$$

where $h_k = h\, h\circ T \cdots h\circ T^{k-1}$. The statistics of such observations arise naturally in dynamical systems with Euclidean symmetry [79]. Regarding statistical properties, there exist results in the literature when the base is uniformly expanding [27, 36, 70, 71] and when the base is non-uniformly expanding [16, 28, 39, 71]. In this thesis, our attention is focussed on the latter. In particular, we give a new proof of the results in [39] which leads to much stronger conclusions. Our aim in the next few paragraphs is to give a "birds-eye" view of the thesis. We begin with the new results; more details are given in the corresponding chapter outlines, which give precise statements of the main theorems.

Chapter 4: Gordin's method [37] for studying the statistical properties of deterministic dynamical systems has seen extensive development in both the probability literature [45, 56, 68, 84] and the dynamical systems literature [63, 96, 97]. Roughly speaking, the idea is to decompose observables as the sum of a martingale and an asymptotically negligible coboundary, which allows one to utilise results from the martingale literature. Recently, Korepanov, Kosloff, and Melbourne [60] introduced a new version of this method for Hölder observables of non-uniformly expanding dynamical systems. We extend the ideas of [60] to the compact group extension setting, working with equivariant Hölder observables as described above. An immediate consequence is the central limit theorem (CLT) and weak invariance principle (WIP), recovering [39, Theorem 1.10]. As well as being more elementary, our method of proof has numerous advantages. In [39], the result is first proved for compact group extensions of the induced uniformly expanding system, and then after further arguments this is lifted to the original system. Our approach bypasses such induced limit theorems and directly applies to the original system. Moreover, we obtain optimal moment estimates which are not readily available in the literature.

Chapter 5: Another advantage of our methods is that we can readily decompose the square of the martingale in the decomposition described above, which allows us to control sums of squares as is often required in more sophisticated limit laws. As an application, we prove an almost sure invariance principle (ASIP) by appealing to the results in [22]. This is a powerful statistical property, which implies the CLT, WIP, and various other probabilistic results (see [84, Chapter 1]). The ASIP was originally introduced in [93, 94], and has been proved for various dynamical systems [25, 42, 47, 59, 72, 74]. However, our results appear to be the first available for compact group extensions with a non-uniformly expanding base, although results do exist when the base is uniformly hyperbolic [36]. Our error rates improve on these results in the uniformly expanding setting.

Chapter 6: Our approach also allows explicit control on various constants associated with T and h, making the method useful for studying Birkhoff sums of the form $\sum_{k=0}^{n-1} \phi_n \circ T_{h^{(n)}}$, where the observables ϕ_n and cocyles $h^{(n)}$ vary with n. Under mild conditions on the cocycles, we show that both the CLT and WIP hold. Such Birkhoff sums arise naturally in homogenisation problems, in which deterministic systems with multiple timescales converge to a stochastic differential equation [83]. As an application, we give a homogenisation result of discrete fast-slow dynamical systems with additive noise, where the fast dynamics are generated by a family of compact group extensions.

Both Chapter 2 and Chapter 3 serve as an exposition of known results, and can be seen as necessary preparation for the rest of the thesis. We briefly describe the contents of these chapters below.

Chapter 2: We first establish notation and introduce basic notions in ergodic theory. This is followed by a review of some classical probability theory, such as the convergence of probability measures and martingale theory. We next introduce the Koopman and transfer operators, and review the relevant tools from spectral theory and harmonic analysis which we require to study these operators. Finally, we introduce Gordin's method and give statements of the statistical limit theorems which we utilise.

Chapter 3: We give a precise definition of non-uniformly expanding dynamical systems, as well as prove the existence of an ergodic absolutely continuous invariant probability measure. This is done using the method of Young [101, 102], in which we represent a non-uniformly expanding dynamical system as a tower over its induced uniformly expanding system. By first constructing an ergodic absolutely continuous invariant probability measure on the base of the tower, this is then extended to the whole tower using standard arguments which we review.

Remark 1.0.1. *The main new results in this thesis are the moment estimates in Chapter 4, the ASIP in Chapter 5, and the CLT, WIP and homogenisation result in Chapter 6.*

Chapter 2

Preliminaries

2.1 Notation

Throughout, we use d, i, j, k, l, m, n to denote integer-valued indices. That is, if we write $n \geq 1$, we implicitly mean $n \in \mathbb{Z}$ with $n \geq 1$. This nomenclature is not reserved for other indices. For example, if we write $p > 1$, we mean $p \in (1, \infty)$. For $d \geq 1$, by $x \in \mathbb{R}^d$ we mean the column vector $x = (x_1, ..., x_d)^T$. Fix a norm $|\cdot|$ on $\mathbb{R}^d$. For $\Sigma \in \mathbb{R}^{d,d}$, where $\mathbb{R}^{d,d}$ denotes the set of $d \times d$ real matrices, we denote by $\|\Sigma\| = \inf\{C \geq 0 : |\Sigma x| \leq C|x| \text{ for all } x \in \mathbb{R}^d\}$ the corresponding operator norm. We let $\mathrm{O}(d) = \{\Sigma \in \mathbb{R}^{d,d} \mid \det \Sigma \neq 0 \text{ and } \Sigma^T = \Sigma^{-1}\}$ denote the group of $d \times d$ orthogonal matrices with binary product matrix multiplication.

Let $(X, \mathcal{B}, \mu)$ be a probability space. For $1 \leq p < \infty$ we denote by $L^p(X; \mathbb{R}^d)$ the space of measurable functions $f \colon X \to \mathbb{R}^d$ such that

$$|f|_p = \left(\int_X |f|^p \, \mathrm{d}\mu \right)^{1/p} < \infty.$$

Similarly, we denote by $L^\infty(X; \mathbb{R}^d)$ the space of essentially bounded functions $f \colon X \to \mathbb{R}^d$. That is, measurable functions $f \colon X \to \mathbb{R}^d$ such that

$$|f|_\infty = \inf \big\{ C : |f| \leq C \text{ almost surely} \big\} < \infty.$$

For all $1 \le p \le \infty$, we have that $L^p(X;\mathbb{R}^d)$ is a Banach space when equipped with $|\cdot|_p$. In the case $d = 1$, we write $L^p(X)$. Naturally, the above extends to $\mathbb{R}^{d,d}-$valued functions by replacing the norm $|\cdot|$ with the operator norm $\|\cdot\|$.

Let (X, d) be a bounded metric space and $\eta \in (0, 1]$. We say that $v\colon X \to \mathbb{R}^d$ is $\eta-$Hölder if

$$|v|_\eta = \sup_{\substack{x,y\in X\\ x\neq y}} \frac{\big|v(x) - v(y)\big|}{d(x, y)^\eta} < \infty.$$

When $\eta = 1$, we refer to v as Lipschitz and write the above semi-norm as $\mathrm{Lip}(v)$. We define $C^\eta(X;\mathbb{R}^d)$ to be the space of $\mathbb{R}^d-$valued $\eta-$Hölder functions. This defines a Banach space when equipped with the norm $\|v\|_\eta = |v|_\infty + |v|_\eta$. When $\eta = 1$, we write the above norm as $\|v\|_{\mathrm{Lip}} = |v|_\infty + \mathrm{Lip}(v)$. In the case $d = 1$, we write $C^\eta(X)$. Again, the above extends to $\mathbb{R}^{d,d}-$valued functions by replacing the norm $|\cdot|$ with the operator norm $\|\cdot\|$.

For $\Sigma \in \mathbb{R}^{d,d}$, we denote by $\mathcal{N}(0,\Sigma)$ the (multivariate) normal distribution with mean 0 and covariance matrix Σ. If Σ is singular, then $\mathcal{N}(0,\Sigma)$ is degenerate. That is, $\mathcal{N}(0,\Sigma)$ is supported on a space of dimension less than d. If Σ is non-singular, then we say that $\mathcal{N}(0,\Sigma)$ is non-degenerate, and its probability density function is given by

$$f_{0,\Sigma}(x) = \frac{1}{(2\pi)^{d/2}(\det\Sigma)^{1/2}} \exp\bigg(-\frac{1}{2}x^T\Sigma^{-1}x\bigg).$$

Throughout, we use $\ll$ and "big O" notation interchangeably, writing $a_n \ll b_n$ or $a_n = O(b_n)$ if there is a constant $C > 0$ such that $a_n \le Cb_n$ for all $n \ge 1$. We write $a_n = o(b_n)$ if $\lim_{n\to\infty} a_n/b_n = 0$.

2.2 Ergodicity and basic constructions

We begin by introducing some basic notions in ergodic theory (see for example [98]). For our purposes, a dynamical system is a probability space $(X, \mathcal{B}, \mu)$ equipped with a measurable map $T\colon X \to X$. We write this as the quadruple $(X, \mathcal{B}, \mu, T)$ and refer to T as the dynamical system. The evolution of the system is studied by considering iterates $T^n = T \circ T \circ \cdots \circ T$ for $n \geq 1$. We say that T (or μ) is *non-singular* if for all $B \in \mathcal{B}$, we have $\mu(B) = 0$ if and only if $\mu(T^{-1}(B)) = 0$. We say that T (or μ) is *invariant*, or T is *measure preserving*, if $\mu(T^{-1}(B)) = \mu(B)$ for all $B \in \mathcal{B}$. If T is measure preserving, we say that T (or μ) is *ergodic* if for all $B \in \mathcal{B}$ with $T^{-1}(B) = B$, one has $\mu(B) \in \{0, 1\}$. Equivalently, T (or μ) is ergodic if and only if $v \in L^2(X)$ with $v \circ T = v$ $\mu-$almost surely implies v is constant $\mu-$almost surely. Any two distinct ergodic measures are mutually singular. That is, if μ_1 and μ_2 are two distinct ergodic measures, then there exists $B \in \mathcal{B}$ such that $\mu_1(B) = \mu_2(X \setminus B) = 1$. Ergodicity is a key component in deducing statistical properties of deterministic dynamical systems. The most well-known result in this vein is the celebrated Birkhoff ergodic theorem [11], which says that almost everywhere, the time average and space average coincide for ergodic systems.

Theorem 2.2.1 (Birkhoff's ergodic theorem). *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving dynamical system. If $v \in L^1(X)$, then*

$$\frac{1}{n} \sum_{k=0}^{n-1} v \circ T^k \to \int_X v \, \mathrm{d}\mu \qquad a.s.$$

In Section 4.6, we require the following consequence of Birkhoff's ergodic theorem. Due to its elementary nature, we state and prove it here. We first give a lemma.

Lemma 2.2.2. *Let $b > 0$. Suppose $(a_n)_{n \geq 0} \subset \mathbb{R}$ with $\lim_{n \to \infty} n^{-b} a_n = 0$. Then $\lim_{n \to \infty} n^{-b} \max_{0 \leq k \leq n} |a_k| = 0$.*

Proof. Suppose first that there exists $C \geq 0$ such that $|a_n| \leq C$ for all $n \geq 0$. Then $n^{-b} \max_{0 \leq k \leq n} |a_k| \leq C n^{-b} \to 0$, proving the result in this case.

Suppose now that $|a_n| \to \infty$. Note that for each $n \geq 0$, there exists $0 \leq k_n \leq n$ such that $\max_{0 \leq k \leq n} |a_k| = |a_{k_n}|$. It is immediate that $(k_n)_{n \geq 0}$ is non-decreasing and $k_n \to \infty$. Therefore

$$n^{-b} \max_{0 \leq k \leq n} |a_k| = \frac{|a_{k_n}|}{n^b} = \frac{k_n^b}{n^b} \cdot \frac{|a_{k_n}|}{k_n^b} \leq \frac{|a_{k_n}|}{k_n^b} \to 0,$$

completing the proof. □

Corollary 2.2.3. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving dynamical system. Suppose $p \geq 1$ and $v \in L^p(X)$. Then $\max_{0\leq k\leq n} |v \circ T^k| = o(n^{1/p})$ almost surely.*

Proof. In view of Lemma 2.2.2, it suffices to show that $v \circ T^n = o(n^{1/p})$ almost surely. By Birkhoff's ergodic theorem, we have

$$\frac{1}{n}\sum_{k=0}^{n-1} v^p \circ T^k \to \int v^p \, \mathrm{d}\mu \quad \text{a.s.}$$

Therefore

$$\begin{aligned} \frac{v^p \circ T^n}{n} &= \frac{1}{n}\sum_{k=0}^{n} v^p \circ T^k - \frac{1}{n}\sum_{k=0}^{n-1} v^p \circ T^k \\ &= \frac{n+1}{n} \cdot \frac{1}{n+1}\sum_{k=0}^{n} v^p \circ T^k - \frac{1}{n}\sum_{k=0}^{n-1} v^p \circ T^k \\ &\to 0 \quad \text{a.s.} \end{aligned}$$

The result follows. □

A measure preserving dynamical system $(X, \mathcal{B}, \mu, T)$ is said to be *mixing* if for all $v, w \in L^2(X)$, we have

$$\lim_{n\to\infty} \int_X v \circ T^n \, w \, \mathrm{d}\mu = \int_X v \, \mathrm{d}\mu \int_X w \, \mathrm{d}\mu.$$

We say that T (or μ) is *weak mixing* if for all $A, B \in \mathcal{B}$, we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \left|\mu(T^{-k}(A) \cap B) - \mu(A)\mu(B)\right| = 0.$$

It is standard that mixing implies weak mixing and weak mixing implies ergodicity. Weak mixing can be equivalently characterised as follows: T (or μ) is weak mixing if and only if $v \in L^2(X)$ with $v \circ T = e^{i\omega} v$ $\mu-$almost surely for $\omega \in [0, 2\pi)$ implies v is constant $\mu-$almost surely.

We next give some constructions which are used throughout the thesis. We begin with return times and their corresponding induced transformations, which allow us to focus on specific regions of the state space of our underlying dynamical system.

Definition 2.2.4. *Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system and $Y \in \mathcal{B}$ be such that $\mu(Y) > 0$. We call a measurable function $\tau \colon Y \to \mathbb{Z}^+$ a return time if $T^{\tau(y)}y \in Y$ for all $y \in Y$. We define the induced transformation $F \colon Y \to Y$ by $Fy = T^{\tau(y)}y$.*

Remark 2.2.5. *We make the following observations:*

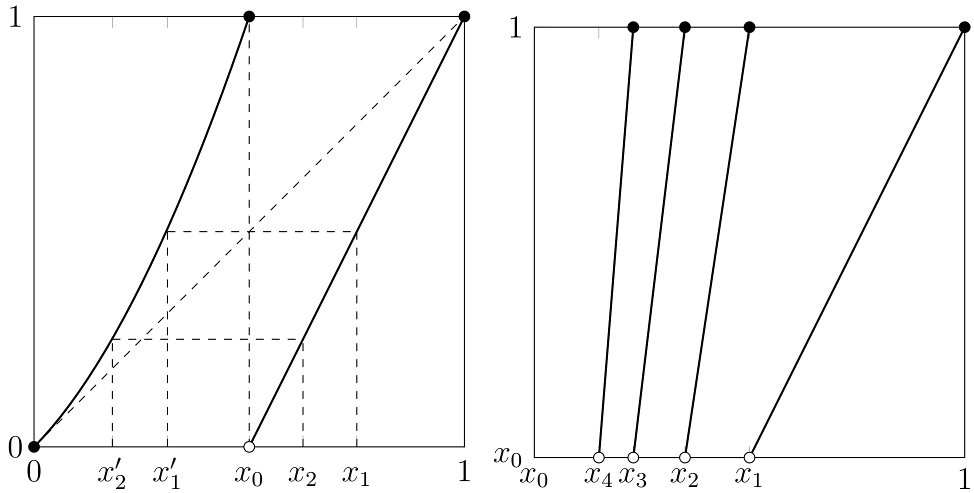
(i) *As will be seen in Chapter 3, it is often the case that inducing yields a dynamical system with better global properties than the original system, which makes it easier to analyse. Moreover, interesting conclusions about the original system can often be obtained by analysing the induced system [66, 76].*

(ii) *In the context of measure-preserving dynamical systems, a classical example of a return time is the first return $\tau \colon Y \to \mathbb{Z}^+$ defined by $\tau(y) = \inf\{n \geq 1 \mid T^n y \in Y\}$. This is well-defined by the Poincaré recurrence theorem [98, Theorem 1.4] and integrable by Kac's lemma [13, Theorem 3.2.4].*

Example 2.2.6. *Let $X = [0, 1]$ and $\gamma \in [0, 1)$. The intermittent map [65] is defined as*

$$T(x) = \begin{cases} x(1 + 2^\gamma x^\gamma) & \text{if } x \in [0, 1/2], \\ 2x - 1 & \text{if } x \in (1/2, 1]. \end{cases}$$

Let $Y = (1/2, 1]$. We construct the first return to Y and corresponding induced transformation. The first few steps of the construction are illustrated in Figure 2.1 (A). Let $x_0 = 1/2$ and observe that $T^{-1}x_0 \in \{x_1', 3/4\}$ for some $x_1' \in (0, 1/2)$. We set $x_1 = 3/4$ and $a_1 = (x_1, 1]$. Similarly, $T^{-1}x_1' \in \{x_2', x_2\}$ for some $x_2' \in (0, x_1')$ and $x_2 \in (1/2, x_1)$. Set $a_2 = (x_2, x_1]$. We can continue this process inductively to generate a partition $\alpha = (a_n)_{n \geq 1}$ of Y. Define $\tau \colon Y \to \mathbb{Z}^+$ by $\tau(y) = n$ if $y \in a_n$. It is immediate that τ is the first return to Y. Moreover, $F = T^\tau$ restricts to a bijection from a_n onto Y, as is shown in Figure 2.1 (B) for the first 4 branches.



(A) Inducing scheme (B) Induced transformation

Figure 2.1: Intermittent map

We next introduce compact group extensions of dynamical systems, which belong to a class of systems called *skew products* (see for example [2, 81]), where the first coordinate is determined by some given dynamical system and the second coordinate is determined by compact group translations. We first make a remark regarding representations of compact connected Lie groups.

Remark 2.2.7. *Let G be a compact connected Lie group with Haar measure ν, and suppose that $(\pi, \mathbb{R}^d)$ is a representation of G for some $d \geq 1$. By Weyl's unitarian trick [34, Proposition 6.1.1], there exists a $\pi(G)$ – invariant inner product $[\cdot,\cdot]$ on $\mathbb{R}^d$. By fixing an orthonormal basis of $\mathbb{R}^d$ with respect to this inner product, we may suppose that $\pi\colon G \to \mathrm{O}(d)$. From here on, unless otherwise stated, we write $\pi(G)$ as G and by $g \cdot x$ we denote multiplication of the matrix $g \in G$ with $x \in \mathbb{R}^d$. We use throughout that $\|g\| = 1$ for all $g \in G$, where $\|\cdot\|$ is the operator norm corresponding to the norm induced by $[\cdot,\cdot]$.*

Definition 2.2.8. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving dynamical system and $G \subset \mathrm{O}(d)$ be a closed subgroup. For $\eta \in (0, 1]$, we call $h \in C^\eta(X; G)$ an η – Hölder cocycle. For such $h\colon X \to G$, we define the compact group extension*

$T_h\colon X \times G \to X \times G$ by $T_h(x, g) = (Tx, gh(x))$.

Remark 2.2.9. *Let ν denote the Haar measure on G and consider the product probability measure $m = \mu \times \nu$. Then m is T_h –invariant. Indeed, for $\phi \in L^1(X \times G)$, we have*

$$\int_{X\times G} \phi \circ T_h(x,g)\,\mathrm{d}m = \int_X \int_G \phi\big(Tx, gh(x)\big)\,\mathrm{d}\nu\,\mathrm{d}\mu = \int_X \int_G \phi(Tx,g)\,\mathrm{d}\nu\,\mathrm{d}\mu$$
$$= \int_G \int_X \phi(Tx,g)\,\mathrm{d}\mu\,\mathrm{d}\nu = \int_G \int_X \phi(x,g)\,\mathrm{d}\mu\,\mathrm{d}\nu = \int_{X\times G} \phi(x,g)\,\mathrm{d}m.$$

For large classes of dynamical systems, ergodicity of compact group extensions is typical (see for example [35]). To conclude this section, we give some examples of compact group extensions for which ergodicity fails, even when the underlying dynamical system is assumed to be ergodic.

Example 2.2.10. *Let $T\colon X \to X$ be ergodic. Suppose $G = \mathbb{R}/2\pi\mathbb{Z}$ with binary product addition modulo 2π, so $\nu = \mathrm{Leb}$. Suppose further that we have a constant cocycle $h = 2\pi c$, where $c \in \mathbb{Q} \cap [0, 1)$. Note that $T_h(x, \theta) = (Tx, R_c\theta)$, where $R_c\colon G \to G$ defined by $R_c\theta = \theta + c \bmod 2\pi$ is the circle rotation by angle $2\pi c$. Write $c = p/q$ where $p \geq 0, q \geq 1$, and $hcf(p, q) = 1$. For*

$$H = \bigcup_{k=0}^{q-1} \left[\frac{2k\pi}{q}, \frac{(2k+1)\pi}{q}\right],$$

note that $R_c^{-1}(H) = H$ and $\nu(H) = 1/2$. It follows that $T_h^{-1}(X \times H) = X \times H$ and $m(X \times H) = \nu(H) \notin \{0, 1\}$, so that m is not ergodic.

Example 2.2.11. *Let $T\colon X \to X$ be ergodic. Suppose G is a non-abelian compact connected Lie group with compatible bi-invariant metric d. Suppose further that for some $h \in G$, we have $h(x) = h$ for all $x \in X$. Let $H = \overline{\langle h\rangle}$. Since H is closed, it is a Lie subgroup of G. Moreover, by an approximation argument, H is abelian. Therefore $\dim H < \dim G$, so that $\nu(H) = 0$. Consider the thickening of H by $\epsilon > 0$ small, given by $B_\epsilon(H) = \{g \in G \mid \inf_{h' \in H} d(g, h') < \epsilon\}$. Then $g \in B_\epsilon(H)$ if and only if $gh \in B_\epsilon(H)$, so that $T_h^{-1}(X \times B_\epsilon(H)) = X \times B_\epsilon(H)$. However, $m(X \times B_\epsilon(H)) = \nu(B_\epsilon(H)) \notin \{0, 1\}$, and so ergodicity fails.*

2.3 Convergence of probability measures

As noted in Remark 1.0.1, the main results in this thesis are of a probabilistic nature. In this section, we give an overview of the relevant probability theory required to formulate these.

We begin with a brief recap of weak convergence (see for example [10]). Suppose we have a probability space $(X, \mathcal{B}, \mu)$ and let $\mathcal{M}$ be a metric space. We say that $W\colon X \to \mathcal{M}$ is a *random element* of $\mathcal{M}$ if it is measurable as a function from $(X, \mathcal{B}) \to (\mathcal{M}, \mathcal{B}(\mathcal{M}))$, where $\mathcal{B}(\mathcal{M})$ denotes the Borel $\sigma-$algebra of $\mathcal{M}$. If $\mathcal{M} = \mathbb{R}$ or $\mathbb{C}$, we call W a *random variable*, and if $\mathcal{M} = \mathbb{R}^d$ for $d > 1$, we call W a *random vector*. The probability *distribution* or *law* of W is the pushforward measure $\mu_W = \mu \circ W^{-1}$ induced by W on $\mathcal{B}(\mathcal{M})$. Let $(W_n)_{n\geq 0}$ be a sequence of random elements taking values in $\mathcal{M}$. We say that the sequence of probability measures μ_{W_n} converges *weakly* to μ_W, written $\mu_{W_n} \to_w \mu_W$, or that the random elements W_n converge *weakly* to W, written $W_n \to_w W$, if

$$\lim_{n\to\infty} \int_{\mathcal{M}} f \,\mathrm{d}\mu_{W_n} = \int_{\mathcal{M}} f \,\mathrm{d}\mu_W \quad \text{for all } f \in C_b(\mathcal{M}),$$

where $C_b(\mathcal{M})$ is the set of all bounded, continuous, real-valued functions on $\mathcal{M}$. Equivalently, $W_n \to_w W$ if

$$\lim_{n\to\infty} \mathbb{E}\big[f(W_n)\big] = \mathbb{E}\big[f(W)\big] \quad \text{for all } f \in C_b(\mathcal{M}),$$

where $\mathbb{E}$ denotes the expectation with respect to the underlying probability measure μ.

Remark 2.3.1. *We also refer to weak convergence as convergence in distribution, and use these interchangeably throughout. Almost sure convergence, L^p-convergence, and convergence in probability all imply weak convergence.*

One of the most classical results regarding weak convergence is the Lindeberg-Lévy central limit theorem [9, 19, 31]. We recall this below.

Theorem 2.3.2 (Lindeberg-Lévy central limit theorem). *Let $(X_k)_{k\geq 0}$ be a sequence of independent and identically distributed random vectors with $\mathbb{E}[X_k] = 0$ and $\mathbb{E}[X_k X_k^T] = \Sigma \in \mathbb{R}^{d,d}$ for all $k \geq 0$. Then*

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} X_k \to_w \mathcal{N}(0, \Sigma).$$

We next state the definition of Brownian motion. From here on, we use $\sim$ to denote equivalence in distribution. Recall that a stochastic process is a family of measurable functions.

Definition 2.3.3. *Let $W = (W(t))_{t\geq 0}$ be an $\mathbb{R}^d$ –valued stochastic process with $W(0) = 0$. We say that W is a Brownian motion on $\mathbb{R}^d$ with mean 0 and covariance matrix $\Sigma \in \mathbb{R}^{d,d}$ if the following hold:*

(i) *For all $t \geq 0$, we have $W(t) \sim \mathcal{N}(0, t\Sigma)$.*

(ii) *For $0 = t_0 < t_1 < t_2 < \cdots < t_n$, the random variables $W(t_1)$, $W(t_2) - W(t_1)$,…, $W(t_n) - W(t_{n-1})$ are independent.*

(iii) *For all $t > s \geq 0$, we have $W(t) - W(s) \sim W(t - s)$.*

(iv) *$t \mapsto W(t)$ is almost surely continuous.*

The existence of Brownian motion was first shown by Wiener [99]. We next recall the space of càdlàg functions [82]. Let $d \geq 1$ and $E \subset \mathbb{R}$. We denote by $D(E; \mathbb{R}^d)$ the space of functions $f : E \to \mathbb{R}^d$ which are right continuous and admit left limits. Let Λ denote the set of strictly increasing continuous bijections from E to itself. Define the Skorokhod metric s on $D(E; \mathbb{R}^d)$ by

$$s(f, g) = \inf_{\lambda \in \Lambda} \max \left\{ |\lambda - I|_\infty, |f - g \circ \lambda|_\infty \right\}.$$

This makes $D(E; \mathbb{R}^d)$ into a complete separable metric space [91]. It is immediate from the definition that $s(f, g) \leq |f - g|_\infty$. If $E = [0, \infty)$, we have by [92] that weak convergence in $D(E; \mathbb{R}^d)$ is equivalent to weak convergence in $D([0, T]; \mathbb{R}^d)$ for all $T > 0$.

The next result we state is Donsker's invariance principle [29, 33]. This can be seen as a generalisation of the central limit theorem to stochastic processes. Donsker's invariance principle says that Brownian motion is the limit of suitably rescaled random walks. More precisely:

Theorem 2.3.4 (**Donsker's invariance principle**). *Let $(X_k)_{k\geq 0}$ be a sequence of independent and identically distributed random vectors with $\mathbb{E}[X_k] = 0$ and $\mathbb{E}[X_k X_k^T] = \Sigma \in \mathbb{R}^{d,d}$ for all $k \geq 0$. For $n \geq 1$, define the random elements $W_n \colon X \to D([0,\infty);\mathbb{R}^d)$ by $W_n(t) = n^{-1/2}\sum_{k=0}^{[nt]-1} X_k$ for $t \geq 0$. Then $W_n \to_w W$ in $D([0,\infty);\mathbb{R}^d)$, where W denotes the Brownian motion with mean 0 and covariance matrix Σ.*

Remark 2.3.5. *In later chapters, we prove an analogue of Donsker's invariance principle for sequences of random vectors which are not independent in general. To make this distinction clear, we call such a result a weak invariance principle.*

The next result is the continuous mapping theorem [10, 45, 67], which says that continuous functions between metric spaces preserve limits even if their arguments are sequences of random elements. We then use this to show how Donsker's invariance principle implies the Lindeberg-Lévy central limit theorem.

Theorem 2.3.6 (**Continuous mapping theorem**). *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces and consider random elements $(W_n)_{n\geq 0}$ and W taking values in $\mathcal{X}$. If $h \colon \mathcal{X} \to \mathcal{Y}$ is continuous and $W_n \to_w W$, then $h(W_n) \to_w h(W)$.*

Remark 2.3.7. *Let W_n and W be as in Theorem 2.3.4, and note that $W_n \to_w W$ in $D([0,1];\mathbb{R}^d)$. Take $\mathcal{X} = D([0,1];\mathbb{R}^d)$ and $\mathcal{Y} = \mathbb{R}^d$ in Theorem 2.3.6, and let $h \colon \mathcal{X} \to \mathcal{Y}$ be defined by $h(f) = f(1)$. Let Λ denote the set of all strictly increasing bijections from $[0,1] \to [0,1]$, and note that $\lambda(1) = 1$ for all $\lambda \in \Lambda$. Letting s denote the Skorokhod metric on $\mathcal{X}$, we have for $f, g \in \mathcal{X}$ that*

$$s(f,g) = \inf_{\lambda\in\Lambda} \max\left\{|\lambda - I|_\infty, |f - g\circ\lambda|_\infty\right\} \geq \left|f(1) - g(1)\right| = \left|h(f) - h(g)\right|,$$

so that h is continuous. Therefore $n^{-1/2}\sum_{k=0}^{n-1} X_k = h(W_n) \to_w h(W) \sim \mathcal{N}(0,\Sigma)$, recovering the Lindeberg-Lévy central limit theorem.

We next recall the Cramér-Wold device [9, 21], which relates distributional convergence of random vectors to that of its one-dimensional projections.

Theorem 2.3.8 (Cramér-Wold). *Suppose $(X_n)_{n\geq 0}$ and X are d – dimensional random vectors. Then $X_n \to_w X$ if and only if $c \cdot X_n \to_w c \cdot X$ for all $c \in \mathbb{R}^d$.*

The next result is an analogue of Slutsky's theorem for random elements, from [10]. Roughly speaking, this allows us to deduce weak convergence of a sequence of random elements from that of another, "closely related" sequence of random elements. More precisely:

Theorem 2.3.9. *Let $(\mathcal{X}, s)$ be a separable metric space. Suppose that $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are sequences of random elements of $\mathcal{X}$. If $X_n \to_w X$ and $s(X_n, Y_n) \to 0$ in probability, then $Y_n \to_w X$.*

We now recall the definition of uniform integrability. We restrict to what is needed for the thesis. For a more detailed exposition, see [31, 87].

Definition 2.3.10. *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $(M_n)_{n\geq 0}$ be a sequence of random variables. The collection $(M_n)_{n\geq 0}$ is said to be uniformly integrable if*

$$\lim_{K\to\infty} \sup_{n\geq 0} \int_X |M_n| \mathbb{1}_{\{|M_n|\geq K\}} \, \mathrm{d}\mu = 0.$$

We require the following sufficient condition for uniform integrability.

Proposition 2.3.11. *Suppose $(M_n)_{n\geq 0}$ is a sequence of random variables. If $\sup_{n\geq 0} |M_n|_p < \infty$ for some $p > 1$, then $(M_n)_{n\geq 0}$ is uniformly integrable.*

The next result says that weak convergence and uniform integrability implies convergence of moments.

Proposition 2.3.12. *Let $(X, \mathcal{B}, \mu)$ be a probability space and suppose $(M_n)_{n\geq 0}$ and M are random vectors with $M_n \to_w M$. If $(|M_n|^p)_{n\geq 0}$ is uniformly integrable for some $p > 1$, then for all $0 < q \leq p$ we have*

$$\lim_{n\to\infty} \int_X |M_n|^q \, \mathrm{d}\mu_n = \int_X |M|^q \, \mathrm{d}\mu.$$

To conclude this section, we recall the notion of tightness, which roughly speaking, prevents the escape of mass to infinity. A more detailed exposition can be found in [10].

Definition 2.3.13. *Let* $(X, \mathcal{B}, \mu)$ *be a probability space and let* $(W_n)_{n\geq 0}$ *be a sequence of random elements of a metric space* $\mathcal{M}$. *We say that* $(W_n)_{n\geq 0}$ *is tight if for every* $\epsilon > 0$, *there exists a compact set* $K \subset \mathcal{M}$ *such that* $\mu(W_n \in K) > 1 - \epsilon$ *for all* $n \geq 0$.

The fundamental result relating tightness to weak convergence is the following:

Theorem 2.3.14 (**Prokhorov's theorem**). *Let* $(W_n)_{n\geq 0}$ *be a sequence of random elements of a complete separable metric space. Then* $(W_n)_{n\geq 0}$ *is tight if and only if every subsequence of* $(W_n)_{n\geq 0}$ *contains a weakly convergent subsubsequence.*

Remark 2.3.15. *Our formulation of Prokhorov's theorem follows immediately from the classical formulation in [85].*

2.4 Martingale theory

As was mentioned in the introduction, much of our approach relies heavily on martingale techniques. In this section, we give a minimal, for the purpose of this thesis, review of martingale theory. Much of this is classical and can be found in [9, 10, 31, 100]. In what follows, $(X, \mathcal{B}, \mu)$ is our underlying probability space. We begin by recalling the definition of conditional expectation [58] and some basic properties which we require.

Theorem 2.4.1. *Let* $Y \in L^1(X)$ *and* $\mathcal{A} \subset \mathcal{B}$ *be a* σ – *algebra on* X. *Then there exists a random variable* Z *such that*

(i) $Z \in L^1(X)$.

(ii) Z *is* $\mathcal{A}$ – *measurable.*

(iii) $\mathbb{E}[Y \mathbb{1}_A] = \mathbb{E}[Z \mathbb{1}_A]$ *for all* $A \in \mathcal{A}$.

Such a random variable Z is unique up to sets of measure zero, and is denoted by $\mathbb{E}[Y \,|\, \mathcal{A}]$.

Proposition 2.4.2. *Let $Y, Z \in L^1(X)$ and $\mathcal{A} \subset \mathcal{B}$ be a σ – algebra. The following hold true:*

(i) *If $Y \geq 0$, then $\mathbb{E}[Y \,|\, \mathcal{A}] \geq 0$ almost surely.*

(ii) *For all $c \in \mathbb{R}$, we have $\mathbb{E}[cY + Z \,|\, \mathcal{A}] = c\mathbb{E}[Y \,|\, \mathcal{A}] + \mathbb{E}[Z \,|\, \mathcal{A}]$ almost surely.*

(iii) *If Y is $\mathcal{A}$ – measurable, then $\mathbb{E}[YZ \,|\, \mathcal{A}] = Y\mathbb{E}[Z \,|\, \mathcal{A}]$ almost surely.*

(iv) $\mathbb{E}[\mathbb{E}[Y \,|\, \mathcal{A}]] = \mathbb{E}[Y]$.

(v) $|\mathbb{E}[Y \,|\, \mathcal{A}]|_p \leq |Y|_p$ *for all $1 \leq p < \infty$.*

(vi) *If $T\colon X \to X$ is measure preserving, then $\mathbb{E}[Y \circ T \,|\, T^{-1}(\mathcal{A})] = \mathbb{E}[Y \,|\, \mathcal{A}] \circ T$ almost surely.*

We next define martingales. Recall that a family of σ – algebras $(\mathcal{B}_n)_{n \geq 0}$ on X is called a *filtration* if $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset \mathcal{B}$ for all $n \geq 0$.

Definition 2.4.3. *A sequence of random variables $(M_n)_{n \geq 0}$ defined on $(X, \mathcal{B}, \mu)$ is called a martingale with respect to the filtration $(\mathcal{B}_n)_{n \geq 0}$ if for all $n \geq 0$, we have*

(i) *M_n is $\mathcal{B}_n$ – measurable.*

(ii) *$M_n \in L^1(X)$.*

(iii) *$\mathbb{E}[M_{n+1} \,|\, \mathcal{B}_n] = M_n$ almost surely.*

We call $(M_n)_{n \geq 0}$ a martingale difference sequence with respect to the filtration $(\mathcal{B}_n)_{n \geq 0}$ if it satisfies (i), (ii), and the additional property that

(iii') *$\mathbb{E}[M_{n+1} \,|\, \mathcal{B}_n] = 0$ almost surely.*

Remark 2.4.4. *The above definition naturally extends to random vectors and random matrices by the requirement that each component is a martingale (respectively martingale difference sequence) in the above sense.*

The next proposition describes the connection between a martingale and a martingale difference sequence.

Proposition 2.4.5. *The following hold true:*

(i) *If $(M_n)_{n\geq 0}$ is a martingale with respect to the filtration $(\mathcal{B}_n)_{n\geq 0}$, then $(Y_n)_{n\geq 0}$ defined by $Y_0 = 0$ and $Y_n = M_n - M_{n-1}$ for $n \geq 1$ is a martingale difference sequence with respect to $(\mathcal{B}_n)_{n\geq 0}$.*

(ii) *If $(Y_n)_{n\geq 0}$ is a martingale difference sequence with respect to $(\mathcal{B}_n)_{n\geq 0}$, then $(M_n)_{n\geq 0}$ defined by $M_n = \sum_{k=0}^n Y_k$ is a martingale with respect to $(\mathcal{B}_n)_{n\geq 0}$.*

Proof. Measurability with respect to the filtration $(\mathcal{B}_n)_{n\geq 0}$ and integrability is immediate in both cases. For (i), note that if $(M_n)_{n\geq 0}$ is a martingale, then for all $n \geq 0$ we have

$$\mathbb{E}\big[Y_{n+1}\,\big|\mathcal{B}_n\big] = \mathbb{E}\big[M_{n+1} - M_n\,\big|\mathcal{B}_n\big] = \mathbb{E}\big[M_{n+1}\,\big|\mathcal{B}_n\big] - \mathbb{E}\big[M_n\,\big|\mathcal{B}_n\big] = M_n - M_n = 0,$$

where the second equality uses Proposition 2.4.2 (ii), and the third equality uses the definition of a martingale and Proposition 2.4.2 (iii). Similarly for (ii), if $(Y_n)_{n\geq 0}$ is a martingale difference sequence, then for all $n \geq 0$ we have

$$\mathbb{E}\big[M_{n+1}\,\big|\mathcal{B}_n\big] = \mathbb{E}\left[Y_{n+1} + \sum_{k=0}^{n} Y_k\,\middle|\mathcal{B}_n\right] = \mathbb{E}\big[Y_{n+1}\,\big|\mathcal{B}_n\big] + \sum_{k=0}^{n} Y_k = M_n,$$

completing the proof. □

The next proposition tells us that any martingale difference sequence satisfies a useful orthogonality property. For our purposes, we state this in the random vector setting.

Proposition 2.4.6. *Let $d \geq 1$. Suppose $(Y_n)_{n\geq 0}$ is a martingale difference sequence taking values in $\mathbb{R}^d$ with respect to the filtration $(\mathcal{B}_n)_{n\geq 0}$. Then $\mathbb{E}[Y_i Y_j^T] = 0$ for $i \neq j$.*

Proof. Suppose first that $j < i$. We have from Proposition 2.4.2 (iv) and (iii) that

$$\mathbb{E}[Y_i Y_j^T] = \mathbb{E}\big[\mathbb{E}\big[Y_i Y_j^T \,\big|\, \mathcal{B}_{i-1}\big]\big] = \mathbb{E}\big[\mathbb{E}\big[Y_i \,\big|\, \mathcal{B}_{i-1}\big] Y_j^T\big] = 0.$$

Similarly, when $i < j$, we have

$$\mathbb{E}[Y_i Y_j^T] = \mathbb{E}\big[\mathbb{E}\big[Y_i Y_j^T \,\big|\, \mathcal{B}_{j-1}\big]\big] = \mathbb{E}\big[Y_i \,\mathbb{E}\big[Y_j^T \,\big|\, \mathcal{B}_{j-1}\big]\big] = \mathbb{E}\big[Y_i \big(\mathbb{E}\big[Y_j \,\big|\, \mathcal{B}_{j-1}\big]\big)^T\big] = 0.$$

This completes the proof. □

We now give some classical inequalities which we require for estimating moments in Section 4.8. We begin with Doob's L^p-inequality [30].

Theorem 2.4.7 (Doob's L^p-inequality). *Suppose $(M_n)_{n\geq 0}$ is a martingale. Then for $p > 1$, we have*

$$\big|\max_{0\leq j\leq n} |M_j|\big|_p \leq \frac{p}{p-1} |M_n|_p \quad \textit{for all } n \geq 0.$$

We next state Burkholder's inequality [17].

Theorem 2.4.8 (Burkholder's inequality). *Suppose $(M_n)_{n\geq 0}$ is a martingale. For each $p > 1$, there exists $C(p) > 0$ such that*

$$|M_n|_p \leq C(p) \left| \left(\sum_{k=0}^{n} |M_k - M_{k-1}|^2 \right)^{1/2} \right|_p \quad \textit{for all } n \geq 0.$$

We conclude with Rio's inequality [86]. The formulation given here is due to [23] (see also [78]).

Theorem 2.4.9 (Rio's inequality). *Let $(X_n)_{n\geq 0} \subset L^p(X)$ be a sequence of random variables for some $p \geq 2$ with $\mathbb{E}[X_n] = 0$ for all $n \geq 0$. Suppose $(\mathcal{B}_n)_{n\geq 0}$ is a filtration for which X_n is $\mathcal{B}_n-$measurable for all $n \geq 0$. Define*

$$b_{\ell,n} = \max_{0\leq \ell \leq m \leq n} \left| X_\ell \sum_{k=\ell}^{m} \mathbb{E}\big[X_k \,\big|\, \mathcal{B}_\ell\big] \right|_{p/2}.$$

Under this setup, there exists $C(p) > 0$ such that

$$\left| \max_{0\le j\le n} \left| \sum_{k=0}^{j} X_k \right| \right|_p \le C(p) \left(\sum_{\ell=0}^{n} b_{\ell,n} \right)^{1/2} \qquad \text{for all } n \ge 0.$$

2.5 The Koopman and transfer operators

The Koopman and transfer operators form key tools when studying the statistical properties of deterministic dynamical systems. In this section, we define these operators and give some immediate consequences which are used throughout. Much of this section is classical, and can be found in many ergodic theory books (see for example [4, 13]). Throughout, we suppose that our underlying probability space is given by $(X, \mathcal{B}, \mu)$.

Definition 2.5.1. *Let $T\colon X \to X$ be a transformation. The Koopman operator $\mathcal{U}\colon L^1(X) \to L^1(X)$ for T is defined by $\mathcal{U}v = v \circ T$.*

Definition 2.5.2. *Let $T\colon X \to X$ be a non-singular transformation. The transfer operator $\mathcal{P}\colon L^1(X) \to L^1(X)$ for T is defined as follows: For $v \in L^1(X)$, we define $\mathcal{P}v$ to be the unique element in $L^1(X)$ which satisfies*

$$\int_X \mathcal{P}v\, w \,\mathrm{d}\mu = \int_X v\, \mathcal{U}w \,\mathrm{d}\mu \qquad \text{for all } w \in L^\infty(X).$$

We next give some basic properties of these operators.

