On the Rates of Convergence for a Deterministic Weak Invariance Principle

Author:
Nicolò Paviato – 1899104

Supervisor:
Prof. Ian Melbourne

Warwick Mathematics Institute

This dissertation has been submitted to the University of Warwick to fulfil the requirements of the module MA917 - Summer Research Project

Abstract

This project regards rates of convergence for Donsker's Weak Invariance Principle applied to dynamical systems. The only known reference to this matter is the one of Antoniou and Melbourne, which uses an estimate based on the Skorokhod embedding theorem. The first chapter expands the calculations of their paper in a simplified setting, giving rates of $O(n^{-(\frac{1}{4}-\gamma)})$ for every $\gamma > 0$ in the Prokhorov metric for discrete-time systems. Chapter two contains an original result, showing a martingale-coboundary decomposition for Hölder observables in a continuous-time setting. The aim of this project is to set the foundation to apply techniques of Courbot, in order to improve the estimates and get also multi-dimensional rates.

Acknowledgements

I am very grateful for the opportunity of writing this project after my first year at the University of Warwick, which has been full of nice academical and personal experiences. I thank the MASDOC CDT, EPSRC and the University of Warwick for providing the funding for my studies.

Many thanks go to Professor Ian Melbourne who constantly and thoroughly helped me, proposing interesting topics and guiding me through the beginning of my research. I am really happy to start my PhD with him, looking forward to getting the best out of our collaboration.

Contents

Introduction

Great interest is given to statistical properties of dynamical systems. Examples are decay of correlations and rates of mixing, which are useful instruments to study long-term behaviour. Another powerful tool from probability theory is the Central Limit Theorem (**CLT**) which also finds an application to deterministic cases. Let $T\colon \Lambda \to \Lambda$ be a dynamical system endowed with a probability measure μ, invariant under the transformation T. An *observable* $v : \Lambda \to \mathbb{R}$ is a measurable function on the probability space (Λ, μ) and therefore can be seen as a random variable. By mean of Birkhoff sums and because T is measure-preserving, we can interpret

$$v_n = \sum_{j=0}^{n} v \circ T^n \qquad\qquad n \in \mathbb{N}$$

as a sum of dependent, identically distributed real random variables. We note that for every initial point $x \in \Lambda$ the elements $v_n(x)$ are totally determined and therefore we are applying probability theory to something not intrinsically random.

If we assume exponential decay of correlations for T and take a regular observable v, we often have convergence in distribution of $\frac{v_n}{\sqrt{n}}$ to a Gaussian random variable. It is shown in [13] that the latter is true taking for example the doubling map $T\colon [0,1] \to [0,1]$, $Tx = 2x$ $(mod\ 1)$ and $v\colon [0,1] \to \mathbb{R}$ Lipschitz. The main tool for the proof is a decomposition of v which is done through the transformation T, writing v as

$$v = m + \chi \circ T - \chi. \tag{1}$$

Here, χ is a L^∞-function defined by a converging series and it is called the *coboundary* of v. The function m is the *martingale part*; this name is justified because by Birkhoff sums m can generate a martingale. That is why we call (1) the **martingale-coboundary decomposition** of v: its power lies in reading the observable v as a sum of something meaningful for the CLT, plus a negligible coboundary term.

In a similar way we consider a functional CLT, i.e. Donsker's Weak Invariance Principle (**WIP**). The latter states convergence in distribution of a sequence of continuous paths $W_n \in \mathscr{C}[0,1]$ to a Wiener process. Unlike the general statement found in Billingsley [3], we use the dependent Birkhoff sums v_n defined above to generate

$$W_n(t) = \frac{v_{nt}}{\sqrt{n}} \qquad \text{for} \quad t = \frac{k}{n} \quad \text{and} \quad k \in \{0, \dots, n\},$$

with linear interpolation between the nodes k/n. The sequence $\{W_n\}_{n \in \mathbb{N}^+}$ generates a family of pushforward laws on $\mathscr{C}[0,1]$ which converge to a Wiener measure. The latter is the element of $\mathscr{M}_1(\mathscr{C}[0,1])$ with finite-dimensional distributions on $\mathbb{R}^n$ coinciding with the ones generated by a Brownian motion. The decomposition (1) plays again an important role in the proof of the WIP, as can be seen in [13] for the doubling map.