Proposition 2.5.3. *Let $T\colon X \to X$ be a non-singular transformation with associated Koopman and transfer operators $\mathcal{U}$ and $\mathcal{P}$ respectively. The following hold true:*

(i) *$\mathcal{U}, \mathcal{P}\colon L^1(X) \to L^1(X)$ are linear operators with $\mathcal{U}1 = 1$.*

(ii) *$\int_X \mathcal{P}v \,\mathrm{d}\mu = \int_X v \,\mathrm{d}\mu$ for all $v \in L^1(X)$.*

(iii) *$\int_X \mathcal{P}^n v\, w \,\mathrm{d}\mu = \int_X v\, \mathcal{U}^n w \,\mathrm{d}\mu$ for all $n \ge 1$, $v \in L^1(X)$, and $w \in L^\infty(X)$.*

(iv) *$|\mathcal{U}w|_\infty \le |w|_\infty$ for all $w \in L^\infty(X)$ and $|\mathcal{P}v|_1 \le |v|_1$ for all $v \in L^1(X)$.*

If T is measure preserving, then in addition to the above, we have

(v) $\mathcal{P}1 = 1$.

(vi) $\int_X \mathcal{U}v \, \mathrm{d}\mu = \int_X v \, \mathrm{d}\mu$ *for all* $v \in L^1(X)$.

(vii) $\mathcal{P}\mathcal{U}v = v$ *and* $\mathcal{U}\mathcal{P}v = \mathbb{E}[v \,|\, T^{-1}(\mathcal{B})]$ *for all* $v \in L^1(X)$.

(viii) $|\mathcal{U}v|_p = |v|_p$ *and* $|\mathcal{P}v|_p \le |v|_p$ *for all* $1 \le p \le \infty$ *and* $v \in L^p(X)$.

Proof. We prove (iv), (vii), and (viii). The other assertions are immediate from the definitions of $\mathcal{U}$, $\mathcal{P}$, and invariance of the measure. We begin with (iv). Since $T(X) \subset X$, we have $|\mathcal{U}w|_\infty = |w \circ T|_\infty \le |w|_\infty$ for all $w \in L^\infty(X)$. Next note that for $v \in L^1(X)$, we have

$$\begin{aligned} |\mathcal{P}v|_1 &= \sup_{\substack{w \in L^\infty(Y) \\ |w|_\infty \le 1}} \left\{ \left| \int_Y \mathcal{P}v \, w \, \mathrm{d}\mu \right| \right\} = \sup_{\substack{w \in L^\infty(Y) \\ |w|_\infty \le 1}} \left\{ \left| \int_Y v \, \mathcal{U}w \, \mathrm{d}\mu \right| \right\} \\ &\le \sup_{\substack{w \in L^\infty(Y) \\ |w|_\infty \le 1}} \left\{ \int_Y |v||\mathcal{U}w| \, \mathrm{d}\mu \right\} \le |v|_1, \end{aligned}$$

proving (iv).

We next prove (vii). We have for $v \in L^1(X)$ and $w \in L^\infty(X)$ that

$$\int_X \mathcal{P}(\mathcal{U}v) \, w \, \mathrm{d}\mu = \int_X \mathcal{U}v \, \mathcal{U}w \, \mathrm{d}\mu = \int_X \mathcal{U}(vw) \, \mathrm{d}\mu = \int_X vw \, \mathrm{d}\mu.$$

Since w is arbitrary, $\mathcal{P}\mathcal{U}v = v$ as required. To show that $\mathcal{U}\mathcal{P}v = \mathbb{E}[v \,|\, T^{-1}(\mathcal{B})]$, we first note that $\mathcal{U}\mathcal{P}v$ is $T^{-1}(\mathcal{B})-$measurable. Indeed, for $A \subset \mathbb{R}$ Borel measurable, we have

$$(\mathcal{U}\mathcal{P}v)^{-1}(A) = (\mathcal{P}v \circ T)^{-1}(A) = T^{-1}\big((\mathcal{P}v)^{-1}(A)\big) \in T^{-1}(\mathcal{B})$$

since $\mathcal{P}v \in L^1(X)$ is measurable. Moreover, we have for $T^{-1}(B) \in T^{-1}(\mathcal{B})$ that

$$\begin{aligned} \int_{T^{-1}(B)} \mathcal{U}\mathcal{P}v \, \mathrm{d}\mu &= \int_X (\mathcal{P}v \circ T) \, \mathbb{1}_{T^{-1}(B)} \, \mathrm{d}\mu = \int_X (\mathcal{P}v \circ T) \, (\mathbb{1}_B \circ T) \, \mathrm{d}\mu \\ &= \int_X \mathcal{P}v \, \mathbb{1}_B \, \mathrm{d}\mu = \int_X v \, \mathbb{1}_B \circ T \, \mathrm{d}\mu = \int_{T^{-1}(B)} v \, \mathrm{d}\mu, \end{aligned}$$

which proves (vii).

For (viii), suppose first that $1 \le p < \infty$ and $v \in L^p(X)$. By $\mu-$invariance of T, we have $|\mathcal{U}v|_p = |v|_p$. To show that $|\mathcal{P}v|_p \le |v|_p$, note that by Proposition 2.4.2 (v), we have $|\mathbb{E}[v \,|\, T^{-1}(\mathcal{B})]|_p \le |v|_p$. Therefore

$$\begin{aligned}|\mathcal{P}v|_p^p &= \int_X |\mathcal{P}v|^p \,\mathrm{d}\mu = \int_X |\mathcal{U}\mathcal{P}v|^p \,\mathrm{d}\mu = \int_X \big|\mathbb{E}\big[v \,\big|\, T^{-1}(\mathcal{B})\big]\big|^p \,\mathrm{d}\mu \\ &\le \int_X |v|^p \,\mathrm{d}\mu = |v|_p^p,\end{aligned}$$

and the result follows. We now suppose that $p = \infty$ and $v \in L^\infty(X)$. Note that for all $M \ge 0$, we have $(|v| \circ T)^{-1}((M,\infty)) = T^{-1}(|v|^{-1}((M,\infty)))$. Therefore, by $T-$invariance of μ, we have

$$\begin{aligned}|\mathcal{U}v|_\infty &= \inf\big\{M \ge 0 \mid \mu\big((|v| \circ T)^{-1}\big((M,\infty)\big)\big) = 0\big\} \\ &= \inf\big\{M \ge 0 \mid \mu\big(|v|^{-1}\big((M,\infty)\big)\big) = 0\big\} = |v|_\infty.\end{aligned}$$

To show $|\mathcal{P}v|_\infty \le |v|_\infty$, note that since $\pm v \le |v|_\infty$, we have by Proposition 2.4.2 (i) that $|\mathbb{E}[v \,|\, T^{-1}(\mathcal{B})]| \le |v|_\infty$. Therefore

$$|\mathcal{P}v|_\infty = |\mathcal{U}\mathcal{P}v|_\infty = \big|\mathbb{E}\big[v \,\big|\, T^{-1}(\mathcal{B})\big]\big|_\infty \le |v|_\infty,$$

completing the proof. □

Example 2.5.4. *Consider the doubling map $T\colon [0,1] \to [0,1]$ given by $Tx = 2x \mod 1$, as shown in Figure 2.2. The Lebesgue measure μ on $[0,1]$ is ergodic and invariant for T [13]. Let $\mathcal{P}$ denote the transfer operator for T. Note that for $v \in L^1([0,1])$ and $w \in L^\infty([0,1])$, we have*

$$\begin{aligned}\int_0^1 \mathcal{P}v\, w \,\mathrm{d}\mu &= \int_0^1 v(x)\, w(Tx) \,\mathrm{d}x = \int_0^{1/2} v(x)\, w(2x) \,\mathrm{d}x + \int_{1/2}^1 v(x)\, w(2x-1) \,\mathrm{d}x \\ &= \frac{1}{2}\int_0^1 v\left(\frac{y}{2}\right) w(y) \,\mathrm{d}y + \frac{1}{2}\int_0^1 v\left(\frac{y+1}{2}\right) w(y) \,\mathrm{d}y,\end{aligned}$$

so that

$$\mathcal{P}v(x) = \frac{1}{2}v\left(\frac{x}{2}\right) + \frac{1}{2}v\left(\frac{x+1}{2}\right).$$

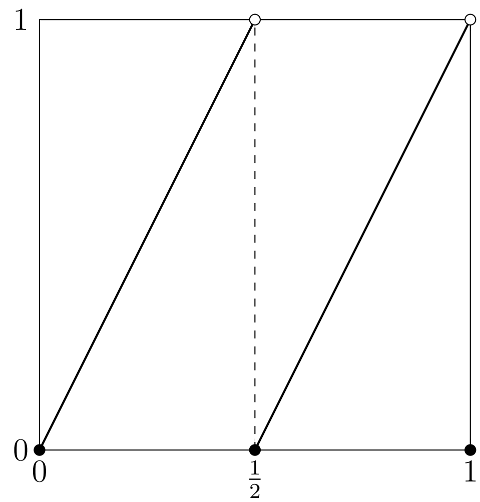


Figure 2.2: Doubling map

Recall that a measure μ on $\mathcal{B}$ is absolutely continuous with respect to a measure ρ on $\mathcal{B}$ if $\rho(B) = 0$ implies $\mu(B) = 0$ for all $B \in \mathcal{B}$. We denote by $\mathrm{d}\mu/\,\mathrm{d}\rho$ the Radon-Nikodym derivative of μ with respect to ρ. One of the major uses of the transfer operator is the construction of absolutely continuous invariant probability measures. Moreover, as we see in Section 3.3, it allows us to establish various useful properties of these measures and their densities. On this note, we conclude this section by giving a criterion to identify such densities.

Proposition 2.5.5. *Let $T\colon X \to X$ be a non-singular transformation with transfer operator $\mathcal{P}\colon L^1(X) \to L^1(X)$ and underlying probability measure ρ. Suppose $f \in L^1(X)$ with $|f|_1 = 1$ and $f \geq 0$. Then $\mathcal{P}f = f$ if and only if the measure μ given by $\mathrm{d}\mu = f\,\mathrm{d}\rho$ is $T-$invariant.*

Proof. Let us suppose that μ is $T-$invariant. Then for any $B \in \mathcal{B}$, noting that $\mathbb{1}_{T^{-1}(B)} = \mathbb{1}_B \circ T$, we have

$$\int_B f\,\mathrm{d}\rho = \int_{T^{-1}(B)} f\,\mathrm{d}\rho = \int_X f\,\mathbb{1}_B \circ T\,\mathrm{d}\rho = \int_X \mathcal{P}f\,\mathbb{1}_B\,\mathrm{d}\rho = \int_B \mathcal{P}f\,\mathrm{d}\rho.$$

Since $B \in \mathcal{B}$ is arbitrary, $\mathcal{P}f = f$.

Conversely, let us suppose that $\mathcal{P}f = f$. Then by the above calculation, for any $B \in \mathcal{B}$ we have

$$\int_{T^{-1}(B)} f \,\mathrm{d}\rho = \int_B \mathcal{P}f \,\mathrm{d}\rho = \int_B f \,\mathrm{d}\rho,$$

so that μ is $T-$invariant. $\square$

Remark 2.5.6. *As one would expect, the Koopman and transfer operators extend component-wise to $\mathbb{R}^d$ and $\mathbb{R}^{d,d}-$valued observables.*

2.6 Peter-Weyl theorem

In this section we state the Peter-Weyl theorem [34], which gives us an explicit orthonormal basis of $L^2(G)$ when G is a compact group. This is required in Section 4.3, when we study the transfer operator for compact group extensions.

We begin with some preliminary definitions, which can be found in [14, 34]. Let (π, V) be a finite-dimensional complex representation of G with inner product $[\cdot,\cdot]$ on V. If G preserves $[\cdot,\cdot]$, then we say that π is a *unitary representation*. We say that π is *irreducible* if the only closed subspaces $W \subset V$ such that $\pi(G)W \subset W$ are $W = \{0\}$ and $W = V$. Finally, two representations (π, V_π) and (ρ, V_ρ) of a compact group G are *equivalent* if there exists an isomorphism $\mathcal{A}\colon V_\pi \to V_\rho$ such that $\mathcal{A}\pi(g) = \rho(g)\mathcal{A}$.

Theorem 2.6.1 (Peter-Weyl theorem). *Let Σ denote the equivalence classes of irreducible unitary representations of the compact group G. Suppose that a representative π is chosen from each equivalence class, and let $u^{(\pi)}_{i,j}(g) = [\pi(g)e_i, e_j]$ denote the matrix coefficients of π in an orthonormal basis of V. Letting $d^{(\pi)} = \dim(V_\pi)$ denote the degree of the representation π, we have that the set of functions*

$$\left\{\sqrt{d^{(\pi)}}u^{(\pi)}_{i,j} \mid \pi \in \Sigma, 1 \le i, j \le d^{(\pi)}\right\}$$

form an orthonormal basis of $L^2(G)$.

Example 2.6.2. *Consider the group $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. In this case, the irreducible representations are one-dimensional and given by $\pi_n(e^{i\theta}) = e^{in\theta}$. There is a single matrix coefficient for each representation, which is given by the function $u_n(e^{i\theta}) = e^{in\theta}$. The Peter-Weyl theorem then states that these functions form an orthonormal basis of $L^2(\mathbb{T})$ — a standard result from Fourier theory.*

2.7 Separation of spectrum

Probabilistic limit laws for deterministic dynamical systems often follow from good spectral properties of the associated transfer operator (see for example [43]). In this section, we give a brief overview of the spectral theory we require throughout the thesis. A major reference on this topic is [50].

Let X be a Banach space over the complex scalar field $\mathbb{C}$ and let I be the identity operator on X. Let $\mathcal{P}\colon X \to X$ be a bounded linear operator. The *spectrum* of $\mathcal{P}$, denoted $\sigma(\mathcal{P})$, is defined by $\sigma(\mathcal{P}) = \{z \in \mathbb{C} \mid zI - \mathcal{P} \text{ is not invertible.}\}$. The *spectral radius* of $\mathcal{P}$, denoted $r(\mathcal{P})$, is defined by $r(\mathcal{P}) = \sup\{|z| : z \in \sigma(\mathcal{P})\}$. The spectral radius formula says that $r(\mathcal{P}) = \lim_{n\to\infty} \|\mathcal{P}^n\|^{1/n}$, where $\|\cdot\|$ denotes the operator norm.

The following theorem from [50] is required in Section 4.5. Informally, it says that if we can separate the spectrum of some bounded linear operator, then the underlying Banach space can also be separated in a convenient way.

Theorem 2.7.1. *Let $\mathcal{B}(X)$ denote the space of bounded linear operators on the Banach space X and let $\mathcal{P} \in \mathcal{B}(X)$. Suppose $\sigma(\mathcal{P}) = \sigma_{in} \cup \sigma_{ext}$, where σ_{in} and σ_{ext} are compact and disjoint. Let $\mathcal{C}$ be a smooth closed curve which does not intersect $\sigma(\mathcal{P})$, and which contains σ_{in} in its interior and σ_{ext} in its exterior. Then the following hold true:*

(i) $\pi = \frac{1}{2\pi} \int_{\mathcal{C}} (zI - \mathcal{P})^{-1}\, \mathrm{d}z \in \mathcal{B}(X)$ *is a projection. That is,* $\pi^2 = \pi$, *so that* $X = \operatorname{Im}\pi \oplus \ker \pi$.

(ii) $\pi \circ \mathcal{P} = \mathcal{P} \circ \pi$, *so that* $\mathcal{P}(\operatorname{Im}\pi) \subset \operatorname{Im}\pi$ *and* $\mathcal{P}(\ker \pi) \subset \ker \pi$.

(iii) $\sigma(\mathcal{P}|_{\operatorname{Im}\pi}) = \sigma_{in}$ *and* $\sigma(\mathcal{P}|_{\ker\pi}) = \sigma_{ext}$.

Definition 2.7.2. *We refer to* π *defined in (i) above as the spectral projection of* σ_{in}*. It is independent of the choice of* $\mathcal{C}$*, since* $(zI - \mathcal{P})^{-1}$ *is holomorphic in* $\mathbb{C} \setminus \sigma(\mathcal{P})$.

Remark 2.7.3. *If* $z \in \sigma(\mathcal{P})$ *is an isolated eigenvalue, then we can separate the spectrum and define the projection* π_z *as above. The multiplicity of* z *is defined as* $\dim(\operatorname{Im}\pi_z)$*. If this is finite, then* $\operatorname{Im}\pi_z$ *coincides with the generalised eigenspace corresponding to* z.

Let us denote by $\sigma_{\text{ess}}(\mathcal{P})$ the *essential spectrum* of $\mathcal{P}$. That is, $\sigma_{\text{ess}}(\mathcal{P})$ consists of those points in $\sigma(\mathcal{P})$ which are not isolated eigenvalues of finite multiplicity. We denote by $r_{\text{ess}}(\mathcal{P}) = \sup\{|z| : z \in \sigma_{\text{ess}}(\mathcal{P})\}$ the *essential spectral radius* of $\mathcal{P}$. For any $\epsilon > 0$, one can think of $\mathcal{P}$ as a finite matrix outside of $\{z \in \mathbb{C} : |z| \leq r_{\text{ess}}(\mathcal{P}) + \epsilon\}$ and something more complicated inside this set. To conclude this section, we introduce the following criteria [46] (c.f. [26, 62, 95]) which allows us to estimate $r_{\text{ess}}(\mathcal{P})$. The formulation given here is adapted from [64].

Proposition 2.7.4. *Suppose we have the following setup:*

(i) *Two Banach spaces* $X_1 \subset X_2$ *with norms* $\|\cdot\|_{X_1}$ *and* $\|\cdot\|_{X_2}$ *satisfying* $\|\cdot\|_{X_2} \leq \|\cdot\|_{X_1}$.

(ii) *An operator* $\mathcal{P}\colon X_1 \to X_1$ *and constants* $C > 0$ *and* $\theta \in (0,1)$ *satisfying*

$$\|\mathcal{P}^n v\|_{X_2} \leq C\|v\|_{X_2}$$

and

$$\|\mathcal{P}^n v\|_{X_1} \leq C\big(\theta^n\|v\|_{X_1} + \|v\|_{X_2}\big)$$

for all $v \in X_1$ *and all* $n \geq 1$.

(iii) *The unit ball of* X_1 *is relatively compact in* X_2.

Then $r_{\text{ess}}(\mathcal{P}) \leq \theta$.

Remark 2.7.5. *We refer to the inequality* $\|\mathcal{P}^n v\|_{X_1} \leq C(\theta^n\|v\|_{X_1} + \|v\|_{X_2})$ *as a Lasota-Yorke inequality.*

2.8 Martingale-coboundary decomposition

In this section, we give an abstract formulation of the main method used in the thesis. It is assumed that $(X, \mathcal{B}, \mu, T)$ is an ergodic measure preserving dynamical system with associated transfer operator $\mathcal{P}$ and $d \geq 1$. We call an integrable function $v\colon X \to \mathbb{R}^d$ an *observable*. We consider the sequence of functions $(v \circ T^k)_{k \geq 0}$, which resemble a stochastic process when our underlying dynamical system is "sufficiently chaotic".

Remark 2.8.1. *Note that for $k \geq 0$ and $A \subset \mathbb{R}^d$ Borel measurable, we have $(v \circ T^k)^{-1}(A) = T^{-k}(v^{-1}(A))$. Therefore, by μ – invariance of T, we have*

$$\mu\big((v \circ T^k)^{-1}(A)\big) = \mu\big((v \circ T^{k-1})^{-1}(A)\big) = \cdots = \mu\big(v^{-1}(A)\big),$$

so that the sequence $(v \circ T^k)_{k \geq 0}$ is identically distributed. However, in general, this sequence is not independent.

Birkhoff's ergodic theorem tells us the sequence $(v \circ T^k)_{k \geq 0}$ satisfies the strong law of large numbers. It is natural to ask whether we can deduce stronger information about the limiting behaviour of this sequence, such as a central limit theorem or weak invariance principle. To answer such questions, we introduce the method of Gordin [37], which decomposes observables into a sum of a martingale and an asymptotically negligible coboundary.

Definition 2.8.2. *We say that an observable $v \in L^1(X; \mathbb{R}^d)$ admits a martingale-coboundary decomposition if there exist $m \in L^1(X; \mathbb{R}^d)$ and $\chi\colon X \to \mathbb{R}^d$ measurable such that $v = m + \chi \circ T - \chi$ and $m \in \ker \mathcal{P}$.*

We next justify why we call $v = m + \chi \circ T - \chi$ a martingale-coboundary decomposition.

Proposition 2.8.3. *For $n \geq 1$ and $1 \leq k \leq n$, define $\mathcal{B}_{n,k} = T^{-(n-k)}(\mathcal{B})$. Suppose $m \in L^1(X; \mathbb{R}^d)$ with $\mathcal{P}m = 0$. Then $(m \circ T^{n-k}, \mathcal{B}_{n,k})_{k=1}^n$ is a sequence of martingale differences.*

Proof. Fix $n \geq 1$. Note that $T^{-1}(\mathcal{B}) \subset \mathcal{B}$. Therefore for $1 \leq k \leq n-1$, we have

$$\mathcal{B}_{n,k} = T^{-(n-k)}(\mathcal{B}) \subset T^{-(n-k)+1}(\mathcal{B}) = T^{-(n-(k+1))}(\mathcal{B}) = \mathcal{B}_{n,k+1}.$$

Now, observe that if $A \subset \mathbb{R}^d$ is Borel measurable, then

$$(m \circ T^{n-k})^{-1}(A) = T^{-(n-k)}\big(m^{-1}(A)\big) \in \mathcal{B}_{n,k}, \tag{2.8.1}$$

so that $m \circ T^{n-k}$ is $\mathcal{B}_{n,k}-$measurable. Integrability is immediate by invariance of T. Finally, letting $\mathcal{U}$ denote the Koopman operator for T, we have from Proposition 2.4.2 (vi) and Proposition 2.5.3 (vii) that

$$\mathbb{E}\big[m \circ T^{n-k} \,\big|\, \mathcal{B}_{n,k-1}\big] = \mathbb{E}\big[m \,\big|\, T^{-1}(\mathcal{B})\big] \circ T^{n-k} = (\mathcal{U}\mathcal{P}m) \circ T^{n-k} = 0,$$

where the final equality follows from the fact that $m \in \ker \mathcal{P}$. □

Remark 2.8.4. *By similar arguments, one can show that $(T^{-n}(\mathcal{B}))_{n\geq 0}$ is a non-increasing filtration, $m\circ T^n$ is $T^{-n}(\mathcal{B})$ – measurable, and $\mathbb{E}[m\circ T^n \mid T^{-(n+1)}(\mathcal{B})] = 0$ for all $n \geq 0$ almost surely.*

Example 2.8.5. *Recall from Example 2.5.4 that the doubling map $T\colon [0,1] \to [0,1]$ defined by $Tx = 2x \bmod 1$ has transfer operator $\mathcal{P}\colon L^1([0,1]) \to L^1([0,1])$ given by*

$$\mathcal{P}v(x) = \frac{1}{2}v\left(\frac{x}{2}\right) + \frac{1}{2}v\left(\frac{x+1}{2}\right) \quad \textit{for } v \in L^1([0,1]).$$

We show that if $v\colon [0,1] \to \mathbb{R}$ is Lipschitz with $\int_0^1 v(x)\,\mathrm{d}x = 0$, then v admits a martingale-coboundary decomposition. Note first that for $x, y \in [0,1]$, we have

$$\begin{aligned}\big|\mathcal{P}v(x) - \mathcal{P}v(y)\big| &\leq \frac{1}{2}\left|v\left(\frac{x}{2}\right) - v\left(\frac{y}{2}\right)\right| + \frac{1}{2}\left|v\left(\frac{x+1}{2}\right) - v\left(\frac{y+1}{2}\right)\right| \\ &\leq \frac{1}{2}\operatorname{Lip}(v)\left|\frac{x}{2} - \frac{y}{2}\right| + \frac{1}{2}\operatorname{Lip}(v)\left|\frac{x+1}{2} - \frac{y+1}{2}\right| \\ &= \frac{1}{2}\operatorname{Lip}(v)|x-y|,\end{aligned}$$

so that $\operatorname{Lip}(\mathcal{P}v) \leq \operatorname{Lip}(v)/2$. Inductively, one has

$$\operatorname{Lip}(\mathcal{P}^n v) \leq \frac{1}{2^n}\operatorname{Lip}(v) \quad \textit{for all } n \geq 1. \tag{2.8.2}$$

Next, note from Proposition 2.5.3 (ii) that $\int_0^1 \mathcal{P}^n v(y)\,\mathrm{d}y = \int_0^1 v(y)\,\mathrm{d}y = 0$ *for all* $n \geq 1$*. Therefore, for* $x \in [0,1]$ *we have*

$$\begin{aligned}\big|\mathcal{P}^n v(x)\big| &= \left|\mathcal{P}^n v(x) - \int_0^1 \mathcal{P}^n v(y)\,\mathrm{d}y\right| \\ &\leq \int_0^1 \big|\mathcal{P}^n v(x) - \mathcal{P}^n v(y)\big|\,\mathrm{d}y \\ &\leq \frac{1}{2^n}\operatorname{Lip}(v),\end{aligned}$$

where the final inequality uses (2.8.2) *and the fact that* $|x-y| \leq 1$ *for* $x, y \in [0,1]$*. It follows that* $|\mathcal{P}^n v|_\infty \leq \operatorname{Lip}(v)/2^n$ *for all* $n \geq 1$*, and so*

$$\|\mathcal{P}^n v\|_{\mathrm{Lip}} = \operatorname{Lip}(\mathcal{P}^n v) + |P^n v|_\infty \leq \frac{1}{2^{n-1}}\operatorname{Lip}(v) \quad \text{for all } n \geq 1.$$

Define $\chi, m\colon [0,1] \to \mathbb{R}$ *by*

$$\chi = \sum_{n=1}^{\infty} \mathcal{P}^n v$$

and

$$m = v + \chi - \chi \circ T.$$

Note that

$$\|\chi\|_{\mathrm{Lip}} \leq \sum_{n=1}^{\infty} \|\mathcal{P}^n v\|_{\mathrm{Lip}} \leq \sum_{n=1}^{\infty} \frac{1}{2^{n-1}}\operatorname{Lip}(v) = 2\operatorname{Lip}(v),$$

so that χ *is well-defined and Lipschitz. Also note that*

$$|m|_\infty \leq |v|_\infty + |\chi|_\infty + |\chi \circ T|_\infty \leq |v|_\infty + 2|\chi|_\infty \leq \|v\|_{\mathrm{Lip}} + 2\|\chi\|_{\mathrm{Lip}} < \infty,$$

so that $m \in L^\infty([0,1])$*. Finally, we have from Proposition 2.5.3 (vii) that* $\mathcal{P}(\chi \circ T) = \chi$*, so that*

$$\mathcal{P}m = \mathcal{P}v + \mathcal{P}\chi - \mathcal{P}(\chi \circ T) = \mathcal{P}v + \sum_{n=2}^{\infty} \mathcal{P}^n v - \sum_{n=1}^{\infty} \mathcal{P}^n v = 0.$$

Thus v *admits a martingale-coboundary decomposition as claimed.*

We next move on to martingale limit theorems. We begin with a central limit theorem for sequences of martingale differences. By Theorem 2.3.8, the proof is an immediate consequence of [32] (see also [69]).

Theorem 2.8.6. *Suppose that* $m \in L^2(X;\mathbb{R}^d)$ *with* $\mathcal{P}m = 0$. *Write* $\Sigma = \int_X mm^T \, \mathrm{d}\mu \in \mathbb{R}^{d,d}$. *Then*

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} m \circ T^k \to_w \mathcal{N}(0, \Sigma).$$

The next result is a weak invariance principle from [10] for sequences of martingale differences. The form given here is a special case of [60, Theorem A.1], where the authors allow the probability space and underlying dynamics to vary with each iterate. For our purposes, we keep fixed the probability space but let the transformation vary with each iterate.

Theorem 2.8.7. *Let* $(X, \mathcal{B}, \mu, T_n)$ *be a sequence of ergodic measure preserving transformations with Koopman and transfer operators* $\mathcal{U}_n$ *and* $\mathcal{P}_n$ *respectively. Suppose that* $m_n \in L^2(X;\mathbb{R}^d)$ *with* $\mathcal{P}_n m_n = 0$. *For* $n \geq 1$, *define the random elements* $M_n \colon X \to D([0,\infty);\mathbb{R}^d)$ *by* $M_n(t) = n^{-1/2} \sum_{k=0}^{[nt]-1} m_n \circ T_n^k$ *for* $t \geq 0$. *Suppose that the family* $(|m_n|^2)_{n\geq 0}$ *is uniformly integrable, and suppose there exists a constant matrix* $\Sigma \in \mathbb{R}^{d,d}$ *such that for each* $t \geq 0$, *we have*

$$\frac{1}{n} \sum_{k=0}^{[nt]-1} \mathcal{U}_n \mathcal{P}_n (m_n m_n^T) \circ T_n^k \to_w t\Sigma.$$

Then $M_n \to_w W$ *in* $D([0,\infty);\mathbb{R}^d)$, *where* W *is the Brownian motion with mean* 0 *and covariance matrix* Σ.

Remark 2.8.8. *The proof of Theorem 2.8.7 in [60] is a standard argument in probability theory. Namely, by Prokhorov's theorem, weak convergence of* M_n *to* W *is equivalent to showing convergence of the associated finite-dimensional distributions and tightness of* $(M_n)_{n\geq 0}$ *in* $D([0,\infty);\mathbb{R}^d)$ *(see [10, Example 5.1]).*

To conclude this section, we state the results from [22, Section 2] which we require. These are known as *almost sure invariance principles.* We return to the setup of $(X, \mathcal{B}, \mu, T)$ being an ergodic measure preserving transformation with transfer operator $\mathcal{P}$, and take $d = 1$. Let $m \in L^p(X)$ for some $p \geq 2$ and let $\sigma^2 = \int_X m^2 \,\mathrm{d}\mu$. Suppose that $m \in \ker \mathcal{P}$. By Remark 2.8.4, $(\mathcal{B}_n)_{n\geq 0} = (T^{-n}(\mathcal{B}))_{n\geq 0}$ is a non-increasing filtration for which $m \circ T^n$ is $\mathcal{B}_n-$measurable and $\mathbb{E}[m \circ T^n \,|\, \mathcal{B}_{n+1}] = 0$ for all $n \geq 0$ almost surely.

Theorem 2.8.9. *If $p = 2$, then there exists a probability space supporting a sequence of random variables $(S_n)_{n\geq 1}$ with the same joint distributions as $(\sum_{k=0}^{n-1} m \circ T^k)_{n\geq 1}$ and a sequence of independent and identically distributed random variables $(Z_n)_{n\geq 1}$ with distribution $\mathcal{N}(0, \sigma^2)$, such that*

$$\sup_{1\leq k\leq n} \left| S_k - \sum_{j=1}^{k} Z_j \right| = o\big((n \log\log n)^{1/2}\big) \qquad a.s.$$

Theorem 2.8.10. *Suppose that $2 < p < 4$ and*

$$\sum_{k=0}^{n-1} \Big(\mathbb{E}\big[m^2 \circ T^k \,\big|\, \mathcal{B}_{k+1}\big] - \sigma^2 \Big) = o(n^{2/p}) \qquad a.s.$$

Then there exists a probability space supporting a sequence of random variables $(S_n)_{n\geq 1}$ with the same joint distributions as $(\sum_{k=0}^{n-1} m \circ T^k)_{n\geq 1}$ and a sequence of independent and identically distributed random variables $(Z_n)_{n\geq 1}$ with distribution $\mathcal{N}(0, \sigma^2)$, such that

$$\sup_{1\leq k\leq n} \left| S_k - \sum_{j=1}^{k} Z_j \right| = o(n^{1/p} (\log n)^{1/2}) \qquad a.s.$$

Theorem 2.8.11. *Suppose that $p \geq 4$ and*

$$\sum_{k=0}^{n-1} \Big(\mathbb{E}\big[m^2 \circ T^k \,\big|\, \mathcal{B}_{k+1}\big] - \sigma^2 \Big) = O\big((n \log\log n)^{1/2}\big) \qquad a.s.$$

Then there exists a probability space supporting a sequence of random variables $(S_n)_{n\geq 1}$ with the same joint distributions as $(\sum_{k=0}^{n-1} m \circ T^k)_{n\geq 1}$ and a sequence of *independent and identically distributed random variables* $(Z_n)_{n\geq 1}$ *with distribution* $\mathcal{N}(0,\sigma^2)$, *such that*

$$\sup_{1\leq k\leq n}\left|S_k-\sum_{j=1}^{k}Z_j\right|=O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})\qquad a.s.$$

Chapter 3

Non-uniformly expanding dynamical systems

3.1 Outline

In this chapter, we introduce non-uniformly expanding dynamical systems and show that they admit an ergodic absolutely continuous invariant probability measure. Roughly speaking, these can be thought of systems for which expansion holds only on a subset of the state space, and its onset is non-uniform in time. A precise definition and some examples which are referred to throughout are given in Section 3.2. By an inducing scheme, one obtains a piecewise uniformly expanding dynamical system which belong to a class of maps called Gibbs-Markov maps [1]. In Section 3.3 we prove the existence of an ergodic absolutely continuous invariant probability measure for such maps, whose densities have "good" regularity. The approach we use is by now standard [61, 101, 102], however introduces techniques which are utilised throughout. Finally, in Section 3.4, we introduce the notion of Young towers [101, 102]. By representing our non-uniformly expanding system T as a Young tower over its induced Gibbs-Markov map, we see how to extend the measure constructed in Section 3.3 to obtain an ergodic absolutely continuous invariant probability measure for T.

3.2 Definition and examples

Let (X, d) be a bounded metric space with Borel probability measure ρ and let $T\colon X \to X$ be a non-singular transformation. Without loss of generality, we suppose that $\operatorname{diam}(X) = \sup\{d(x,y) \mid x, y \in X\} = 1$. Let $Y \subset X$ be a subset of positive measure, and let α be a countable measurable partition of Y (mod 0) with $\rho(a) > 0$ for all $a \in \alpha$. We suppose that there is an integrable return time $\tau\colon Y \to \mathbb{Z}^+$ which is constant on each $a \in \alpha$. We also suppose there are constants $\lambda > 1$, $\eta \in (0, 1]$, and $C_0, C_1 \geq 1$ such that for all $a \in \alpha$, the following hold:

(i) The map $F = T^\tau$ restricts to a bijection from a onto Y with measurable inverse.

(ii) $d(Fx, Fy) \geq \lambda\, d(x, y)$ for all $x, y \in a$.

(iii) $d(T^\ell x, T^\ell y) \leq C_0\, d(Fx, Fy)$ for all $x, y \in a$ and $0 \leq \ell < \tau(a)$.

(iv) $\zeta_0 = \frac{\mathrm{d}\rho}{\mathrm{d}\rho \circ F}$ satisfies $|\log \zeta_0(x) - \log \zeta_0(y)| \leq C_1\, d(Fx, Fy)^\eta$ for all $x, y \in a$.

Such a dynamical system $T\colon X \to X$ is called *non-uniformly expanding*, with induced map $F = T^\tau\colon Y \to Y$. We refer to condition (ii) as *expansiveness* and condition (iv) as *bounded distortion*.

Example 3.2.1. *The simplest class of examples are (piecewise) uniformly expanding maps, which are non-uniformly expanding with $\tau = 1$. The doubling map from Example 2.5.4 is uniformly expanding. Another example is the Gauss map $T\colon [0,1] \to [0,1]$ which is defined by*

$$T(x) = \begin{cases} 1/x - [1/x] & \text{if } x \in (0,1], \\ 0 & \text{if } x = 0, \end{cases}$$

where $[1/x]$ denotes the integer part of $1/x$. This is shown in Figure 3.1.

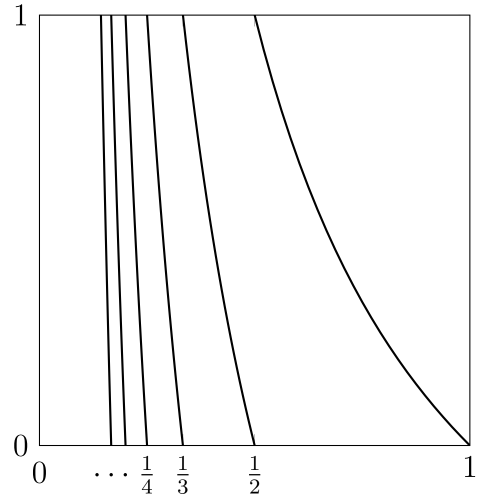


Figure 3.1: Gauss map

Example 3.2.2. *Let $X = [0,1]$ and $\gamma \in [0,1)$. Recall the intermittent map in Example 2.2.6 defined by*

$$T(x) = \begin{cases} x(1+2^\gamma x^\gamma) & \text{if } x \in [0,1/2], \\ 2x-1 & \text{if } x \in (1/2,1]. \end{cases}$$

This is non-uniformly expanding. From Figure 2.1, we see that uniform expansion occurs everywhere except at the neutral fixed point at 0. In Example 2.2.6 we explicitly constructed an inducing scheme as described above with first return time τ and partition $\alpha = (a_n)_{n\geq 1}$ of $Y = (1/2,1]$.

Example 3.2.3. *Let $X = [-1,1]$ and $a \in [0,2]$. Unimodal maps $T\colon X \to X$ defined by $T(x) = 1 - ax^2$ which satisfy the Collet-Eckmann condition [20] are non-uniformly expanding [101]. Recall the Collett-Eckman condition says that there are constants $b, c > 0$ such that $|(T^n)'(1)| \geq ce^{bn}$ for all $n \geq 1$. By [5, 49] this condition holds for a set of parameters a with positive Lebesgue measure.*

Example 3.2.4. *In [97], Viana introduced a C^3 – open class of multidimensional non-uniformly expanding maps. To be definite, we restrict attention to maps on*

$X = \mathbb{R}/\mathbb{Z} \times \mathbb{R}$. *Let* $\lambda \in \mathbb{Z}$ *with* $\lambda \geq 16$ *and define* $T_0\colon X \to X$ *by*

$$T_0(\theta, y) = (\lambda\theta \bmod 1, a_0 + a\sin(2\pi\theta) - y^2),$$

where $a_0 \in (1,2)$ *is chosen so that* 0 *is a preperiodic point for the quadratic map* $y \mapsto a_0 - y^2$, *and* a *is sufficiently small. Then* C^3 *–maps sufficiently close to* T_0 *are non-uniformly expanding [3].*

3.3 Existence of ergodic absolutely continuous invariant probability measures for Gibbs-Markov maps

Let (Y, ρ_Y) be a probability space and suppose that α is a countable measurable partition of Y with $\rho_Y(a) > 0$ for all $a \in \alpha$. For $n \geq 1$ and $a_0, \dots, a_{n-1} \in \alpha$, we define the n *– cylinder*

$$[a_0, \dots, a_{n-1}] = \bigcap_{k=0}^{n-1} F^{-k}(a_k)$$

and we let α_n denote the partition of Y into n–cylinders. It is assumed that the partition separates points in Y, meaning if $x, y \in Y$ with $x \neq y$, there exists $n \geq 1$ such that x and y lie in distinct n–cylinders. Fix $\gamma \in (0,1)$ and define the *symbolic metric* $d_\gamma(x,y) = \gamma^{n(x,y)}$ for $x, y \in Y$, where the separation time $n(x,y)$ is the greatest $n \geq 0$ such that x and y lie in the same n–cylinder. We say that $F\colon Y \to Y$ is *Gibbs-Markov* if there exists $C_1 > 0$ such that for all $a \in \alpha$, the following hold:

(i) $F|_a\colon a \to Y$ is a bijection with measurable inverse.

(ii) $\zeta_0 = \frac{d\rho_Y}{d\rho_Y \circ F}$ satisfies $|\log\zeta_0(x) - \log\zeta_0(y)| \leq C_1\, d_\gamma(Fx, Fy)$ for all $x, y \in a$.

Remark 3.3.1. *If* $T\colon X \to X$ *is non-uniformly expanding as in Section 3.2, then the induced map* $F\colon Y \to Y$ *is Gibbs-Markov for* $\gamma \in [\lambda^{-\eta}, 1)$. *This follows from* *the fact that if $x, y \in Y$, then $d(x,y)^\eta \leq d_\gamma(x,y)$. To see this, note first that if $n(x,y) = 0$, then*

$$d(x,y)^\eta \leq \operatorname{diam}(X)^\eta = 1 = d_\gamma(x,y).$$

Next note that if $n(x,y) = n > 0$, then by expansiveness of F on partition elements, we have

$$d(x,y) \leq \lambda^{-n}\, d(F^n x, F^n y) \leq \lambda^{-n} \leq \gamma^{n/\eta} = d_\gamma(x,y)^{1/\eta},$$

as claimed.

For the rest of this section, we fix $\gamma \in (0,1)$ and let $F\colon Y \to Y$ be Gibbs-Markov as above. We aim to construct an absolutely continuous invariant probability measure for F by utilising Proposition 2.5.5. We begin by giving a pointwise expression of iterates of the transfer operator for F.

Proposition 3.3.2. *Let $\mathcal{P}\colon L^1(Y) \to L^1(Y)$ denote the transfer operator for F. Then for $n \geq 1$, $V \in L^1(Y)$, and $y \in Y$, we have*

$$\mathcal{P}^n V(y) = \sum_{a \in \alpha_n} (\zeta_0)_n(y_a) V(y_a),$$

where $(\zeta_0)_n = \zeta_0\, \zeta_0 \circ F \cdots \zeta_0 \circ F^{n-1}$ and y_a is the unique element in a such that $F^n y_a = y$.

Proof. We proceed by induction. For $W \in L^\infty(Y)$, we have

$$\begin{aligned}
\int_Y \mathcal{P}V\, W \,\mathrm{d}\rho_Y &= \int_Y V\, W \circ F \,\mathrm{d}\rho_Y = \sum_{a \in \alpha} \int_a V(y)\, W(Fy) \zeta_0(y) \,\mathrm{d}\rho_Y \circ F \\
&= \sum_{a \in \alpha} \int_a \zeta_0(F^{-1}y) V(F^{-1}y)\, W(y) \,\mathrm{d}\rho_Y = \sum_{a \in \alpha} \int_Y \zeta_0(y_a) V(y_a)\, W(y) \,\mathrm{d}\rho_Y \\
&= \int_Y \Bigl(\sum_{a \in \alpha} \zeta_0(y_a) V(y_a) \Bigr) W(y) \,\mathrm{d}\rho_Y,
\end{aligned}$$

where the third equality uses a change of variables. Since $W \in L^\infty(Y)$ is arbitrary, the base case follows.

Let $n > 1$ and suppose that for all $V \in L^1(Y)$ and $y \in Y$, we have

$$\mathcal{P}^{n-1}V(y) = \sum_{a \in \alpha_{n-1}} (\zeta_0)_{n-1}(y_a)V(y_a).$$

Then noting that $\mathcal{P}^n = \mathcal{P}^{n-1} \circ \mathcal{P}$, we have

$$\begin{aligned}\mathcal{P}^n V(y) &= \sum_{a \in \alpha_{n-1}} (\zeta_0)_{n-1}(y_a)\mathcal{P}V(y_a) = \sum_{a \in \alpha_{n-1}} (\zeta_0)_{n-1}(y_a)\Big(\sum_{b \in \alpha} \zeta_0(y_{a_b})V(y_{a_b})\Big) \\ &= \sum_{a \in \alpha_{n-1}} \sum_{b \in \alpha} (\zeta_0)_{n-1}(Fy_{a_b})\zeta_0(y_{a_b})V(y_{a_b}) = \sum_{a \in \alpha_n} (\zeta_0)_n(y_a)V(y_a),\end{aligned}$$

where the final equality follows from relabelling. This completes the proof. □

We next introduce a space of observables which is used throughout the thesis. For $d \geq 1$ and $\gamma \in (0,1)$, we say that an observable $V \colon Y \to \mathbb{R}^d$ is Lipschitz if $\|V\|_\gamma = |V|_\gamma + |V|_\infty < \infty$, where

$$|V|_\gamma = \sup_{\substack{x,y \in Y \\ x \neq y}} \frac{\big|V(x) - V(y)\big|}{d_\gamma(x,y)}.$$

The space of Lipschitz observables $F_\gamma(Y;\mathbb{R}^d)$ is a Banach space under the norm $\|\cdot\|_\gamma$. When $d = 1$, we write $F_\gamma(Y)$. It is immediate that if $\gamma_1 \leq \gamma_2$, then $F_{\gamma_1}(Y;\mathbb{R}^d) \subset F_{\gamma_2}(Y;\mathbb{R}^d)$. The next proposition is a standard observation of the symbolic metric d_γ, which is useful for showing that various observables are Lipschitz.

Proposition 3.3.3. *Let $x, y \in Y$. Then $d_\gamma(F^j x, F^j y) = \gamma^{-j}\, d_\gamma(x,y)$ for all $0 \leq j \leq n(x,y)$.*

Proof. Write $n(x,y) = n$ and observe that $n(F^j x, F^j y) = n - j$. Therefore

$$d_\gamma(F^j x, F^j y) = \gamma^{n-j} = \gamma^{-j}\gamma^n = \gamma^{-j}\, d_\gamma(x,y),$$

as claimed. □

Let $\mathcal{P}_{\rho_Y} \colon L^1(Y) \to L^1(Y)$ denote the transfer operator for F with respect to the measure ρ_Y. To construct an absolutely continuous invariant probability measure, we consider the Cesàro averages

$$R_n = \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{P}_{\rho_Y}^j 1 \tag{3.3.1}$$

and show they have a convergent subsequence in $L^1(Y)$. To do this, we work via the auxiliary Banach space $F_\gamma(Y)$. Our first step is to verify that $(R_n)_{n\geq 1} \subset F_\gamma(Y)$. This is an immediate consequence of the next two lemmas.

Lemma 3.3.4. *There exists $C > 1$ such that for all $n \geq 1$, $a \in \alpha_n$, and $x, y \in a$, we have $(\zeta_0)_n(x) \leq C\rho_Y(a)$ and $|(\zeta_0)_n(x) - (\zeta_0)_n(y)| \leq C\rho_Y(a)\, d_\gamma(F^n x, F^n y)$.*

Proof. First note that $\log(\zeta_0)_n = \sum_{k=0}^{n-1} \log(\zeta_0 \circ F^k)$ for all $n \geq 1$. Therefore, for all $a \in \alpha_n$ and $x, y \in a$, we have

$$\begin{aligned}
\big|\log(\zeta_0)_n(x) - \log(\zeta_0)_n(y)\big| &\leq \sum_{k=0}^{n-1} \big|\log\big(\zeta_0(F^k x)\big) - \log\big(\zeta_0(F^k y)\big)\big| \\
&\leq C_1 \sum_{k=1}^{n} d_\gamma(F^k x, F^k y) = C_1 \sum_{k=1}^{n} \gamma^{n-k}\, d_\gamma(F^n x, F^n y) \\
&\leq C_1 \sum_{k=0}^{\infty} \gamma^k\, d_\gamma(F^n x, F^n y) = \left(\frac{C_1}{1-\gamma}\right) d_\gamma(F^n x, F^n y),
\end{aligned}$$

where the first equality uses Proposition 3.3.3. It follows that

$$\log\left(\frac{(\zeta_0)_n(x)}{(\zeta_0)_n(y)}\right) \leq \left(\frac{C_1}{1-\gamma}\right) d_\gamma(F^n x, F^n y).$$

Therefore

$$\frac{(\zeta_0)_n(x)}{(\zeta_0)_n(y)} \leq \exp\left(\left(\frac{C_1}{1-\gamma}\right) d_\gamma(F^n x, F^n y)\right) \leq \exp\left(\frac{C_1}{1-\gamma}\right) =: c < \infty,$$

where the second inequality uses that $\operatorname{diam}_\gamma(Y) = \sup\{d_\gamma(x,y) \mid x, y \in Y\} = 1$. Now, since $C_1 > 0$ and $\gamma \in (0,1)$, we have $c > 1$. Taking the supremum over $x \in a$ and then taking the infimum over $y \in a$ gives us

$$\sup_a (\zeta_0)_n \leq c \inf_a (\zeta_0)_n. \tag{3.3.2}$$

Therefore

$$\rho_Y(a) = \int_Y \mathbb{1}_a \, \mathrm{d}\rho_Y = \int_Y \mathcal{P}^n_{\rho_Y} \mathbb{1}_a \, \mathrm{d}\rho_Y \geq \inf_Y \mathcal{P}^n_{\rho_Y} \mathbb{1}_a = \inf_a (\zeta_0)_n \geq c^{-1} \sup_a (\zeta_0)_n,$$

proving the first estimate.

For the second estimate, we first note that for $s, t > 0$, we have $|s - t| \leq \max\{s, t\} |\log s - \log t|$. Indeed, supposing without loss of generality that $s > t$, we have

$$s - t = \int_t^s \mathrm{d}x = s \int_t^s \frac{\mathrm{d}x}{s} \leq s \int_t^s \frac{\mathrm{d}x}{x} = s(\log s - \log t).$$

Therefore

$$\begin{aligned} \big|(\zeta_0)_n(x) - (\zeta_0)_n(y)\big| &\leq \sup_a (\zeta_0)_n \, \big|\log(\zeta_0)_n(x) - \log(\zeta_0)_n(y)\big| \\ &\leq \left(\frac{cC_1}{1-\gamma}\right) \rho_Y(a) \, d_\gamma(F^n x, F^n y). \end{aligned}$$

The result now follows with $C = \max\{c, (1-\gamma)^{-1} c C_1\} > 1$. □

Lemma 3.3.5. *For all $n \geq 1$, we have $\mathcal{P}^n_{\rho_Y} 1 \in F_\gamma(Y)$ with $\|\mathcal{P}^n_{\rho_Y} 1\|_\gamma \leq 2C + 1$, where C is as in Lemma 3.3.4.*

Proof. For $x, y \in Y$, and $x_a, y_a \in a$ with $F^n x_a = x$, $F^n y_a = y$, we have that

$$\begin{aligned} \big|\mathcal{P}^n_{\rho_Y} 1(x) - \mathcal{P}^n_{\rho_Y} 1(y)\big| &\leq \sum_{a \in \alpha_n} \big|(\zeta_0)_n(x_a) - (\zeta_0)_n(y_a)\big| \\ &\leq C \sum_{a \in \alpha_n} \rho_Y(a) \, d_\gamma(F^n x_a, F^n y_a) \\ &= C \, d_\gamma(x, y), \end{aligned}$$

where the second inequality follows from Lemma 3.3.4. Therefore $|\mathcal{P}^n_{\rho_Y}|_\gamma \leq C$. Now, from Proposition 2.5.3 (iv), we have $|\mathcal{P}^n_{\rho_Y} 1|_1 \leq 1$. For $x \in Y$, it follows that

$$\begin{aligned} \big|\mathcal{P}^n_{\rho_Y} 1(x)\big| &\leq \left|\mathcal{P}^n_{\rho_Y} 1(x) - \int_Y \mathcal{P}^n_{\rho_Y} 1(y) \, \mathrm{d}\rho_Y(y)\right| + |\mathcal{P}^n_{\rho_Y} 1|_1 \\ &\leq \left| \int_Y \big(\mathcal{P}^n_{\rho_Y} 1(x) - \mathcal{P}^n_{\rho_Y} 1(y)\big) \, \mathrm{d}\rho_Y(y)\right| + 1 \\ &\leq |\mathcal{P}^n_{\rho_Y} 1|_\gamma \, \mathrm{diam}_\gamma(Y) + 1 = |\mathcal{P}^n_{\rho_Y} 1|_\gamma + 1. \end{aligned}$$

Therefore $|\mathcal{P}^n_{\rho_Y} 1|_\infty \leq |\mathcal{P}^n_{\rho_Y} 1|_\gamma + 1 \leq C + 1$, and the result follows. □

Proposition 3.3.6. *For $n \geq 1$, let R_n be as in* (3.3.1). *Then $R_n \in F_\gamma(Y)$ with $\|R_n\|_\gamma \leq 2C + 1$, where C is as in Lemma 3.3.4.*

Proof. Using Lemma 3.3.5, we have

$$\|R_n\|_\gamma \leq \frac{1}{n} \sum_{j=0}^{n-1} \|\mathcal{P}_{\rho_Y}^j 1\|_\gamma \leq 2C + 1,$$

as claimed. □

It remains to verify that $(R_n)_{n\geq 1}$ has a convergent subsequence in $L^1(Y)$. This follows directly from the next proposition. For later purposes, we state and prove this result in the general setting $d \geq 1$.

Proposition 3.3.7. *The unit ball of $F_\gamma(Y;\mathbb{R}^d)$ is compact in $L^1(Y;\mathbb{R}^d)$.*

Proof. We show that every $(V_n) \subset F_\gamma(Y;\mathbb{R}^d)$ with $\|V_n\|_\gamma \leq 1$ has a subsequence which converges in $L^1(Y;\mathbb{R}^d)$ to an element of the unit ball of $F_\gamma(Y;\mathbb{R}^d)$. We begin by finding a pointwise convergence subsequence.

For $a \in \cup_{k\geq 1}\alpha_k$, let $y_a \in a$ be a representative of a. Since $|V_n|_\infty \leq \|V_n\|_\gamma \leq 1$, the sequence $V_n(y_a)$ is bounded, so by Bolzano-Weierstrass, has a convergent subsequence. Since $\cup_{k\geq 1}\alpha_k$ is countable, we may suppose via a diagonal argument and by relabelling that $V_n(y_a)$ converges for all $a \in \cup_{k\geq 1}\alpha_k$. We show that (V_n) converges pointwise on Y by showing it is pointwise Cauchy. Fix $\epsilon > 0$. Let k be sufficiently large so that for all $a \in \alpha_k$, we have $\operatorname{diam}_\gamma(a) = \gamma^k < \epsilon$. Let $y \in Y$, and suppose $y \in a$ for some $a \in \alpha_k$. Then for $n, m \geq 1$, we have

$$\begin{aligned} \big|V_n(y) - V_m(y)\big| &\leq \big|V_n(y) - V_n(y_a)\big| + \big|V_n(y_a) - V_m(y_a)\big| + \big|V_m(y_a) - V_m(y)\big| \\ &\leq \big(|V_n|_\gamma + |V_m|_\gamma\big)\, d_\gamma(y, y_a) + \big|V_n(y_a) - V_m(y_a)\big| \\ &< 2\epsilon + \big|V_n(y_a) - V_m(y_a)\big|. \end{aligned}$$

Now, since V_n converges at y_a, there exists $N_a \geq 1$ such that $|V_n(y_a) - V_m(y_a)| < \epsilon$ for $n, m \geq N_a$. It follows that $|V_n(y) - V_m(y)| < 3\epsilon$ for $n, m \geq N_a$. Therefore, there exists $V \colon Y \to \mathbb{R}^d$ such that $V_n \to V$ pointwise.

We next show that $V \in F_\gamma(Y;\mathbb{R}^d)$ with $\|V\|_\gamma \le 1$. Note that for $y \in Y$, using pointwise convergence of V_n to V, we have

$$\big|V(y)\big| = \liminf_{n\to\infty} \big|V_n(y)\big| \le \liminf_{n\to\infty} |V_n|_\infty,$$

so that $|V|_\infty \le \liminf_{n\to\infty} |V_n|_\infty$. Moreover, for $x, y \in Y$, we have

$$\frac{\big|V_n(x) - V_n(y)\big|}{d_\gamma(x,y)} \le |V_n|_\gamma \quad \text{and} \quad \frac{\big|V_n(x) - V_n(y)\big|}{d_\gamma(x,y)} \to \frac{\big|V(x) - V(y)\big|}{d_\gamma(x,y)},$$

so that

$$\frac{\big|V(x) - V(y)\big|}{d_\gamma(x,y)} \le \liminf |V_n|_\gamma.$$

Therefore $|V|_\gamma \le \liminf_{n\to\infty} |V_n|_\gamma$. It follows that

$$\|V\|_\gamma \le \liminf_{n\to\infty} |V_n|_\infty + \liminf_{n\to\infty} |V_n|_\gamma \le \liminf_{n\to\infty} \big(|V_n|_\infty + |V_n|_\gamma\big) \le 1,$$

as claimed.

To conclude, we show that $V_n \to V$ in $L^1(Y;\mathbb{R}^d)$. Note that $|V_n - V| \le 2$ for all $n \ge 1$. Combining this with pointwise convergence, it follows that $V_n \to V$ in $L^1(Y;\mathbb{R}^d)$ by the dominated convergence theorem. This completes the proof. □

We now have the machinery to construct the required measure μ_Y on Y.

Proposition 3.3.8. *There exists a unique ergodic F–invariant probability measure μ_Y on Y which is equivalent to ρ_Y. Moreover, $\mathrm{d}\mu_Y/\,\mathrm{d}\rho_Y \in F_\gamma(Y)$ and $\log(\mathrm{d}\mu_Y/\,\mathrm{d}\rho_Y) \in F_\gamma(Y)$.*

Proof. For $n \ge 1$, let R_n be as in (3.3.1). From Proposition 3.3.6, $(R_n) \subset F_\gamma(Y)$ is bounded, and so by Proposition 3.3.7 there exists a subsequence $(R_{n_k}) \subset (R_n)$ and $R \in F_\gamma(Y)$ such that $R_{n_k} \to R$ in $L^1(Y)$. We show that R is the required density by appealing to Proposition 2.5.5.

First note that $R \ge 0$. Indeed, $L^1(Y)$ convergence of R_{n_k} to R implies there is a subsequence $(R_{n_{k_\ell}}) \subset (R_{n_k})$ such that $R_{n_{k_\ell}} \to R$ almost surely. Since $R_{n_{k_\ell}} \ge 0$ for all ℓ, non-negativity of R follows. Moreover, note that

$$\left|\int_Y R\,\mathrm{d}\rho_Y - 1\right| = \left|\int_Y R\,\mathrm{d}\rho_Y - \int_Y R_{n_k}\,\mathrm{d}\rho_Y\right| \le |R - R_{n_k}|_1 \to 0,$$

so that $\int_Y R\,\mathrm{d}\rho_Y = 1$. Now, $\mathcal{P}_{\rho_Y}$ is a bounded operator on $L^1(Y)$ by Proposition 2.5.3 (iv). In particular,

$$\mathcal{P}_{\rho_Y} R = \lim_{k\to\infty} \mathcal{P}_{\rho_Y} R_{n_k} = \lim_{k\to\infty} \frac{1}{n_k}\sum_{j=1}^{n_k} \mathcal{P}_{\rho_Y}^j 1 = \lim_{k\to\infty}\left(\frac{1}{n_k}\sum_{j=0}^{n_k}\mathcal{P}_{\rho_Y}^j 1 - \frac{1}{n_k}\right) = R,$$

so that R is an invariant density.

Define $\mathrm{d}\mu_Y = R\,\mathrm{d}\rho_Y$. By absolute continuity, to prove ergodicity of μ_Y, it suffices to show that any $F-$invariant set B with $\rho_Y(B) > 0$ has $\rho_Y(B) = 1$. Note that for any $\epsilon > 0$, there exist $k \geq 1$ sufficiently large and $a \in \alpha_k$ such that

$$\frac{\rho_Y(B\cap a)}{\rho_Y(a)} > 1 - \epsilon.$$

Next, noting that $(\xi_0)_k = \mathrm{d}\rho_Y/(\mathrm{d}\rho_Y \circ F^k)$, we have

$$\rho_Y\big(F^k(B\cap a)\big) = \int_{B\cap a} \mathrm{d}\rho_Y \circ F^k = \int_{B\cap a} (\xi_0)_k^{-1}\,\mathrm{d}\rho_Y \geq \frac{\rho_Y(B\cap a)}{C\rho_Y(a)},$$

where the final inequality follows from Lemma 3.3.4. Moreover, observe that $\rho_Y(F^k(a)) = \rho_Y(Y) = 1 < C$. Finally, since $B = F^{-1}(B)$, we have $F^k(B) \subset B$. Combining the above, we conclude that

$$\rho_Y(B) \geq \rho_Y\big(F^k(B)\big) \geq \rho_Y\big(F^k(B\cap a)\big) = \frac{\rho_Y\big(F^k(B\cap a)\big)}{\rho_Y\big(F^k(a)\big)} > \frac{\rho_Y(B\cap a)}{\rho_Y(a)} > 1-\epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have that $\rho_Y(B) = 1$, and ergodicity of μ_Y follows.

To prove that μ_Y is equivalent to ρ_Y, we show that R is bounded away from 0. Note that as $\int_Y R\,\mathrm{d}\rho_Y = 1$, there exists $y_0 \in Y$ with $R(y_0) = c > 0$. Since R is continuous with respect to d_γ, there exists $k \geq 1$ sufficiently large such that $R > c/2$ on the $k-$cylinder a containing y_0. It follows that for any $y \in Y$, we have

$$R(y) = \mathcal{P}_{\rho_Y}^k R(y) = \sum_{b\in\alpha_k} (\zeta_0)_k(y_b)R(y_b) > (\zeta_0)_k(y_a)\frac{c}{2} \geq \frac{c}{2}\inf_a (\zeta_0)_k =: M.$$

We show that $M > 0$. This is equivalent to showing $\inf_a(\zeta_0)_k > 0$. Suppose for contradiction that $\inf_a(\zeta_0)_k = 0$. Then from (3.3.2) we have $\sup_a(\zeta_0)_k = 0$, so that $(\zeta_0)_k|_a = 0$. It follows that

$$\rho_Y(a) = \int_a (\zeta_0)_k \,\mathrm{d}\rho_Y \circ F^k = 0,$$

yielding a contradiction. Therefore $M > 0$ and R is bounded away from 0 as required.

We now show that $\log R$ is Lipschitz. Note that since R is positive, we have for $x, y \in Y$ that

$$\frac{R(x) - R(y)}{R(y)} > -1.$$

Therefore

$$\begin{aligned} \big|\log R(x) - \log R(y)\big| &= \left|\log\left(\frac{R(x)}{R(y)}\right)\right| = \left|\log\left(1 + \frac{R(x) - R(y)}{R(y)}\right)\right| \\ &\leq \left|\frac{R(x) - R(y)}{R(y)}\right| \leq \left|\frac{R(x) - R(y)}{M}\right| \leq \frac{|R|_\gamma}{M}\, d_\gamma(x, y), \end{aligned}$$

proving the claim.

Finally, uniqueness follows from the fact that any two distinct ergodic measures are mutually singular. □

Let $\zeta = \mathrm{d}\mu_Y/(\mathrm{d}\mu_Y \circ F)$ and for $n \geq 1$, let $\zeta_n = \zeta\, \zeta \circ F \cdots \zeta \circ F^{n-1}$. To conclude this section, we give some estimates for ζ_n analogous to Lemma 3.3.4 which will be of use to us throughout.

Proposition 3.3.9. *There exists $D > 1$ such that for all $n \geq 1$, $a \in \alpha_n$, and $x, y \in a$, we have $\zeta_n(x) \leq D\mu_Y(a)$ and $|\zeta_n(x) - \zeta_n(y)| \leq D\mu_Y(a)\, d_\gamma(F^n x, F^n y)$.*

Proof. The proof of these estimates is identical to that of Lemma 3.3.4, once we prove that ζ has bounded distortion. Therefore, we show that there exists $C > 0$ such that for all $a \in \alpha$ and $x, y \in a$, we have

$$\big|\log \zeta(x) - \log \zeta(y)\big| \leq C\, d_\gamma(Fx, Fy).$$

To do this, first note that

$$\zeta = \frac{\mathrm{d}\mu_Y}{\mathrm{d}\mu_Y \circ F} = \left(\frac{\mathrm{d}\rho_Y}{\mathrm{d}\rho_Y \circ F}\right)\left(\frac{\mathrm{d}\mu_Y}{\mathrm{d}\rho_Y}\right)\left(\frac{\mathrm{d}\rho_Y \circ F}{\mathrm{d}\mu_Y \circ F}\right).$$

Let $g = \log(\mathrm{d}\mu_Y / \mathrm{d}\rho_Y)$. Then

$$\log\left(\frac{\mathrm{d}\rho_Y \circ F}{\mathrm{d}\mu_Y \circ F}\right) = -\log\left(\frac{\mathrm{d}\mu_Y \circ F}{\mathrm{d}\rho_Y \circ F}\right) = -\log\left(\left(\frac{\mathrm{d}\mu_Y}{\mathrm{d}\rho_Y}\right) \circ F\right) = -\log\left(\frac{\mathrm{d}\mu_Y}{\mathrm{d}\rho_Y}\right) \circ F$$
$$= -g \circ F.$$

Therefore, recalling $\zeta_0 = \mathrm{d}\rho_Y/(\mathrm{d}\rho_Y \circ F)$, we have

$$\log \zeta = \log \zeta_0 + g - g \circ F.$$

Now, from Proposition 3.3.8, we have $|g|_\gamma < \infty$. Moreover, ζ_0 has bounded distortion. Therefore, for all $a \in \alpha$ and $x, y \in a$ we have

$$\begin{aligned}\big|\log \zeta(x) - \log \zeta(y)\big| &\leq \big|\log \zeta_0(x) - \log \zeta_0(y)\big| + \big|g(x) - g(y)\big| + \big|g(Fx) - g(Fy)\big| \\ &\leq C_1\, d_\gamma(Fx, Fy) + |g|_\gamma\, d_\gamma(x, y) + |g|_\gamma\, d_\gamma(Fx, Fy) \\ &= \big(C_1 + |g|_\gamma \gamma + |g|_\gamma\big)\, d_\gamma(Fx, Fy),\end{aligned}$$

where we use Proposition 3.3.3 in the final step. This completes the proof. □

3.4 Representation as Young tower over induced Gibbs-Markov map

Let (X, d) be a bounded metric space with Borel probability measure ρ and let $T\colon X \to X$ be a non-singular transformation. Let $Y \subset X$ with $\rho(Y) > 0$ and suppose that there is a Gibbs-Markov map $F\colon Y \to Y$ and integrable return time $\tau\colon Y \to \mathbb{Z}^+$ which is constant on each partition element, such that $F = T^\tau$. By the results in the previous section, there exists a unique ergodic absolutely continuous $F-$invariant probability measure μ_Y on Y which is equivalent to $\rho|_Y$. Using this, we describe how to construct a unique ergodic $T-$invariant probability measure μ on X which is equivalent to ρ.

We begin by defining the Young tower [101, 102] for T. Let $\Delta = \{(y, \ell) \in Y \times \mathbb{Z} \mid 0 \leq \ell \leq \tau(y) - 1\}$. The *tower map* $f\colon \Delta \to \Delta$ for T is defined by

$$f(y, \ell) = \begin{cases} (y, \ell + 1) & \text{if } 0 \leq \ell \leq \tau(y) - 2, \\ (Fy, 0) & \text{if } \ell = \tau(y) - 1. \end{cases}$$

The projection $\pi\colon \Delta \to X$ given by $\pi(y,\ell) = T^\ell y$ defines a semi-conjugacy from f to T. That is, π is surjective and satisfies $\pi \circ f = T \circ \pi$. The ℓth level of the tower is the set $\Delta_\ell = \{(y,\ell) \in \Delta\}$. Naturally, the base of the tower (i.e. the $0th$ level) identifies with $Y \subset X$, and the ℓth level of the tower is a copy of $\{\tau > \ell\}$. Define the probability measures μ_Δ and μ on Δ and X respectively by

$$\mu_\Delta = \frac{\mu_Y \times \{counting\}}{\int_Y \tau \, \mathrm{d}\mu}$$

and $\mu = \pi_* \mu_\Delta$.

Proposition 3.4.1. *The measure μ_Δ is f – ergodic and invariant. Moreover, μ is the unique ergodic T – invariant probability measure which is equivalent to ρ.*

Proof. We prove the first statement. The second statement follows from the fact that π is a semi-conjugacy and Proposition 3.3.8. For invariance of μ_Δ, let $\hat{v} \in L^2(\Delta)$ and note that

$$\begin{aligned}\int_\Delta \hat{v} \circ f \, \mathrm{d}\mu_\Delta &= |\tau|_1^{-1} \int_Y \left(\hat{v}(Fy, 0) + \sum_{\ell=0}^{\tau(y)-2} \hat{v}(y, \ell+1) \right) \mathrm{d}\mu_Y \\ &= |\tau|_1^{-1} \int_Y \left(\hat{v}(y, 0) + \sum_{\ell=1}^{\tau(y)-1} \hat{v}(y, \ell) \right) \mathrm{d}\mu_Y = \int_\Delta \hat{v} \, \mathrm{d}\mu_\Delta,\end{aligned}$$

where we use F – invariance of μ_Y and relabelling in the second equality.

For ergodicity of μ_Δ, let $\hat{v} \in L^2(\Delta)$ be such that $\hat{v} \circ f = \hat{v}$ μ_Δ – almost surely. Then $\hat{v} \circ f(\cdot, \ell) = \hat{v}(\cdot, \ell)$ for all $0 \le \ell \le \tau - 1$ μ_Y – almost surely. In particular, for almost every $y \in Y$, we have

$$\hat{v}(y,0) = \hat{v}(y,1) = \cdots = \hat{v}(y, \tau(y)-1) = \hat{v}(Fy, 0). \tag{3.4.1}$$

Therefore, by ergodicity of μ_Y, we have that $\hat{v}(\cdot, 0)$ is constant μ_Y – almost surely. To show that this implies $\hat{v}$ is constant μ_Δ – almost surely, take any y' in the set of full measure for which (3.4.1) holds, and note that for $0 \le \ell \le \tau(y') - 1$, we have $\hat{v}(y', \ell) = \hat{v}(y', 0) = c$ for some constant c. It follows that $\hat{v}$ is constant μ_Δ – almost surely, proving ergodicity. ∎

Remark 3.4.2. *In the case* $\operatorname{hcf}\{\tau(a) \mid a \in \alpha\} = 1$*, we have from [101, 102, Theorem 1] that* μ *is mixing.*

Example 3.4.3. *Figure 3.2 shows the Young tower construction and corresponding tower map for the intermittent map introduced in Example 2.2.6. As we can see, the dynamics of the tower map* f *are governed as follows: Each* $x \in (x_0, 1]$ *moves up the tower until it reaches the top level above* x*, after which it is bijectively returned to the base by the induced transformation* F*.*

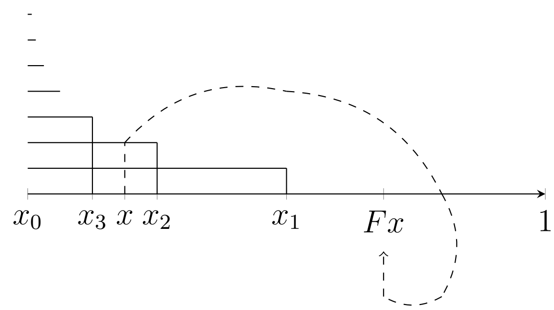


Figure 3.2: Young tower and tower map for the intermittent map

To conclude this chapter, we give a pointwise expression of the transfer operator for f.

Proposition 3.4.4. *Let* $\mathcal{L}\colon L^1(\Delta) \to L^1(\Delta)$ *be defined by*

$$\mathcal{L}\hat{v}(y,\ell) = \begin{cases} \sum_{a\in\alpha} \zeta(y_a)\hat{v}(y_a, \tau(y_a) - 1) & \text{if } \ell = 0, \\ \hat{v}(y, \ell - 1) & \text{if } \ell \geq 1. \end{cases} \tag{3.4.2}$$

Then $\mathcal{L}$ *is the transfer operator for* f*.*

Proof. Let us write $v(y) = \hat{v}(y, \tau(y) - 1)$, and recall from Proposition 3.3.2 that

$\mathcal{P}v(y) = \sum_{a\in\alpha}\zeta(y_a)\hat{v}(y_a, \tau(y_a)-1)$. Let $\hat{w} \in L^\infty(\Delta)$, and note that

$$\begin{aligned}\int_\Delta \hat{v}\,\hat{w}\circ f\,\mathrm{d}\mu_\Delta &= |\tau|_1^{-1}\int_Y \left(v(y)\hat{w}(Fy,0) + \sum_{\ell=0}^{\tau(y)-2}\hat{v}(y,\ell)\hat{w}(y,\ell+1)\right)\mathrm{d}\mu_Y \\ &= |\tau|_1^{-1}\int_Y \left(\mathcal{P}v(y)\hat{w}(y,0) + \sum_{\ell=1}^{\tau(y)-1}\hat{v}(y,\ell-1)\hat{w}(y,\ell)\right)\mathrm{d}\mu_Y \\ &= |\tau|_1^{-1}\int_Y \sum_{\ell=0}^{\tau(y)-1}\mathcal{L}\hat{v}(y,\ell)\,\hat{w}(y,\ell)\,\mathrm{d}\mu_Y = \int_\Delta \mathcal{L}\hat{v}\,\hat{w}\,\mathrm{d}\mu_\Delta.\end{aligned}$$

The result follows. □

Chapter 4

Primary martingale-coboundary decomposition

4.1 Outline

Let (X, d) be a bounded metric space and $T\colon X \to X$ be non-uniformly expanding with partition α, return time $\tau\colon Y \to \mathbb{Z}^+$, induced map $F\colon Y \to Y$, and constants $\lambda > 1$, $\eta \in (0, 1]$, and $C_0, C_1 \geq 1$ as in Section 3.2. Let μ and μ_Y denote the ergodic invariant Borel probability measures on X and Y respectively which were constructed in Chapter 3. Let G be a compact connected Lie group with Haar measure ν, and suppose that $(\pi, \mathbb{R}^d)$ is a representation of G for some $d \geq 1$. As in Remark 2.2.7, we fix a $G-$invariant inner product $[\cdot, \cdot]$ on $\mathbb{R}^d$ and view G as a closed subgroup of $\mathrm{O}(d)$. We study the compact group extension $T_h\colon X \times G \to X \times G$ defined by $T_h(x, g) = (Tx, gh(x))$, where $h \in C^\eta(X; G)$. The probability measure $m = \mu \times \nu$ is T_h-invariant and is assumed to be ergodic.

Remark 4.1.1. *Ergodicity of m is typical in the following sense, as in [35, Theorem 1.5]. The set of Hölder cocycles $h\colon X \to G$ for which m is not ergodic lies in a closed subspace of infinite codimension in the space of all Hölder cocycles.*

We consider equivariant observables $\phi\colon X \times G \to \mathbb{R}^d$ of the form $\phi(x, g) = g \cdot v(x)$, where $v \in C^\eta(X; \mathbb{R}^d)$ with $\int_{X \times G} \phi \, \mathrm{d}m = 0$. In this chapter, we construct our primary martingale-coboundary decomposition for a lifted version of the observable ϕ and then apply the results of Section 2.4 and Section 2.8. We extend the approach of Korepanov, Kosloff, and Melbourne [60], who applied this method to lifted Hölder observables of non-uniformly expanding maps. As an application, we recover [39, Theorem 1.10]. Moreover, we obtain optimal moment estimates which do not seem to be readily available in the literature.

Before stating the main results of this chapter (Theorem 4.1.2, Theorem 4.1.3, and Theorem 4.1.5), we introduce induced versions of the function $v\colon X \to \mathbb{R}^d$. Define $V\colon Y \to \mathbb{R}^d$ and $V^*\colon Y \to \mathbb{R}$ by

$$V(y) = \sum_{\ell=0}^{\tau(y)-1} h_\ell(y) \cdot v(T^\ell y) \tag{4.1.1}$$

and

$$V^*(y) = \max_{0\le k\le \tau(y)-1} \left| \sum_{\ell=0}^{k} h_\ell(y) \cdot v(T^\ell y) \right| \tag{4.1.2}$$

respectively, where $h_\ell = h\, h \circ T \cdots h \circ T^{\ell-1}$. It is immediate that $|V| \le \tau |v|_\infty$ and $|V^*| \le \tau |v|_\infty$. In particular, $\tau \in L^p(Y)$ for $p > 1$ implies $V \in L^p(Y;\mathbb{R}^d)$ and $V^* \in L^p(Y)$.

Theorem 4.1.2. *Suppose $\tau \in L^p(Y)$ for some $p > 1$. If $V \in L^2(Y;\mathbb{R}^d)$, then there exists $\Sigma \in \mathbb{R}^{d,d}$ such that $g\Sigma = \Sigma g$ for all $g \in G$ and*

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \phi \circ T_h^k \to_w \mathcal{N}(0,\Sigma).$$

For $n \ge 1$, define the random elements $W_n\colon X \times G \to D([0,\infty);\mathbb{R}^d)$ by $W_n(t) = n^{-1/2} \sum_{k=0}^{[nt]-1} \phi \circ T_h^k$ for $t \ge 0$.

Theorem 4.1.3. *Suppose $\tau \in L^p(Y)$ for some $p > 1$. If $V^* \in L^2(Y)$, then $W_n \to_w W$ in $D([0,\infty);\mathbb{R}^d)$, where W is a d-dimensional Brownian motion with mean 0 and covariance matrix Σ, where Σ is as in Theorem 4.1.2.*

Remark 4.1.4. *We make the following observations regarding Theorem 4.1.2 and Theorem 4.1.3.*

(i) *By [103], we have strong distributional convergence. That is, the associated weak convergence holds for any probability measure that is absolutely continuous with respect to m.*

(ii) *By [70], we obtain strong distributional convergence for the measure $\mu \times \delta_{g_0}$ for any $g_0 \in G$ fixed, where δ_{g_0} denotes the Dirac measure at g_0. That is, the above results hold for $\phi \circ T_h^k(\cdot, g_0)$ for $g_0 \in G$ fixed and any probability measure absolutely continuous with respect to μ.*

(iii) *The covariance matrix Σ is typically non-singular in the following sense, as in [79] (see also [36, Section 5]). The set of Hölder functions $v\colon X \to \mathbb{R}^d$ for which $\det \Sigma = 0$ lies in a closed subspace of infinite codimension in the set of all Hölder functions.*

Theorem 4.1.5. *Suppose $\tau \in L^p(Y)$ for some $p > 1$. There exists a constant $C > 0$ independent of v, h, and n such that:*

(i) *If $1 < p < 2$, then*

$$\left\| \max_{1 \le j \le n} \left| \sum_{k=0}^{j-1} \phi \circ T_h^k \right| \right\|_p \le C n^{1/p} \|v\|_\eta \|h\|_\eta \quad \textit{for all } n \ge 1.$$

(ii) *If $p \ge 2$, then*

$$\left\| \max_{1 \le j \le n} \left| \sum_{k=0}^{j-1} \phi \circ T_h^k \right| \right\|_{2(p-1)} \le C n^{1/2} \|v\|_\eta \|h\|_\eta \quad \textit{for all } n \ge 1.$$

Remark 4.1.6. *Optimality of these estimates can be seen as in [73, Section 3] and [77, Remark 3.7].*

The structure of the chapter is as follows: In Section 4.2, we represent the transformation $T_h\colon X \times G \to X \times G$ as an extension of a Young tower over an induced transformation. In Section 4.3, Section 4.4, and Section 4.5, we develop the necessary theory to derive a martingale-coboundary decomposition for V. This is constructed in Section 4.6, and by further arguments we obtain our primary martingale-coboundary decomposition for the lifted version of ϕ. In Section 4.7 we give the proofs of Theorem 4.1.2 and Theorem 4.1.3, and in Section 4.8 we prove Theorem 4.1.5 as well as characterising the covariance matrix Σ in terms of the observable ϕ. Finally, in Section 4.9 we give some examples for which our results hold.

From here on, unless otherwise stated, we implicitly consider complex-valued function spaces and the complexified action of G on $\mathbb{C}^d$, which allows us to utilise the results from Section 2.6 and Section 2.7. We write $\phi = g \cdot v$ as shorthand for $\phi(x, g) = g \cdot v(x)$. To simplify results, by C we denote various constants which depend continuously on $\lambda > 1$, $C_0, C_1 \geq 1$, $\eta \in (0, 1]$, $p > 1$, and $D > 1$, where p is the integrability of τ and D is as in Proposition 3.3.9.

4.2 Compact group extension of Young tower

In this section, we give a tower representation of T_h on which we derive our primary martingale-coboundary decomposition. Define the return time $\tau\colon Y \times G \to \mathbb{Z}^+$ by $\tau(y, g) = \tau(y)$. Define the induced cocycle $H\colon Y \to G$ by $H = h_\tau = h\, h \circ T \cdots h \circ T^{\tau-1}$ and the induced compact group extension $F_H = T_h^\tau\colon Y \times G \to Y \times G$ by $F_H(y, g) = (Fy, gH(y))$, with ergodic invariant probability measure $m_Y = \mu_Y \times \nu$.

Let $\Delta \times G = \{(y, g, \ell) \in Y \times G \times \mathbb{Z} \mid 0 \leq \ell \leq \tau(y) - 1\}$. We define the *tower map* $f_H\colon \Delta \times G \to \Delta \times G$ for T_h by

$$f_H(y, g, \ell) = \begin{cases} (y, g, \ell + 1) & \text{if } 0 \leq \ell \leq \tau(y) - 2, \\ (F_H(y, g), 0) & \text{if } \ell = \tau(y) - 1. \end{cases}$$

Remark 4.2.1. *There are two equivalent ways to view the above construction. The first is to view it as a tower over the compact group extension $F_H\colon Y \times G \to Y \times G$ with height τ. Alternatively, one can view this as the compact* *group extension of the tower map* $f\colon \Delta \to \Delta$ *defined in Section 3.4, with cocycle* $\hat{H}\colon \Delta \to G$ *given by*

$$\hat{H}(y,\ell) = \begin{cases} I_d & \text{if } 0 \le \ell \le \tau(y)-2, \\ H(y) & \text{if } \ell = \tau(y)-1. \end{cases}$$

Proposition 4.2.2. *The following hold true:*

(i) *The projection* $\pi_H\colon \Delta \times G \to X \times G$ *given by* $\pi_H(y,g,\ell) = T_h^\ell(y,g)$ *defines a semi-conjugacy from* f_H *to* T_h.

(ii) *The probability measure* $m_\Delta = \mu_\Delta \times \nu$ *is* f_H *–ergodic and invariant, and* $m = \mu \times \nu$ *satisfies* $m = (\pi_H)_* m_\Delta$.

Proof. For (i), we have that $\pi_H \circ f_H = T_h \circ \pi_H$ is immediate from the definitions. Now, $\pi_H(y,g,\ell) = T_h^\ell(y,g) = (T^\ell y, gh_\ell(y)) = (\pi(y,\ell), gh_\ell(y))$, where $\pi\colon \Delta \to X$ is the semi-conjugacy from Section 3.4. Let $(x,g) \in X \times G$ and $(y,\ell) \in \Delta$ be such that $\pi(y,\ell) = x$. Then $\pi_H(y, gh_\ell(y)^{-1}, \ell) = (\pi(y,\ell), gh_\ell(y)^{-1}h_\ell(y)) = (x,g)$, so that π_H is surjective. This proves (i).

For (ii), ergodicity and invariance of m_Δ follow from a similar argument as in Proposition 3.4.1. To prove that $m = (\pi_H)_* m_\Delta$, we let $\phi \in L^1(X \times G; \mathbb{R}^d)$ and show that $\int_{\Delta\times G} \phi \circ \pi_H \, \mathrm{d}m_\Delta = \int_{X\times G} \phi \, \mathrm{d}m$. Recalling that $\mu = \pi_* \mu_\Delta$, we have

$$\begin{aligned}
&\int_{\Delta\times G} \phi \circ \pi_H(y,g,\ell)\, \mathrm{d}m_\Delta(y,g,\ell) = \int_\Delta \int_G \phi\big(\pi(y,\ell), gh_\ell(y)\big)\, \mathrm{d}\nu(g)\, \mathrm{d}\mu_\Delta(y,\ell) \\
&= \int_\Delta \int_G \phi\big(\pi(y,\ell), g\big)\, \mathrm{d}\nu(g)\, \mathrm{d}\mu_\Delta(y,\ell) = \int_G \int_\Delta \phi\big(\pi(y,\ell), g\big)\, \mathrm{d}\mu_\Delta(y,\ell)\, \mathrm{d}\nu(g) \\
&= \int_G \int_X \phi(x,g)\, \mathrm{d}\mu(x)\, \mathrm{d}\nu(g) = \int_{X\times G} \phi(x,g)\, \mathrm{d}m(x,g),
\end{aligned}$$

completing the proof. □

For $\phi \in L^1(X \times G; \mathbb{R}^d)$, we define the *lifted observable* $\hat{\phi} \in L^1(\Delta \times G; \mathbb{R}^d)$ by $\hat{\phi} = \phi \circ \pi_H$. The next proposition says that distributional results for $(\phi \circ T_h^k)_{k\ge0}$ are equivalent to distributional results for $(\hat{\phi} \circ f_H^k)_{k\ge0}$.

Proposition 4.2.3. *Suppose $T_h\colon X\times G\to X\times G$ and $f_H\colon \Delta\times G\to\Delta\times G$ are as above. Let $\phi\in L^1(X\times G;\mathbb{R}^d)$ and $\hat\phi=\phi\circ\pi_H\in L^1(\Delta\times G;\mathbb{R}^d)$. Then $(\phi\circ T_h^k)_{k\geq 0}\sim(\hat\phi\circ f_H^k)_{k\geq 0}$.*

Proof. We show that the finite dimensional distributions of $(\phi\circ T_h^k)_{k\geq0}$ and $(\hat\phi\circ f_H^k)_{k\geq0}$ coincide. Since a stochastic process is determined by such distributions, the result then follows. Fix $n\geq 1$ and let $A=A_0\times\cdots\times A_n$ be a product of Borel subsets of $\mathbb{R}^d$. Then

$$
\begin{aligned}
m_\Delta\big((\hat\phi\circ f_H^k)_{k=0}^n\in A\big) &= m_\Delta\Bigg(\bigcap_{k=0}^n\{\hat\phi\circ f_H^k\in A_k\}\Bigg)\\
&= m_\Delta\Bigg(\bigcap_{k=0}^n\{\phi\circ T_h^k\circ\pi_H\in A_k\}\Bigg) = m_\Delta\Bigg(\bigcap_{k=0}^n \pi_H^{-1}\big((\phi\circ T_h^k)^{-1}(A_k)\big)\Bigg)\\
&= m_\Delta\Bigg(\pi_H^{-1}\Bigg(\bigcap_{k=0}^n(\phi\circ T_h^k)^{-1}(A_k)\Bigg)\Bigg) = m\Bigg(\bigcap_{k=0}^n(\phi\circ T_h^k)^{-1}(A_k)\Bigg)\\
&= m\big((\phi\circ T_h^k)_{k=0}^n\in A\big),
\end{aligned}
$$

completing the proof. □

4.3 Twisted transfer operators

For the rest of the thesis, we let $f\colon\Delta\to\Delta$ and $f_H\colon\Delta\times G\to\Delta\times G$ denote the tower maps for $F\colon Y\to Y$ and $F_H\colon Y\times G\to Y\times G$ respectively. In this section, we introduce twisted versions of the transfer operators for F and f, and show how they relate to the transfer operators for F_H and f_H respectively. This is done by utilising the results of Section 2.6.