These results are still true under more general hypotheses and find a wide range of applications, for example in the convergence of fast-slow systems to solutions of stochastic differential equations, as done in Antoniou's PhD thesis [1]. Both CLT and WIP are true for non-uniformly expanding dynamical systems, using Hölder observables (see [2, 8, 9]). In particular for the WIP, Antoniou and Melbourne have used a generalized martingale-coboundary decomposition described in [8], finding in [2] a rate of convergence of $O(n^{-(\frac{1}{4}+r)})$ in the Prokhorov metric on $\mathscr{M}_1(\mathscr{C}[0,1])$, where $r > 0$ depends on the regularity of the transformation. We remark that [2] is heretofore the only reference available on rates for WIP applied to deterministic systems. The applied method is based on the Skorokhod embedding theorem used by Kubilius in [11], and therefore cannot yield a better rate than $O(n^{-\frac{1}{4}})$. In order to improve this result and especially find rates for the multi-dimensional case, it is reasonable to start working in the continuum and apply continuous-time martingales techniques; there are several papers in the probability literature concerning this matter and the most promising rates are found in Courbot [6, 7]. References to papers with inferior rates can be found in Courbot [5, 10]. This project aims to set the foundation for the application of these new techniques in a later PhD thesis, which will lead to new rates of convergence for the deterministic case.

The first chapter of this project takes inspiration from [13]: it is an exposition of existing work of [1, 2] in a simplified setting, proving a rate of convergence for the WIP of $O(n^{-(\frac{1}{4}-\gamma)})$ for every $\gamma > 0$, valid for a discrete dynamical system. In section 1.1 we recall definitions and properties of the Koopman and transfers operators, assuming some hypotheses on their actions upon Hölder observables. Then in section 1.2 we obtain two martingale-coboundary decompositions: one for an observable v, and writing $v = m + \chi \circ T - \chi$ as in (1), the other for the observable $\breve{v} = m^2 - \int m^2$. After showing some properties of the respective martingale parts, section 1.3 uses Burkholder's inequality [4] to prove some estimates which leave us in the right setting to apply ideas from [2]. The main estimate from probability theory is borrowed from Kubilius [11], which shows the decay of the Prokhorov distance between a Brownian motion and an additional sequence of processes (that is close enough to W_n). Following this line of reasoning, we reach the desired rate in section 1.4.

The material found in chapter 2 is original work and is meant to be the starting point for the improvement and extension of latter results. We consider here a semigroup of measure-preserving transformation $\{T_t\}_{t\geq 0}$ that replaces the iterated discrete-time map T. We then develop this new setting, showing properties of the families of Koopman and transfer operators. An essential step to move forward is to use Bochner integration in place of Birkhoff sums: therefore we recall shortly the theory, referring to [12, 15]. Given v Hölder continuous with zero mean, section 2.3 extends the decomposition (1) to $v_t = \int_0^t v \circ T_s \,\mathrm{d}s$ with

$$v_t = m_t + \chi \circ T_t - \chi, \tag{2}$$

with new martingale and coboundary parts. Finally we show how m_t can be interpreted, generating a continuous-time martingale array similar to what obtained in the discrete case. The decomposition (2) will be used for the next project to apply estimates of [6, 7], eventually showing new rates for non-uniformly expanding dynamical systems and vector-valued observables $v\colon \Lambda \to \mathbb{R}^d$.