Let $\mathcal{P}\colon L^1(Y;\mathbb{R}^d)\to L^1(Y;\mathbb{R}^d)$ denote the transfer operator for F with respect to the measure μ_Y. We define the *twisted transfer operator* $\mathcal{P}_H\colon L^1(Y;\mathbb{R}^d)\to L^1(Y;\mathbb{R}^d)$ for F by $\mathcal{P}_HV=\mathcal{P}(H^{-1}\cdot V)$. We begin by analysing how P_H behaves under iteration.

Proposition 4.3.1. *Let $V\in L^1(Y;\mathbb{R}^d)$ and $n\geq 1$. Then $\mathcal{P}_H^nV=\mathcal{P}^n(H_n^{-1}\cdot V)$, where $H_0=I_d$ and $H_n=H\,H\circ F\cdots H\circ F^{n-1}$.*

Proof. We proceed via induction. The base case is immediate by definition. Assume the result holds for $n-1 \geq 1$. Then for $W \in L^\infty(Y;\mathbb{R}^d)$, we have

$$\begin{aligned}
&\int_Y \big[\mathcal{P}_H^n V, W\big]\,\mathrm{d}\mu_Y = \int_Y \big[\mathcal{P}_H(\mathcal{P}_H^{n-1}V), W\big]\,\mathrm{d}\mu_Y \\
&= \int_Y \big[\mathcal{P}\big(H^{-1}\cdot\mathcal{P}^{n-1}(H_{n-1}^{-1}\cdot V)\big), W\big]\,\mathrm{d}\mu_Y = \int_Y \big[\mathcal{P}^{n-1}(H_{n-1}^{-1}\cdot V), H\cdot W\circ F\big]\,\mathrm{d}\mu_Y \\
&= \int_Y \big[H_{n-1}^{-1}\cdot V, (H\cdot W\circ F)\circ F^{n-1}\big]\,\mathrm{d}\mu_Y \\
&= \int_Y \big[H_{n-1}^{-1}\cdot V, H\circ F^{n-1}\cdot W\circ F^n\big]\,\mathrm{d}\mu_Y = \int_Y \big[H_n^{-1}\cdot V, W\circ F^n\big]\,\mathrm{d}\mu_Y \\
&= \int_Y \big[\mathcal{P}^n(H_n^{-1}\cdot V), W\big]\,\mathrm{d}\mu_Y.
\end{aligned}$$

Since W is arbitrary, the result follows. □

Remark 4.3.2. *For $n \geq 1$, let α_n denote the partition of Y into n–cylinders. Let $\zeta = \mathrm{d}\mu_Y/(\mathrm{d}\mu_Y\circ F)$ and denote $\zeta_n = \zeta\,\zeta\circ F\cdots\zeta\circ F^{n-1}$. Given $y \in Y$ and $a \in \alpha_n$, let y_a denote the unique element in a such that $F^n y_a = y$. Then for $V \in L^1(Y;\mathbb{R}^d)$, we have*

$$\mathcal{P}_H^n V(y) = \sum_{a\in\alpha_n} \zeta_n(y_a)H_n(y_a)^{-1}\cdot V(y_a)$$

by Proposition 4.3.1 and Proposition 3.3.2.

For the main results of this section, we require two preliminary lemmas. Let us fix the representation $\pi\colon G \to \mathrm{O}(d)$ and choose coordinates so that $[x,y] = \sum_{k=1}^d x_k y_k$ for $x, y \in \mathbb{R}^d$.

Lemma 4.3.3. *Let $\Phi\colon G \to \mathbb{R}^d$ be defined by $\Phi(g) = \pi(g)V$ where $V \in \mathbb{R}^d$. Then for any $\Psi \in L^2(G;\mathbb{R}^d)$, there exists $W \in \mathbb{R}^d$ such that $\int_G [\Phi,\Psi]\,\mathrm{d}\nu = \int_G [\Phi,\pi W]\,\mathrm{d}\nu$.*

Proof. Since π is finite dimensional, we may suppose without loss that π is irreducible. First, note that for $W \in \mathbb{R}^d$, we have

$$\int_G \big[\Phi(g),\pi(g)W\big]\,\mathrm{d}\nu = \int_G \big[\pi(g)V,\pi(g)W\big]\,\mathrm{d}\nu = \big[V,W\big] = \sum_{j=1}^d V_j W_j. \tag{4.3.1}$$

Next, by Theorem 2.6.1, we have

$$\Psi(g) = \sum_{\substack{\rho\in\Sigma \\ 1\le i,j\le d^\rho}} u^\rho_{i,j}(g) Z^\rho_{i,j},$$

where $Z^\rho_{i,j} \in \mathbb{R}^{d^\rho}$. Hence by orthogonality, it follows that

$$\begin{aligned}\int_G \big[\Phi(g), \Psi(g)\big]\,\mathrm{d}\nu &= \sum_{\substack{\rho\in\Sigma \\ 1\le i,j\le d^\rho}} \int_G \big[\pi(g)V, u^\rho_{i,j}(g)Z^\rho_{i,j}\big]\,\mathrm{d}\nu \\ &= \sum_{1\le i,j\le d} \int_G \big[\pi(g)V, u^\pi_{i,j}(g)Z^\pi_{i,j}\big]\,\mathrm{d}\nu.\end{aligned}$$

Denote the coordinates of $Z^\pi_{i,j}$ by $(Z^\pi_{i,j})_k$ for $1 \le k \le d$. The coordinates of $\pi(g)V$ are given by $(\pi(g)V)_k = \sum_{\ell=1}^d u^\pi_{k,\ell}(g)V_\ell$. Hence, continuing the calculation and using orthogonality once more, we have

$$\begin{aligned}\int_G \big[\Phi(g), \Psi(g)\big]\,\mathrm{d}\nu &= \sum_{1\le i,j\le d} \sum_{k=1}^d \int_G \big(\pi(g)V\big)_k u^\pi_{i,j}(g)(Z^\pi_{i,j})_k\,\mathrm{d}\nu \\ &= \sum_{1\le i,j\le d} \sum_{1\le k,\ell\le d} \int_G u^\pi_{k,\ell}(g)V_\ell\, u^\pi_{i,j}(g)(Z^\pi_{i,j})_k\,\mathrm{d}\nu \\ &= d \sum_{1\le i,j\le d} V_j (Z^\pi_{i,j})_i.\end{aligned}$$

Comparing with the right-hand side of (4.3.1), we obtain the solution

$$W_j = d\sum_{i=1}^d (Z^\pi_{i,j})_i, \tag{4.3.2}$$

completing the proof. □

Lemma 4.3.4. *Let $\Phi\colon Y\times G \to \mathbb{R}^d$ be defined by $\Phi(y,g) = \pi(g)V(y)$, where $V \in L^2(Y;\mathbb{R}^d)$. Then for any $\Psi \in L^2(Y\times G;\mathbb{R}^d)$, there exists $W \in L^2(Y;\mathbb{R}^d)$ such that*

(i) $\int_{Y\times G} [\Phi, \Psi]\,\mathrm{d}m_Y = \int_{Y\times G} [\Phi, \pi W]\,\mathrm{d}m_Y$.

(ii) *If* $\Psi'(y,g)=\Psi(Fy,g)$*, then* $\int_{Y\times G}[\Phi,\Psi']\,\mathrm{d}m_Y=\int_{Y\times G}[\Phi,\pi W\circ F]\,\mathrm{d}m_Y$.

Proof. Suppose without loss of generality that π is irreducible. By Fubini's theorem and Lemma 4.3.3, we have

$$\begin{aligned}\int_{Y\times G}\big[\Phi,\Psi\big]\,\mathrm{d}m_Y &= \int_Y\left(\int_G\big[\Phi,\Psi\big]\,\mathrm{d}\nu\right)\mathrm{d}\mu_Y\\ &= \int_Y\left(\int_G\big[\pi(g)V(y),\pi(g)W(y)\big]\,\mathrm{d}\nu\right)\mathrm{d}\mu_Y\\ &= \int_{Y\times G}\big[\pi(g)V(y),\pi(g)W(y)\big]\,\mathrm{d}m_Y\end{aligned}$$

for some $W\colon Y\to\mathbb{R}^d$. We next verify that $W\in L^2(Y;\mathbb{R}^d)$. Note that

$$\Psi(y,g)=\sum_{\substack{\rho\in\Sigma\\1\le i,j\le d^\rho}}u^\rho_{i,j}(g)Z^\rho_{i,j}(y). \tag{4.3.3}$$

We see from (4.3.2) that $W_j(y)=d\sum_{i=1}^d(Z^\pi_{i,j}(y))_i$ for $1\le j\le d$. Therefore, it suffices to show that $(Z^\pi_{i,j})_i\in L^2(Y)$ for $1\le i,j\le d$. Observe that

$$\big(Z^\pi_{i,j}(y)\big)_i=d\int_G\Psi_i(y,g)u^\pi_{i,j}(g)\,\mathrm{d}\nu,$$

where Ψ_i denotes the *ith* component of Ψ. Therefore

$$\big|\big(Z^\pi_{i,j}(y)\big)_i\big|\le\int_G\big|d\Psi_i(y,g)u^\pi_{i,j}(g)\big|\,\mathrm{d}\nu\le\big|\sqrt{d}\Psi_i(y,\cdot)\big|_2\big|\sqrt{d}u^\pi_{i,j}\big|_2=\big|\sqrt{d}\Psi_i(y,\cdot)\big|_2,$$

so that

$$\begin{aligned}\int_Y\big|\big(Z^\pi_{i,j}(y)\big)_i\big|^2\,\mathrm{d}\mu_Y &\le\int_Y\big|\sqrt{d}\Psi_i(y,\cdot)\big|_2^2\,\mathrm{d}\mu_Y=d\int_Y\left(\int_G|\Psi_i|^2(y,g)\,\mathrm{d}\nu\right)\mathrm{d}\mu_Y\\ &=d\int_{Y\times G}|\Psi_i|^2(y,g)\,\mathrm{d}m_Y<\infty.\end{aligned}$$

This proves (i).

For (ii), note that

$$\Psi'(y,g)=\Psi(Fy,g)=\sum_{\substack{\rho\in\Sigma\\1\le i,j\le d^\rho}}u^\rho_{i,j}(g)Z^\rho_{i,j}(Fy).$$

The result follows from (4.3.2). □

Proposition 4.3.5. *Let $\mathcal{M}\colon L^1(Y\times G;\mathbb{R}^d)\to L^1(Y\times G;\mathbb{R}^d)$ denote the transfer operator for F_H. Suppose $\Phi\colon Y\times G\to\mathbb{R}^d$ is given by $\Phi=g\cdot V$, where $V\in L^1(Y;\mathbb{R}^d)$. Then $\mathcal{M}\Phi=g\cdot\mathcal{P}_H V$.*

Proof. We first prove the result when $V\in L^2(Y;\mathbb{R}^d)$. We note that

$$\begin{aligned}&\int_{Y\times G}\big[\mathcal{M}\Phi,\Psi\big]\,\mathrm{d}m_Y=\int_{Y\times G}\big[\Phi,\Psi\circ F_H\big]\,\mathrm{d}m_Y\\&=\int_{Y\times G}\big[g\cdot V(y),\Psi\big(Fy,gH(y)\big)\big]\,\mathrm{d}m_Y=\int_{Y\times G}\big[gH(y)^{-1}\cdot V(y),\Psi(Fy,g)\big]\,\mathrm{d}m_Y,\end{aligned}$$

where the final equality follows from invariance of the Haar measure. Let $W\in L^2(Y;\mathbb{R}^d)$ be as in Lemma 4.3.4. Continuing the calculation above, we have

$$\begin{aligned}\int_{Y\times G}\big[\mathcal{M}\Phi,\Psi\big]\,\mathrm{d}m_Y&=\int_{Y\times G}\big[g\cdot H(y)^{-1}\cdot V(y),g\cdot W(Fy)\big]\,\mathrm{d}m_Y\\&=\int_Y\big[H(y)^{-1}\cdot V(y),W(Fy)\big]\,\mathrm{d}\mu_Y=\int_Y\big[\mathcal{P}_H V,W\big]\,\mathrm{d}\mu_Y\\&=\int_{Y\times G}\big[g\cdot\mathcal{P}_H V,g\cdot W\big]\,\mathrm{d}m_Y=\int_{Y\times G}\big[g\cdot\mathcal{P}_H V,\Psi\big]\,\mathrm{d}m_Y,\end{aligned}$$

where the first equality follows from Lemma 4.3.4 (ii) and the final equality follows from Lemma 4.3.4 (i). The result for $V\in L^2(Y;\mathbb{R}^d)$ follows.

To complete the proof, we use the density of $L^2(Y;\mathbb{R}^d)$ in $L^1(Y;\mathbb{R}^d)$. Suppose $\Phi=g\cdot V$ where $V\in L^1(Y;\mathbb{R}^d)$. Let $V_n\in L^2(Y;\mathbb{R}^d)$ with $V_n\to V$ in $L^1(Y;\mathbb{R}^d)$ and set $\Phi_n=g\cdot V_n$. From the first part of the proof, we have $\mathcal{M}\Phi_n=g\cdot\mathcal{P}_H V_n$ for all $n\geq 1$. Now, by Proposition 2.5.3 (viii), we have that $\mathcal{M}$ is a bounded operator on $L^1(Y\times G;\mathbb{R}^d)$. In addition, for $W\in L^1(Y;\mathbb{R}^d)$, we have

$$|\mathcal{P}_H W|_1=\big|\mathcal{P}(H^{-1}\cdot W)\big|_1\leq|H^{-1}\cdot W|_1=|W|_1,$$

so that $\mathcal{P}_H$ is a bounded operator on $L^1(Y;\mathbb{R}^d)$. It follows that

$$\mathcal{M}\Phi=\lim_{n\to\infty}\mathcal{M}\Phi_n=g\cdot\lim_{n\to\infty}\mathcal{P}_H V_n=g\cdot\mathcal{P}_H V,$$

completing the proof. □

Let $\mathcal{L}\colon L^1(\Delta;\mathbb{R}^d) \to L^1(\Delta;\mathbb{R}^d)$ denote the transfer operator for f and $\hat{H}\colon \Delta \to G$ be defined as in Remark 4.2.1 by

$$\hat{H}(y,\ell) = \begin{cases} I_d & \text{if } 0 \le \ell \le \tau(y) - 2, \\ H(y) & \text{if } \ell = \tau(y) - 1. \end{cases}$$

The twisted transfer operator $\mathcal{L}_H\colon L^1(\Delta;\mathbb{R}^d) \to L^1(\Delta;\mathbb{R}^d)$ for f is defined by $\mathcal{L}_H \hat{v} = \mathcal{L}(\hat{H}^{-1} \cdot \hat{v})$.

Remark 4.3.6. *Given $y \in Y$ and $a \in \alpha$, let y_a denote the unique element in a such that $Fy_a = y$. For $\hat{v} \in L^1(\Delta;\mathbb{R}^d)$, it follows from Proposition 3.4.4 that*

$$\mathcal{L}_H \hat{v}(y,\ell) = \begin{cases} \sum_{a\in\alpha} \zeta(y_a) H(y_a)^{-1} \cdot \hat{v}(y_a, \tau(y_a) - 1) & \text{if } \ell = 0, \\ \hat{v}(y, \ell - 1) & \text{if } \ell \ge 1, \end{cases}$$

where $\zeta = \mathrm{d}\mu_Y / (\mathrm{d}\mu_Y \circ F)$.

An identical proof to that of Proposition 4.3.5 can be done to conclude the following:

Proposition 4.3.7. *Let $\hat{\mathcal{L}}\colon L^1(\Delta \times G;\mathbb{R}^d) \to L^1(\Delta \times G;\mathbb{R}^d)$ denote the transfer operator for f_H. Suppose $\hat{\phi}\colon \Delta \times G \to \mathbb{R}^d$ is given by $\hat{\phi} = g \cdot \hat{v}$, where $\hat{v} \in L^1(\Delta;\mathbb{R}^d)$. Then $\hat{\mathcal{L}}\hat{\phi} = g \cdot \mathcal{L}_H \hat{v}$.*

4.4 Basic properties of V

Throughout this section, we let $V\colon Y \to \mathbb{R}^d$ be as in (4.1.1). We begin by introducing the notion of locally Lipschitz functions. Let $a \in \alpha$, $\gamma \in (0,1)$, and $W\colon Y \to \mathbb{R}^d$. We adopt a convenient abuse of notation and define

$$|\mathbb{1}_a W|_\gamma = \sup_{\substack{x,y\in a \\ x \ne y}} \frac{\big|W(x) - W(y)\big|}{d_\gamma(x,y)}.$$

We say that W is *locally Lipschitz* and write $W \in F_\gamma^{\mathrm{loc}}(Y;\mathbb{R}^d)$ if $\|\mathbb{1}_a W\|_\gamma = |\mathbb{1}_a W|_\gamma + |\mathbb{1}_a W|_\infty < \infty$ for all $a \in \alpha$. The above definition extends to subsets of

$\mathbb{R}^{d,d}$, and so it makes sense to speak of locally Lipschitz cocyles. We restrict to $\gamma \in [\lambda^{-\eta}, 1)$, which makes $F\colon Y \to Y$ Gibbs-Markov and enables us to use the results of Section 3.3. Throughout, we use that $d(x,y)^\eta \leq d_\gamma(x,y)$ for $x, y \in Y$. We first verify that the induced cocycle H and induced observable V are locally Lipschitz.

Lemma 4.4.1. *Suppose $a \geq 1$ and $x, b \geq 0$ with $x \leq a$ and $x \leq b$. Then $x \leq ab^\epsilon$ for all $\epsilon \in (0,1]$.*

Proof. If $b \leq 1$, then $x \leq b \leq b^\epsilon \leq ab^\epsilon$. If $b > 1$, then $x \leq a \leq ab^\epsilon$. □

Proposition 4.4.2. *Let $\epsilon \in (0,1]$. The following hold true:*

(i) $H \in F^{\mathrm{loc}}_{\gamma^\epsilon}(Y; G)$ *with* $\|1_a H\|_{\gamma^\epsilon} \leq C\tau(a)^\epsilon \|h\|_\eta$ *for all* $a \in \alpha$

(ii) $V \in F^{\mathrm{loc}}_{\gamma^\epsilon}(Y; \mathbb{R}^d)$ *with* $\|1_a V\|_{\gamma^\epsilon} \leq C\tau(a)^{1+\epsilon} \|v\|_\eta \|h\|_\eta$ *for all* $a \in \alpha$.

Proof. We begin by making the following observation: For $n \geq 1$ and $x, y \in X$, we have

$$\big\|h_n(x) - h_n(y)\big\| \leq \sum_{k=0}^{n-1} \big\|h(T^k x) - h(T^k y)\big\|. \tag{4.4.1}$$

Indeed, when $n = 1$ then the result is trivial. If we assume (4.4.1) holds for $n - 1 \geq 1$, then

$$\begin{aligned}
&\big\|h_n(x) - h_n(y)\big\| = \big\|h_{n-1}(x)h(T^{n-1}x) - h_{n-1}(y)h(T^{n-1}y)\big\| \\
&\leq \big\|h_{n-1}(x) - h_{n-1}(y)\big\|\big\|h(T^{n-1}x)\big\| + \big\|h_{n-1}(y)\big\|\big\|h(T^{n-1}x) - h(T^{n-1}y)\big\| \\
&= \big\|h_{n-1}(x) - h_{n-1}(y)\big\| + \big\|h(T^{n-1}x) - h(T^{n-1}y)\big\| \\
&\leq \left(\sum_{k=0}^{n-2} \big\|h(T^k x) - h(T^k y)\big\|\right) + \big\|h(T^{n-1}x) - h(T^{n-1}y)\big\| \\
&= \sum_{k=0}^{n-1} \big\|h(T^k x) - h(T^k y)\big\|,
\end{aligned}$$

yielding (4.4.1) for n.

We now proceed with the proof of (i). Fix $a \in \alpha$. For $x, y \in a$ and $1 \le \ell \le \tau(a)$, we have $\|h_\ell(x) - h_\ell(y)\| \le \|h_\ell(x)\| + \|h_\ell(y)\| = 2$. Moreover, from (4.4.1), we have

$$\begin{aligned}\big\|h_\ell(x) - h_\ell(y)\big\| &\le \sum_{k=0}^{\ell-1} \big\|h(T^k x) - h(T^k y)\big\| \le \sum_{k=0}^{\ell-1} |h|_\eta\, d(T^k x, T^k y)^\eta \\ &\le \ell |h|_\eta C_0^\eta\, d(Fx, Fy)^\eta \le \ell |h|_\eta C_0^\eta\, d_\gamma(Fx, Fy) \\ &\le \tau(a)\|h\|_\eta C_0^\eta \gamma^{-1}\, d_\gamma(x, y), \end{aligned} \tag{4.4.2}$$

where we use Proposition 3.3.3 in the final inequality. Applying Lemma 4.4.1 gives

$$\|h_\ell(x) - h_\ell(y)\| \le 2\big(\tau(a)\|h\|_\eta C_0^\eta \gamma^{-1}\big)^\epsilon\, d_\gamma(x, y)^\epsilon \ll \tau(a)^\epsilon \|h\|_\eta\, d_{\gamma^\epsilon}(x, y). \tag{4.4.3}$$

Taking $\ell = \tau(a)$ gives us $|1\!\!1_a H|_{\gamma^\epsilon} \ll \tau(a)^\epsilon \|h\|_\eta$. In addition, we have $\|H(y)\| = 1 \le \tau(a)^\epsilon$ for all $y \in a$ so that $|1\!\!1_a H|_\infty \le \tau(a)^\epsilon$. This completes the proof of (i).

For (ii), note that for $a \in \alpha$ and $x, y \in a$, we have

$$\begin{aligned}\big|V(x) - V(y)\big| &= \Bigg| \sum_{\ell=0}^{\tau(a)-1} h_\ell(x) \cdot v(T^\ell x) - \sum_{\ell=0}^{\tau(a)-1} h_\ell(y) \cdot v(T^\ell y)\Bigg| \\ &\le \sum_{\ell=0}^{\tau(a)-1} \big|h_\ell(x) \cdot \big(v(T^\ell x) - v(T^\ell y)\big)\big| + \sum_{\ell=0}^{\tau(a)-1} \big|\big(h_\ell(x) - h_\ell(y)\big) \cdot v(T^\ell y)\big| \\ &= \sum_{\ell=0}^{\tau(a)-1} \big|v(T^\ell x) - v(T^\ell y)\big| + \sum_{\ell=0}^{\tau(a)-1} \big|\big(h_\ell(x) - h_\ell(y)\big) \cdot v(T^\ell y)\big| \\ &\le \sum_{\ell=0}^{\tau(a)-1} \big|v(T^\ell x) - v(T^\ell y)\big| + \sum_{\ell=0}^{\tau(a)-1} \big\|h_\ell(x) - h_\ell(y)\big\| |v|_\infty \end{aligned}$$

Next, we see that

$$\begin{aligned}\sum_{\ell=0}^{\tau(a)-1} \big|v(T^\ell x) - v(T^\ell y)\big| &\le \sum_{\ell=0}^{\tau(a)-1} |v|_\eta\, d(T^\ell x, T^\ell y)^\eta \le \sum_{\ell=0}^{\tau(a)-1} |v|_\eta C_0^\eta\, d(Fx, Fy)^\eta \\ &\le \tau(a)|v|_\eta C_0^\eta \gamma^{-\epsilon}\, d_{\gamma^\epsilon}(x, y), \end{aligned}$$

where the final inequality follows from the argument in (4.4.2). Moreover, we have from (4.4.3) that

$$\sum_{\ell=0}^{\tau(a)-1} \big\|h_\ell(x)-h_\ell(y)\big\||v|_\infty \ll \sum_{\ell=0}^{\tau(a)-1} \tau(a)^\epsilon \|h\|_\eta |v|_\infty\, d_{\gamma^\epsilon}(x,y) = \tau(a)^{1+\epsilon}\|h\|_\eta\, d_{\gamma^\epsilon}(x,y).$$

Therefore

$$|1\!\!1_a V|_{\gamma^\epsilon} \ll \tau(a)^{1+\epsilon}\|v\|_\eta\|h\|_\eta.$$

In addition,

$$\big|V(y)\big| \le \sum_{\ell=0}^{\tau(a)-1} \big|h_\ell(y)\cdot v(T^\ell y)\big| = \sum_{\ell=0}^{\tau(a)-1} \big|v(T^\ell y)\big| \le \tau(a)|v|_\infty,$$

and so

$$|1\!\!1_a V|_\infty \le \tau(a)|v|_\infty. \tag{4.4.4}$$

The result follows. □

Recall the twisted transfer operator $\mathcal{P}_H\colon L^1(Y;\mathbb{R}^d) \to L^1(Y;\mathbb{R}^d)$ defined by $\mathcal{P}_H W = \mathcal{P}(H^{-1}\cdot W)$. The next proposition shows that $\mathcal{P}_H$ has a smoothing effect on V, with regularity depending on the integrability of τ.

Proposition 4.4.3. *Suppose the return time $\tau \in L^p(Y)$ for $p > 1$.*

(i) *If $p \ge 2$, then $P_H V \in F_\gamma(Y;\mathbb{R}^d)$ with $\|P_H V\|_\gamma \le C\|v\|_\eta\|h\|_\eta$.*

(ii) *If $1 < p < 2$, then $P_H V \in F_{\gamma^{p-1}}(Y;\mathbb{R}^d)$ with $\|P_H V\|_{\gamma^{p-1}} \le C\|v\|_\eta\|h\|_\eta$.*

Proof. We begin by estimating the sup norm. For $y \in Y$, we have from Remark 4.3.2, Proposition 3.3.9, and (4.4.4) that

$$\begin{aligned}\big|P_H V(y)\big| &\le \sum_{a\in\alpha} \zeta(y_a)\big|H(y_a)^{-1}\cdot V(y_a)\big| \le D\sum_{a\in\alpha}\mu_Y(a)|1\!\!1_a V|_\infty \\ &\le D\sum_{a\in\alpha}\mu_Y(a)\tau(a)|v|_\infty = D|\tau|_1|v|_\infty.\end{aligned}$$

Therefore $|P_H V|_\infty \ll |v|_\infty$.

To estimate the Lipschitz semi-norm, note that for $x, y \in Y$, we have

$$\begin{aligned}
&\big|P_H V(x) - P_H V(y)\big| \\
&= \Big|\sum_{a\in\alpha} \zeta(x_a) H(x_a)^{-1} \cdot V(x_a) - \sum_{a\in\alpha} \zeta(y_a) H(y_a)^{-1} \cdot V(y_a)\Big| \\
&\le \sum_{a\in\alpha} \big|\zeta(x_a) - \zeta(y_a)\big| \big|H(x_a)^{-1} \cdot V(x_a)\big| \\
&\quad + \sum_{a\in\alpha} \zeta(y_a) \big|\big(H(x_a)^{-1} - H(y_a)^{-1}\big) \cdot V(x_a)\big| \\
&\quad + \sum_{a\in\alpha} \zeta(y_a) \big|H(y_a)^{-1} \cdot \big(V(x_a) - V(y_a)\big)\big| =: I + II + III. \qquad (4.4.5)
\end{aligned}$$

Proof of (i). We look at I, II, and III in turn. The terms involving ζ are dealt with using Proposition 3.3.9. For I, we have from (4.4.4) that

$$I \le D|v|_\infty \sum_{a\in\alpha} \mu_Y(a)\tau(a)\, d_\gamma(Fx_a, Fy_a) \ll |v|_\infty\, d_\gamma(x, y).$$

We now look at II. Note that from orthogonality, we have

$$\begin{aligned}
\big\|H(x_a)^{-1} - H(y_a)^{-1}\big\| &= \big\|H(x_a)^T - H(y_a)^T\big\| = \big\|(H(x_a) - H(y_a))^T\big\| \\
&= \big\|H(x_a) - H(y_a)\big\|. \qquad (4.4.6)
\end{aligned}$$

Combining this with (4.4.4) gives

$$\begin{aligned}
\big|\big(H(x_a)^{-1} - H(y_a)^{-1}\big) \cdot V(x_a)\big| &\le \big\|H(x_a)^{-1} - H(y_a)^{-1}\big\| |1_a V|_\infty \\
&= \big\|H(x_a) - H(y_a)\big\| \tau(a) |v|_\infty. \qquad (4.4.7)
\end{aligned}$$

It follows from Proposition 4.4.2 (i) that

$$II \ll |v|_\infty \|h\|_\eta \sum_{a\in\alpha} \mu_Y(a)\tau(a)^2\, d_\gamma(x_a, y_a) \ll |v|_\infty \|h\|_\eta\, d_\gamma(x, y),$$

where the final estimate uses Proposition 3.3.3 and the fact that $p \ge 2$. Similarly for III, we use Proposition 4.4.2 (ii) to deduce that

$$\begin{aligned}
III &= \sum_{a\in\alpha} \zeta(y_a) \big|V(x_a) - V(y_a)\big| \ll \|v\|_\eta \|h\|_\eta \sum_{a\in\alpha} \mu_Y(a)\tau(a)^2\, d_\gamma(x_a, y_a) \\
&\ll \|v\|_\eta \|h\|_\eta\, d_\gamma(x, y).
\end{aligned}$$

Combining the above, we conclude that $P_H V \in F_\gamma(Y)$ with $\|P_H V\|_\gamma \ll \|v\|_\eta \|h\|_\eta$ when $p \geq 2$.

Proof of (ii). We again want to estimate the terms in (4.4.5). We deal with the terms involving ζ by using Proposition 3.3.9. For I, proceeding as in (i) gives

$$I \ll |v|_\infty \, d_\gamma(x,y) \leq |v|_\infty \, d_{\gamma^{p-1}}(x,y).$$

For II, we have from Proposition 4.4.2 (i) that

$$\big\|H(x_a) - H(y_a)\big\| \ll \tau(a)^{p-1} \|h\|_\eta \, d_{\gamma^{p-1}}(x_a, y_a).$$

Combining this with (4.4.7) and Proposition 3.3.3 gives

$$II \ll |v|_\infty \|h\|_\eta \sum_{a \in \alpha} \mu_Y(a) \tau(a)^p \, d_{\gamma^{p-1}}(x_a, y_a) \ll |v|_\infty \|h\|_\eta \, d_{\gamma^{p-1}}(x,y).$$

Finally, for III, Proposition 4.4.2 (ii) gives

$$\big|V(x_a) - V(y_a)\big| \ll \tau(a)^p \|v\|_\eta \|h\|_\eta \, d_{\gamma^{p-1}}(x_a, y_a),$$

and combining this with Proposition 3.3.3 yields

$$\begin{aligned} III = \sum_{a \in \alpha} \zeta(y_a) \big|V(x_a) - V(y_a)\big| &\ll \|v\|_\eta \|h\|_\eta \sum_{a \in \alpha} \mu_Y(a) \tau(a)^p \, d_{\gamma^{p-1}}(x_a, y_a) \\ &\ll \|v\|_\eta \|h\|_\eta \, d_{\gamma^{p-1}}(x,y). \end{aligned}$$

It follows that $|P_H V|_{\gamma^{p-1}} \ll \|v\|_\eta \|h\|_\eta$ when $1 < p < 2$. □

Remark 4.4.4. *From here on, we restrict to* $\gamma \in [\max\{\lambda^{-\eta}, \lambda^{-\eta(p-1)}\}, 1)$. *In particular,* $P_H V \in F_\gamma(Y; \mathbb{R}^d)$ *for all* $p > 1$.

4.5 Spectral properties of $\mathcal{P}_H$

In this section, we obtain a spectral decomposition of the twisted transfer operator $\mathcal{P}_H$ when acting on $F_\gamma(Y; \mathbb{R}^d) \subset L^1(Y; \mathbb{R}^d)$. Once this is done, we give some consequences which are required throughout. We begin by giving a Lasota-Yorke inequality. For this, we need the pointwise expression of $\mathcal{P}_H^n$ given in Remark 4.3.2. For $V \in L^1(Y;\mathbb{R}^d)$ and $y \in Y$, we recall that this is given by

$$\mathcal{P}_H^n V(y) = \sum_{a \in \alpha_n} \zeta_n(y_a) H_n(y_a)^{-1} \cdot V(y_a),$$

where y_a is the unique element in a such that $F^n y_a = y$.

Proposition 4.5.1. *It holds true that $\mathcal{P}_H \colon F_\gamma(Y;\mathbb{R}^d) \to F_\gamma(Y;\mathbb{R}^d)$. Moreover,*

(i) $|\mathcal{P}_H^n V|_1 \leq |V|_1$ *for all* $n \geq 1$ *and* $V \in F_\gamma(Y;\mathbb{R}^d)$.

(ii) $\|\mathcal{P}_H^n V\|_\gamma \leq C\|h\|_\eta(\gamma^n \|V\|_\gamma + |V|_1)$ *for all* $n \geq 1$ *and* $V \in F_\gamma(Y;\mathbb{R}^d)$.

Proof. Fix $V \in F_\gamma(Y;\mathbb{R}^d)$. By Proposition 2.5.3 (viii), we have

$$|\mathcal{P}_H^n V|_1 = \big|\mathcal{P}^n(H_n^{-1} \cdot V)\big|_1 \leq |H_n^{-1} \cdot V|_1 = |V|_1.$$

This proves (i).

We now look at (ii). Recall $\|\mathcal{P}_H^n V\|_\gamma = |\mathcal{P}_H^n V|_\gamma + |\mathcal{P}_H^n V|_\infty$. Note that for $x \in Y$, we have

$$\begin{aligned}\big|\mathcal{P}_H^n V(x)\big| &\leq \left|\mathcal{P}_H^n V(x) - \int_Y \mathcal{P}_H^n V(y)\,\mathrm{d}\mu_Y(y)\right| + |\mathcal{P}_H^n V|_1 \\ &\leq \left|\int_Y \big(\mathcal{P}_H^n V(x) - \mathcal{P}_H^n V(y)\big)\,\mathrm{d}\mu_Y(y)\right| + |V|_1 \\ &\leq |\mathcal{P}_H^n V|_\gamma \operatorname{diam}_\gamma(Y) + |V|_1 = |\mathcal{P}_H^n V|_\gamma + |V|_1, \end{aligned} \tag{4.5.1}$$

so that $|\mathcal{P}_H^n V|_\infty \leq |\mathcal{P}_H^n V|_\gamma + |V|_1$. Therefore, it suffices to prove the estimate for $|\mathcal{P}_H^n V|_\gamma$. For $x, y \in Y$, we have

$$\begin{aligned}&\big|\mathcal{P}_H^n V(x) - \mathcal{P}_H^n V(y)\big| \leq \sum_{a \in \alpha_n} \big|\zeta_n(x_a) - \zeta_n(y_a)\big|\big|H_n(x_a)^{-1} \cdot V(x_a)\big| \\ &+ \sum_{a \in \alpha_n} \zeta_n(y_a)\big|\big(H_n(x_a)^{-1} - H_n(y_a)^{-1}\big) \cdot V(x_a)\big| \\ &+ \sum_{a \in \alpha_n} \zeta_n(y_a)\big|H_n(y_a)^{-1} \cdot \big(V(x_a) - V(y_a)\big)\big| =: I + II + III.\end{aligned}$$

All the terms involving ζ_n are dealt with using Proposition 3.3.9. Repeating the argument of (4.5.1) for V and $x_a \in a$, we have $|V(x_a)| \le \gamma^n |V|_\gamma + |V|_1$, where we use that $\operatorname{diam}_\gamma(a) = \gamma^n$ for $a \in \alpha_n$. Therefore

$$|\mathbb{1}_a V|_\infty \le \gamma^n |V|_\gamma + |V|_1. \tag{4.5.2}$$

It follows that

$$I \le D \sum_{a\in\alpha_n} \mu_Y(a)\big(\gamma^n|V|_\gamma + |V|_1\big)\, d_\gamma(F^n x_a, F^n y_a) = D\big(\gamma^n|V|_\gamma + |V|_1\big)\, d_\gamma(x,y).$$

We next study II. Note that since $H_n\colon Y \to G$, we have from (4.4.6) that

$$\big|\big(H_n(x_a)^{-1} - H_n(y_a)^{-1}\big)\cdot V(x_a)\big| \le \big\|H_n(x_a) - H_n(y_a)\big\| |\mathbb{1}_a V|_\infty. \tag{4.5.3}$$

Moreover, we have

$$\big\|H_n(x_a) - H_n(y_a)\big\| \le \sum_{k=0}^{n-1} \big\|H(F^k x_a) - H(F^k y_a)\big\|.$$

Indeed, the proof of (4.4.1) goes through identically. Therefore

$$\begin{aligned}
\big\|H_n(x_a) - H_n(y_a)\big\| &\le \sum_{k=0}^{n-1} \big\|H(F^k x_a) - H(F^k y_a)\big\| \\
&\le \sum_{k=0}^{n-1} |\mathbb{1}_{F^k(a)} H|_\gamma\, d_\gamma(F^k x_a, F^k y_a) = \sum_{k=0}^{n-1} |\mathbb{1}_{F^k(a)} H|_\gamma \gamma^{k-n}\, d_\gamma(F^n x_a, F^n y_a) \\
&= \sum_{k=0}^{n-1} |\mathbb{1}_{F^k(a)} H|_\gamma \gamma^{k-n}\, d_\gamma(x,y),
\end{aligned}$$

where we use Proposition 3.3.3 in the first equality. Combining this with (4.5.2) and (4.5.3), we have

$$II \le D\big(\gamma^n|V|_\gamma + |V|_1\big) \sum_{a\in\alpha_n} \mu_Y(a) \sum_{k=0}^{n-1} |\mathbb{1}_{F^k(a)} H|_\gamma \gamma^{k-n}\, d_\gamma(x,y). \tag{4.5.4}$$

Now,

$$\begin{aligned}
\sum_{a\in\alpha_n}\mu_Y(a)\sum_{k=0}^{n-1}|\mathbb{1}_{F^k(a)}H|_\gamma\gamma^{k-n} &= \sum_{k=0}^{n-1}\sum_{b\in\alpha_{n-k}}\sum_{\substack{a\in\alpha_n\\F^k(a)=b}}\mu_Y(a)|\mathbb{1}_{F^k(a)}H|_\gamma\gamma^{k-n}\\
&= \sum_{k=0}^{n-1}\gamma^{k-n}\sum_{b\in\alpha_{n-k}}|\mathbb{1}_bH|_\gamma\sum_{\substack{a\in\alpha_n\\F^k(a)=b}}\mu_Y(a) = \sum_{k=0}^{n-1}\gamma^{k-n}\sum_{b\in\alpha_{n-k}}|\mathbb{1}_bH|_\gamma\mu_Y(b)\\
&\le \sum_{k=0}^{n-1}\gamma^{k-n}\sum_{a\in\alpha}|\mathbb{1}_aH|_\gamma\mu_Y(a) \le \frac{1}{1-\gamma}\sum_{a\in\alpha}|\mathbb{1}_aH|_\gamma\mu_Y(a)\\
&\ll \sum_{a\in\alpha}\tau(a)\|h\|_\eta\mu_Y(a) \ll \|h\|_\eta,
\end{aligned}$$

where we use Proposition 4.4.2 (i) and integrability of τ in the final line. Combining this with (4.5.4) gives us

$$II \ll \|h\|_\eta\big(\gamma^n|V|_\gamma + |V|_1\big)\, d_\gamma(x,y).$$

For III, we have

$$\begin{aligned}
III &= \sum_{a\in\alpha_n}\zeta_n(y_a)\big|V(x_a)-V(y_a)\big| \le D\sum_{a\in\alpha_n}\mu_Y(a)|V|_\gamma\, d_\gamma(x_a,y_a)\\
&= D\sum_{a\in\alpha_n}\mu_Y(a)|V|_\gamma\gamma^n\, d_\gamma(F^nx_a,F^ny_a) = D\gamma^n|V|_\gamma\, d_\gamma(x,y),
\end{aligned}$$

where we use Proposition 3.3.3 in the second equality. Therefore $|\mathcal{P}_H^nV|_\gamma \ll \|h\|_\eta(\gamma^n\|V\|_\gamma + |V|_1)$, proving (ii). □

Corollary 4.5.2. *Let $\mathcal{B}(F_\gamma(Y;\mathbb{R}^d))$ denote the space of bounded linear operators on $F_\gamma(Y;\mathbb{R}^d)$. There exists $r\in(\gamma,1)$ such that*

$$\pi = \frac{1}{2\pi i}\int_{\partial B_r(0)}(zI-\mathcal{P}_H)^{-1}\,\mathrm{d}z \in \mathcal{B}\big(F_\gamma(Y;\mathbb{R}^d)\big)$$

defines a projection, so that $F_\gamma(Y;\mathbb{R}^d) = E_0\oplus E_1$ where

(i) *$E_0 = \operatorname{Im}\pi$ and $E_1 = \ker\pi$ are closed and $\mathcal{P}_H$ –invariant.*

(ii) $\dim E_1 < \infty$ *and all eigenvalues of* $\mathcal{P}_H|_{E_1}$ *lie on the unit circle.*

(iii) $r(\mathcal{P}_H|_{E_0}) < r$.

Proof. We begin by verifying the hypotheses of Proposition 2.7.4 in turn. Clearly $|\cdot|_1 \le \|\cdot\|_\gamma$. Proposition 4.5.1 proves the second hypothesis. Finally, the unit ball of $F_\gamma(Y;\mathbb{R}^d)$ is compact in $L^1(Y;\mathbb{R}^d)$ by Proposition 3.3.7. Therefore $r_{\mathrm{ess}}(\mathcal{P}_H) \le \gamma$. We next show that $r(\mathcal{P}_H) \le 1$. For $V \in F_\gamma(Y;\mathbb{R}^d)$ and $n \ge 1$, we have from Proposition 4.5.1 (ii) that

$$\|\mathcal{P}_H^n V\|_\gamma \ll \gamma^n \|V\|_\gamma + |V|_1 \ll \|V\|_\gamma.$$

Letting $\|\cdot\|$ denote the operator norm, we have the existence of $c > 0$ such that

$$\|\mathcal{P}_H^n\| \le c \quad \text{for all } n \ge 1. \tag{4.5.5}$$

Therefore,

$$r(\mathcal{P}_H) = \lim_{n\to\infty} \|\mathcal{P}_H^n\|^{1/n} \le \lim_{n\to\infty} c^{1/n} = 1.$$

Let us choose $r \in (\gamma, 1)$ sufficiently large so that $\sigma(\mathcal{P}_H) \cap \{z \in \mathbb{C} : r \le |z| \le 1\} \subset \{z \in \mathbb{C} : |z| = 1\}$. The result then follows from Theorem 2.7.1. □

The next lemma shows that on restriction to E_0, the operators $\mathcal{P}_H^k$ decay exponentially.

Lemma 4.5.3. *There exists* $C > 0$ *such that* $\|\mathcal{P}_H|_{E_0}^n\| \le Cr^n$ *for all* $n \ge 1$.

Proof. We first note by Corollary 4.5.2 (iii) that $r(\mathcal{P}_H|_{E_0}) < r$. Now, $r(\mathcal{P}_H|_{E_0}) = \lim_{n\to\infty} \|\mathcal{P}_H|_{E_0}^n\|^{1/n}$, so there exists $N \ge 1$ such that $\|\mathcal{P}_H|_{E_0}^n\| \le r^n$ for all $n \ge N$. Let $C' > 0$ be large enough such that for all $1 \le n \le N-1$, we have $\|\mathcal{P}_H|_{E_0}^n\| \le C'r^n$. The result follows. □

We next examine how $\mathcal{P}_H$ behaves on E_1. Since $\dim E_1 < \infty$, we can write E_1 as the direct sum of its generalised eigenspaces. The next proposition removes the possibility of generalised eigenfunctions existing.

Lemma 4.5.4. *The generalised eigenspaces of* $\mathcal{P}_H|_{E_1}$ *are ordinary eigenspaces.*

Proof. First note that $\mathcal{P}_H$ acts on E_1 as a matrix since E_1 is finite dimensional. Let $\omega \in [0, 2\pi)$ be such that $e^{i\omega}$ is an eigenvalue for $\mathcal{P}_H$, with corresponding generalised eigenspace G_ω and Jordan block $\boldsymbol{J}_\omega$. We show that $\boldsymbol{J}_\omega$ is of degree 1. Suppose for contradiction that $\boldsymbol{J}_\omega$ is of degree $j > 1$. We consider the restriction of $\mathcal{P}_H$ to G_ω, which is represented by $\boldsymbol{J_\omega}$. Iterating this Jordan block, we have

$$\boldsymbol{J}^n_{\boldsymbol{\omega}} = \begin{pmatrix} (e^{i\omega})^n & \binom{n}{1}(e^{i\omega})^{n-1} & \cdots & \binom{n}{j-2}(e^{i\omega})^{n-j+2} & \binom{n}{j-1}(e^{i\omega})^{n-j+1} \\ & (e^{i\omega})^n & \cdots & \binom{n}{j-3}(e^{i\omega})^{n-j+3} & \binom{n}{j-2}(e^{i\omega})^{n-j+2} \\ & & \ddots & \vdots & \vdots \\ & & & (e^{i\omega})^n & \binom{n}{1}(e^{i\omega})^{n-1} \\ & & & & (e^{i\omega})^n \end{pmatrix}.$$

To work out the operator norm of $\mathcal{P}_H|_{G_\omega}$, note that since all norms on finite dimensional spaces are equivalent, we may work without loss of generality with the ℓ^1-norm. Consider the vector $\mathbf{e} = (e_1, \dots e_j)^T \in \mathbb{R}^d$ defined by

$$e_i = \begin{cases} 1 & \text{if } i = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Then $|\mathbf{e}|_{\ell_1} = 1$, and so we have $\|\boldsymbol{J}^n_{\boldsymbol{\omega}}\| \geq |\boldsymbol{J}^n_{\boldsymbol{\omega}}\mathbf{e}|_{\ell^1} = n+1$. On the other hand, we have $\|\boldsymbol{J}^n_{\boldsymbol{\omega}}\| = \|\mathcal{P}_H|^n_{G_\omega}\| \leq \|\mathcal{P}^n_H\| \leq c$ for all $n \geq 1$ from (4.5.5). Taking n sufficiently large yields a contradiction. It follows that all Jordan blocks corresponding to eigenvalues on the unit circle are of degree 1, completing the proof. □

Remark 4.5.5. *Let $F_{\gamma,0}(Y;\mathbb{R}^d) = \{V \in F_\gamma(Y;\mathbb{R}^d) \mid \int_{Y\times G} g\cdot V \,\mathrm{d}m_Y = 0\}$ This is a closed subspace of $F_\gamma(Y;\mathbb{R}^d)$ and hence a Banach space when equipped with $\|\cdot\|_\gamma$. Let $\mathcal{M}$ denote the transfer operator for F_H. Note that if $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$, we have from Proposition 4.3.5 and Proposition 2.5.3 (ii) that*

$$\int_{Y\times G} g\cdot P_H V \,\mathrm{d}m_Y = \int_{Y\times G} \mathcal{M}(g\cdot V)\,\mathrm{d}m_Y = \int_{Y\times G} g\cdot V\,\mathrm{d}m_Y = 0. \qquad (4.5.6)$$

In particular, $\mathcal{P}_H\colon F_{\gamma,0}(Y;\mathbb{R}^d) \to F_{\gamma,0}(Y;\mathbb{R}^d)$. In addition, the previous arguments in this section go through and we get a similar decomposition to that of Corollary 4.5.2.

To conclude this section, we show that 1 is not an eigenvalue for $\mathcal{P}_H$ when viewed as an operator on $F_{\gamma,0}(Y;\mathbb{R}^d)$. To do this, we consider the $L^2(Y;\mathbb{R}^d)$–adjoint of $\mathcal{P}_H$, denoted $\mathcal{U}_H\colon L^2(Y;\mathbb{R}^d) \to L^2(Y;\mathbb{R}^d)$. We call this the *twisted Koopman operator* for F, and it is defined by $\mathcal{U}_H V = H\cdot V\circ F$ for $V \in L^2(Y;\mathbb{R}^d)$. It satisfies

$$\int_Y \big[\mathcal{U}_H V, W\big]\,\mathrm{d}\mu_Y = \int_Y \big[V, \mathcal{P}_H W\big]\,\mathrm{d}\mu_Y \quad \text{for all } W \in L^2(Y;\mathbb{R}^d).$$

We make the standard observation that $\mathcal{P}_H\mathcal{U}_H(V) = \mathcal{P}(H^{-1}\cdot(H\cdot V\circ F)) = \mathcal{P}(\mathcal{U}V) = V$.

Lemma 4.5.6. *The twisted transfer operator $\mathcal{P}_H\colon F_{\gamma,0}(Y;\mathbb{R}^d) \to F_{\gamma,0}(Y;\mathbb{R}^d)$ has no eigenvalue at* 1*.*

Proof. Suppose $\mathcal{P}_H V = V$ for some $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$. Then

$$\begin{aligned}
&\int_Y |\mathcal{U}_H V - V|^2\,\mathrm{d}\mu_Y = \int_Y \big[\mathcal{U}_H V - V, \mathcal{U}_H V - V\big]\,\mathrm{d}\mu_Y\\
&= \int_Y \big[\mathcal{U}_H V, \mathcal{U}_H V\big]\,\mathrm{d}\mu_Y - \int_Y \big[\mathcal{U}_H V, V\big]\,\mathrm{d}\mu_Y - \int_Y \big[V, \mathcal{U}_H V\big]\,\mathrm{d}\mu_Y + \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y\\
&= \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y - \int_Y \big[V, \mathcal{P}_H V\big]\,\mathrm{d}\mu_Y - \int_Y \big[\mathcal{P}_H V, V\big]\,\mathrm{d}\mu_Y + \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y\\
&= \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y - \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y - \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y + \int_Y \big[V, V\big]\,\mathrm{d}\mu_Y\\
&= 0.
\end{aligned} \tag{4.5.7}$$

It follows that $\mathcal{U}_H V = V$. Define $\Psi\colon Y\times G \to \mathbb{R}^d$ by $\Psi = g\cdot V$. Then

$$\Psi\circ F_H = gH\cdot V\circ F = g\cdot(H\cdot V\circ F) = g\cdot\mathcal{U}_H V = g\cdot V = \Psi.$$

By ergodicity of m_Y, it follows that Ψ is constant m_Y–almost surely. Since $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$, we have $\int_{Y\times G}\Psi\,\mathrm{d}m_Y = 0$ so that $\Psi = 0$ m_Y–almost surely. Therefore $V = 0$ μ_Y–almost surely. ∎

4.6 Construction of the primary decomposition

Let $\phi = g \cdot v$ where $v \in C^\eta(X;\mathbb{R}^d)$, and suppose that $\int_{X\times G} \phi \, \mathrm{d}m = 0$. Define the lifted observable $\hat{\phi} = \phi \circ \pi_H \colon \Delta \times G \to \mathbb{R}^d$, where π_H is the semi-conjugacy as in Section 4.2. Note that for $(y, g, \ell) \in \Delta \times G$, we have

$$\hat{\phi}(y, g, \ell) = \phi\big(T_h^\ell(y, g)\big) = \phi\big(T^\ell y, gh_\ell(y)\big) = gh_\ell(y) \cdot v(T^\ell y) = g \cdot \big(h_\ell(y) \cdot v(T^\ell y)\big). \tag{4.6.1}$$

In this section we construct our primary martingale-coboundary decomposition for $\hat{\phi}$. We begin by deriving such a decomposition for V, where V is as in (4.1.1), as well as giving information about the regularity of the components.

Proposition 4.6.1. *Suppose $\tau \in L^p(Y)$ with $p > 1$. There exist $J, M \colon Y \to \mathbb{R}^d$ with*

$$V = M + \mathcal{U}_H J - J \quad \text{and } M \in \ker \mathcal{P}_H. \tag{4.6.2}$$

Moreover, $J \in F_\gamma(Y;\mathbb{R}^d)$ with $\|J\|_\gamma \le C\|v\|_\eta \|h\|_\eta$ and $M \in L^p(Y)$ with $|M|_p \le C\|v\|_\eta\|h\|_\eta$.

Proof. In view of Remark 4.5.5, we consider $\mathcal{P}_H \colon F_{\gamma,0}(Y;\mathbb{R}^d) \to F_{\gamma,0}(Y;\mathbb{R}^d)$. Let $F_{\gamma,0}(Y;\mathbb{R}^d) = E_0 \oplus E_1$ be the spectral decomposition from Corollary 4.5.2. By Corollary 4.5.2 (ii), Lemma 4.5.4, and Lemma 4.5.6, the spectrum of $\mathcal{P}_H|_{E_1}$ consists of finitely many eigenvalues $e^{i\omega_k}$ for $1 \le k \le j$, where each $\omega_k \in (0, 2\pi)$. Now, recall from Proposition 4.4.3 that $\mathcal{P}_H V \in F_\gamma(Y;\mathbb{R}^d)$. Moreover, as in (4.5.6), we have

$$\begin{aligned}
\int_{Y\times G} g \cdot \mathcal{P}_H V \, \mathrm{d}m_Y &= \int_{Y\times G} g \cdot V \, \mathrm{d}m_Y = \int_{Y\times G} \sum_{\ell=0}^{\tau(y)-1} gh_\ell(y) \cdot v(T^\ell y) \, \mathrm{d}m_Y \\
&= \int_{Y\times G} \sum_{\ell=0}^{\tau(y)-1} \phi(y, g, \ell) \, \mathrm{d}m_Y = |\tau|_1 \int_{\Delta\times G} \hat{\phi} \, \mathrm{d}m_\Delta = 0,
\end{aligned}$$

so that $\mathcal{P}_H V \in F_{\gamma,0}(Y;\mathbb{R}^d)$. Therefore, we have the decomposition

$$\mathcal{P}_H V = W_0 + \sum_{k=1}^{j} W_k,$$

where $W_0 \in E_0$ and $\mathcal{P}_H W_k = e^{i\omega_k} W_k$ for $1 \le k \le j$. Moreover, we have from Proposition 4.4.3 that

$$\|W_0\|_\gamma = \big\|\pi(\mathcal{P}_H V)\big\|_\gamma \le \|\pi\| \|\mathcal{P}_H V\|_\gamma \ll \|v\|_\eta \|h\|_\eta,$$

where π is the spectral projection defined in Corollary 4.5.2. Similarly, $\|W_k\|_\gamma \ll \|v\|_\eta \|h\|_\eta$ for $1 \le k \le j$. By Lemma 4.5.3, we have $\|\mathcal{P}_H^n W_0\|_\gamma \ll r^n \|W_0\|_\gamma$. Define $J_0 \colon Y \to \mathbb{R}^d$ by

$$J_0 = \sum_{n=0}^{\infty} \mathcal{P}_H^n W_0.$$

Note that

$$\|J_0\|_\gamma \le \sum_{n=0}^{\infty} \|\mathcal{P}_H^n W_0\|_\gamma \ll \sum_{n=0}^{\infty} r^n \|v\|_\eta \|h\|_\eta \ll \|v\|_\eta \|h\|_\eta, \tag{4.6.3}$$

so that J_0 is well-defined. Now define $J, M \colon Y \to \mathbb{R}^d$ by

$$J = J_0 + \sum_{k=1}^{j} \frac{e^{-i\omega_k} W_k}{e^{-i\omega_k} - 1}$$

and

$$M = V + J - \mathcal{U}_H J.$$

Noting that $\mathcal{P}_H(\mathcal{U}_H J) = J$, we have

$$\begin{aligned} \mathcal{P}_H M &= \mathcal{P}_H V + \mathcal{P}_H J - J \\ &= \mathcal{P}_H V + \mathcal{P}_H J_0 - J_0 + \sum_{k=1}^{j} \left(\frac{W_k}{e^{-i\omega_k} - 1} - \frac{e^{-i\omega_k} W_k}{e^{-i\omega_k} - 1} \right) \\ &= \mathcal{P}_H V - W_0 - \sum_{k=1}^{j} W_k = 0, \end{aligned} \tag{4.6.4}$$

so that $M \in \ker \mathcal{P}_H$.

We now look at the regularity of J and M. First note that since $e^{i\omega_k} \ne 1$ for all $1 \le k \le j$, we have

$$\left\| \frac{e^{-i\omega_k} W_k}{e^{-i\omega_k} - 1} \right\|_\gamma = \left\| \frac{W_k}{1 - e^{i\omega_k}} \right\| \ll \|W_k\|_\gamma \ll \|v\|_\eta \|h\|_\eta.$$

Combining this with (4.6.3), it follows that $\|J\|_\gamma \ll \|v\|_\eta \|h\|_\eta$. For M, first note that $|\mathcal{U}_H J|_p \le |\mathcal{U}_H J|_\infty = |J|_\infty$, where the final equality follows from Proposition 2.5.3 (viii). Therefore

$$|M|_p \le |V|_p + |J|_p + |\mathcal{U}_H J|_p \le |\tau|_p |v|_\infty + 2|J|_\infty \ll \|v\|_\eta \|h\|_\eta, \tag{4.6.5}$$

as claimed. □

Define the induced observable $\Phi\colon Y \times G \to \mathbb{R}^d$ by

$$\Phi(y,g) = \sum_{k=0}^{\tau(y)-1} \hat{\phi}(y,g,k),$$

and note that

$$\Phi(y,g) = \sum_{k=0}^{\tau(y)-1} g h_k(y) \cdot v(T^k y) = g \cdot V(y).$$

Using (4.6.2), we get a similar decomposition for Φ on $Y \times G$. Define $\chi, \Psi\colon Y \times G \to \mathbb{R}^d$ by $\chi = g \cdot J$ and $\Psi = g \cdot M$. Note that

$$\chi\big(F_H(y,g)\big) = \chi\big(Fy, gH(y)\big) = g \cdot \big(H(y) \cdot J(Fy)\big) = g \cdot \mathcal{U}_H J(y),$$

so that

$$\Phi = g \cdot V = g \cdot M + g \cdot \mathcal{U}_H J - g \cdot J = \Psi + \chi \circ F_H - \chi. \tag{4.6.6}$$

We lift this to the tower extension $\Delta \times G$ by following the approach in [60]. Define $\hat{\chi}, \hat{\psi}\colon \Delta \times G \to \mathbb{R}^d$ by

$$\hat{\chi}(y,g,\ell) = \chi(y,g) + \sum_{k=0}^{\ell-1} \hat{\phi}(y,g,k) \tag{4.6.7}$$

and

$$\hat{\psi}(y,g,\ell) = \begin{cases} 0 & \text{if } \ell \le \tau(y) - 2, \\ \Psi(y,g) & \text{if } \ell = \tau(y) - 1. \end{cases} \tag{4.6.8}$$

The next proposition shows explicitly how the regularity of $\hat{\chi}$ and $\hat{\psi}$ is dictated by the integrability of the return time τ.

Proposition 4.6.2. *Suppose $\tau \in L^p(Y)$.*

(i) *If $p > 1$, then $\hat{\psi} \in L^p(\Delta \times G; \mathbb{R}^d)$ with $|\hat{\psi}|_p \le C\|v\|_\eta \|h\|_\eta$.*

(ii) *If $1 < p < 2$ and $V \in L^2(Y; \mathbb{R}^d)$, then $\hat{\psi} \in L^2(\Delta \times G; \mathbb{R}^d)$ with $|\hat{\psi}|_2 \le C(|V|_2 + \|v\|_\eta \|h\|_\eta)$.*

(iii) *If $p \ge 2$, then $\hat{\chi} \in L^{p-1}(\Delta \times G; \mathbb{R}^d)$ with $|\hat{\chi}|_{p-1} \le C\|v\|_\eta \|h\|_\eta$.*

Proof. For (i), we have

$$\begin{aligned}\int_{\Delta\times G} |\hat{\psi}|^p \, \mathrm{d}m_\Delta &= |\tau|_1^{-1} \int_{Y\times G} \sum_{\ell=0}^{\tau(y)-1} \big|\hat{\psi}(y,g,\ell)\big|^p \, \mathrm{d}m_Y \\ &= |\tau|_1^{-1} \int_{Y\times G} \big|\Psi(y,g)\big|^p \, \mathrm{d}m_Y = |\tau|_1^{-1} \int_Y \big|M(y)\big|^p \, \mathrm{d}\mu_Y \le |M|_p^p \\ &\ll \|v\|_\eta^p \|h\|_\eta^p, \end{aligned} \tag{4.6.9}$$

where the final estimate uses Proposition 4.6.1. Therefore $|\hat{\psi}|_p \ll \|v\|_\eta \|h\|_\eta$ as claimed.

For (ii), note that since $V \in L^2(Y; \mathbb{R}^d)$, we have from Proposition 4.6.1 that $|M|_2 \le |V|_2 + |\mathcal{U}_H J|_2 + |J|_2 \le |V|_2 + 2|J|_\infty \ll |V|_2 + \|v\|_\eta \|h\|_\eta$. The result now follows from (4.6.9) with $p = 2$.

We now look at (iii). Let $(y,g,\ell) \in \Delta \times G$. We have using Proposition 4.6.1 that

$$\begin{aligned}\big|\hat{\chi}(y,g,\ell)\big| &\le \big|\chi(y,g)\big| + \sum_{k=0}^{\ell-1} \big|\hat{\phi}(y,g,k)\big| \le |J|_\infty + \ell |v|_\infty \ll (1+\ell)\|v\|_\eta \|h\|_\eta \\ &\le \tau(y) \|v\|_\eta \|h\|_\eta. \end{aligned} \tag{4.6.10}$$

Therefore

$$\begin{aligned}\int_{\Delta\times G} \big|\hat{\chi}(y,g,\ell)\big|^{p-1} \, \mathrm{d}m_\Delta &\ll \int_{\Delta\times G} \tau(y)^{p-1} \|v\|_\eta^{p-1} \|h\|_\eta^{p-1} \, \mathrm{d}m_\Delta \\ &= |\tau|_1^{-1} \int_{Y\times G} \sum_{\ell=0}^{\tau(y)-1} \tau(y)^{p-1} \|v\|_\eta^{p-1} \|h\|_\eta^{p-1} \, \mathrm{d}m_Y \\ &= |\tau|_1^{-1} \int_Y \tau(y)^p \|v\|_\eta^{p-1} \|h\|_\eta^{p-1} \, \mathrm{d}\mu_Y \ll \|v\|_\eta^{p-1} \|h\|_\eta^{p-1}, \end{aligned} \tag{4.6.11}$$

from which the result follows. ∎

The next proposition shows that $\hat{\phi}$ admits a martingale-coboundary decomposition. Recall that $\hat{\mathcal{L}}\colon L^1(\Delta \times G;\mathbb{R}^d) \to L^1(\Delta \times G;\mathbb{R}^d)$ denotes the transfer operator for f_H.

Proposition 4.6.3 (Primary martingale-coboundary decomposition). *Let $\hat{\phi}, \hat{\chi}$, and $\hat{\psi}$ be as in* (4.6.1),(4.6.7), *and* (4.6.8) *respectively. Then*

$$\hat{\phi} = \hat{\psi} + \hat{\chi} \circ f_H - \hat{\chi} \qquad \textit{and } \hat{\psi} \in \ker \hat{\mathcal{L}}.$$

Proof. Fix $(y, g, \ell) \in \Delta \times G$. We first suppose that $\ell \le \tau(y) - 2$. Then $\hat{\psi}(y, g, \ell) = 0$. In particular, we have

$$\begin{aligned}\hat{\chi} \circ f_H(y, g, \ell) - \hat{\chi}(y, g, \ell) &= \hat{\chi}(y, g, \ell + 1) - \hat{\chi}(y, g, \ell) \\ &= \hat{\phi}(y, g, \ell) = \hat{\phi}(y, g, \ell) - \hat{\psi}(y, g, \ell).\end{aligned}$$

Now suppose that $\ell = \tau(y) - 1$. Then

$$\begin{aligned}\hat{\chi} \circ f_H(y, g, \ell) - \hat{\chi}(y, g, \ell) &= \hat{\chi}\big(F_H(y, g), 0\big) - \hat{\chi}(y, g, \ell) \\ &= \chi\big(F_H(y, g)\big) - \chi(y, g) - \sum_{k=0}^{\ell-1} \hat{\phi}(y, g, k) \\ &= \Phi(y, g) - \Psi(y, g) - \sum_{k=0}^{\ell-1} \hat{\phi}(y, g, k) \\ &= \hat{\phi}(y, g, \ell) - \hat{\psi}(y, g, \ell),\end{aligned}$$

where the third equality uses (4.6.6). This proves the first statement.

We next show that $\hat{\psi} \in \ker \hat{\mathcal{L}}$. Write $\hat{\psi} = g \cdot \hat{m}$, where

$$\hat{m}(y, \ell) = \begin{cases} 0 & \text{if } \ell \le \tau(y) - 2, \\ M(y) & \text{if } \ell = \tau(y) - 1. \end{cases} \tag{4.6.12}$$

For $\ell \ge 1$, we have from Proposition 4.3.7 and Remark 4.3.6 that

$$\hat{\mathcal{L}}\hat{\psi}(y, g, \ell) = g \cdot \mathcal{L}_H \hat{m}(y, \ell) = g \cdot \hat{m}(y, \ell - 1) = 0.$$

Similarly, for $\ell = 0$, we have

$$\hat{\mathcal{L}}\hat{\psi}(y,g,0) = g \cdot \mathcal{L}_H \hat{m}(y,0) = g \cdot \sum_{a\in\alpha} \zeta(y_a) H(y_a)^{-1} \cdot M(y_a) = g \cdot \mathcal{P}_H M(y) = 0,$$

where the final equality follows from (4.6.2). $\square$

To conclude this section, we show that $\max_{0\le k\le n} |\hat{\chi} \circ f_H^k|$ converges to 0 almost surely when suitably normalised, with exponent depending on the integrability of τ. Recall $V^*\colon Y \to \mathbb{R}$ defined as in (4.1.2) by

$$V^*(y) = \max_{0\le k\le \tau(y)-1} \left| \sum_{\ell=0}^{k} h_\ell(y) \cdot v(T^\ell y) \right|.$$

Proposition 4.6.4. *Suppose $\tau \in L^p(Y)$ with $p > 1$ and let $\hat{\chi}$ be as in (4.6.7). Then*

$$\max_{0\le k\le n} |\hat{\chi} \circ f_H^k| = o(n^{1/p}) \quad a.s.$$

Moreover, if $1 < p < 2$ and $V^* \in L^2(Y)$, then

$$\max_{0\le k\le n} |\hat{\chi} \circ f_H^k| = o(n^{1/2}) \quad a.s.$$

Proof. Fix $(y,g,\ell) \in \Delta \times G$ and suppose first that $p > 1$. Recall that we define $H_0 = I_d$ and $H_j = H\, H \circ F \cdots H \circ F^{j-1}$. For any $n \ge 0$ and $0 \le k' \le n$ there exist $j \in \{0,\dots,k'\}$ and $\ell' \in \{0,\dots,\tau(F^j y) - 1\}$ such that $f_H^{k'}(y,g,\ell) = (F^j y, gH_j(y), \ell')$. Therefore, we have from (4.6.10) that

$$\begin{aligned} \big|\hat{\chi}\big(f_H^{k'}(y,g,\ell)\big)\big| &= \big|\hat{\chi}\big(F^j y, gH_j(y), \ell'\big)\big| \ll \|v\|_\eta \|h\|_\eta \tau(F^j y) \\ &\le \|v\|_\eta \|h\|_\eta \max_{0\le k\le n} \tau(F^k y). \end{aligned}$$

Since $\tau \in L^p(Y)$, it follows from Corollary 2.2.3 that

$$\max_{0\le k'\le n} \big|\hat{\chi} \circ f_H^{k'}(y,g,\ell)\big| \ll \|v\|_\eta \|h\|_\eta \max_{0\le k\le n} \tau(F^k y) = o(n^{1/p}) \quad \text{a.s.} \tag{4.6.13}$$

Suppose now that $1 < p < 2$ and $V^* \in L^2(Y)$. Note that

$$\big|\hat{\chi}(y,g,\ell)\big| \le \big|\chi(y,g)\big| + \left| \sum_{k=0}^{\ell-1} \hat{\phi}(y,g,k) \right|. \tag{4.6.14}$$

Now,

$$\Big|\sum_{k=0}^{\ell-1}\hat{\phi}(y,g,k)\Big| = \Big|\sum_{k=0}^{\ell-1}\phi\big(T^k y, gh_k(y)\big)\Big| = \Big|\sum_{k=0}^{\ell-1} gh_k(y)\cdot v(T^k y)\Big|$$
$$= \Big|\sum_{k=0}^{\ell-1} h_k(y)\cdot v(T^k y)\Big| \le V^*(y).$$

Therefore, continuing (4.6.14), we have

$$\big|\hat{\chi}(y,g,\ell)\big| \le |J|_\infty + V^*(y) \ll \|v\|_\eta \|h\|_\eta + V^*(y).$$

For any $n \ge 0$ and $0 \le k' \le n$, there exist $j \in \{0,\dots,k'\}$ and $\ell' \in \{0,\dots,\tau(F^j y)-1\}$ such that $f_H^{k'}(y,g,\ell) = (F^j y, gH_j(y), \ell')$. Therefore

$$\big|\hat{\chi}\big(f_H^{k'}(y,g,\ell)\big)\big| \ll \|v\|_\eta\|h\|_\eta + V^*(F^j y) \le \|v\|_\eta\|h\|_\eta + \max_{0\le k\le n} V^*(F^k y).$$

Since $V^* \in L^2(Y)$, we have from Corollary 2.2.3 that

$$\max_{0\le k'\le n} |\hat{\chi}\circ f_H^{k'}| \ll \|v\|_\eta\|h\|_\eta + \max_{0\le k\le n} V^*\circ F^k = o(n^{1/2}) \quad \text{a.s.} \tag{4.6.15}$$

as claimed. □

4.7 Proofs of Theorem 4.1.1 and Theorem 4.1.2

We now proceed with the proofs of Theorem 4.1.2 and Theorem 4.1.3. From Proposition 4.2.3, it suffices to prove the results for the lifted observable $\hat{\phi}$. For the remainder of the chapter, we let $\hat{\chi}$ and $\hat{\psi}$ be as in (4.6.7) and (4.6.8) respectively.

Proof of Theorem 4.1.2. Note that since $p > 1$ and $V \in L^2(Y;\mathbb{R}^d)$, we have $\hat{\psi} \in L^2(\Delta\times G;\mathbb{R}^d)$ by Proposition 4.6.2. Moreover, $\hat{\psi} \in \ker\hat{\mathcal{L}}$ by Proposition 4.6.3. Defining

$$\Sigma = \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta \in \mathbb{R}^{d,d}, \tag{4.7.1}$$

it follows from Theorem 2.8.6 that $n^{-1/2}\sum_{k=0}^{n-1}\hat{\psi}\circ f_H^k \to_w \mathcal{N}(0,\Sigma)$.

Write $S_n\hat{\phi} = \sum_{k=0}^{n-1} \hat{\phi} \circ f_H^k$ and $S_n\hat{\psi} = \sum_{k=0}^{n-1} \hat{\psi} \circ f_H^k$. We show that $|n^{-1/2}(S_n\hat{\phi} - S_n\hat{\psi})| \to 0$ in probability. The result then follows from Theorem 2.3.9. Note that for any $p > 1$, we have by (4.6.11) that $\mathbb{E}[|\hat{\chi}|^{p-1}] < \infty$. Now, for any $\epsilon > 0$, we have by Markov's inequality that

$$\begin{aligned} m_\Delta\big(\big|n^{-1/2}(S_n\hat{\phi} - S_n\hat{\psi})\big| > \epsilon\big) &\leq \epsilon^{-(p-1)} n^{-(p-1)/2}\mathbb{E}\big[|\hat{\chi} \circ f_H^n - \chi|^{p-1}\big] \\ &\leq 2\epsilon^{-(p-1)} n^{-(p-1)/2}\mathbb{E}\big[|\hat{\chi}|^{p-1}\big] \to 0, \end{aligned}$$

where we use f_H-invariance of m_Δ in the second inequality. Therefore, it holds true that $n^{-1/2}(S_n\hat{\phi} - S_n\hat{\psi}) \to 0$ in probability, and the result follows.

It remains to verify that Σ commutes with the action of G on $\mathbb{R}^d$. Since $\hat{\psi} = g\cdot\hat{m}$, where $\hat{m}$ is defined as in (4.6.12), we have for all $a \in G$ that $\hat{\psi}(y, ag, \ell) = ag \cdot \hat{m}(y, \ell) = a \cdot (g \cdot \hat{m}(y, \ell)) = a \cdot \hat{\psi}(y, g, \ell)$. It follows from invariance of the Haar measure that

$$\int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta = a\bigg(\int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\bigg)a^T = a\bigg(\int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\bigg)a^{-1}, \quad (4.7.2)$$

so that $\Sigma = a\Sigma a^{-1}$ for all $a \in G$. This completes the proof. □

Proof of Theorem 4.1.3. For $n \geq 1$, define the random elements $\hat{M}_n, \hat{W}_n \colon \Delta\times G \to D([0,\infty); \mathbb{R}^d)$ by $\hat{M}_n(t) = n^{-1/2} \sum_{k=0}^{[nt]-1} \hat{\phi} \circ f_H^k$ and $\hat{W}_n(t) = n^{-1/2} \sum_{k=0}^{[nt]-1} \hat{\psi} \circ f_H^k$ for $t \geq 0$. Let Σ be as in as in (4.7.1) and let W be the Brownian motion with mean 0 and covariance matrix Σ.

We first show that $\hat{M}_n \to_w W$ in $D([0,\infty); \mathbb{R}^d)$. Since $V^* \in L^2(Y)$, we have $V \in L^2(Y; \mathbb{R}^d)$. Therefore $\hat{\psi} \in L^2(\Delta \times G; \mathbb{R}^d)$ with $\hat{\mathcal{L}}\hat{\psi} = 0$, as argued in the proof of Theorem 4.1.2. Let $\hat{\mathcal{U}} \colon L^1(\Delta \times G; \mathbb{R}^d) \to L^1(\Delta \times G; \mathbb{R}^d)$ denote the Koopman operator for f_H. Since m_Δ is f_H-invariant, we have from Proposition 2.5.3 (vii) that

$$\Sigma = \int_{\Delta\times G} \hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T) dm_\Delta.$$

Therefore, for $t \geq 0$ we have by Birkhoff's ergodic theorem that

$$\frac{1}{n}\sum_{k=0}^{[nt]-1} \hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T) \circ f_H^k = t \cdot \frac{[nt]}{nt} \cdot \frac{1}{[nt]} \sum_{k=0}^{[nt]-1} \hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T) \circ f_H^k \to t\Sigma \quad \text{a.s.}$$

It follows from Theorem 2.8.7 that $\hat{M}_n \to_w W$ in $D([0,\infty);\mathbb{R}^d)$.

We next show that this implies convergence of $\hat{W}_n$. Observe that for any $T>0$, we have from [92] that $\hat{M}_n \to_w W$ in $D([0,T];\mathbb{R}^d)$. Moreover, note that

$$\begin{aligned}
&\sup_{t\in[0,T]} \big|\hat{W}_n(t) - \hat{M}_n(t)\big| = \frac{1}{\sqrt{n}} \sup_{t\in[0,T]} \Big| \sum_{k=0}^{[nt]-1} \hat{\chi}\circ f_H^{k+1} - \hat{\chi}\circ f_H^k \Big| \\
&= \frac{1}{\sqrt{n}} \sup_{t\in[0,T]} |\hat{\chi}\circ f_H^{[nt]} - \hat{\chi}| \le \frac{2}{\sqrt{n}} \max_{0\le k\le [nT]} |\hat{\chi}\circ f_H^k| \\
&= \frac{2\sqrt{[nT]}}{\sqrt{n}} \frac{1}{\sqrt{[nT]}} \max_{0\le k\le [nT]} |\hat{\chi}\circ f_H^k| \to 0 \quad \text{a.s.}
\end{aligned}$$

by Proposition 4.6.4. Therefore, setting s to be the Skorokhod metric defined on $D([0,T];\mathbb{R}^d)$, we have $s(\hat{W}_n, \hat{M}_n) \to 0$ almost surely. It follows from Theorem 2.3.9 that $\hat{W}_n \to_w W$ in $D([0,T];\mathbb{R}^d)$. Since $T>0$ is arbitrary, we have from [92] that $\hat{W}_n \to_w W$ in $D([0,\infty);\mathbb{R}^d)$, as required. □

4.8 Moment estimates and covariance matrix

In this section we obtain a uniform estimate on $|\max_{1\le k\le n}|\hat{\chi}\circ f_H^k - \hat{\chi}||_p$. As an application, we prove Theorem 4.1.5 and give a characterisation of the covariance matrix $\Sigma = \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \, dm_\Delta$ in terms of the underlying observable ϕ. We follow [60, Section 2]. We begin with a lemma.

Lemma 4.8.1. *Suppose $\tau \in L^p(Y)$ where $p>1$. Then*

$$\big|\max_{1\le k\le n} |\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p \le C\|v\|_\eta \|h\|_\eta \big(n^{1/p} |1_{\{\tau\ge n\}}\tau|_p + a + n^{1/p}|1_{\{\tau\ge a\}}\tau|_p\big)$$

for all $a\ge 0$ and $n\ge 1$.

Proof. The proof follows [60, Proposition 2.7]. For $n\ge 0$, let $A_n = \{(y,g,\ell)\in \Delta\times G \mid 0\le \ell < \tau(y)-n\}$. We have

$$\big|\max_{1\le k\le n} |\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p \le \big|1_{A_n} \max_{1\le k\le n} |\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p + \big|1_{\Delta\times G\setminus A_n} \max_{1\le k\le n} |\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p. \tag{4.8.1}$$

We look at the terms on the right hand side in turn. For $c \geq 0$, define $t_c = |\mathbb{1}_{\{\tau \geq c\}}\tau|_p$. Then

$$\begin{aligned}\sum_{k\geq n} k^{p-1}\mu_Y(\tau > k) &\leq \sum_{k\geq n}\sum_{j\geq k} k^{p-1}\mu_Y(\tau = j) = \sum_{j\geq n}\mu_Y(\tau = j)\sum_{k=n}^{j} k^{p-1} \\ &\leq \sum_{j\geq n} j^p \mu_Y(\tau = j) = t_n^p. \end{aligned} \tag{4.8.2}$$

For $n \geq 0$, we denote the nth level of the tower by $\Delta_n = \{(y, g, \ell) \in \Delta\times G \mid \ell = n\}$. Then

$$\begin{aligned} m_\Delta(\Delta_n) &= m_\Delta\big(\big\{(y, g, \ell) \in Y \times G \times \mathbb{Z} \mid 0 \leq \ell \leq \tau(y) - 1, \ell = n\big\}\big) \\ &= |\tau|_1^{-1}\mu_Y(\tau > n). \end{aligned}$$

For $k \geq 0$, let

$$A_n^k = \big\{(y, g, \ell) \in A_n \mid \ell = k\big\} = \big\{(y, g, k) \in Y \times G \times \mathbb{Z} \mid 0 \leq k < \tau(y) - n\big\}.$$

Note that $f_H^n(y, g, \ell) \in \Delta_{n+k}$ if and only if $\ell = k$ and $k < \tau(y) - n$, so that $A_n^k = f_H^{-n}(\Delta_{n+k})$. Therefore, we have

$$\begin{aligned} m_\Delta(A_n) &= m_\Delta\bigg(\bigcup_{k\geq 0} A_n^k\bigg) = m_\Delta\bigg(\bigcup_{k\geq 0} f_H^{-n}(\Delta_{n+k})\bigg) = m_\Delta\bigg(\bigcup_{k\geq n} f_H^{-n}(\Delta_k)\bigg) \\ &= m_\Delta\bigg(f_H^{-n}\bigg(\bigcup_{k\geq n}\Delta_k\bigg)\bigg) = m_\Delta\bigg(\bigcup_{k\geq n}\Delta_k\bigg) = |\tau|_1^{-1}\sum_{k\geq n}\mu_Y(\tau > k). \end{aligned}$$

It follows from (4.8.2) that

$$n^{p-1}m_\Delta(A_n) = n^{p-1}|\tau|_1^{-1}\sum_{k\geq n}\mu_Y(\tau > k) \leq \sum_{k\geq n} k^{p-1}\mu_Y(\tau > k) \leq t_n^p.$$

Now, if $(y, g, \ell) \in A_n$, then for $1 \leq k \leq n$ we have

$$\begin{aligned} \big|(\hat{\chi}\circ f_H^k - \hat{\chi})(y, g, \ell)\big| &= \big|\hat{\chi}(y, g, \ell + k) - \hat{\chi}(y, g, \ell)\big| \leq \sum_{j=\ell}^{\ell+k-1}\big|\hat{\phi}(y, g, j)\big| \\ &\leq k|v|_\infty \leq n|v|_\infty. \end{aligned}$$

Therefore

$$\Big|1_{A_n} \max_{1\le k\le n} |\hat\chi\circ f_H^k - \hat\chi|\Big|_p \le n|v|_\infty m_\Delta(A_n)^{1/p} = n^{1/p}|v|_\infty\big[n^{p-1}m_\Delta(A_n)\big]^{1/p}$$
$$\le n^{1/p}|v|_\infty t_n \ll n^{1/p}\|v\|_\eta\|h\|_\eta t_n.$$

We now look at the second term on the right hand side in (4.8.1). Recall from (4.6.13) that

$$\max_{0\le k\le n}\big|\hat\chi\circ f_H^k(y,g,\ell)\big| \ll \|v\|_\eta\|h\|_\eta \max_{0\le k\le n}\tau(F^k y).$$

Let $\tau_a = 1_{\{\tau\ge a\}}\tau$. Then $\tau^p \le a^p + \tau_a^p$. It follows that

$$\begin{aligned}&\|v\|_\eta^{-p}\|h\|_\eta^{-p}\max_{1\le k\le n}\big|\hat\chi\big(f_H^k(y,g,\ell)\big)-\hat\chi(y,g,\ell)\big|^p\\ &\le 2^p\|v\|_\eta^{-p}\|h\|_\eta^{-p}\max_{0\le k\le n}\big|\hat\chi\big(f_H^k(y,g,\ell)\big)\big|^p\\ &\ll \max_{0\le k\le n}\tau^p(F^k y)\le a^p+\sum_{k=0}^n\tau_a^p(F^k y).\end{aligned}$$

Suppose that $\hat\gamma\colon \Delta\times G\to\mathbb{R}$ has the form $\hat\gamma(y,g,\ell)=U(y)$ where $U\colon Y\to\mathbb{R}$. Then since $\Delta\times G\setminus A_n=\{(y,g,\ell)\in Y\times G\times\mathbb{Z}\mid \max\{0,\tau(y)-n\}\le\ell\le\tau(y)-1\}$, we have

$$\begin{aligned}&\int_{\Delta\times G\setminus A_n}|\hat\gamma|dm_\Delta\\ &=|\tau|_1^{-1}\int_Y\Bigg(\sum_{\ell=\tau(y)-n}^{\tau(y)-1}1_{\{\tau>n\}}(y)\big|U(y)\big|+\sum_{\ell=0}^{\tau(y)-1}1_{\{\tau\le n\}}(y)\big|U(y)\big|\Bigg)\,d\mu_Y\\ &=|\tau|_1^{-1}\int_Y\min\{\tau,n\}|U|\,d\mu_Y\le\int_Y\min\{\tau,n\}|U|\,d\mu_Y.\end{aligned}$$

Now, recall that $\mathcal{P}$ is the transfer operator for F. We have using Proposition 3.3.2 and Proposition 3.3.9 that

$$\big|\mathcal{P}^k\tau(y)\big|\le\sum_{a\in\alpha_k}\big|\zeta_k(y_a)\tau(y_a)\big|\le D\sum_{a\in\alpha_k}\mu_Y(a)\tau(y_a)=D|\tau|_1<\infty$$

for all $k \geq 1$, since τ is integrable. Therefore $|\mathcal{P}^k\tau|_\infty < \infty$. Combining the above, we have

$$
\begin{aligned}
&\|v\|_\eta^{-p}\|h\|_\eta^{-p}\int_{\Delta\times G\setminus A_n} \max_{1\leq k\leq n} |\hat{\chi}\circ f_H^k - \hat{\chi}|^p dm_\Delta \\
&\ll a^p + \sum_{k=0}^{n}\int_{\Delta\times G\setminus A_n} \tau_a^p(F^k y) dm_\Delta \\
&\leq a^p + \sum_{k=0}^{n}\int_Y \min\{\tau, n\}\tau_a^p\circ F^k \,\mathrm{d}\mu_Y \leq a^p + \int_Y n\tau_a^p \,\mathrm{d}\mu_Y + \sum_{k=1}^{n}\int_Y \tau\tau_a^p\circ F^k di\mu_Y \\
&= a^p + n|\tau_a^p|_1 + \sum_{k=1}^{n}|\mathcal{P}^k\tau\cdot\tau_a^p|_1 \ll a^p + n|\tau_a^p|_1 = a^p + nt_a^p.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\big|\mathbb{1}_{\Delta\times G\setminus A_n} \max_{1\leq k\leq n}|\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p &\ll \|v\|_\eta\|h\|_\eta(a^p + nt_a^p)^{1/p} \\
&\leq \|v\|_\eta\|h\|_\eta(a + n^{1/p}t_a),
\end{aligned}
$$

where the final inequality holds since $a, t_a \geq 0$ and $p > 1$. □

Proposition 4.8.2. *Suppose $\tau \in L^p(Y)$ where $p > 1$. Then*

$$\big|\max_{1\leq k\leq n}|\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p \leq Cn^{1/p}\|v\|_\eta\|h\|_\eta \quad \text{for all } n \geq 1. \tag{4.8.3}$$

Moreover,

$$\big|\max_{1\leq k\leq n}|\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p = o(n^{1/p}). \tag{4.8.4}$$

Proof. For the first statement, taking $a = 0$ in Lemma 4.8.1 gives us

$$
\begin{aligned}
\big|\max_{1\leq k\leq n}|\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p &\ll \|v\|_\eta\|h\|_\eta\big(n^{1/p}|\mathbb{1}_{\{\tau\geq n\}}\tau|_p + n^{1/p}|\tau|_p\big) \\
&\ll \|v\|_\eta\|h\|_\eta n^{1/p}.
\end{aligned}
$$

We now prove the second statement. For $c \geq 0$, write $t_c = |\mathbb{1}_{\{\tau\geq c\}}\tau|_p$. Let $q > p$ and note that $t_n \leq t_{n^{1/q}}$. Taking $a = n^{1/q}$ in Lemma 4.8.1, we get

$$
\begin{aligned}
\big|\max_{1\leq k\leq n}|\hat{\chi}\circ f_H^k - \hat{\chi}|\big|_p &\ll \|v\|_\eta\|h\|_\eta(n^{1/p}t_n + n^{1/q} + n^{1/p}t_{n^{1/q}}) \\
&\ll \|v\|_\eta\|h\|_\eta(n^{1/q} + n^{1/p}t_{n^{1/q}}) \\
&= \|v\|_\eta\|h\|_\eta\big(n^{1/q} + n^{1/p}|\mathbb{1}_{\{\tau\geq n^{1/q}\}}\tau|_p\big).
\end{aligned}
$$

It follows that

$$\frac{1}{n^{1/p}}\Big|\max_{1\le k\le n}|\hat{\chi}\circ f_H^k-\hat{\chi}|\Big|_p \ll \|v\|_\eta\|h\|_\eta\left(\frac{n^{1/q}}{n^{1/p}}+|1_{\{\tau\ge n^{1/q}\}}\tau|_p\right)\to 0,$$

where the convergence of the first term on the right hand side follows since $q>p$, and the convergence of the second term on the right hand side follows from the dominated convergence theorem. □

We next give moment estimates for $\sum_{k=0}^{n-1}\hat{\psi}\circ f_H^k$ and prove Theorem 4.1.5.