Chapter 1

Rates of Convergence for Discrete Dynamical Systems

This chapter considers a dynamical system $T\colon \Lambda \to \Lambda$ endowed with a probability measure μ, for which Hölder observables satisfy the Weak Invariance Principle. Starting with a short review of important tools like Koopman and transfer operators, it shows with Proposition 1.4 how the transformation T induces a useful decomposition for Hölder functions, which is the central tool to apply general results from probability. As a matter of fact, we will see in section 1.3 how the first part of this decomposition carries "martingale" properties. This chapter takes inspiration from [13], expanding its calculation under the further assumption of exponential decay of Hölder norm for the iterated transfer operator.

Using Burkhorlder's inequality [4], we prove a series of inequalities connected to the decomposition cited above, which lead eventually to the starting point to follow the ideas exposed in Antoniou and Melbourne [2]. Following the steps of their paper (for which we reported and adapted the proofs), we are able to prove for the WIP a rate of $O(n^{-(\frac{1}{4}-\gamma)})$ for every $\gamma > 0$, using the Prokhorov metric.

1.1 Koopman and Transfer Operators

Consider a probability space $(\Lambda, \mathscr{B}, \mu)$ and a measure-preserving transformation $T\colon \Lambda \to \Lambda$, i.e. such that the pushforward measure is $T_*\mu = \mu$. The bounded linear map

$$U\colon L^1(\Lambda) \to L^1(\Lambda) \qquad Uw := w \circ T, \tag{1.1}$$

is called the *Koopman operator* induced by T. It is known [13] that there exists a unique bounded linear map $P\colon L^1(\Lambda) \to L^1(\Lambda)$ such that

$$\int_\Lambda Pv\ w \,\mathrm{d}\mu = \int_\Lambda v\ Uw \,\mathrm{d}\mu, \tag{1.2}$$

for every $v \in L^1(\Lambda)$ and $w \in L^\infty(\Lambda)$. P is known as the *transfer operator* (or *Perron-Frobenius operator*). As a matter of fact $\|Uv\|_\infty \leq \|v\|_\infty$ which implies $UL^\infty \subseteq L^\infty$.

Remark 1.1. The operators U and P preserve integrals: for every $v \in L$ we have that

- $\int_\Lambda Uv \,\mathrm{d}\mu = \int_\Lambda v \circ T \,\mathrm{d}\mu = \int_\Lambda v \,\mathrm{d}\mu$, because T is measure-preserving;
- $\int_\Lambda Pv \,\mathrm{d}\mu = \int_\Lambda v \; U1 \,\mathrm{d}\mu = \int_\Lambda v \,\mathrm{d}\mu$, where 1 is the constant function.

Lemma 1.2. *Let U and P be the Koopman and transfer operators of a measure-preserving transformation $T\colon \Lambda \to \Lambda$. Then*

(a) $PU = Id|_{L^1}$*;*

(b) $UP = \mathbb{E}\big[\,\cdot\,|T^{-1}\mathscr{B}\big]$*;*

(c) $UP = Id|_{L^1}$ *if T is invertible.*

Proof. (a) By definition of P we have that for every $v \in L^1(\Lambda)$ and $w \in L^\infty(\Lambda)$

$$\int_\Lambda PUv \; w \,\mathrm{d}\mu = \int_\Lambda Uv \; Uw \,\mathrm{d}\mu = \int_\Lambda (vw) \circ T \,\mathrm{d}\mu = \int_\Lambda vw \,\mathrm{d}\mu.$$

(b) To check the integral condition for any $v \in L^1(\Lambda)$ and $B \in \mathscr{B}$, we use the invariance of T, the fact that $\mathbb{1}_B \circ T = \mathbb{1}_{T^{-1}B}$ and definition of P:

$$\begin{aligned}\int_{T^{-1}B} UPv \,\mathrm{d}\mu &= \int_\Lambda (Pv \circ T)\mathbb{1}_{T^{-1}B} \,\mathrm{d}\mu = \int_\Lambda (Pv \; \mathbb{1}_B) \circ T \,\mathrm{d}\mu \\ &= \int_\Lambda Pv \; \mathbb{1}_B \,\mathrm{d}\mu = \int_\Lambda v(\mathbb{1}_B \circ T) \,\mathrm{d}\mu = \int_{T^{-1}B} v \,\mathrm{d}\mu.\end{aligned}$$

Now, for fixed $B \in \mathscr{B}(\mathbb{R})$ and $v \in L^1(\Lambda)$, we see that UPv is $T^{-1}\mathscr{B}$-measurable,

$$(UPv)^{-1}B = (Pv \circ T)^{-1}B = T^{-1}\big((Pv)^{-1}B\big) \in T^{-1}\mathscr{B},$$

because $Pv \in L^1(\Lambda)$ is Borel measurable.

(c) If T is invertible, so is U with inverse $U^{-1}v = v \circ T^{-1}$.
Then we have $UP = UPUU^{-1} = UU^{-1} = Id$.

□

Lemma 1.3. *Take two probability spaces $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$, $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ and a measure-preserving map $\pi\colon \Omega_1 \to \Omega_2$. Let $X \in L^1(\Lambda)(\Omega_2)$ and $\mathcal{G} \subset \mathcal{A}$ a sub σ-algebra. Then*

$$\mathbb{E}\big[X \circ \pi | \pi^{-1}\mathcal{G}\big] = \mathbb{E}\big[X|\mathcal{G}\big] \circ \pi.$$

Proof. We need to verify the $\pi^{-1}\mathcal{G}$-measurability of the right-hand side and that it satisfies the integral property of the conditional expectation. If $A \in \mathcal{A}$, then $\mathbb{E}[X|\mathcal{G}]^{-1}A \in \mathcal{G}$ by the definition, and so the first condition follows. Then the integral condition

$$\int_{\pi^{-1}A} \mathbb{E}\big[X|\mathcal{G}\big] \circ \pi \,\mathrm{d}\mathbb{P}_1 = \int_A \mathbb{E}\big[X|\mathcal{G}\big] \,\mathrm{d}\mathbb{P}_2 = \int_A X \,\mathrm{d}\mathbb{P}_2 = \int_{\pi^{-1}A} X \circ \pi \,\mathrm{d}\mathbb{P}_1$$

proves that $\mathbb{E}[X|\mathcal{G}]\circ\pi$ is a version of the conditional expectation on the left-hand side. □

1.2 Two Martingale-Coboundary Decompositions

Consider now Λ to be a bounded metric space with distance d_Λ, $\mathscr{B}$ is its Borel σ-algebra, and μ a probability measure on $(\Lambda, \mathscr{B})$. Take $\alpha \in (0,1]$ and define the space of Hölder functions

$$\mathscr{C}^\alpha := \left\{ v \colon \Lambda \to \mathbb{R} \;\Big|\; \sup_{s\neq t} \frac{|v(s)-v(t)|}{d_\Lambda(s,t)^\alpha} < \infty \right\}$$

which is contained in $L^\infty(\Lambda)$ because $diam(\Lambda) < \infty$. It is known that $\mathscr{C}^\alpha$ is a Banach space using the norm

$$\|v\|_\alpha := \|v\|_\infty + \sup_{s\neq t} \frac{|v(s)-v(t)|}{d_\Lambda(s,t)^\alpha} = \|v\|_\infty + \mathcal{H}(v),$$

where $\mathcal{H}(v)$ is the Hölder constant for v. Let $\mathscr{C}^\alpha_0 := \left\{ v \in \mathscr{C}^\alpha | \int_\Lambda v \,\mathrm{d}\mu = 0 \right\}$. Take a measure-preserving transformation $T\colon \Lambda \to \Lambda$ and consider the related Koopman (1.1) and transfer (1.2) operators.