Corollary 4.8.3. *Suppose $\tau\in L^p(Y)$.*

(i) *If $1<p<2$, then*

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1}\hat{\psi}\circ f_H^k\right|\right|_p\le Cn^{1/p}\|v\|_\eta\|h\|_\eta\quad \textit{for all } n\ge 1.$$

(ii) *If $p\ge 2$, then*

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1}\hat{\psi}\circ f_H^k\right|\right|_p\le Cn^{1/2}\|v\|_\eta\|h\|_\eta\quad \textit{for all } n\ge 1.$$

Proof. First note that from Proposition 4.6.3, Proposition 2.8.3, and Proposition 2.4.5 (ii), we have that $(\sum_{k=1}^{j}\hat{\psi}\circ f_H^{n-k})_{j=1}^n$ is a martingale. Therefore, for $p>1$ we have

$$\left|\max_{1\le j\le n}\left|\sum_{k=1}^{j}\hat{\psi}\circ f_H^{n-k}\right|\right|_p\ll\left|\sum_{k=1}^{n}\hat{\psi}\circ f_H^{n-k}\right|_p\ll\left|\left(\sum_{k=1}^{n}|\hat{\psi}\circ f_H^{n-k}|^2\right)^{1/2}\right|_p,\qquad(4.8.5)$$

where we use Theorem 2.4.7 in the first inequality and Theorem 2.4.8 in the second inequality. We analyse (4.8.5) in the cases $1<p<2$ and $p\ge 2$ separately.

Suppose first that $1<p<2$. Note that for $a_1,\dots a_n\ge 0$, we have that $(\sum_{k=1}^n a_k^2)^{p/2}\le\sum_{k=1}^n a_k^p$. Therefore

$$\begin{aligned}\left|\left(\sum_{k=1}^{n}|\hat{\psi}\circ f_H^{n-k}|^2\right)^{1/2}\right|_p&=\left(\int_{\Delta\times G}\left(\sum_{k=1}^{n}|\hat{\psi}\circ f_H^{n-k}|^2\right)^{p/2}\mathrm{d}m_\Delta\right)^{1/p}\\&\le\left(\int_{\Delta\times G}\sum_{k=1}^{n}|\hat{\psi}\circ f_H^{n-k}|^p\,\mathrm{d}m_\Delta\right)^{1/p}=n^{1/p}|\hat{\psi}|_p\ll n^{1/p}\|v\|_\eta\|h\|_\eta,\end{aligned}\qquad(4.8.6)$$

where we use Proposition 4.6.2 (i) in the final estimate. Therefore, noting that

$$\sum_{k=0}^{j-1} \hat{\psi} \circ f_H^k = \sum_{k=1}^{n} \hat{\psi} \circ f_H^{n-k} - \sum_{k=1}^{n-j} \hat{\psi} \circ f_H^{n-k}, \tag{4.8.7}$$

we obtain

$$\begin{aligned}\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \hat{\psi} \circ f_H^k\right|\right|_p &\le \left|\max_{1\le j\le n}\left|\sum_{k=1}^{n} \hat{\psi} \circ f_H^{n-k}\right|\right|_p + \left|\max_{1\le j\le n}\left|\sum_{k=1}^{n-j} \hat{\psi} \circ f_H^{n-k}\right|\right|_p \\ &\ll n^{1/p}\|v\|_\eta \|h\|_\eta,\end{aligned}$$

proving (i).

Suppose now that $p \ge 2$. We have

$$\begin{aligned}&\left|\left(\sum_{k=1}^{n} |\hat{\psi} \circ f_H^{n-k}|^2\right)^{1/2}\right|_p = \left|\sum_{k=1}^{n} |\hat{\psi} \circ f_H^{n-k}|^2\right|_{p/2}^{1/2} \le \left(\sum_{k=1}^{n} \left||\hat{\psi} \circ f_H^{n-k}|^2\right|_{p/2}\right)^{1/2} \\ &= n^{1/2}\left||\hat{\psi}|^2\right|_{p/2}^{1/2} = n^{1/2}|\hat{\psi}|_p \ll n^{1/2}\|v\|_\eta \|h\|_\eta.\end{aligned}$$

Again it follows from (4.8.7) that

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \hat{\psi} \circ f_H^k\right|\right|_p \ll n^{1/2}\|v\|_\eta \|h\|_\eta,$$

proving (ii). □

Proof of Theorem 4.1.5. First note that from Proposition 4.2.3, we have for $q > 1$ that

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \phi \circ T_h^k\right|\right|_q = \left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \hat{\phi} \circ f_H^k\right|\right|_q .$$

This allows us to appeal to the martingale-coboundary decomposition in Proposition 4.6.3. Note that $\sum_{k=0}^{j-1} \hat{\phi} \circ f_H^k = (\sum_{k=0}^{j-1} \hat{\psi} \circ f_H^k) + \hat{\chi} \circ f_H^j - \hat{\chi}$. For $p > 1$, it follows from Corollary 4.8.3 and (4.8.3) that

$$\begin{aligned}\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \hat{\phi} \circ f_H^k\right|\right|_p &\le \left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \hat{\psi} \circ f_H^k\right|\right|_p + \left|\max_{1\le j\le n}|\hat{\chi} \circ f_H^j - \hat{\chi}|\right|_p \\ &\ll n^{1/p}\|v\|_\eta \|h\|_\eta,\end{aligned}$$

proving (i).

Suppose now that $p \geq 2$. For $0 \leq k \leq n$, let $\mathcal{B}_k = f_H^{-(n-k)}(\mathcal{B})$, where $\mathcal{B}$ is the underlying $\sigma-$algebra on $\Delta \times G$. We aim to apply Theorem 2.4.9 to $X_k = \hat{\phi} \circ f_H^{n-k}$. First note that by an identical calculation to (2.8.1), X_k is $\mathcal{B}_k-$measurable for all $0 \leq k \leq n$. From Proposition 4.6.3, Proposition 2.8.3, and Proposition 2.4.5 (ii), we have that $(\sum_{k=0}^{j} \hat{\psi} \circ f_H^{n-k})_{j=0}^{n}$ is a martingale with respect to $(\mathcal{B}_j)_{j=0}^{n}$. For $0 \leq \ell \leq m \leq n$, it follows that

$$\begin{aligned}\sum_{k=\ell}^{m} \mathbb{E}\big[\hat{\psi} \circ f_H^{n-k} \,\big|\, \mathcal{B}_\ell\big] &= \mathbb{E}\Bigg[\sum_{k=\ell}^{m} \hat{\psi} \circ f_H^{n-k} \,\Bigg|\, \mathcal{B}_\ell\Bigg] \\ &= \mathbb{E}\Bigg[\sum_{k=0}^{m} \hat{\psi} \circ f_H^{n-k} - \sum_{k=0}^{\ell-1} \hat{\psi} \circ f_H^{n-k} \,\Bigg|\, \mathcal{B}_\ell\Bigg] \\ &= \sum_{k=0}^{\ell} \hat{\psi} \circ f_H^{n-k} - \sum_{k=0}^{\ell-1} \hat{\psi} \circ f_H^{n-k} = \hat{\psi} \circ f_H^{n-\ell}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\sum_{k=\ell}^{m} \mathbb{E}\big[X_k \,\big|\, \mathcal{B}_\ell\big] &= \sum_{k=\ell}^{m} \mathbb{E}\big[\hat{\psi} \circ f_H^{n-k} + \hat{\chi} \circ f_H^{n-(k-1)} - \hat{\chi} \circ f_H^{n-k} \,\big|\, \mathcal{B}_\ell\big] \\ &= \hat{\psi} \circ f_H^{n-\ell} + \mathbb{E}\big[\hat{\chi} \circ f_H^{n-(\ell-1)} \,\big|\, \mathcal{B}_\ell\big] - \hat{\chi} \circ f_H^{n-m}.\end{aligned}$$

Now, by Proposition 2.4.2 (v) and invariance of f_H, we have that

$$\big|\mathbb{E}\big[\hat{\chi} \circ f_H^{n-(\ell-1)} \,\big|\, \mathcal{B}_\ell\big]\big|_{p-1} \leq |\hat{\chi}|_{p-1}.$$

By Proposition 4.6.2 (i) and (iii), it follows that

$$\max_{0 \leq \ell \leq m \leq n} \left| \sum_{k=\ell}^{m} \mathbb{E}\big[X_k \,\big|\, \mathcal{B}_\ell\big] \right|_{p-1} \ll \|v\|_\eta \|h\|_\eta.$$

Therefore

$$\begin{aligned}\max_{0 \leq \ell \leq m \leq n} \left| X_\ell \sum_{k=\ell}^{m} \mathbb{E}\big[X_k \,\big|\, \mathcal{B}_\ell\big] \right|_{p-1} &\leq |\hat{\phi}|_\infty \max_{0 \leq \ell \leq m \leq n} \left| \sum_{k=\ell}^{m} \mathbb{E}\big[X_k \,\big|\, \mathcal{B}_\ell\big] \right|_{p-1} \\ &\ll \|v\|_\eta^2 \|h\|_\eta^2.\end{aligned}$$

Applying Theorem 2.4.9 gives

$$\Big| \max_{0\le j\le n} \Big| \sum_{k=0}^{j} X_k \Big| \Big|_{2(p-1)} \ll \Big(\sum_{\ell=0}^{n} \max_{0\le \ell\le m\le n} \Big| X_\ell \sum_{k=\ell}^{m} \mathbb{E}\big[X_k \,\big|\, \mathcal{B}_\ell\big] \Big|_{p-1} \Big)^{1/2}$$
$$\ll n^{1/2} \|v\|_\eta \|h\|_\eta.$$

To complete the proof, we note that $\sum_{k=0}^{j-1} \hat\phi\circ f_H^k = \sum_{k=0}^{n} X_k - \sum_{k=0}^{n-j} X_k$ to obtain

$$\Big| \max_{1\le j\le n} \Big| \sum_{k=0}^{j-1} \hat\phi\circ f_H^k \Big| \Big|_{2(p-1)} \le \Big| \max_{1\le j\le n} \Big| \sum_{k=0}^{n} X_k \Big| \Big|_{2(p-1)} + \Big| \max_{1\le j\le n} \Big| \sum_{k=0}^{n-j} X_k \Big| \Big|_{2(p-1)}$$
$$\ll n^{1/2} \|v\|_\eta \|h\|_\eta,$$

as required. □

Under the assumption $V \in L^2(Y;\mathbb{R}^d)$, we can strengthen the estimate in Corollary 4.8.3 (i).

Corollary 4.8.4. *Suppose $\tau \in L^p(Y)$ where $1 < p < 2$ and $V \in L^2(Y;\mathbb{R}^d)$. Then*

$$\Big| \max_{1\le j\le n} \Big| \sum_{k=0}^{j-1} \hat\psi\circ f_H^k \Big| \Big|_2 \le C n^{1/2} \big(|V|_2 + \|v\|_\eta \|h\|_\eta \big) \quad \textit{for all } n \ge 1.$$

Proof. We have from (4.8.5) that

$$\Big| \max_{1\le j\le n} \Big| \sum_{k=1}^{j} \hat\psi\circ f_H^{n-k} \Big| \Big|_2 \ll \Big| \Big(\sum_{k=1}^{n} |\hat\psi\circ f_H^{n-k}|^2 \Big)^{1/2} \Big|_2 = \Big| \sum_{k=1}^{n} |\hat\psi\circ f_H^{n-k}|^2 \Big|_1^{1/2}$$
$$\le \Big(\sum_{k=1}^{n} \big| |\hat\psi\circ f_H^{n-k}|^2 \big|_1 \Big)^{1/2} = n^{1/2} |\hat\psi|_2 \ll n^{1/2} \big(|V|_2 + \|v\|_\eta \|h\|_\eta \big),$$

where we use Proposition 4.6.2 (ii) in the final estimate. Combining this with (4.8.7) gives

$$\Big| \max_{1\le j\le n} \Big| \sum_{k=0}^{j-1} \hat\psi\circ f_H^k \Big| \Big|_2 \le \Big| \max_{1\le j\le n} \Big| \sum_{k=1}^{n} \hat\psi\circ f_H^{n-k} \Big| \Big|_2 + \Big| \max_{1\le j\le n} \Big| \sum_{k=1}^{n-j} \hat\psi\circ f_H^{n-k} \Big| \Big|_2$$
$$\ll n^{1/2} \big(|V|_2 + \|v\|_\eta \|h\|_\eta \big),$$

completing the proof. □

To conclude the section, we show that for τ sufficiently regular, the covariance matrix $\Sigma = \int_{\Delta\times G} \hat\psi\hat\psi^T \,\mathrm{d}m_\Delta$ can be characterised in terms of the observable ϕ.

Corollary 4.8.5. *Suppose $\tau \in L^p(Y)$ with $p \geq 2$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \int_{X\times G} \left(\sum_{k=0}^{n-1} \phi \circ T_h^k\right)\left(\sum_{k=0}^{n-1} \phi \circ T_h^k\right)^T \mathrm{d}m = \int_{\Delta\times G} \hat\psi\hat\psi^T \,\mathrm{d}m_\Delta.$$

Proof. Write $S_n\hat\phi = \sum_{k=0}^{n-1} \hat\phi \circ f_H^k$ and $S_n\hat\psi = \sum_{k=0}^{n-1} \hat\psi \circ f_H^k$. From Proposition 4.2.3, we have

$$\int_{X\times G} \left(\sum_{k=0}^{n-1} \phi \circ T_h^k\right)\left(\sum_{k=0}^{n-1} \phi \circ T_h^k\right)^T \mathrm{d}m = \int_{\Delta\times G} (S_n\hat\phi)(S_n\hat\phi)^T \,\mathrm{d}m_\Delta.$$

Moreover, we have from Proposition 4.6.3 and Proposition 2.8.3 that $(\hat\psi \circ f_H^{n-k})_{k=1}^n$ is a martingale difference sequence, so that

$$\int_{\Delta\times G} (S_n\hat\psi)(S_n\hat\psi)^T \,\mathrm{d}m_\Delta = n \int_{\Delta\times G} \hat\psi\hat\psi^T \,\mathrm{d}m_\Delta$$

by Proposition 2.4.6. From Theorem 4.1.5 (ii) and Corollary 4.8.3 (ii), we have $|S_n\hat\phi|_2 \leq |S_n\hat\phi|_{2(p-1)} \ll n^{1/2}\|v\|_\eta\|h\|_\eta$ and $|S_n\hat\psi|_2 \leq |S_n\hat\psi|_p \ll n^{1/2}\|v\|_\eta\|h\|_\eta$. Therefore

$$\begin{aligned}
&\frac{1}{n}\left\| \int_{\Delta\times G} (S_n\hat\phi)(S_n\hat\phi)^T \,\mathrm{d}m_\Delta - \int_{\Delta\times G} (S_n\hat\psi)(S_n\hat\psi)^T \,\mathrm{d}m_\Delta \right\| \\
&\leq \frac{1}{n}\big|(S_n\hat\phi)(S_n\hat\phi)^T - (S_n\hat\psi)(S_n\hat\psi)^T\big|_1 \\
&\leq \frac{1}{n}\Big(\big|S_n\hat\phi\big((S_n\hat\phi)^T - (S_n\hat\psi)^T\big)\big|_1 + \big|(S_n\hat\psi - S_n\hat\phi)(S_n\hat\psi)^T\big|_1\Big) \\
&\leq \frac{1}{n}\Big(|S_n\hat\phi|_2\big|(S_n\hat\phi)^T - (S_n\hat\psi)^T\big|_2 + |S_n\hat\psi - S_n\hat\phi|_2\big|(S_n\hat\psi)^T\big|_2\Big) \\
&= \frac{1}{n}\big(|S_n\hat\phi|_2 + |S_n\hat\psi|_2\big)|S_n\hat\phi - S_n\hat\psi|_2 = \frac{1}{n}\big(|S_n\hat\phi|_2 + |S_n\hat\psi|_2\big)|\hat\chi \circ f_H^n - \hat\chi|_2 \\
&\ll \frac{\|v\|_\eta\|h\|_\eta}{\sqrt{n}}|\hat\chi \circ f_H^n - \hat\chi|_2 \to 0,
\end{aligned} \tag{4.8.8}$$

where the final convergence uses that $p \geq 2$ with (4.8.4). □

Remark 4.8.6. *In the case that $T_h \colon X \times G \to X \times G$ is mixing sufficiently fast (for example, polynomially of order n^{-b} for $b > 1$), one has the characterisation*

$$\Sigma = \int_{X\times G} \phi\, \phi^T \,\mathrm{d}m + \sum_{k=1}^{\infty} \int_{X\times G} \phi \circ T_h^k\, \phi^T \,\mathrm{d}m + \sum_{k=1}^{\infty} \int_{X\times G} \phi\, (\phi \circ T_h^k)^T \,\mathrm{d}m.$$

See for example [36, Section 3.2].

Remark 4.8.7. *Let $\hat{\phi} \colon \Delta \times G \to \mathbb{R}^d$ be such that $\int_{\Delta\times G} \hat{\phi} \,\mathrm{d}m_\Delta = 0$. Suppose further that*

(i) $\hat{\phi} \in L^\infty(\Delta \times G; \mathbb{R}^d)$.

(ii) $\sum_{\ell=0}^{\tau(y)-1} \hat{\phi}(y, g, \ell) = g \cdot V(y)$ *for some $V \colon Y \to \mathbb{R}^d$.*

(iii) *There exists $\gamma \in [\lambda^{-\eta}, 1)$ such that $\mathcal{P}_H V \in F_\gamma(Y; \mathbb{R}^d)$.*

For such observables, we can see that all the results in this chapter go through with $\|v\|_\eta \|h\|_\eta$ replaced by $|\hat{\phi}|_\infty + \|\mathcal{P}_H V\|_\gamma$. Since $\mathbb{R}^{d,d} \cong \mathbb{R}^{d^2}$, the results also go through for matrix-valued observables satisfying the above.

4.9 Examples

We now give some examples of where our results apply. In the literature, it is common to find estimates of the return time tail $\mu_Y(\tau > n)$. We begin by showing how these estimates relate to the integrability of τ.

Lemma 4.9.1. *Suppose $\tau \colon Y \to \mathbb{Z}^+$ is a return time and $p > 1$.*

(i) *If there exists $b > 0$ such that $\mu_Y(\tau > n) = O(n^{-b})$, then $\tau \in L^p(Y)$ if and only if $b > p$.*

(ii) *If there exist $b > 0$ and $0 < c \le 1$ such that $\mu_Y(\tau > n) = O(e^{-bn^c})$, then $\tau \in L^p(Y)$.*

Proof. We have

$$\int_Y |\tau|^p \,\mathrm{d}\mu_Y = \int_0^\infty \mu_Y(\tau^p > x)\,\mathrm{d}x = \sum_{n=0}^\infty \int_n^{n+1} \mu_Y(\tau^p > x)\,\mathrm{d}x$$
$$\le \sum_{n=0}^\infty \int_n^{n+1} \mu_Y(\tau^p > n)\,\mathrm{d}x = \sum_{n=0}^\infty \mu_Y(\tau > n^{1/p}).$$

The results follow. □

Note that if $\tau \in L^p(Y)$ for $p > 1$, then we immediately have $V \in L^p(Y;\mathbb{R}^d)$ and $V^* \in L^p(Y)$. We apply Lemma 4.9.1 to the examples in Section 3.2.

(i) Uniformly expanding maps are non-uniformly expanding with $\tau = 1 \in L^p(Y)$ for all $1 < p \le \infty$, and so our results apply.

(ii) The intermittent maps with parameter $\gamma \in (0,1)$ as in Example 2.2.6 satisfy $\mu_Y(\tau > n) = O(n^{-1/\gamma})$ [102]. It follows from Lemma 4.9.1 (i) that $\tau \in L^2(Y)$ if and only if $\gamma \in [0,1/2)$, and our results immediately apply in this case. When $\gamma \in [1/2,1)$, we have $\tau \in L^p(Y)$ for $1 < p < 2$. In [39, Theorem 4.1], it is shown under some mild conditions on v and h at the neutral fixed point 0 that $V^* \in L^2(Y)$, and so our results apply.

(iii) The unimodal maps (along Collet-Eckmann parameters) in Example 3.2.3 satisfy $\mu_Y(\tau > n) = O(e^{-dn})$ for some $d > 0$ [101]. Therefore $\tau \in L^p(Y)$ for all $p > 1$ by Lemma 4.9.1 (ii), and our results apply.

(iv) The Viana maps considered in Example 3.2.4 satisfy $\mu_Y(\tau > n) = O(e^{-bn^{1/2}})$ for some $b > 0$ [41]. Therefore $\tau \in L^p(Y)$ for all $p > 1$ by Lemma 4.9.1 (ii), and our results apply.

Chapter 5

Secondary martingale-coboundary decomposition

5.1 Outline

We continue with the setup of Chapter 4. Moreover, we suppose throughout that $p \geq 2$. In this chapter, we construct a secondary martingale-coboundary decomposition as in [60]. This decomposes the square of the martingale in our primary martingale-coboundary decomposition, which allows us to control sums of squares as is often required in more sophisticated limit laws. As an application, we prove the following almost sure invariance principle for the one-dimensional projections of our underlying observable. This is done by appealing to the results of Cuny and Merlevède [22] which were stated in Section 2.8.

Theorem 5.1.1. *Suppose $\tau \in L^p(Y)$ for some $p \geq 2$. Define $\Sigma \in \mathbb{R}^{d,d}$ by*

$$\Sigma = \lim_{n\to\infty} \int_{X\times G} \left(\sum_{k=0}^{n-1} \phi \circ T_h^k \right) \left(\sum_{k=0}^{n-1} \phi \circ T_h^k \right)^T \mathrm{d}m$$

and let $c \in \mathbb{R}^d$ with $c^T \Sigma c > 0$. Then there exists a probability space supporting a *sequence of random variables* $(S_n)_{n\geq 1}$ *with the same joint distributions as* $(\sum_{k=0}^{n-1}(c\cdot\phi)\circ T_h^k)_{n\geq 1}$ *and a sequence of independent and identically distributed random variables* $(Z_n)_{n\geq 1}$ *with distribution* $\mathcal{N}(0, c^T\Sigma c)$*, such that almost surely,*

$$\sup_{1\leq k\leq n}\left|S_k - \sum_{j=1}^{k} Z_j\right| = \begin{cases} o\big((n\log\log n)^{1/2}\big) & \text{if } p = 2, \\ o(n^{1/p}(\log n)^{1/2}) & \text{if } p\in(2,4), \\ O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}) & \text{if } p\geq 4. \end{cases}$$

Remark 5.1.2. *The rates in the almost sure invariance principle have additional powerful implications (see [6] and references therein). We note that:*

(i) *For* $p > 2$*, Theorem 5.1.1 recovers Theorem 4.1.2 and Theorem 4.1.3. This is done by a Cramér-Wold argument, see for example [36, Corollary 2.7 and Corollary 2.8].*

(ii) *When* $p = 2$*, the given rate is not sufficient for deducing these results. However, this rate does imply the law of the iterated logarithm and functional law of the iterated logarithm. For completeness, these are given in Section 5.3.*

(iii) *The given rate when* $p = 2$ *can also be obtained for* $1 < p < 2$ *when we assume* V *as defined in* (4.1.1) *lies in* $L^2(Y;\mathbb{R}^d)$*. This is easily seen from the proof.*

Remark 5.1.3. *In our proof, we utilise results of Cuny and Merlevède [22], who use a Skorokhod embedding of reverse martingales in Brownian motion [89] to obtain the almost sure invariance principle for sequences of reverse martingale differences. The best rate achievable via this approach is* $O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})$ *[55]. We obtain this rate when* $p\geq 4$*, which improves on the rate* $O(n^{1/4+\delta})$ *for any* $\delta > 0$ *given in [36] in the case of a uniformly expanding base.*

The structure of the chapter is as follows: In Section 5.2, we define a certain matrix-valued observable in terms of the martingale from the primary martingale-coboundary decomposition, and then verify the conditions of Remark 4.8.7 to give us our secondary martingale-coboundary decomposition. In Section 5.3, we use this to prove Theorem 5.1.1 as well as proving the law of the iterated logarithm and functional law of the iterated logarithm. We remark that the results in this chapter apply to the examples considered in Section 4.9.

5.2 Construction of the secondary decomposition

In this section we suppose that $\tau \in L^p(Y)$ with $p > 2$. In particular, we may choose $\gamma \in [\lambda^{-\eta}, 1)$ as in Remark 4.4.4. Recall our observable $\phi \colon X \times G \to \mathbb{R}^d$ given by $\phi = g \cdot v$ where $v \in C^\eta(X; \mathbb{R}^d)$ and $\int_{X\times G} \phi \, \mathrm{d}m = 0$. In the notation of Chapter 4, we have the decompositions

$$\begin{aligned} \hat{\phi} &= \hat{\psi} + \hat{\chi} \circ f_H - \hat{\chi} && \text{where } \hat{\psi} \in \ker \hat{\mathcal{L}}, \\ \Phi &= \Psi + \chi \circ F_H - \chi, && \text{and} \\ V &= M + \mathcal{U}_H J - J && \text{where } M \in \ker \mathcal{P}_H \end{aligned}$$

from Section 4.6. We can extend the transfer operator $\hat{\mathcal{L}}$ for f_H to an operator on $L^1(\Delta\times G; \mathbb{R}^{d,d})$ by acting component-wise. Similarly, we let $\hat{\mathcal{U}} \colon L^1(\Delta\times G; \mathbb{R}^{d,d}) \to L^1(\Delta \times G; \mathbb{R}^{d,d})$ denote the Koopman operator for f_H. Since $\hat{\psi}\hat{\psi}^T \in L^1(\Delta \times G; \mathbb{R}^{d,d})$, we can define $\tilde{\phi} \colon \Delta \times G \to \mathbb{R}^{d,d}$ by

$$\tilde{\phi} = \hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T) - \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \, \mathrm{d}m_\Delta. \tag{5.2.1}$$

Note by Proposition 2.5.3 (vii) that

$$\int_{\Delta\times G} \tilde{\phi} \, \mathrm{d}m_\Delta = \int_{\Delta\times G} \hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T) \, \mathrm{d}m_\Delta - \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \, \mathrm{d}m_\Delta = 0.$$

We aim to apply Remark 4.8.7 to $\tilde{\phi}$. Our first step in this direction is to show that M is locally Lipschitz.

Proposition 5.2.1. *If $\epsilon \in (0, 1]$, then $M \in F^{\mathrm{loc}}_{\gamma^\epsilon}(Y; \mathbb{R}^d)$ with $\|\mathbb{1}_a M\|_{\gamma^\epsilon} \leq C\tau(a)^{1+\epsilon}\|v\|_\eta \|h\|_\eta$ for all $a \in \alpha$.*

Proof. Let $a \in \alpha$ and note that $\mathbb{1}_a M = \mathbb{1}_a(V + J - \mathcal{U}_H J)$. Since $\gamma^\epsilon \in [\gamma, 1)$, it follows from Proposition 4.6.1 that $\|J\|_{\gamma^\epsilon} \ll \|v\|_\eta \|h\|_\eta$. Moreover, from Proposition 4.4.2 (ii), we have $\|\mathbb{1}_a V\|_{\gamma^\epsilon} \ll \tau(a)^{1+\epsilon}\|v\|_\eta\|h\|_\eta$. We show $\|\mathbb{1}_a \mathcal{U}_H J\|_{\gamma^\epsilon} \ll \tau(a)^\epsilon \|v\|_\eta \|h\|_\eta$, from which the result follows. Note that $\|\mathbb{1}_a H\|_{\gamma^\epsilon} \ll \tau(a)^\epsilon \|h\|_\eta$ from Proposition 4.4.2 (i). For $x, y \in a$, we have

$$\begin{aligned}
\big|\mathcal{U}_H J(x) - \mathcal{U}_H J(y)\big| &= \big|H(x) \cdot J(Fx) - H(y) \cdot J(Fy)\big| \\
&\le \big\|H(x) - H(y)\big\|\big|J(Fx)\big| + \big\|H(x)\big\|\big|J(Fx) - J(Fy)\big| \\
&\ll \tau(a)^\epsilon \|h\|_\eta |J|_\infty \, d_{\gamma^\epsilon}(x, y) + |J|_{\gamma^\epsilon} \, d_{\gamma^\epsilon}(Fx, Fy) \\
&\ll \tau(a)^\epsilon \|v\|_\eta \|h\|_\eta \, d_{\gamma^\epsilon}(x, y),
\end{aligned}$$

so that $|\mathbb{1}_a \mathcal{U}_H J|_{\gamma^\epsilon} \ll \tau(a)^\epsilon \|v\|_\eta \|h\|_\eta$. To estimate the sup norm, note that

$$|\mathbb{1}_a \mathcal{U}_H J|_\infty \le |J|_\infty \ll \|v\|_\eta \|h\|_\eta.$$

Therefore $\|\mathbb{1}_a \mathcal{U}_H J\|_{\gamma^\epsilon} \ll \tau(a)^\epsilon \|v\|_\eta \|h\|_\eta$, completing the proof. □

Define the induced observable $\tilde{\Phi} \colon Y \times G \to \mathbb{R}^{d,d}$ by

$$\tilde{\Phi}(y, g) = \sum_{\ell=0}^{\tau(y)-1} \tilde{\phi}(y, g, \ell). \tag{5.2.2}$$

Remark 4.8.7 (ii) requires us to show that $\tilde{\Phi}$ is equivariant. In the next proposition, we define a suitable action on $\mathbb{R}^{d,d}$, as well as showing it is invariant with respect to the Frobenius inner product. Recall this is given by $\langle A, B\rangle = \mathrm{tr}(A^T B)$ for $A, B \in \mathbb{R}^{d,d}$, where $\mathrm{tr}(\cdot)$ denotes the trace.

Proposition 5.2.2. *Define $\star \colon G \times \mathbb{R}^{d,d} \to \mathbb{R}^{d,d}$ by $g \star A = gAg^T$. This is a continuous linear action of G on $\mathbb{R}^{d,d}$. Moreover, for all $g \in G$ and $A, B \in \mathbb{R}^{d,d}$, we have $\langle g \star A, g \star B\rangle = \langle A, B\rangle$.*

Proof. That $\star$ defines a continuous linear action follows from the definition. For the second statement, we have from properties of the trace and orthogonality that

$$\begin{aligned}
\langle g \star A, g \star B\rangle &= \langle gAg^T, gBg^T\rangle = \mathrm{tr}(gA^T g^T gBg^T) = \mathrm{tr}(gA^T Bg^T) \\
&= \mathrm{tr}(g^T gA^T B) = \mathrm{tr}(A^T B) = \langle A, B\rangle,
\end{aligned}$$

as claimed. □

Let us denote by $\mathcal{L}_H\colon L^1(\Delta;\mathbb{R}^{d,d}) \to L^1(\Delta;\mathbb{R}^{d,d})$ the twisted transfer operator for f with respect to $\star$, defined for $\tilde{v} \in L^1(\Delta;\mathbb{R}^{d,d})$ by $\mathcal{L}_H(\tilde{v}) = \mathcal{L}(\hat{H}^{-1} \star \tilde{v})$, where $\hat{H}\colon \Delta \to G$ is as in Remark 4.2.1 and $\mathcal{L}\colon L^1(\Delta;\mathbb{R}^{d,d}) \to L^1(\Delta;\mathbb{R}^{d,d})$ is the transfer operator for f. We can immediately apply Proposition 4.3.7 to conclude:

Proposition 5.2.3. *Suppose* $\tilde{\phi}\colon \Delta \times G \to \mathbb{R}^{d,d}$ *is given by* $\tilde{\phi} = g \star \tilde{v}$*, where* $\tilde{v} \in L^1(\Delta;\mathbb{R}^{d,d})$*. Then* $\hat{\mathcal{L}}\tilde{\phi} = g \star \mathcal{L}_H\tilde{v}$*.*

We set out some notation for the rest of the section. We fix $\|\cdot\|$ to be the Frobenius norm on $\mathbb{R}^{d,d}$. That is, $\|A\| = (\mathrm{tr}(A^TA))^{1/2}$ for $A \in \mathbb{R}^{d,d}$. Let $\mathcal{U}, \mathcal{P}\colon L^1(Y;\mathbb{R}^{d,d}) \to L^1(Y;\mathbb{R}^{d,d})$ denote the Koopman and transfer operators for F. We denote by $\mathcal{U}_H, \mathcal{P}_H\colon L^1(Y;\mathbb{R}^{d,d}) \to L^1(Y;\mathbb{R}^{d,d})$ the twisted Koopman and twisted transfer operators for F with respect to $\star$. These are defined for $\tilde{V} \in L^1(Y;\mathbb{R}^{d,d})$ by $\mathcal{U}_H(\tilde{V}) = H \star \mathcal{U}\tilde{V}$ and $\mathcal{P}_H(\tilde{V}) = \mathcal{P}(H^{-1} \star \tilde{V})$. The next two propositions verify the points of Remark 4.8.7. We first give a lemma.

Lemma 5.2.4. *For* $(y, g, \ell) \in \Delta \times G$*, we have*

$$\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T)(y,g,\ell) = \begin{cases} 0 & \text{if } \ell \le \tau(y) - 2, \\ g \star \mathcal{U}_H\mathcal{P}_H(MM^T)(y) & \text{if } \ell = \tau(y) - 1. \end{cases}$$

Proof. First note from (4.6.8) that $\hat{\psi}(y,g,\ell) = g \cdot \hat{m}(y,\ell)$, where

$$\hat{m}(y,\ell) = \begin{cases} 0 & \text{if } \ell \le \tau(y) - 2, \\ M(y) & \text{if } \ell = \tau(y) - 1. \end{cases}$$

Therefore $\hat{\psi}\hat{\psi}^T = (g \cdot \hat{m})(g \cdot \hat{m})^T = g(\hat{m}\hat{m}^T)g^T = g \star (\hat{m}\hat{m}^T)$. Suppose $\ell \le \tau(y) - 2$. We have from Proposition 5.2.3 that

$$\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T)(y,g,\ell) = \hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T)(y,g,\ell+1) = g \star \mathcal{L}_H(\hat{m}\hat{m}^T)(y,\ell+1).$$

Moreover, from Remark 4.3.6 we have

$$\mathcal{L}_H(\hat{m}\hat{m}^T)(y,\ell+1) = \hat{m}\hat{m}^T(y,\ell) = 0,$$

which proves the claim when $\ell \le \tau(y)-2$. Suppose now that $\ell = \tau(y)-1$. We have from Proposition 5.2.3 that

$$\begin{aligned}\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat\psi\hat\psi^T)(y,g,\ell) &= \hat{\mathcal{L}}(\hat\psi\hat\psi^T)(F_H(y,g),0) = \hat{\mathcal{L}}(\hat\psi\hat\psi^T)(Fy, gH(y),0)\\ &= gH(y)\star\mathcal{L}_H(\hat m\hat m^T)(Fy,0).\end{aligned} \tag{5.2.3}$$

Now, writing $(Fy)_a$ for the unique element in $a\in\alpha$ such that $F(Fy)_a = Fy$, we have from Remark 4.3.6 and Remark 4.3.2 that

$$\begin{aligned}\mathcal{L}_H(\hat m\hat m^T)(Fy,0) &= \sum_{a\in\alpha}\zeta\big((Fy)_a\big)H\big((Fy)_a\big)^{-1}\star\hat m\hat m^T\big((Fy)_a,\tau(a)-1\big)\\ &= \sum_{a\in\alpha}\zeta\big((Fy)_a\big)H\big((Fy)_a\big)^{-1}\star MM^T\big((Fy)_a\big)\\ &= \mathcal{P}_H(MM^T)(Fy),\end{aligned}$$

so that continuing from (5.2.3), we have

$$\begin{aligned}\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat\psi\hat\psi^T)(y,g,\ell) &= gH(y)\star\mathcal{P}_H(MM^T)(Fy) = g\star H(y)\star\mathcal{P}_H(MM^T)(Fy)\\ &= g\star\mathcal{U}_H\mathcal{P}_H(MM^T)(y),\end{aligned}$$

completing the proof. □

Proposition 5.2.5. *The observable $\tilde\phi$ defined in (5.2.1) satisfies $\tilde\phi\in L^\infty(Y;\mathbb{R}^{d,d})$ with $|\tilde\phi|_\infty \le C\|v\|_\eta^2\|h\|_\eta^2$.*

Proof. For $y\in Y$ and $g\in G$, we have from Proposition 5.2.2 that

$$\big\|g\star\mathcal{U}_H\mathcal{P}_H(MM^T)(y)\big\| = \big\|gH(y)\star\mathcal{U}\mathcal{P}_H(MM^T)(y)\big\| = \big\|\mathcal{U}\mathcal{P}_H(MM^T)(y)\big\|.$$

Moreover, note that

$$\left\|\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\right\| \le \int_{\Delta\times G}\|\hat\psi\hat\psi^T\|\,\mathrm{d}m_\Delta \le |\hat\psi|_2^2. \tag{5.2.4}$$

Therefore, from Lemma 5.2.4 we have

$$|\tilde\phi|_\infty \le \big|\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat\psi\hat\psi^T)\big|_\infty + \left|\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\right|_\infty \le \big|\mathcal{P}_H(MM^T)\big|_\infty + |\hat\psi|_2^2.$$

From Proposition 4.6.1 and (4.4.4), we note that

$$\begin{aligned} |\mathbb{1}_a M|_\infty &\le |\mathbb{1}_a V|_\infty + |\mathbb{1}_a \mathcal{U}_H J|_\infty + |\mathbb{1}_a J|_\infty \le |\mathbb{1}_a V|_\infty + 2|J|_\infty \\ &\ll \tau(a)\|v\|_\eta \|h\|_\eta. \end{aligned} \tag{5.2.5}$$

Therefore

$$|\mathbb{1}_a M M^T|_\infty \ll \tau(a)^2 \|v\|_\eta^2 \|h\|_\eta^2. \tag{5.2.6}$$

For $y \in Y$, it follows from Remark 4.3.2, Proposition 3.3.9, and Proposition 5.2.2 that

$$\begin{aligned} \big\|\mathcal{P}_H(MM^T)(y)\big\| &\le D \sum_{a\in\alpha} \mu_Y(a) \big\|H(y_a) \star MM^T(y_a)\big\| = D \sum_{a\in\alpha} \mu_Y(a)\big\|MM^T(y_a)\big\| \\ &\ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)^2 \|v\|_\eta^2 \|h\|_\eta^2 \ll \|v\|_\eta^2\|h\|_\eta^2, \end{aligned} \tag{5.2.7}$$

and so $|\mathcal{P}_H(MM^T)|_\infty \ll \|v\|_\eta^2\|h\|_\eta^2$. Finally, Proposition 4.6.2 (i) gives $|\hat\psi|_2^2 \ll \|v\|_\eta^2\|h\|_\eta^2$, and the result follows. □

Proposition 5.2.6. *Let $\tau \in L^p(Y)$ with $p > 2$ and $\tilde\Phi$ be as in (5.2.2). Then $\tilde\Phi = g \star \tilde V$, where $\tilde V \colon Y \to \mathbb{R}^{d,d}$ satisfies the following:*

(i) *If $p \ge 3$, then $\mathcal{P}_H \tilde V \in F_\gamma(Y; \mathbb{R}^{d,d})$ with $\|\mathcal{P}_H\tilde V\|_\gamma \le C\|v\|_\eta^2\|h\|_\eta^2$.*

(ii) *If $2 < p < 3$, then $\mathcal{P}_H\tilde V \in F_{\gamma^{p-2}}(Y;\mathbb{R}^{d,d})$ with $\|\mathcal{P}_H\tilde V\|_{\gamma^{p-2}} \le C\|v\|_\eta^2\|h\|_\eta^2$.*

Proof. First note that for $(y, g) \in Y \times G$, we have from Lemma 5.2.4 that

$$\begin{aligned} \tilde\Phi(y,g) &= \sum_{\ell=0}^{\tau(y)-1} \tilde\phi(y,g,\ell) = \sum_{\ell=0}^{\tau(y)-1} \Big(\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat\psi\hat\psi^T)(y,g,\ell) - \int_{\Delta\times G} \hat\psi\hat\psi^T \, \mathrm{d}m_\Delta \Big) \\ &= g \star \mathcal{U}_H \mathcal{P}_H(MM^T)(y) - \tau(y) \int_{\Delta\times G} \hat\psi\hat\psi^T\,\mathrm{d}m_\Delta. \end{aligned}$$

Now, from (4.7.2) we have

$$\int_{\Delta\times G} \hat\psi\hat\psi^T\,\mathrm{d}m_\Delta = g\Big(\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big)g^T = g\star\Big(\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big).$$

It follows from linearity of $\star$ that

$$\begin{aligned}\tilde{\Phi}(y,g) &= g \star \mathcal{U}_H \mathcal{P}_H(MM^T)(y) - \tau(y) \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta \\ &= g \star \mathcal{U}_H \mathcal{P}_H(MM^T)(y) - g \star \left(\tau(y) \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right) \\ &= g \star \left(\mathcal{U}_H \mathcal{P}_H(MM^T)(y) - \tau(y) \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right) = g \star \tilde{V}(y),\end{aligned}$$

where we define $\tilde{V} \colon Y \to \mathbb{R}^{d,d}$ by

$$\tilde{V}(y) = \mathcal{U}_H \mathcal{P}_H(MM^T)(y) - \tau(y) \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta.$$

We now prove (i) and (ii) in turn.