Let us assume two working hypotheses:

(A) $U\mathscr{C}^\alpha \subseteq \mathscr{C}^\alpha$ and there exists $K > 0$ such that $\|Uv\|_\alpha \leq K\|v\|_\alpha$ for every $v \in \mathscr{C}^\alpha$;

(B) $P\mathscr{C}^\alpha \subseteq \mathscr{C}^\alpha$ and there exist $\gamma \in (0,1)$, $C > 0$ such that for every $v \in \mathscr{C}^\alpha_0(\Lambda)$, $n \in \mathbb{N}^+$

$$\|P^n v\|_\alpha \leq C\gamma^n \|v\|_\alpha.$$

Since μ is a probability measure, the inclusions $\mathscr{C}^\alpha \subset L^\infty \subset L^p$ are continuous for every $p \geq 1$, i.e.

$$\|v\|_p \leq \|v\|_\infty \leq \|v\|_\alpha,$$

for every Hölder function v. Furthermore, integral invariance expressed in Remark 1.1 together with (A) and (B) yield that $U\mathscr{C}^\alpha_0 \subseteq \mathscr{C}^\alpha_0$ and $P\mathscr{C}^\alpha_0 \subseteq \mathscr{C}^\alpha_0$.

Assumption (A) is reasonable: as a matter of fact, if T is Lipschitz with constant L, we have for every $v \in \mathscr{C}^\alpha$

$$\big|Uv(x) - Uv(y)\big| = \big|v\big(T(x)\big) - v\big(T(y)\big)\big| \le \mathcal{H}(v)\big|T(x) - T(y)\big|^\alpha \le \mathcal{H}(v)\ L^\alpha\ \big|x - y\big|^\alpha$$

for every $x, y \in \Lambda$. This yields that $\mathcal{H}(Uv) \le L^\alpha\ \mathcal{H}(v)$ and therefore

$$\|Uv\|_\alpha = \|Uv\|_\infty + \mathcal{H}(Uv) \le \|v\|_\infty + L^\alpha\ \mathcal{H}(v) \le \max\big(1, L^\alpha\big)\|v\|_\alpha.$$

By (B) we obtain a first martingale-coboundary decomposition for an observable $v \in \mathscr{C}_0^\alpha$.

Proposition 1.4. Let $v \in \mathscr{C}_0^\alpha$. Then there exist $m, \chi \in \mathscr{C}^\alpha$ such that

$$v = m + \chi \circ T - \chi, \tag{1.3}$$

with $m \in \ker P \cap \mathscr{C}_0^\alpha$. Furthermore there exist constants $C', C'' > 0$ (independent of v) such that

$$\|\chi\|_\alpha \le C'\|v\|_\alpha, \qquad \|m\|_\alpha \le C''\|v\|_\alpha.$$

Proof. Thanks to (B) we know that $\sum_{n=1}^\infty \|P^n v\|_\alpha \le C\|v\|_\alpha \sum_{n=1}^\infty \gamma^n < \infty$. This shows that the sequence $\big\{\sum_{n=1}^k P^n v\big\}_{k>0}$ is Cauchy and so by completeness we define the function

$$\chi \coloneqq \sum_{n=1}^\infty P^n v \in \mathscr{C}^\alpha. \tag{1.4}$$

We now define m to our needs:

$$m \coloneqq v - \chi \circ T + \chi. \tag{1.5}$$

By (A) we get that $\chi \circ T = U\chi \in \mathscr{C}^\alpha$; therefore $m \in \mathscr{C}^\alpha$ because it is a sum of Hölder functions.
Invariance of T yields $m \in \mathscr{C}_0^\alpha$:

$$\int_\Lambda m \,\mathrm{d}\mu = \int_\Lambda v \,\mathrm{d}\mu - \int_\Lambda \chi \circ T \,\mathrm{d}\mu + \int_\Lambda \chi \,\mathrm{d}\mu = 0.$$

By Lemma 1.2 (a) $PU = Id$ and since $\chi \circ T = U\chi$, then $m \in \ker P$:

$$Pm = Pv - PU\chi + P\chi = Pv - \sum_{n=1}^\infty P^n v + \sum_{n=2}^\infty P^n v = 0.$$