Proof of (i). Note that

$$\mathcal{P}_H \tilde{V} = \mathcal{P}_H(MM^T) - \mathcal{P}_H\left(\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right).$$

We first estimate $|\mathcal{P}_H \tilde{V}|_\infty$. From (5.2.4) and Proposition 4.6.2 (i), we have

$$\left\|\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right\| \ll \tau \|v\|_\eta^2 \|h\|_\eta^2. \tag{5.2.8}$$

For $y \in Y$, arguing as in (5.2.7) gives

$$\begin{aligned}&\left\|\mathcal{P}_H\left(\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right)(y)\right\| \le D \sum_{a\in\alpha} \mu_Y(a) \left\|H(y_a) \star \tau(a) \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right\| \\ &= D \sum_{a\in\alpha} \mu_Y(a) \left\|\tau(a) \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right\| \ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)\|v\|_\eta^2 \|h\|_\eta^2 \ll \|v\|_\eta^2 \|h\|_\eta^2,\end{aligned}$$

and so

$$\left|\mathcal{P}_H\left(\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right)\right|_\infty \ll \|v\|_\eta^2 \|h\|_\eta^2.$$

It follows from (5.2.7) that

$$|\mathcal{P}_H \tilde{V}|_\infty \le \left|\mathcal{P}_H(MM^T)\right|_\infty + \left|\mathcal{P}_H\left(\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \,\mathrm{d}m_\Delta\right)\right|_\infty \ll \|v\|_\eta^2 \|h\|_\eta^2.$$

We next analyse the Lipschitz semi-norm of $\mathcal{P}_H\tilde{V}$. Note that $|\mathcal{P}_H\tilde{V}|_\gamma \leq |\mathcal{P}_H(MM^T)|_\gamma + |\mathcal{P}_H(\tau\int_{\Delta\times G}\hat{\psi}\hat{\psi}^T\,\mathrm{d}m_\Delta)|_\gamma$. We first estimate $|\mathcal{P}_H(MM^T)|_\gamma$. Let $x, y \in Y$. Note by Remark 4.3.2 that

$$\begin{aligned}
&\big\|\mathcal{P}_H(MM^T)(x) - \mathcal{P}_H(MM^T)(y)\big\| \\
&\leq \sum_{a\in\alpha}\big|\zeta(x_a) - \zeta(y_a)\big|\big\|H(x_a)^{-1}\star MM^T(x_a)\big\| \\
&\quad + \sum_{a\in\alpha}\zeta(y_a)\big\|\big(H(x_a)^{-1} - H(y_a)^{-1}\big)\star MM^T(x_a)\big\| \\
&\quad + \sum_{a\in\alpha}\zeta(y_a)\big\|H(y_a)^{-1}\star\big(MM^T(x_a) - MM^T(y_a)\big)\big\| =: I + II + III. \qquad (5.2.9)
\end{aligned}$$

All the terms involving ζ are dealt with using Proposition 3.3.9. For I, note that from Proposition 5.2.2 and (5.2.6), we have

$$\big\|H(x_a)^{-1}\star MM^T(x_a)\big\| = \big\|MM^T(x_a)\big\| \ll \tau(a)^2\|v\|_\eta^2\|h\|_\eta^2.$$

Therefore

$$I \ll \sum_{a\in\alpha}\mu_Y(a)\tau(a)^2\|v\|_\eta^2\|h\|_\eta^2\,d_\gamma(x,y) \ll \|v\|_\eta^2\|h\|_\eta^2\,d_\gamma(x,y).$$

For II, first note that

$$\big\|(H(x_a)^{-1} - H(y_a)^{-1})^T\big\| = \big\|(H(x_a)^T - H(y_a)^T)^T\big\| = \big\|H(x_a) - H(y_a)\big\| \leq 2.$$

In addition, note from (4.4.6) that

$$\big\|H(x_a)^{-1} - H(y_a)^{-1}\big\| = \big\|H(x_a) - H(y_a)\big\|.$$

Therefore, by sub-multiplicativity of the Frobenius norm, we have that

$$\begin{aligned}
&\big\|\big(H(x_a)^{-1} - H(y_a)^{-1}\big)\star MM^T(x_a)\big\| \\
&\leq \big\|H(x_a)^{-1} - H(y_a)^{-1}\big\|\big\|MM^T(x_a)\big\|\big\|(H(x_a)^{-1} - H(y_a)^{-1})^T\big\| \\
&\leq 2\big\|H(x_a) - H(y_a)\big\|\big\|MM^T(x_a)\big\| \\
&\ll \big\|H(x_a) - H(y_a)\big\|\tau(a)^2\|v\|_\eta^2\|h\|_\eta^2, \qquad (5.2.10)
\end{aligned}$$

where the final estimate uses (5.2.6). Moreover, from Proposition 4.4.2 (i), we have

$$\big\|H(x_a) - H(y_a)\big\| \ll \tau(a)\|h\|_\eta\, d_\gamma(x_a, y_a).$$

Therefore

$$\begin{aligned} II &\ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)^3\|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x_a, y_a) \\ &\le \sum_{a\in\alpha} \mu_Y(a)\tau(a)^3\|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x, y) \ll \|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x, y), \end{aligned}$$

where the second inequality uses Proposition 3.3.3 and the final inequality uses that $p \ge 3$. For III, we have from Proposition 5.2.1 and (5.2.5) that

$$\begin{aligned} &\big\|MM^T(x_a) - MM^T(y_a)\big\| \\ &\le \big\|M(x_a)\big(M^T(x_a) - M^T(y_a)\big)\big\| + \big\|\big(M(x_a) - M(y_a)\big)M^T(y_a)\big\| \\ &\le \big(\big|M(x_a)\big| + \big|M(y_a)\big|\big)\big|M(x_a) - M(y_a)\big| \\ &\ll \tau(a)^3\|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x_a, y_a) \\ &\le \tau(a)^3\|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x, y), \end{aligned}$$

where the final inequality uses Proposition 3.3.3. Combining this with Proposition 5.2.2 gives

$$\begin{aligned} III &\ll \sum_{a\in\alpha} \mu_Y(a)\big\|MM^T(x_a) - MM^T(y_a)\big\| \\ &\ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)^3\|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x, y) \\ &\ll \|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x, y). \end{aligned}$$

Therefore $|\mathcal{P}_H(MM^T)|_\gamma \ll \|v\|_\eta^2\|h\|_\eta^2$. To complete the proof of (i), we analyse $|\mathcal{P}_H(\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T\, \mathrm{d}m_\Delta)|_\gamma$. Since the arguments are similar to the above, we omit some details. For $x, y \in Y$, we have that

$$\begin{aligned}
&\Big\|\mathcal{P}_H\Big(\tau\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big)(x)-\mathcal{P}_H\Big(\tau\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big)(y)\Big\|\\
&\le\sum_{a\in\alpha}\big|\zeta(x_a)-\zeta(y_a)\big|\Big\|\tau(a)\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big\|\\
&\quad+\sum_{a\in\alpha}\zeta(y_a)\big\|H(x_a)-H(y_a)\big\|\Big\|\tau(a)\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big\|\\
&\quad+\sum_{a\in\alpha}\zeta(y_a)\Big\|\tau(a)\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta-\tau(a)\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta\Big\| =: I+II+III.
\end{aligned}\tag{5.2.11}$$

As before, we use Proposition 3.3.9 to deal with the terms involving ζ. Now, $I \ll \|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x,y)$ by (5.2.8). Similarly, $II \ll \|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x,y)$, where we use Proposition 4.4.2 (i), (5.2.8), and $p \ge 2$. Finally, $III = 0$ is immediate. Therefore $|\mathcal{P}_H(\tau\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta)|_\gamma \ll \|v\|_\eta^2\|h\|_\eta^2$ as claimed.

Proof of (ii). The estimate for the sup norm remains unchanged. As before, $|\mathcal{P}_H\tilde V|_{\gamma^{p-2}} \le |\mathcal{P}_H(MM^T)|_{\gamma^{p-2}} + |\mathcal{P}_H(\tau\int_{\Delta\times G}\hat\psi\hat\psi^T\,\mathrm{d}m_\Delta)|_{\gamma^{p-2}}$. We begin by analysing (5.2.9). We deal with the terms involving ζ by using Proposition 3.3.9. Proceeding as in (i) for I gives

$$I \ll \|v\|_\eta^2\|h\|_\eta^2\, d_\gamma(x,y) \le \|v\|_\eta^2\|h\|_\eta^2\, d_{\gamma^{p-2}}(x,y).$$

We now analyse II. From Proposition 4.4.2 (i), we have

$$\big\|H(x_a)-H(y_a)\big\| \ll \tau(a)^{p-2}\|h\|_\eta\, d_{\gamma^{p-2}}(x_a,y_a).$$

Combining this with (5.2.10) gives

$$\begin{aligned}
II &\ll \sum_{a\in\alpha}\mu_Y(a)\tau(a)^p\|v\|_\eta^2\|h\|_\eta^2\, d_{\gamma^{p-2}}(x_a,y_a)\\
&\le \sum_{a\in\alpha}\mu_Y(a)\tau(a)^p\|v\|_\eta^2\|h\|_\eta^2\, d_{\gamma^{p-2}}(x,y)\\
&\ll \|v\|_\eta^2\|h\|_\eta^2\, d_{\gamma^{p-2}}(x,y).
\end{aligned}$$

For III, note that

$$\begin{aligned}\big\|MM^T(x_a) - MM^T(y_a)\big\| &\le \big(\big|M(x_a)\big| + \big|M(y_a)\big|\big)\big|M(x_a) - M(y_a)\big| \\ &\ll \tau(a)^p \|v\|_\eta^2 \|h\|_\eta^2 \, d_{\gamma^{p-2}}(x_a, y_a) \\ &\le \tau(a)^p \|v\|_\eta^2 \|h\|_\eta^2 \, d_{\gamma^{p-2}}(x, y),\end{aligned}$$

where the second estimate uses Proposition 5.2.1 and (5.2.5). Therefore

$$\begin{aligned}III &\ll \sum_{a\in\alpha} \mu_Y(a)\big\|MM^T(x_a) - MM^T(y_a)\big\| \\ &\ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)^p \|v\|_\eta^2 \|h\|_\eta^2 \, d_{\gamma^{p-2}}(x, y) \\ &\ll \|v\|_\eta^2 \|h\|_\eta^2 \, d_{\gamma^{p-2}}(x, y).\end{aligned}$$

Combining the above gives us $|\mathcal{P}_H(MM^T)|_{\gamma^{p-2}} \ll \|v\|_\eta^2\|h\|_\eta^2$. To conclude the proof, we note that $|\mathcal{P}_H(\tau \int_{\Delta\times G} \hat{\psi}\hat{\psi}^T \, \mathrm{d}m_\Delta)|_{\gamma^{p-2}} \ll \|v\|_\eta^2\|h\|_\eta^2$ by an identical argument to that in (i). ☐

From Remark 4.8.7, Proposition 5.2.5, and Proposition 5.2.6, we can write

$$\tilde{\phi} = \tilde{\psi} + \tilde{\chi} \circ f_H - \tilde{\chi} \quad \text{where } \tilde{\psi} \in \ker \hat{\mathcal{L}} \tag{5.2.12}$$

for some $\tilde{\psi}, \tilde{\chi} \colon \Delta \times G \to \mathbb{R}^{d,d}$, with the main results of Chapter 4 going through. We refer to (5.2.12) as a *secondary martingale-coboundary decomposition*. To conclude the section, we state the results which we explicitly require. The first result is analogous to Proposition 4.6.2 (i).

Proposition 5.2.7. *Let $\tilde{\psi}$ be as in the secondary martingale-coboundary decomposition* (5.2.12)*. Then $\tilde{\psi} \in L^p(\Delta \times G; \mathbb{R}^{d,d})$ with $|\tilde{\psi}|_p \le C\|v\|_\eta^2\|h\|_\eta^2$.*

The next result is analogous to Proposition 4.6.4.

Proposition 5.2.8. *Let $\tilde{\chi}$ be as in the secondary martingale-coboundary decomposition* (5.2.12)*. Then $\max_{0\le k\le n} \|\tilde{\chi} \circ f_H^k\| = o(n^{1/p})$ almost surely.*

The next result is analogous to Theorem 4.1.5 (ii).

Proposition 5.2.9. *Let $\tilde{\phi}$ be as in* (5.2.1)*. Then*

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1}\tilde{\phi}\circ f_H^k\right|\right|_{2(p-1)} \le Cn^{1/2}\|v\|_\eta^2\|h\|_\eta^2 \quad \textit{for all } n\ge 1.$$

5.3 Almost sure invariance principle and consequences

We begin this section by showing that Theorem 5.1.1 holds for the one-dimensional projections of the observable $\hat{\psi}\colon \Delta\times G\to\mathbb{R}^{d,d}$ from the primary martingale-coboundary decomposition. For this, we require the classical law of the iterated logarithm [54, 57], which states that if $(Z_n)_{n\ge1}$ is a sequence of independent and identically distributed random variables with mean 0 and variance $\sigma^2>0$, then

$$\limsup_{n\to\infty}\frac{\sum_{k=1}^n Z_k}{\sqrt{2n\log\log n}} = \sigma^2 \quad \text{a.s.} \tag{5.3.1}$$

Lemma 5.3.1. *Suppose $\tau\in L^p(Y)$ for some $p\ge 2$. Define $\Sigma\in\mathbb{R}^{d,d}$ by*

$$\Sigma = \lim_{n\to\infty}\int_{X\times G}\left(\sum_{k=0}^{n-1}\phi\circ T_h^k\right)\left(\sum_{k=0}^{n-1}\phi\circ T_h^k\right)^T \mathrm{d}m$$

and let $c\in\mathbb{R}^d$ with $c^T\Sigma c>0$. Then there exists a probability space supporting a sequence of random variables $(S_n')_{n\ge1}$ with the same joint distributions as $(\sum_{k=0}^{n-1}(c\cdot\hat{\psi})\circ f_H^k)_{n\ge1}$ and a sequence of independent and identically distributed random variables $(Z_n)_{n\ge1}$ with distribution $\mathcal{N}(0,c^T\Sigma c)$, such that almost surely,

$$\sup_{1\le k\le n}\left|S_k'-\sum_{j=1}^k Z_j\right| = \begin{cases} o\big((n\log\log n)^{1/2}\big) & \textit{if } p=2,\\ o(n^{1/p}(\log n)^{1/2}) & \textit{if } p\in(2,4),\\ O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}) & \textit{if } p\ge 4.\end{cases}$$

Proof. First note by Corollary 4.8.5 that $\Sigma=\int_{\Delta\times G}\hat{\psi}\hat{\psi}^T\,\mathrm{d}m_\Delta$. Observe that since $\hat{\psi}\in L^p(\Delta\times G;\mathbb{R}^d)$, we have $c\cdot\hat{\psi}\in L^p(\Delta\times G;\mathbb{R})$. Now,

$$(c\cdot\hat{\psi})^2 = (c\cdot\hat{\psi})(c\cdot\hat{\psi}) = (c\cdot\hat{\psi})(\hat{\psi}\cdot c) = c^T\hat{\psi}\hat{\psi}^T c. \tag{5.3.2}$$

Therefore
$$\int_{\Delta\times G}(c\cdot\hat{\psi})^2\,\mathrm{d}m_\Delta = c^T\Sigma c.$$
Note that since $\hat{\psi}\in\ker\hat{\mathcal{L}}$, we have $c\cdot\hat{\psi}\in\ker\hat{\mathcal{L}}$. The almost sure invariance principle (ASIP) with the desired rate when $p=2$ is immediate from Theorem 2.8.9.

Let $\mathcal{B}$ denote the underlying $\sigma-$algebra on $\Delta\times G$ and $\mathcal{B}_k = f_H^{-k}(\mathcal{B})$. For $p>2$, we require almost sure estimates for
$$A_n = \sum_{k=0}^{n-1}\Big(\mathbb{E}\big[(c\cdot\hat{\psi})^2\circ f_H^k \,\big|\, \mathcal{B}_{k+1}\big] - c^T\Sigma c\Big).$$
Note that for $n\geq 1$ and $0\leq k\leq n$, we have
$$\begin{aligned}\mathbb{E}\big[(c\cdot\hat{\psi})^2\circ f_H^k\,\big|\,\mathcal{B}_{k+1}\big] &= \mathbb{E}\big[(c\cdot\hat{\psi})^2\,\big|\,\mathcal{B}_1\big]\circ f_H^k\\ &= \hat{\mathcal{U}}\hat{\mathcal{L}}\big((c\cdot\hat{\psi})^2\big)\circ f_H^k\\ &= \big(c^T\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T)c\big)\circ f_H^k,\end{aligned}$$
where we use Proposition 2.4.2 (vi), Proposition 2.5.3 (vii), and (5.3.2) in the first, second, and third equalities respectively. Therefore, using the secondary martingale-coboundary decomposition (5.2.12), we have
$$\begin{aligned}A_n &= \sum_{k=0}^{n-1}\Big(\big(c^T\hat{\mathcal{U}}\hat{\mathcal{L}}(\hat{\psi}\hat{\psi}^T)c\big)\circ f_H^k - c^T\Sigma c\Big) = \sum_{k=0}^{n-1}(c^T\tilde{\phi}c)\circ f_H^k\\ &= \sum_{k=0}^{n-1}\big(c^T(\tilde{\psi}+\tilde{\chi}\circ f_H-\tilde{\chi})c\big)\circ f_H^k\\ &= c^T(\tilde{\chi}\circ f_H^n-\tilde{\chi})c + \sum_{k=0}^{n-1}(c^T\tilde{\psi}c)\circ f_H^k.\end{aligned}$$
Now, we have from Proposition 5.2.8 that almost surely, $\|\tilde{\chi}\circ f_H^n-\tilde{\chi}\| = o(n^{1/p})\subset o((n\log\log n)^{1/2})$. It follows that
$$c^T(\tilde{\chi}\circ f_H^n-\tilde{\chi})c = o\big((n\log\log n)^{1/2}\big)\quad\text{a.s.}\tag{5.3.3}$$
Next, note from Proposition 5.2.7 we have $\tilde{\psi}\in L^2(\Delta\times G;\mathbb{R}^{d,d})$, so that $c^T\tilde{\psi}c\in L^2(\Delta\times G;\mathbb{R})$. Moreover, since $\tilde{\psi}\in\ker\hat{\mathcal{L}}$, we have $c^T\tilde{\psi}c\in\ker\hat{\mathcal{L}}$. Applying Theorem 2.8.9 gives us an ASIP for $(\sum_{k=0}^{n-1}(c^T\tilde{\psi}c)\circ f_H^k)_{n\geq 1}$ with error rate

$o((n \log\log n)^{1/2})$. Therefore, on a possibly enlarged probability space, it follows from the classical law of the iterated logarithm (5.3.1) that almost surely,

$$\begin{aligned}\sum_{k=0}^{n-1}(c^T\tilde{\psi}c)\circ f_H^k &= \left(\sum_{k=0}^{n-1}(c^T\tilde{\psi}c)\circ f_H^k - \sum_{k=1}^{n} Z_k\right) + \sum_{k=1}^{n} Z_k \\ &= O\big((n\log\log n)^{1/2}\big). \end{aligned} \tag{5.3.4}$$

If necessary, by a coupling argument [24], we can redefine $\tilde{\chi}$ and $\tilde{\psi}$ on a further enlarged probability space such that (5.3.3) and (5.3.4) hold. In particular, we have that

$$A_n = O\big((n\log\log n)^{1/2}\big) \quad \text{a.s.}$$

For $p \geq 4$, the ASIP with the desired rate now follows from Theorem 2.8.11. Suppose $p \in (2,4)$. We have that

$$\frac{A_n}{n^{2/p}} \ll \frac{(n\log\log n)^{1/2}}{n^{2/p}} = n^{1/2-2/p}(\log\log n)^{1/2} \to 0 \quad \text{a.s.}$$

since $\frac{1}{2} - \frac{2}{p} < 0$. Therefore $A_n = o(n^{2/p})$ almost surely, and the result now follows from Theorem 2.8.10. □

Proof of Theorem 5.1.1. Note that from Proposition 4.6.4, we have

$$\max_{1\leq k\leq n}\left|\sum_{j=0}^{k-1}\big((c\cdot\hat{\phi})\circ f_H^j - (c\cdot\hat{\psi})\circ f_H^j\big)\right| = \max_{1\leq k\leq n}\left|c\cdot(\hat{\chi}\circ f_H^k - \hat{\chi})\right| = o(n^{1/p}) \quad \text{a.s.}$$

In view of Lemma 5.3.1, we can enlarge the probability space as in [84, p. 23] (see also [7, Lemma A.1]) to support sequences of random variables $(S_n)_{n\geq1}$ and $(S'_n)_{n\geq1}$ with the same joint distributions as $(\sum_{k=0}^{n-1}(c\cdot\hat{\phi})\circ f_H^k)_{n\geq1}$ and $(\sum_{k=0}^{n-1}(c\cdot\hat{\psi})\circ f_H^k)_{n\geq1}$ respectively, such that $(S'_n)_{n\geq1}$ satisfies the desired estimates and

$$\max_{1\leq k\leq n}|S_k - S'_k| = o(n^{1/p}) \quad \text{a.s.}$$

Now, note that

$$n^{1/p} = \begin{cases} o\big((n\log\log n)^{1/2}\big) & \text{if } p = 2, \\ o(n^{1/p}(\log n)^{1/2}) & \text{if } p \in (2,4), \\ O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}) & \text{if } p \geq 4. \end{cases}$$

Therefore, since

$$\max_{1\le k\le n}\left|S_k-\sum_{j=1}^{k}Z_k\right|\le \max_{1\le k\le n}|S_k-S_k'|+\max_{1\le k\le n}\left|S_k'-\sum_{j=1}^{k}Z_k\right|,$$

we have that $(S_n)_{n\ge1}$ satisfies the desired estimates. Finally, noting that π_H is a semi-conjugacy, it follows from Proposition 4.2.3 that the joint distributions of $(\sum_{k=0}^{n-1}(c\cdot\phi)\circ T_h^k)_{n\ge1}$ coincide with those of $(S_n)_{n\ge1}$. This completes the proof. □

To conclude the chapter, we give some corollaries of Theorem 5.1.1. We start with the law of the iterated logarithm.

Corollary 5.3.2. *In the setting of Theorem 5.1.1, we have*

$$\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\log\log n}}=c^T\Sigma c \qquad a.s.$$

Proof. Since $\tau\in L^p(Y)\subset L^2(Y)$ for $p\ge2$, it suffices to prove the result for $p=2$. Note that if $(x_n)\subset\mathbb{R}$ is such that $\limsup_{n\to\infty}x_n<\infty$ and $(y_n)\subset\mathbb{R}$ is convergent, then

$$\limsup_{n\to\infty}(x_n+y_n)=\limsup_{n\to\infty}x_n+\lim_{n\to\infty}y_n.$$

Therefore, we have from the classical law of the iterated logarithm (5.3.1) and Theorem 5.1.1 that almost surely,

$$\begin{aligned}\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\log\log n}}&=\limsup_{n\to\infty}\frac{\sum_{k=1}^{n}Z_k}{\sqrt{2n\log\log n}}+\lim_{n\to\infty}\frac{S_n-\sum_{k=1}^{n}Z_k}{\sqrt{2n\log\log n}}\\&=c^T\Sigma c,\end{aligned}$$

completing the proof. □

We next move on to the functional law of the iterated logarithm. Let $C[0,1]$ denote the Banach space of real-valued continuous functions with supremum norm. Recall that a function $f\colon[0,1]\to\mathbb{R}$ is absolutely continuous if, with respect to the Lebesgue measure, the derivative f' exists almost everywhere, is integrable, and satisfies

$$f(x) = f(0) + \int_0^x f'(t)\,\mathrm{d}t$$

for all $x \in [0,1]$.

Corollary 5.3.3. *Let K be the set of all real-valued absolutely continuous functions $f\colon [0,1] \to \mathbb{R}$ such that $f(0) = 0$ and $\int_0^1 |f'(t)|^2\,\mathrm{d}t \le 1$. Let $c \in \mathbb{R}^d$ and $(S_n)_{n\ge1}$ be as in Theorem 5.1.1. For $n \ge 3$ and $t \in [0,1]$, define the random elements $f_n \in C[0,1]$ by*

$$f_n(t) = \frac{\big(1 - nt + [nt]\big)S_{[nt]} + \big(nt - [nt]\big)S_{[nt]+1}}{\sqrt{2n\log\log n}}.$$

Then almost surely, $(f_n)_{n\ge3}$ is relatively compact in $C[0,1]$ and its set of limit points is precisely K.

Proof. We argue as in [84, Theorem C]. Note that uniformly in t, we have

$$f_n(t) = \frac{S_{[nt]}}{\sqrt{2n\log\log n}} + O\bigg(\frac{1}{\sqrt{2n\log\log n}}\bigg).$$

Let us write

$$B_n(t) = \frac{\sum_{k=1}^{[nt]} Z_k}{\sqrt{2n\log\log n}}.$$

It follows from Theorem 5.1.1 that almost surely, $|f_n(t) - B_n(t)| \to 0$ uniformly in t. That is, almost surely,

$$|f_n - B_n|_\infty \to 0. \tag{5.3.5}$$

Now, [93, Theorem 1] says that the result holds for the sequence $(B_n)_{n\ge3}$. From (5.3.5), it is immediate that the limit points of $(B_n)_{n\ge3}$ and $(f_n)_{n\ge3}$ coincide, so that the set of limit points of $(f_n)_{n\ge3}$ is precisely K. Similarly, relative compactness of $(f_n)_{n\ge3}$ follows from that of $(B_n)_{n\ge3}$. □

Remark 5.3.4. *As with the central limit theorem and weak invariance principle, we have Corollary 5.3.2 follows from Corollary 5.3.3 (see for example [93, Section*

3]). From Theorem 5.1.1, one can also deduce upper and lower class refinements of the law of the iterated logarithm. See [84, Chapter 1] for statements of these and further consequences.

Chapter 6

Generalisation to sequences of compact group extensions

6.1 Outline

Let (X,d) be a bounded metric space and $T\colon X \to X$ be non-uniformly expanding as in Section 3.2. Let μ denote the ergodic invariant Borel probability measure on X constructed in Section 3.4. We suppose throughout this chapter that the return time τ has integrability $p > 2$. Let G be a compact connected Lie group with Haar measure ν, and suppose that $(\pi, \mathbb{R}^d)$ is a representation of G for some $d \geq 1$. As in Remark 2.2.7, we fix a $G-$invariant inner product $[\cdot,\cdot]$ on $\mathbb{R}^d$ and view G as a closed subgroup of $\mathrm{O}(d)$. We consider the sequence of compact group extensions $T_{h^{(n)}}\colon X \times G \to X \times G$ defined by $T_{h^{(n)}}(x,g) = (Tx, gh^{(n)}(x))$, where $h^{(n)}\colon X \to G$ are $\eta-$Hölder and satisfy $\sup_{n\geq 1} \|h^{(n)}\|_\eta < \infty$. For all $n \geq 1$ the probability measure $m = \mu \times \nu$ is $T_{h^{(n)}}-$invariant and assumed to be ergodic as in Remark 4.1.1. We make the following additional assumptions on the cocycles:

Assumption 6.1.1. *There exist $h^{(\infty)} \in C^\eta(X;G)$ and $C > 0$ such that*

$$|h^{(n)} - h^{(\infty)}|_\infty \leq \frac{C}{n} \quad \textit{for all } n \geq 1. \tag{6.1.1}$$

Moreover, the induced compact group extension $F_{H^{(\infty)}}$ is assumed to be mixing.

Remark 6.1.2. *Mixing is typical in the following sense, as in [35, Theorem 1.5 and Remark 1.6]. The set of Hölder cocycles* $h^{(\infty)}\colon X \to G$ *for which* $F_{H^{(\infty)}}$ *is not mixing lies in a closed subspace of infinite codimension in the space of all Hölder cocycles.*

We consider the sequence of equivariant observables $\phi_n\colon X \times G \to \mathbb{R}^d$ of the form $\phi_n = g \cdot v_n$, where $\int_{X\times G} \phi_n \, \mathrm{d}m = 0$ for all $n \geq 1$ and $v_n\colon X \to \mathbb{R}^d$ satisfies $\sup_{n\geq 1} \|v_n\|_\eta < \infty$. In this chapter, we show how the martingale-coboundary decompositions in Chapter 4 and Chapter 5 apply to Birkhoff sums of the form $\sum_{k=0}^{n-1} \phi_n \circ T^k_{h^{(n)}}$. To state our main results, we make some definitions. For $n \geq 1$, define

$$\Sigma_n = \lim_{k\to\infty} \frac{1}{k} \int_{X\times G} \left(\sum_{j=0}^{k-1} \phi_n \circ T^j_{h^{(n)}} \right) \left(\sum_{j=0}^{k-1} \phi_n \circ T^j_{h^{(n)}} \right)^T \mathrm{d}m = \int_{\Delta\times G} \hat{\psi}_n \hat{\psi}_n^T \, \mathrm{d}m_\Delta. \tag{6.1.2}$$

We also define the random elements $W_n\colon X \times G \to D([0,\infty);\mathbb{R}^d)$ by $W_n(t) = n^{-1/2} \sum_{k=0}^{[nt]-1} \phi_n \circ T^k_{h^{(n)}}$ for $t \geq 0$.

Theorem 6.1.3. *Suppose* $\lim_{n\to\infty} \Sigma_n = \Sigma$ *for some* $\Sigma \in \mathbb{R}^{d,d}$*. Then* $g\Sigma = \Sigma g$ *for all* $g \in G$ *and*

(i) $n^{-1/2} \sum_{k=0}^{n-1} \phi_n \circ T^k_{h^{(n)}} \to_w \mathcal{N}(0,\Sigma)$.

(ii) $W_n \to_w W$ *in* $D([0,\infty);\mathbb{R}^d)$*, where* W *is a Brownian motion with mean* 0 *and covariance matrix* Σ.

Let $\mathcal{W} \subset D([0,\infty);\mathbb{R}^d))$ denote the set of weak subsequential limits of $(W_n)_{n\geq 1}$ and $\mathcal{S} \subset \mathbb{R}^{d,d}$ be the set of limit points of $(\Sigma_n)_{n\geq 1}$.

Theorem 6.1.4. *The following hold true:*

(i) $(W_n)_{n\geq 1}$ *is tight.*

(ii) $W \in \mathcal{W}$ *if and only if* W *is a Brownian motion with mean* 0 *and covariance matrix* $\Sigma \in \mathcal{S}$.

The next result says that any weakly convergent subsequence of $(W_n)_{n\geq 1}$ has corresponding convergence of moments.

Theorem 6.1.5. *Let $W \in \mathcal{W}$ and let $(W_{n_k})_{k\geq 1}$ be such that $W_{n_k} \to_w W$. Then*

$$\lim_{k\to\infty} n_k^{-q/2} \int_{X\times G} \left| \sum_{j=0}^{n_k-1} \phi_{n_k} \circ T^j_{h^{(n_k)}} \right|^q \mathrm{d}m = \mathbb{E}\big[\big|W(1)\big|^q\big] \quad \textit{for all } 0 < q < 2(p-1),$$

where $\mathbb{E}$ denotes the expectation with respect to the underlying probability space on which W is defined.

The structure of the chapter is as follows: In Section 6.2, we verify uniformity of the constants which arise as a result of the primary and secondary martingale-coboundary decompositions from Chapter 4 and Chapter 5 respectively. For this, it suffices to verify uniformity on the inducing set Y, for then orthogonality allows us to deduce uniformity on $Y \times G$, and then by lifting we obtain uniformity on $\Delta \times G$. In Section 6.3, we prove the main results above. Finally, in Section 6.4, we give a homogenisation result of discrete fast-slow dynamical systems with additive noise, where the fast dynamics are generated by a family of compact group extensions with non-uniformly expanding base.

6.2 Uniformity for the primary and secondary decompositions

Since $p > 2$, we choose $\gamma \in [\lambda^{-\eta}, 1)$ as in Remark 4.4.4. For each $n \geq 1$, we define the induced cocycles $H^{(n)} \colon Y \to G$ by $H^{(n)} = h_\tau^{(n)}$ and induced functions $V_n \colon Y \to \mathbb{R}^d$ by

$$V_n(y) = \sum_{\ell=0}^{\tau(y)-1} h_\ell^{(n)}(y) \cdot v_n(T^\ell y), \tag{6.2.1}$$

where $h_\ell^{(n)} = h^{(n)}\, h^{(n)} \circ T \cdots h^{(n)} \circ T^{\ell-1}$. To simplify the results in this section, we let $C > 0$ denote various constants which are independent of $n \geq 1$.

Proposition 6.2.1. *Let $\epsilon \in (0,1]$. The following hold true:*

(i) $\|\mathbb{1}_a H^{(n)}\|_{\gamma^\epsilon} \le C\tau(a)^\epsilon \|h^{(n)}\|_\eta$ *for all* $a \in \alpha$ *and* $n \ge 1$.

(ii) $\|\mathbb{1}_a V_n\|_{\gamma^\epsilon} \le C\tau(a)^{1+\epsilon} \|v_n\|_\eta \|h^{(n)}\|_\eta$ *for all* $a \in \alpha$ *and* $n \ge 1$.

Proof. This follows directly from Proposition 4.4.2, since the given constants depend only on the properties of the underlying dynamical system T. □

The next statement is immediate from Proposition 4.4.3 and Proposition 6.2.1.

Proposition 6.2.2. *The twisted transfer operators* $\mathcal{P}_{H^{(n)}} \colon L^1(Y;\mathbb{R}^d) \to L^1(Y;\mathbb{R}^d)$ *as defined in Section 4.3 satisfy* $\|\mathcal{P}_{H^{(n)}} V_n\|_\gamma \le C\|v_n\|_\eta \|h^{(n)}\|_\eta$ *for all* $n \ge 1$.

In view of Section 4.5, we require some control on the spectra of $\mathcal{P}_{H^{(n)}}$ when viewed as operators on the space $F_{\gamma,0}(Y;\mathbb{R}^d) = \{V \in F_\gamma(Y;\mathbb{R}^d) \mid \int_{Y\times G} g\cdot V \, \mathrm{d}m_Y = 0\}$. We first remove the possibility of eigenvalues lying on the unit circle for $P_{H^{(\infty)}}$.

Lemma 6.2.3. *The twisted transfer operator* $\mathcal{P}_{H^{(\infty)}} \colon F_{\gamma,0}(Y;\mathbb{R}^d) \to F_{\gamma,0}(Y;\mathbb{R}^d)$ *has no eigenvalues on the unit circle.*

Proof. Suppose that $\mathcal{P}_{H^{(\infty)}} V = e^{i\omega} V$ for some $\omega \in [0, 2\pi)$ and $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$. Let $\mathcal{U}_{H^{(\infty)}} \colon L^2(Y;\mathbb{R}^d) \to L^2(Y;\mathbb{R}^d)$ denote the twisted Koopman operator for F with respect to the cocycle $H^{(\infty)}$. By arguing as in (4.5.7), we have $\mathcal{U}_{H^{(\infty)}} V = e^{-i\omega} V$. Define $\Psi \colon Y \times G \to \mathbb{R}^d$ by $\Psi = g \cdot V$. Then

$$\begin{aligned}\Psi \circ F_{H^{(\infty)}} &= gH^{(\infty)} \cdot V \circ F = g \cdot (H^{(\infty)} \cdot V \circ F) = g \cdot \mathcal{U}_{H^{(\infty)}} V = g \cdot e^{-i\omega} V \\ &= e^{-i\omega}\Psi.\end{aligned}$$

Since m_Y is mixing, it is weak mixing, so that Ψ is constant m_Y −almost surely. Since $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$, we have $\int_{Y\times G} \Psi \, \mathrm{d}m_Y = 0$ so that $\Psi = 0$ m_Y −almost surely. Therefore $V = 0$ μ_Y −almost surely. □

We look to appeal to the results of [51]. The next proposition verifies that the sequence of twisted transfer operators $\mathcal{P}_{H^{(n)}}$ satisfies the hypotheses [51, (2)–(5)].

We first define some notation. For any bounded linear operator $\mathcal{Q}\colon F_{\gamma,0}(Y;\mathbb{R}^d) \to F_{\gamma,0}(Y;\mathbb{R}^d)$, let

$$|||\mathcal{Q}||| = \sup\big\{|\mathcal{Q}V|_1 : V \in F_{\gamma,0}(Y;\mathbb{R}^d), \|V\|_\gamma \le 1\big\}.$$

Proposition 6.2.4. *The following hold true:*

(i) $|\mathcal{P}^k_{H^{(n)}}V|_1 \le |V|_1$ *for all* $k, n \ge 1$, *and* $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$.

(ii) $\|\mathcal{P}^k_{H^{(n)}}V\|_\gamma \le C(\gamma^k\|V\|_\gamma + |V|_1)$ *for all* $k, n \ge 1$, *and* $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$.

(iii) *For* $n \ge 1$, *if* $z \in \sigma(\mathcal{P}^k_{H^{(n)}})$ *with* $|z| > \gamma$, *then* z *is not in the residual spectrum*[1] *of* $\mathcal{P}_{H^{(n)}}$.

(iv) $|||\mathcal{P}_{H^{(n)}} - \mathcal{P}_{H^{(\infty)}}||| \le \frac{C}{n}$ *for all* $n \ge 1$.

Proof. We have that (i) follows directly from Proposition 4.5.1 (i). For (ii), we combine Proposition 4.5.1 (ii) with the fact that $\sup_{n\ge1}\|h^{(n)}\|_\eta < \infty$. For (iii), note that by the proof of Corollary 4.5.2, we have $r_{\mathrm{ess}}(\mathcal{P}_{H^{(n)}}) < \gamma$ for all $n \ge 1$. Therefore if $z \in \sigma(\mathcal{P}_{H^{(n)}})$ with $|z| > \gamma$, then z is an isolated eigenvalue of finite multiplicity. In particular, z does not lie in the residual spectrum of $\mathcal{P}_{H^{(n)}}$. It remains to verify (iv). Let $V \in F_{\gamma,0}(Y;\mathbb{R}^d)$ with $\|V\|_\gamma \le 1$. Let $y \in Y$ and $y_a \in a$ be the unique element of $a \in \alpha$ for which $Fy_a = y$. Note that by a similar calculation to (4.4.6), we have

$$\big|\big(H^{(n)}(y_a)^{-1} - H^{(\infty)}(y_a)^{-1}\big)\cdot V(y_a)\big| \le \big\|H^{(n)}(y_a) - H^{(\infty)}(y_a)\big\||V|_\infty.$$

Moreover, by repeating the argument of the proof of (4.4.1), we have

$$\big\|h^{(n)}_{\tau(a)}(y_a) - h^{(\infty)}_{\tau(a)}(y_a)\big\| \le \sum_{k=0}^{\tau(a)-1}\big\|h^{(n)}(T^k y_a) - h^{(\infty)}(T^k y_a)\big\|.$$

[1] For our purposes, we only require that the set of eigenvalues is disjoint from the residual spectrum.

In particular,

$$\begin{aligned}\big\|H^{(n)}(y_a) - H^{(\infty)}(y_a)\big\| &= \big\|h^{(n)}_{\tau(a)}(y_a) - h^{(\infty)}_{\tau(a)}(y_a)\big\| \\ &\le \sum_{k=0}^{\tau(a)-1} \big\|h^{(n)}(T^k y_a) - h^{(\infty)}(T^k y_a)\big\| \\ &\le \tau(a)|h^{(n)} - h^{(\infty)}|_\infty \ll \frac{\tau(a)}{n},\end{aligned}$$

where the final inequality follows from (6.1.1). It follows from Remark 4.3.2 and Proposition 3.3.9 that

$$\begin{aligned}\big|(\mathcal{P}_{H^{(n)}} - \mathcal{P}_{H^{(\infty)}})(V)(y)\big| &\le \sum_{a\in\alpha} \zeta_n(y_a)\big|\big(H^{(n)}(y_a)^{-1} - H^{(\infty)}(y_a)^{-1}\big)\cdot V(y_a)\big| \\ &\ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)|h^{(n)} - h^{(\infty)}|_\infty |V|_\infty \ll \frac{|\tau|_1}{n}.\end{aligned}$$

Therefore $|(\mathcal{P}_{H^{(n)}} - \mathcal{P}_{H^{(\infty)}})(V)|_1 \ll 1/n$, so that $\|\!|\mathcal{P}_{H^{(n)}} - \mathcal{P}_{H^{(\infty)}}\|\!| \ll 1/n$ as required. □

We now show how the results in [51] imply uniform exponential contraction of the operators $\mathcal{P}_{H^{(n)}}$ for n sufficiently large.

Proposition 6.2.5. *There exists $R \in (\gamma, 1)$ and $N \ge 1$ such that for all $n \ge N$, we have $\|\mathcal{P}^k_{H^{(n)}}\| \le CR^k$ for all $k \ge 1$.*

Proof. We first note that since $r_{\mathrm{ess}}(P_{H^{(\infty)}}) < \gamma$ and $P_{H^{(\infty)}}$ has no eigenvalues on the unit circle, there exists $R \in (\gamma, 1)$ such that $\sigma(\mathcal{P}_{H^{(\infty)}}) \subset B_R(0)$. From [51, Theorem 1], there exists $N \ge 1$ sufficiently large such that $\sigma(\mathcal{P}_{H^{(n)}}) \subset B_R(0)$ for all $n \ge N$. Applying [51, Corollary 2 (ii)] completes the proof. □

Recall $\mathcal{U}_{H^{(n)}} \colon L^1(Y;\mathbb{R}^d) \to L^1(Y;\mathbb{R}^d)$ denotes the twisted Koopman operator of F with respect to the cocycle $H^{(n)}$. Note that if $V_n \colon Y \to \mathbb{R}^d$ is defined as in (6.2.1), we have from Proposition 4.6.1 that there exist $J_n \in F_\gamma(Y;\mathbb{R}^d)$ and $M_n \in L^p(Y;\mathbb{R}^d)$ such that $V_n = M_n + \mathcal{U}_{H^{(n)}} J_n - J_n$ with $M_n \in \ker \mathcal{P}_{H^{(n)}}$. In addition, Proposition 4.6.1 also gives the existence of constants $C(n) > 0$ depending on n

such that $\|J_n\|_\gamma \le C(n)\|v_n\|_\eta\|h^{(n)}\|_\eta$ and $|M_n|_p \le C(n)\|v_n\|_\eta\|h^{(n)}\|_\eta$. We now use Proposition 6.2.5 to deduce uniformity.

Proposition 6.2.6. *The following hold true:*

(i) $\|J_n\|_\gamma \le C\|v_n\|_\eta\|h^{(n)}\|_\eta$ *for all* $n \ge 1$.

(ii) $|M_n|_p \le C\|v_n\|_\eta\|h^{(n)}\|_\eta$ *for all* $n \ge 1$.

Proof. Let $N \ge 1$ be as in Proposition 6.2.5. By Proposition 4.6.1, there exists $C' > 0$ sufficiently large such that $\|J_n\|_\gamma \le C'\|v_n\|_\eta\|h^{(n)}\|_\eta$ and $|M_n|_p \le C'\|v_n\|_\eta\|h^{(n)}\|_\eta$ for $1 \le n \le N-1$. Therefore, it suffices to prove uniformity for $n \ge N$. For such n, we have from the proof of Proposition 6.2.5 that $\sigma(\mathcal{P}_{H^{(n)}}) \subset B_R(0)$. Therefore, it follows from the proof of Proposition 4.6.1 that

$$J_n = \sum_{k=1}^{\infty} \mathcal{P}_{H^{(n)}}^k V_n \in F_\gamma(Y;\mathbb{R}^d).$$

From Proposition 6.2.5, we have

$$\|J_n\|_\gamma \le \sum_{k=1}^{\infty} \|\mathcal{P}_{H^{(n)}}^k V_n\|_\gamma \ll \sum_{k=0}^{\infty} R^k \|v_n\|_\eta\|h^{(n)}\|_\eta \ll \|v_n\|_\eta\|h^{(n)}\|_\eta,$$

proving (i). For (ii), we note that $|M_n|_p \ll \|v_n\|_\eta\|h^{(n)}\|_\eta$ by an identical calculation to (4.6.5), where we use the uniform estimate for $\|J_n\|_\gamma$ proven above. □

From the previous proposition, uniform versions of the results in Chapter 4 go through. In particular, letting $\pi_{H^{(n)}}$ denote the semi-conjugacies as in Section 4.2 and $\hat{\phi}_n = \phi_n \circ \pi_{H^{(n)}}$ denote the lifted versions of the equivariant observables ϕ_n, we have a sequence of martingale-coboundary decompositions

$$\hat{\phi}_n = \hat{\psi}_n + \hat{\chi}_n \circ f_{H^{(n)}} - \hat{\chi}_n \quad \text{with } \hat{\psi}_n \in \ker \hat{\mathcal{L}}_n, \tag{6.2.2}$$

where $f_{H^{(n)}}$ denotes the tower map for $T_{h^{(n)}}$ and $\hat{\mathcal{L}}_n$ denotes the transfer operator for $f_{H^{(n)}}$. We next state the uniform results we explicitly require. The first result is a uniform version of Proposition 4.6.2 (i).

Proposition 6.2.7. *It holds true that* $|\hat{\psi}_n|_p \le C\|v_n\|_\eta \|h^{(n)}\|_\eta$ *for all* $n \ge 1$.

The next result is a uniform version of (4.8.4).

Proposition 6.2.8. *It holds true that* $|\max_{1\le k\le n} |\hat{\chi}_n \circ f^k_{H^{(n)}} - \hat{\chi}_n||_p = o(n^{1/p})$.

We next state uniform versions of Corollary 4.8.3 (ii) and Theorem 4.1.5 (ii).

Proposition 6.2.9. *It holds true that*

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \hat{\psi}_n \circ f^k_{H^{(n)}}\right|\right|_p \le C n^{1/2}\|v_n\|_\eta \|h^{(n)}\|_\eta \quad \textit{for all } n \ge 1$$

and

$$\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1} \phi_n \circ T^k_{h^{(n)}}\right|\right|_{2(p-1)} \le C n^{1/2}\|v_n\|_\eta \|h^{(n)}\|_\eta \quad \textit{for all } n \ge 1. \tag{6.2.3}$$

By looking at the proof of Corollary 4.8.5, the following result is immediate from the previous two propositions.

Corollary 6.2.10. *We have the convergence*

$$\Sigma_n = \lim_{k\to\infty} \frac{1}{k}\int_{X\times G} \left(\sum_{j=0}^{k-1} \phi_n \circ T^j_{h^{(n)}}\right)\left(\sum_{j=0}^{k-1} \phi_n \circ T^j_{h^{(n)}}\right)^T \mathrm{d}m = \int_{\Delta\times G} \hat{\psi}_n \hat{\psi}_n^T \,\mathrm{d}m_\Delta$$

uniformly in n.

We now switch focus to the secondary martingale-coboundary decomposition. Define the sequence of observables $\tilde{\phi}_n \colon \Delta \times G \to \mathbb{R}^{d,d}$ by

$$\tilde{\phi}_n = \hat{\mathcal{U}}_n \hat{\mathcal{L}}_n(\hat{\psi}_n \hat{\psi}_n^T) - \int_{\Delta\times G} \hat{\psi}_n \hat{\psi}_n^T \,\mathrm{d}m_\Delta, \tag{6.2.4}$$

where $\hat{\mathcal{U}}_n, \hat{\mathcal{L}}_n \colon L^1(\Delta\times G; \mathbb{R}^{d,d}) \to L^1(\Delta\times G; \mathbb{R}^{d,d})$ denote the Koopman and transfer operators for $f_{H^{(n)}}$ respectively. By the results in Section 5.2, we have the secondary martingale-coboundary decomposition for $\tilde{\phi}_n$ given by

$$\tilde{\phi}_n = \tilde{\psi}_n + \tilde{\chi}_n \circ f_{H^{(n)}} - \tilde{\chi}_n \quad \text{with } \tilde{\phi}_n \in \ker \hat{\mathcal{L}}_n.$$

Note that by Remark 4.8.7 and Section 5.2, in order to conclude uniformity of the constants which arise as a result of the secondary martingale-coboundary decomposition, it suffices to verify uniform versions of Proposition 5.2.1, Proposition 5.2.5, and Proposition 5.2.6. By observing the proofs of these results, one can see that this follows directly from the preceding results in this section. We require the following uniform version of Proposition 5.2.9.

Proposition 6.2.11. *Let $\tilde{\phi}_n$ be as in* (6.2.4). *Then*

$$\left| \max_{1\le j\le n} \left| \sum_{k=0}^{j-1} \tilde{\phi}_n \circ f_{H^{(n)}}^k \right| \right|_{2(p-1)} \le C n^{1/2} \|v_n\|_\eta^2 \|h^{(n)}\|_\eta^2 \quad \textit{for all } n \ge 1.$$

6.3 Proofs of the main results

Lemma 6.3.1. *The sequence $(|\hat{\psi}_n|^2)_{n\ge1}$ is uniformly integrable.*

Proof. Let $n \ge 1$ and recall from (4.6.8) that

$$\hat{\psi}_n(y,g,\ell) = \begin{cases} 0 & \text{if } \ell \le \tau(y) - 2, \\ g \cdot M_n(y) & \text{if } \ell = \tau(y) - 1, \end{cases}$$

where M_n is as in Proposition 4.6.1. We begin by showing that $(|M_n|^2)_{n\ge1}$ is uniformly integrable. For this, note by Proposition 6.2.6 (ii) that

$$\begin{aligned} \int_Y \left(|M_n|^2\right)^{p/2} \mathrm{d}\mu_Y = \int_Y |M_n|^p \, \mathrm{d}\mu_Y &\ll \|v_n\|_\eta^p \|h^{(n)}\|_\eta^p \\ &\le \big(\sup_{j\ge1} \|v_j\|_\eta\big)^p \big(\sup_{j\ge1} \|h^{(j)}\|_\eta\big)^p < \infty. \end{aligned}$$

Therefore $(|M_n|^2)_{n\ge1}$ is $L^{p/2}(Y)-$bounded. Since $p > 2$, uniform integrability follows from Proposition 2.3.11.

We next show that this implies uniform integrability of $(|\hat{\psi}_n|^2)_{n\ge1}$. Indeed, fix $\epsilon > 0$ and let $K > 0$ be such that

$$\int_Y |M_n|^2 \mathbb{1}_{\{|M_n|^2 \ge K\}} \, \mathrm{d}\mu_Y \le \epsilon \quad \text{for all } n \ge 1.$$

We have

$$\begin{aligned}\int_{\Delta\times G}|\hat\psi_n|^2 1_{\{|\hat\psi_n|^2\ge K\}}\,\mathrm{d}m_\Delta &= |\tau|_1^{-1}\int_Y |M_n|^2 1_{\{|M_n|^2\ge K\}}\,\mathrm{d}\mu_Y \\ &\le \int_Y |M_n|^2 1_{\{|M_n|^2\ge K\}}\,\mathrm{d}\mu_Y \le \epsilon \quad \text{for all } n\ge 1.\end{aligned}$$

This completes the proof. □

Proof of Theorem 6.1.3. We first note that it suffices to prove (ii), for then (i) follows immediately from the same argument as in Remark 2.3.7. For $n \ge 1$, define the random elements $\hat W_n, \hat M_n \colon \Delta\times G \to D([0,\infty);\mathbb{R}^d)$ by $\hat W_n(t) = n^{-1/2}\sum_{k=0}^{[nt]-1}\hat\phi_n\circ f_{H^{(n)}}^k$ and $\hat M_n(t) = n^{-1/2}\sum_{k=0}^{[nt]-1}\hat\psi_n\circ f_{H^{(n)}}^k$ for $t\ge 0$. Note that since each Σ_n commutes with the action of G on $\mathbb{R}^d$, it follows that Σ commutes with the action of G on $\mathbb{R}^d$. Since $\pi_{H^{(n)}}$ is a semi-conjugacy for all $n\ge 1$, it follows from Proposition 4.2.3 that $(\phi_n\circ T_{h^{(n)}}^k)_{k\ge 0} \sim (\hat\phi_n\circ f_{H^{(n)}}^k)_{k\ge 0}$ for all $n\ge 1$, and so it suffices to show that $\hat W_n \to_w W$ in $D([0,\infty);\mathbb{R}^d)$.

We begin by showing that $\hat M_n \to_w W$ in $D([0,\infty);\mathbb{R}^d)$. First note that each $\hat\psi_n$ lies in $L^2(\Delta\times G;\mathbb{R}^d)$ by Proposition 6.2.7. Moreover, $\hat\psi_n \in \ker\hat{\mathcal{L}}_n$ by (6.2.2). By Lemma 6.3.1, the family $(\hat\psi_n|^2)_{n\ge 1}$ is uniformly integrable. Next, for each $t\ge 0$, we have from (6.2.4) that

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{[nt]-1}\hat{\mathcal{U}}_n\hat{\mathcal{L}}_n(\hat\psi_n\hat\psi_n^T)\circ f_{H^{(n)}}^k - t\Sigma &= \frac{1}{n}\sum_{k=0}^{[nt]-1}(\tilde\phi_n\circ f_{H^{(n)}}^k + \Sigma_n) - t\Sigma \\ &= \frac{1}{n}\sum_{k=0}^{[nt]-1}\tilde\phi_n\circ f_{H^{(n)}}^k + \left(\frac{[nt]}{n}\Sigma_n - t\Sigma\right). \quad (6.3.1)\end{aligned}$$

Now, we have from Proposition 6.2.11 that

$$\begin{aligned}\frac{1}{n}\left|\sum_{k=0}^{n-1}\tilde\phi_n\circ f_{H^{(n)}}^k\right|_{2(p-1)} &\le \frac{1}{n}\left|\max_{1\le j\le n}\left|\sum_{k=0}^{j-1}\tilde\phi_n\circ f_{H^{(n)}}^k\right|\right|_{2(p-1)} \\ &\ll n^{-1/2}\|v_n\|_\eta^2\|h^{(n)}\|_\eta^2 \to 0.\end{aligned}$$

Therefore $n^{-1}\sum_{k=0}^{n-1}\tilde{\phi}_n\circ f_{H^{(n)}}^k \to 0$ in probability. Continuing (6.3.1), we have

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{[nt]-1}\tilde{\phi}_n\circ f_{H^{(n)}}^k+\left(\frac{[nt]}{n}\Sigma_n-t\Sigma\right)&=\frac{[nt]}{n}\frac{1}{[nt]}\sum_{k=0}^{[nt]-1}\tilde{\phi}_n\circ f_{H^{(n)}}^k+\left(\frac{[nt]}{nt}t\Sigma_n-t\Sigma\right)\\&\to 0 \quad \text{in probability.}\end{aligned}$$

Therefore, it follows from Theorem 2.8.7 that $\hat{M}_n\to_w W$ in $D([0,\infty);\mathbb{R}^d)$.

We next show that this implies convergence of $\hat{W}_n$. Observe that for any $T>0$, we have from [92] that $\hat{M}_n\to_w W$ in $D([0,T];\mathbb{R}^d)$. Moreover, note that

$$\begin{aligned}\sup_{t\in[0,T]}\left|\hat{W}_n(t)-\hat{M}_n(t)\right|&=\frac{1}{\sqrt{n}}\sup_{t\in[0,T]}\left|\sum_{k=0}^{[nt]-1}\hat{\chi}_n\circ f_{H^{(n)}}^{k+1}-\hat{\chi}_n\circ f_{H^{(n)}}^k\right|\\&=\frac{1}{\sqrt{n}}\sup_{t\in[0,T]}|\hat{\chi}_n\circ f_{H^{(n)}}^{[nt]}-\hat{\chi}_n|\le\frac{1}{n^{1/p}}\max_{1\le k\le[nT]}|\hat{\chi}_n\circ f_{H^{(n)}}^k-\hat{\chi}_n|.\end{aligned}$$

It follows from Proposition 6.2.8 that $\sup_{t\in[0,T]}|\hat{W}_n(t)-\hat{M}_n(t)|\to 0$ in probability. Setting s to be the Skorokhod metric defined on $D([0,T];\mathbb{R}^d)$, we have $s(\hat{W}_n,\hat{M}_n)\to 0$ in probability. Therefore, we have from Theorem 2.3.9 that $\hat{W}_n\to_w W$ in $D([0,T];\mathbb{R}^d)$. Since $T>0$ is arbitrary, we have from [92] that $\hat{W}_n\to_w W$ in $D([0,\infty);\mathbb{R}^d)$, as required. □

Proof of Theorem 6.1.4. Given a matrix $\Sigma\in\mathbb{R}^{d,d}$, let W_Σ denote the Brownian motion with mean 0 and covariance Σ. To prove (i), we show that any subsequence has a further subsequence which is weakly convergent. The result then follows from Theorem 2.3.14. Let (W_{n_k}) be a subsequence. Note that from (5.2.4) and Proposition 6.2.7, we have (Σ_{n_k}) is bounded. Therefore, we can pass to a further subsequence $(W_{n_{k_\ell}})$ along which $\Sigma_{n_{k_\ell}}\to\Sigma$ for some $\Sigma\in\mathcal{S}$. By Theorem 6.1.3 (ii), we have $W_{n_{k_\ell}}\to_w W_\Sigma$, and tightness follows.

For (ii), let $W\in\mathcal{W}$ and suppose (W_{n_k}) is a subsequence of (W_n) such that $W_{n_k}\to_w W$. From the argument in (i), we can pass to a further subsequence $(W_{n_{k_\ell}})$ along which $W_{n_{k_\ell}}\to_w W_\Sigma$ for some $\Sigma\in\mathcal{S}$. By weak convergence of W_{n_k} to W, it follows that $W\sim W_\Sigma$ and any weak limit is of the required form. Suppose now that $\Sigma\in\mathcal{S}$. Then there exists a subsequence (Σ_{n_k}) of (Σ_n) such that $\lim_{k\to\infty}\Sigma_{n_k}=\Sigma$. It follows from Theorem 6.1.3 (ii) that $W_{n_k}\to_w W_\Sigma$, so that $W_\Sigma\in\mathcal{W}$. □

Proof of Theorem 6.1.5. Note that by (6.2.3), we have

$$\begin{aligned} &n_k^{-(p-1)}\int_{X\times G}\Big|\sum_{j=0}^{n_k-1}\phi_{n_k}\circ T^j_{h^{(n_k)}}\Big|^{2(p-1)}\,\mathrm{d}m \\ &\ll n_k^{-(p-1)}n_k^{2(p-1)/2}\|v_{n_k}\|_\eta^{2(p-1)}\|h^{(n_k)}\|_\eta^{2(p-1)} \\ &\ll \big(\sup_{j\ge 0}\|v_j\|_\eta\big)^{2(p-1)}\big(\sup_{j\ge 0}\|h^{(j)}\|_\eta\big)^{2(p-1)}<\infty. \end{aligned}$$

In particular, if $0<q<2(p-1)$, then $(|W_{n_k}(1)|^q)_{k\ge1}$ is $L^{2(p-1)/q}-$bounded, and hence uniformly integrable by Proposition 2.3.11. The result now follows from Proposition 2.3.12. □

6.4 Application to homogenisation

There is considerable interest in understanding how stochastic behaviour emerges from deterministic systems. One method is via homogenisation, in which deterministic systems with multiple timescales converge to a stochastic differential equation. In particular, there has been much interest in the homogenisation of fast-slow dynamical systems [18, 38, 52, 53, 75]. To conclude the thesis, we give such an application of our results. We first formulate the setup.

As in Section 6.1, we let $T\colon X\to X$ be non-uniformly expanding and G be a compact connected Lie group with fixed representation into $\mathrm{O}(d)$. Consider the family of compact group extensions with base T, given by $T_{h^{(\epsilon)}}\colon X\times G\to X\times G$ for $\epsilon\in[0,\epsilon_0)$, where $\epsilon_0>0$. The cocycles $h^{(\epsilon)}\colon X\to G$ are $\eta-$Hölder and satisfy $\sup_{\epsilon\in[0,\epsilon_0)}\|h^{(\epsilon)}\|_\eta<\infty$. For all $\epsilon\in[0,\epsilon_0)$ the probability measure $m=\mu\times\nu$ is $T_{h^{(\epsilon)}}-$invariant and is assumed to be ergodic as in Remark 4.1.1. We consider the family of equivariant Hölder observables $\phi_\epsilon\colon X\times G\to\mathbb{R}^d$ defined by $\phi_\epsilon=g\cdot v_\epsilon$, where $v_\epsilon\colon X\to\mathbb{R}^d$ satisfies

$$\sup_{\epsilon\in[0,\epsilon_0)}\|v_\epsilon\|_\eta<\infty\quad\text{and}\quad\lim_{\epsilon\to0}|v_\epsilon-v_0|_\infty=0. \tag{6.4.1}$$

We make the following additional assumptions on the cocycles:

Assumption 6.4.1. *There exists $C > 0$ such that*

$$|h^{(\epsilon)} - h^{(0)}|_\infty \le C\epsilon \quad \text{for all } \epsilon \in [0, \epsilon_0). \tag{6.4.2}$$

Moreover, the induced compact group extension $F_{H(0)}$ is assumed to be mixing as in Remark 6.1.2.

Remark 6.4.2. *With this assumption, the results from the previous two sections go through for the family of compact group extensions $T_{h^{(\epsilon)}}$ and equivariant observables ϕ_ϵ.*

We study discrete fast-slow dynamical systems of the form

$$z_\epsilon(n+1) = z_\epsilon(n) + \epsilon^2 a_\epsilon\big(z_\epsilon(n), u_\epsilon(n)\big) + \epsilon\phi_\epsilon\big(u_\epsilon(n)\big), \quad z_\epsilon(0) = \xi_\epsilon. \tag{6.4.3}$$

The slow dynamics $z_\epsilon(n) \in \mathbb{R}^d$ have initial condition ξ_ϵ, and given $u_\epsilon(0) \in X \times G$, the fast dynamics $u_\epsilon(n+1) = T_{h^{(\epsilon)}}(u_\epsilon(n))$ are generated by the family of compact group extensions defined above. Here $a_\epsilon \colon \mathbb{R}^d \times X \times G \to \mathbb{R}^d$ is defined and continuous for $\epsilon \in [0, \epsilon_0)$, and the only source of randomness in the dynamics is the initial condition $u_\epsilon(0)$. We make the following regularity assumptions:

Assumption 6.4.3. *The function a_ϵ and initial condition ξ_ϵ in (6.4.3) satisfy the following:*

(i) $\lim_{\epsilon\to 0} \xi_\epsilon = \xi_0$.

(ii) *There is a constant $L \ge 1$ such that*

$$|a_\epsilon|_\infty \le L \quad \text{and} \quad \mathrm{Lip}(a_\epsilon) = \sup_{z\neq z'} \sup_{(x,g)} \frac{\big|a_\epsilon(z,x,g) - a_\epsilon(z',x,g)\big|}{|z - z'|} \le L$$

for all $\epsilon \in [0, \epsilon_0)$.

(iii) $\lim_{\epsilon\to 0} |a_\epsilon - a_0|_\infty = 0$.

(iv) $a_0(z,x,g) = g \cdot a_0'(z,x)$ *for some* $a_0' \colon \mathbb{R}^d \times X \to \mathbb{R}^d$ *which satisfies*

$$\|a_0'(z,\cdot)\|_\eta < \infty \tag{6.4.4}$$

for all $z \in \mathbb{R}^d$.

We set out some notation for the next theorem. For $z \in \mathbb{R}^d$, let us define $\alpha_z \colon X \times G \to \mathbb{R}^d$ by

$$\alpha_z(x,g) = a_0(z,x,g).$$

Let $P \colon \mathbb{R}^d \to \mathbb{R}^d$ be defined by

$$P(z) = \int_{X \times G} \alpha_z(x,g)\,\mathrm{d}m.$$

Let

$$\Sigma_0 = \lim_{n\to\infty} \frac{1}{n} \int_{X\times G} \left(\sum_{k=0}^{n-1} \phi_0 \circ T^k_{h^{(0)}} \right) \left(\sum_{k=0}^{n-1} \phi_0 \circ T^k_{h^{(0)}} \right)^T \mathrm{d}m$$

and let W denote the Brownian motion with mean 0 and covariance matrix Σ_0. Define $\hat{z}_\epsilon \in D([0,\infty);\mathbb{R}^d)$ by $\hat{z}_\epsilon(t) = z_\epsilon([t\epsilon^{-2}])$ for $t \geq 0$.

Theorem 6.4.4. *P is Lipschitz and $\hat{z}_\epsilon \to_w Z$ in $D([0,\infty);\mathbb{R}^d)$ as $\epsilon \to 0$, where Z is the solution to the integral equation*[2]

$$Z_t = \xi_0 + \int_0^t P(Z_s)\,\mathrm{d}s + W_t, \quad t \geq 0.$$

Proof. By [60, Theorem 6.3 and Remark 6.4] (where x_ϵ, y_ϵ, and X in the cited paper corresponds to our z_ϵ, u_ϵ and Z), it suffices to show:

(i) For all $z \in \mathbb{R}^d$,

$$\lim_{\epsilon\to 0} \int_{X\times G} \left| \epsilon^{1/2} \sum_{k=0}^{[\epsilon^{-1/2}]-1} \alpha_z \circ T^k_{h^{(\epsilon)}} - \int_{X\times G} \alpha_z \,\mathrm{d}m \right| \mathrm{d}m = 0.$$

[2] Equivalently, one can write this as the stochastic differential equation $\mathrm{d}Z = P(Z)\,\mathrm{d}t + \mathrm{d}W$ with initial condition $Z(0) = \xi_0$.

(ii) The family of random elements $W_\epsilon\colon X\times G \to D([0,\infty);\mathbb{R}^d)$ defined for $t\geq 0$ by

$$W_\epsilon(t) = \epsilon \sum_{k=0}^{[t\epsilon^{-2}]-1} \phi_\epsilon \circ T^k_{h^{(\epsilon)}}$$

satisfies $W_\epsilon \to_w W$ in $D([0,\infty);\mathbb{R}^d)$ as $\epsilon \to 0$.

We first prove (i). Fix $z\in\mathbb{R}^d$ and define $\beta_z\colon X\times G\to\mathbb{R}^d$ by $\beta_z(x,g) = \alpha_z(x,g) - \int_{X\times G}\alpha_z\,\mathrm{d}m$. Note that $\int_{X\times G}\beta_z\,\mathrm{d}m = 0$. Moreover, by invariance of the Haar measure, we have

$$\begin{aligned}
&\int_{X\times G}\alpha_z(x,h)\,\mathrm{d}m(x,h) = \int_{X\times G}\alpha_z(x,gh)\,\mathrm{d}m(x,h) = \int_{X\times G} a_0(z,x,gh)\,\mathrm{d}m(x,h)\\
&= \int_{X\times G} gh\cdot a_0'(z,x)\,\mathrm{d}m(x,h) = g\cdot\left(\int_{X\times G} h\cdot a_0'(z,x)\,\mathrm{d}m(x,h)\right)\\
&= g\cdot\left(\int_{X\times G} a_0(z,x,h)\,\mathrm{d}m(x,h)\right) = g\cdot\left(\int_{X\times G}\alpha_z(x,h)\,\mathrm{d}m(x,h)\right).
\end{aligned}$$

Therefore

$$\beta_z(x,g) = g\cdot a_0'(z,x) - g\cdot\left(\int_{X\times G}\alpha_z\,\mathrm{d}m\right) = g\cdot\beta_z'(x),$$

where $\beta_z'\colon X\to\mathbb{R}^d$ is defined by

$$\beta_z'(x) = a_0'(z,x) - \int_{X\times G}\alpha_z\,\mathrm{d}m.$$

Now, by (6.4.4), we have $\|\beta_z'\|_\eta < \infty$. It follows from (6.2.3) that

$$\begin{aligned}
&\int_{X\times G}\left|\sum_{k=0}^{[\epsilon^{-1/2}]-1}\epsilon^{1/2}\beta_z\circ T^k_{h^{(\epsilon)}}\right|\mathrm{d}m \ll \epsilon^{1/2}\big([\epsilon^{-1/2}]\big)^{1/2}\|v_\epsilon\|_\eta\|h^{(\epsilon)}\|_\eta\\
&\leq \epsilon^{1/4}\sup_{\epsilon\in[0,\epsilon_0)}\|v_\epsilon\|_\eta \sup_{\epsilon\in[0,\epsilon_0)}\|h^{(\epsilon)}\|_\eta \to 0 \quad \text{as } \epsilon\to 0.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\int_{X\times G}\left|\epsilon^{1/2}\sum_{k=0}^{[\epsilon^{-1/2}]-1}\alpha_z\circ T_{h^{(\epsilon)}}^k-\int_{X\times G}\alpha_z\,\mathrm{d}m\right|\mathrm{d}m\\
&\le\int_{X\times G}\left|\sum_{k=0}^{[\epsilon^{-1/2}]-1}\epsilon^{1/2}\beta_z\circ T_{h^{(\epsilon)}}^k\right|\mathrm{d}m+\int_{X\times G}\left|\epsilon^{1/2}[\epsilon^{-1/2}]\alpha_z-\alpha_z\right|\mathrm{d}m\\
&\to 0\quad\text{ as }\epsilon\to 0,
\end{aligned}$$

proving (i).

We now prove (ii). For $\epsilon\in[0,\epsilon_0)$, let us define the family of covariance matrices

$$\Sigma_\epsilon=\lim_{n\to\infty}\frac{1}{n}\int_{X\times G}(S_n\phi_\epsilon)(S_n\phi_\epsilon)^T\,\mathrm{d}m,$$

where

$$S_n\phi_\epsilon=\sum_{k=0}^{n-1}\phi_\epsilon\circ T_{h^{(\epsilon)}}^k.$$

By Theorem 6.1.3 (ii), it suffices to show that that $\lim_{\epsilon\to 0}\Sigma_\epsilon=\Sigma_0$. Write

$$I_{\epsilon,n}=\int_{X\times G}(S_n\phi_\epsilon)(S_n\phi_\epsilon)^T\,\mathrm{d}m,$$

and let $\delta>0$. By Corollary 6.2.10, there exists $N\ge 1$ such that $\|N^{-1}I_{\epsilon,N}-\Sigma_\epsilon\|<\delta$ for all $\epsilon\in[0,\epsilon_0)$. Note that

$$\begin{aligned}
&\|I_{\epsilon,N}-I_{0,N}\|\le\left|(S_N\phi_\epsilon)(S_N\phi_\epsilon)^T-(S_N\phi_0)(S_N\phi_0)^T\right|_1\\
&\le\big(|S_N\phi_\epsilon|_2+|S_N\phi_0|_2\big)|S_N\phi_\epsilon-S_N\phi_0|_2\le N\big(|v_\epsilon|_\infty+|v_0|_\infty\big)|S_N\phi_\epsilon-S_N\phi_0|_2.
\end{aligned}$$

We show that this converges to 0 as $\epsilon\to 0$. Note first that $N(|v_\epsilon|_\infty+|v_0|_\infty)\le 2N\sup_{\epsilon\in[0,\epsilon_0)}\|v_\epsilon\|_\eta<\infty$ by (6.4.1). Next note that

$$|S_N\phi_\epsilon-S_N\phi_0|\le\sum_{k=0}^{N-1}|\phi_\epsilon-\phi_0|\circ T_{h^{(\epsilon)}}^k+\sum_{k=0}^{N-1}|\phi_0\circ T_{h^{(\epsilon)}}^k-\phi_0\circ T_{h^{(0)}}^k|=:I+II.$$

We have

$$I\le N|v_\epsilon-v_0|_\infty\to 0\quad\text{ as }\epsilon\to 0$$

by (6.4.1). For II, first observe that for $\epsilon \in [0, \epsilon_0)$, $(x, g) \in X \times G$, and $0 \leq k \leq N-1$, we have

$$\phi_0\big(T^k_{h^{(\epsilon)}}(x, g)\big) = \phi_0\big(T^k x, g h^{(\epsilon)}_k(x)\big) = g h^{(\epsilon)}_k(x) \cdot v_0(T^k x),$$

where $h^{(\epsilon)}_k = h^{(\epsilon)}\, h^{(\epsilon)} \circ T \cdots h^{(\epsilon)} \circ T^{k-1}$. Therefore

$$\sum_{k=0}^{N-1} \big|\phi_0 \circ T^k_{h^{(\epsilon)}}(x, g) - \phi_0 \circ T^k_{h^{(0)}}(x, g)\big| \leq \sum_{k=0}^{N-1} |v_0|_\infty \big\|h^{(\epsilon)}_k(x) - h^{(0)}_k(x)\big\|.$$

Moreover, by an identical argument to the proof of (4.4.1), we have

$$\big\|h^{(\epsilon)}_k(x) - h^{(0)}_k(x)\big\| \leq \sum_{j=0}^{k-1} \big\|h^{(\epsilon)}(T^j x) - h^{(0)}(T^j x)\big\| \leq k|h^{(\epsilon)} - h^{(0)}|_\infty \ll k\epsilon$$

by (6.4.2). It follows that

$$II \ll \sum_{k=0}^{N-1} |v_0|_\infty k\epsilon = \frac{\epsilon(N-1)N}{2} |v_0|_\infty \to 0 \quad \text{ as } \epsilon \to 0.$$

Therefore $|S_N\phi_\epsilon - S_N\phi_0| \to 0$ as $\epsilon \to 0$. In addition,

$$|S_N\phi_\epsilon - S_N\phi_0| \leq |S_N\phi_\epsilon| + |S_N\phi_0| \leq N|v_\epsilon|_\infty + N|v_0|_\infty \leq 2N \sup_{\epsilon \in [0,\epsilon)} \|v_\epsilon\|_\eta < \infty$$

by (6.4.1), and hence $|S_N\phi_\epsilon - S_N\phi_0|_2 \to 0$ as $\epsilon \to 0$ by the dominated convergence theorem. We conclude that $\lim_{\epsilon \to 0} I_{\epsilon,N} = I_{0,N}$. Since

$$\begin{aligned}\|\Sigma_\epsilon - \Sigma_0\| &\leq \|N^{-1}I_{\epsilon,N} - \Sigma_\epsilon\| + N^{-1}\|I_{\epsilon,N} - I_{0,N}\| + \|N^{-1}I_{0,N} - \Sigma_0\| \\ &< 2\delta + N^{-1}\|I_{\epsilon,N} - I_{0,N}\|,\end{aligned}$$

it follows that $\lim_{\epsilon \to 0} \Sigma_\epsilon = \Sigma_0$, as required. □

Bibliography

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory.* Mathematical Surveys and Monographs. Amer. Math. Soc., 1997.

[2] J. Aaronson and M. Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.*, 01(02):193–237, 2001.

[3] J. F. Alves, S. Luzzatto, and V. Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(6):817–839, 2005.

[4] V. Baladi. *Positive Transfer Operators and Decay of Correlations.* Advanced Series in Nonlinear Dynamics. World Scientific, 2000.

[5] M. Benedicks and L. Carleson. On iterations of $1 - ax^2$ on $(-1, 1)$. *Ann. of Math.*, 122(1):1–25, 1985.

[6] I. Berkes, W. Liu, and W. B. Wu. Komlós–Major–Tusnády approximation under dependence. *Ann. Probab.*, 42(2):794–817, 2014.

[7] I. Berkes and W. Philipp. Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.*, 7(1):29–54, 1979.

[8] P. Billingsley. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, 12(5):788–792, 1961.

[9] P. Billingsley. *Probability and Measure.* Wiley Series in Probability and Mathematical Statistics. Wiley, 1986.

[10] P. Billingsley. *Convergence of Probability Measures; 2nd ed.* Wiley Series in Probability and Statistics. Wiley, 1999.

[11] G. D. Birkhoff. Proof of the ergodic theorem. *Proc. Natl. Acad. Sci.*, 17(12):656–660, 1931.

[12] R. E. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Lecture Notes in Mathematics. Springer, 1975.

[13] A. Boyarsky and P. Gora. *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension.* Probability and its Applications. Birkhäuser, 1997.

[14] T. Bröcker and T. Dieck. *Representations of Compact Lie Groups.* Graduate Texts in Mathematics. Springer, 2003.

[15] B. M. Brown. Martingale central limit theorems. *Ann. Math. Statist.*, 42(1):59–66, 1971.

[16] H. Bruin, M. Holland, and I. Melbourne. Subexponential decay of correlations for compact group extensions of non-uniformly expanding systems. *Ergodic Theory Dynam. Systems*, 25(6):1719–1738, 2005.

[17] D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Probab.*, 1(1):19–42, 1973.

[18] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. Multiscale systems, homogenization, and rough paths. In *Probability and Analysis in Interacting Physical Systems*, volume 283 of *Springer Proc. Math. Stat.*, pages 17–48. Springer, 2019.

[19] Y. S. Chow and H. Teicher. *Probability Theory, Independence, Interchangeability, Martingales.* Springer Texts in Statistics. Springer, 1978.

[20] P. Collet and J.-P. Eckmann. Positive Liapunov exponents and absolute continuity for maps of the interval. *Ergodic Theory Dynam. Systems*, 3(1):13–46, 1983.

[21] H. Cramér and H. Wold. Some theorems on distribution functions. *J. Lond. Math. Soc.*, s1-11(4):290–294, 1936.

[22] C. Cuny and F. Merlevède. Strong invariance principles with rate for "reverse" martingale differences and applications. *J. Theoret. Probab.*, 28(1):137–183, 2015.

[23] J. Dedecker and P. Doukhan. A new covariance inequality and applications. *Stochastic Process. Appl.*, 106(1):63–80, 2003.

[24] F. den Hollander. Probability theory: The coupling method. Lecture notes, Leiden University. Available at `https://websites.math.leidenuniv.nl/probability/lecturenotes/CouplingLectures.pdf`, 2012.

[25] M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory Dynam. Systems*, 4(4):541–552, 1984.

[26] W. Doeblin and R. Fortet. Sur des chaînes à liaisons complètes. *Bull. Soc. Math. France*, 65:132–148, 1937.

[27] D. Dolgopyat. On mixing properties of compact group extensions of hyperbolic systems. *Israel J. Math.*, 130(1):157–205, 2002.

[28] D. Dolgopyat. Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.*, 356(4):1637–1689, 2004.

[29] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.*, 6, 1951.

[30] J. L. Doob. *Stochastic Processes.* Wiley Publications in Statistics. Wiley, 1953.

[31] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

[32] A. Dvoretzky. Asymptotic normality for sums of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 2 of *Probab. Theory*, pages 513–535. University of California Press, 1972.

[33] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 2005.

[34] J. Faraut. *Analysis on Lie Groups: An Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2008.

[35] M. Field, I. Melbourne, and A.Török. Stable ergodicity for smooth compact Lie group extensions of hyperbolic basic sets. *Ergodic Theory Dynam. Systems*, 25(2):517–551, 2005.

[36] M. Field, I. Melbourne, and A. Török. Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergodic Theory Dynam. Systems*, 23(1):87–110, 2003.

[37] M. I. Gordin. The central limit theorem for stationary processes. *Soviet Math. Dokl.*, 10:1174–1176, 1969.

[38] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. A*, 469(2156):20130201, 2013.

[39] G. A. Gottwald and I. Melbourne. Central limit theorems and suppression of anomalous diffusion for systems with symmetry. *Nonlinearity*, 29(10):2941, 2016.

[40] S. Gouëzel. Central limit theorem and stable laws for intermittent maps. *Probab. Theory Related Fields*, 128(1):82–122, 2004.

[41] S. Gouëzel. Decay of correlations for nonuniformly expanding systems. *Bull. Soc. Math. France*, 134(1):1–31, 2006.

[42] S. Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.*, 38(4):1639–1671, 2010.

[43] S. Gouëzel. Limit theorems in dynamical systems using the spectral method. In *Hyperbolic Dynamics, Fluctuations and Large Deviations*, volume 89 of *Proc. Sympos. Pure Math.*, page 161–193. Amer. Math. Soc., 2015.

[44] G. Grimmett and D. Welsh. *Probability: An Introduction.* Oxford Science Publications. Oxford University Press, 2014.

[45] P. Hall and C. C. Heyde. *Martingale Limit Theory and its Application.* Probability and Mathematical Statistics. Academic Press, 1980.

[46] H. Hennion. Sur un théorème spectral et son application aux noyaux Lipchitziens. *Proc. Amer. Math. Soc.*, 118(2):627–634, 1993.

[47] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.*, 180(1):119–140, 1982.

[48] I. Ibragimov. A central limit theorem for a class of dependent random variables. *Theory Probab. Appl.*, 8(1):83–89, 1963.

[49] M. V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.*, 81(1):39–88, 1981.

[50] T. Kato. *Perturbation Theory for Linear Operators; 2nd ed.* Grundlehren der mathematischen Wissenschaften. Springer, 1976.

[51] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, Ser. 4, 28(1):141–152, 1999.

[52] D. Kelly and I. Melbourne. Deterministic homogenization for fast-slow systems with chaotic noise. *J. Funct. Anal.*, 272(10):4063–4102, 2014.

[53] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.*, 44(1):479–520, 2016.

[54] A. Khintchine. Über einen Satz der Wahrscheinlichkeitsrechnung. *Fund. Math.*, 6(1):9–20, 1924.

[55] J. Kiefer. On the deviations in the Skorokhod-Strassen approximation scheme. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 13(3):321–332, 1969.

[56] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.

[57] A. Kolmogorov. Über das Gesetz des iterierten Logarithmus. *Math. Ann.*, 101(1):126–135, 1929.

[58] A. N. Kolmogorov, A.T. Bharucha-Reid, and N. Morrison. *Foundations of the Theory of Probability: Second English Edition*. Dover Books on Mathematics. Dover Publications, 2018.

[59] A Korepanov. Rates in almost sure invariance principle for dynamical systems with some hyperbolicity. *Comm. Math. Phys.*, 363(1):173–190, 2018.

[60] A. Korepanov, Z. Kosloff, and I. Melbourne. Martingale–coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 35(4):859–885, 2018.

[61] N. Kryloff and N. Bogoliouboff. La théorie generale de la mesure dans son application a l'étude des systèmes dynamiques de la mecanique nonlineaire. *Ann. of Math.*, 38(1):65–113, 1937.

[62] A. M. Lasota and J. A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488, 1973.

[63] C. Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems*, volume 362 of *Pitman Res. Notes Math. Ser.*, pages 56–75, 1996.

[64] C. Liverani. Invariant measures and their properties. a functional analytic point of view. Lecture notes, University di Roma. Available at `http://axp.mat.uniroma2.it/~liverani/Lavori/live0203.pdf`, 2002.

[65] C. Liverani, B. Saussol, and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems*, 19(3):671–685, 1999.

[66] S. Luzzatto. Introduction to smooth ergodic theory. Lecture notes, SISSA. Available at `http://indico.ictp.it/event/a12289/session/2/contribution/1/material/0/0.pdf`, 2013.

[67] H. B. Mann and A. Wald. On stochastic limit and order relationships. *Ann. Math. Statist.*, 14(3):217–226, 1943.

[68] M Maxwell and M Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.*, 28(2):713–724, 2000.

[69] D. L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probab.*, 2(4):620–628, 1974.

[70] I. Melbourne and M. Nicol. Statistical limit laws for equivariant observations. *Stoch. Dyn.*, 04(01):1–13, 2004.

[71] I. Melbourne and M. Nicol. Statistical properties of endomorphisms and compact group extensions. *J. Lond. Math. Soc.*, 70(2):427–446, 2004.

[72] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.*, 260(1):131–146, 2005.

[73] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.*, 360(12):6661–6676, 2008.

[74] I. Melbourne and M. Nicol. A vector-valued almost sure invariance principle for hyperbolic dynamical systems. *Ann. Probab.*, 37(2):478–505, 2009.

[75] I. Melbourne and A. M. Stuart. A note on diffusion limits of chaotic skew-product flows. *Nonlinearity*, 24(4):1361–1367, 2011.

[76] I. Melbourne and D. Terhesiu. Decay of correlations for non-uniformly expanding systems with general return times. *Ergodic Theory Dynam. Systems*, 34(3):893–918, 2014.

[77] I. Melbourne and A. Török. Convergence of moments for Axiom A and non-uniformly hyperbolic flows. *Ergodic Theory Dynam. Systems*, 32(3):1091–1100, 2012.

[78] F. Merlevède, M. Peligrad, and S. Utev. Recent advances in invariance principles for stationary sequences. *Probab. Surveys*, 3:1–36, 2006.

[79] M. Nicol, I. Melbourne, and P. Ashwin. Euclidean extensions of dynamical systems. *Nonlinearity*, 14(2):275–300, 2001.

[80] W. Parry and M. Pollicott. *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics*, volume 187-188 of *Astérisque*. Soc. Math. France, 1990.

[81] W. Parry and M. Pollicott. Skew products and Livšic theory. In *Representation Theory, Dynamical Systems, and Asymptotic Combinatorics*, volume 217 of *Amer. Math. Soc. Transl. Ser. 2*, pages 139–165, 2006.

[82] K. R. Parthasarathy. *Probability Measures on Metric Spaces*. Chelsea Publishing Series. Amer. Math. Soc., 2005.

[83] G. A. Pavliotis and A. M. Stuart. *Multiscale Methods: Averaging and Homogenization*. Texts in Applied Mathematics. Springer, 2008.

[84] W. Philipp and W.F. Stout. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.*, 161, 1975.

[85] Y. V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.*, 1(2):157–214, 1956.

[86] E. Rio. *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*. Mathématiques et Applications. Springer, 1999.

[87] J. S. Rosenthal. *A First Look at Rigorous Probability Theory*. World Scientific, 2006.

[88] D. Ruelle. *Thermodynamic Formalism*. Encyclopedia of Mathematics and its Applications. Addison Wesley, 1975.

[89] D. J. Scott and R. M. Huggins. On the embedding of processes in Brownian motion and the law of the iterated logarithm for reverse martingales. *Bull. Aust. Math. Soc.*, 27(3):443–459, 1983.

[90] Y. G. Sinaĭ. Gibbs measures in ergodic theory. *Russ. Math. Surv.*, 27(4):21–69, 1972.

[91] A. V. Skorokhod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1(3):261–290, 1956.

[92] C. Stone. Weak convergence of stochastic processes defined on semi-infinite time intervals. *Proc. Amer. Math. Soc.*, 14(5):694–696, 1963.

[93] V. Strassen. An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 3(3):211–226, 1964.

[94] V. Strassen. Almost sure behaviour of sums of independent random variables and martingales. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 2 of *Probab. Theory*, pages 315–343. University of California Press, 1967.

[95] C. T. Ionescu Tulcea and G. Marinescu. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math.*, 52(1):140–147, 1950.

[96] M. Tyran-Kamińska. An invariance principle for maps with polynomial decay of correlations. *Comm. Math. Phys.*, 260(1):1–15, 2005.

[97] M. Viana. Multidimensional nonhyperbolic attractors. *Publ. Math. Inst. Hautes Études Sci.*, 85:63–96, 1997.

[98] P. Walters. *An Introduction to Ergodic Theory.* Graduate Texts in Mathematics. Springer, 2000.

[99] N. Wiener. Differential-space. *J. Math. Phys.*, 2(1-4):131–174, 1923.

[100] D. Williams. *Probability with Martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, 1991.

[101] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.*, 147(3):585–650, 1998.

[102] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110(1):153–188, 1999.

[103] R. Zweimüller. Mixing limit theorems for ergodic transformations. *J. Theoret. Probab.*, 20(4):1059–1071, 2007.