To prove the inequalities we use (1.4) and assumption (B) to get

$$\|\chi\|_\alpha \le \sum_{n=1}^\infty \big\|P^n v\big\|_\alpha \le \|v\|_\alpha\ C \sum_{n=1}^\infty \gamma^n \le C'\|v\|_\alpha, \tag{1.6}$$

with $C' = C\sum_{n=1}^{\infty}\gamma^n$. By the triangle inequality applied on (1.5), together with assumption (A) and finally (1.6),

$$\|m\|_\alpha \le \|v\|_\alpha + \|U\chi\|_\alpha + \|\chi\|_\alpha \le \|v\|_\alpha + KC'\|v\|_\alpha + C'\|v\|_\alpha = C''\|v\|_\alpha, \tag{1.7}$$

with $C'' = 1 + C' + KC'$. □

The formula (1.3) is called *martingale-coboundary decomposition* of v and it is a central tool in proving CLT and WIP. Henceforth we refer to m as the *martingale* part, whereas χ is the *coboundary* of v. The name "martingale" is used because $m \in \mathscr{C}_0^\alpha$ can be used to construct a martingale difference array by mean of backward Birkhoff sums, see Proposition 1.9. Denote respectively with v_n and m_n the Birkhoff sums of v and m. Using (1.3) we can write

$$v_n = \sum_{j=0}^{n-1} v\circ T^j = \sum_{j=0}^{n-1} m\circ T^j + \chi\circ T^n - \chi = m_n + \chi\circ T^n - \chi, \tag{1.8}$$

thanks to a telescopic sum. This representation leads to the following

Proposition 1.5. Let $v \in \mathscr{C}_0^\alpha$ and let $m \in \mathscr{C}_0^\alpha$ be defined in (1.5). Let v_n and m_n denote Birkhoff sums and call $\sigma^2 := \int_\Lambda m^2 \,\mathrm{d}\mu$. Then for every $n \in \mathbb{N}^+$

$$\sigma^2 = \frac{1}{n}\int_\Lambda m_n^2 \,\mathrm{d}\mu = \lim_{k\to\infty}\frac{1}{k}\int_\Lambda v_k^2 \,\mathrm{d}\mu. \tag{1.9}$$

Proof. The first equality of (1.9) is a consequence of an *orthogonality* property of m, i.e. for every $0 \le i < j$ we have that

$$\begin{aligned}\int_\Lambda m\circ T^i m\circ T^j \,\mathrm{d}\mu &= \int_\Lambda \big(m\; m\circ T^{j-i}\big)\circ T^i \,\mathrm{d}\mu \\ = \int_\Lambda m\big(m\circ T^{j-i-1}\big)\circ T \,\mathrm{d}\mu &= \int_\Lambda Pm\; m\circ T^{j-i-1} \,\mathrm{d}\mu = 0,\end{aligned}$$

because $m \in \ker P$ by Proposition 1.4. This property cancels out most terms of the following sum which, together with T-invariance, leads to

$$\int_\Lambda m_n^2 \,\mathrm{d}\mu = \int_\Lambda \sum_{j=0}^{n-1}\big(m\circ T^j\big)^2 \mathrm{d}\mu = n\int_\Lambda m^2 \,\mathrm{d}\mu.$$

The first equality of (1.9) follows.

Applying the equality just proven, together with the triangle inequality and T-invariance to (1.8), we get

$$\big|n^{-\frac{1}{2}}\|v_n\|_2 - \sigma\big| = n^{-\frac{1}{2}}\big|\|v_n\|_2 - \|m_n\|_2\big| \le n^{-\frac{1}{2}}\|v_n - m_n\|_2 \le n^{-\frac{1}{2}}2\|\chi\|_2.$$

Therefore $n^{-\frac{1}{2}}\|v_n\|_2 \to \sigma$ for $n \to \infty$, which proves the second equality of (1.9). $\square$

To have some kind of control over m^2, a *second martingale-coboundary decomposition* is useful.

Proposition 1.6. Let $m \in \mathscr{C}_0^\alpha$ and $\sigma^2 = \int_\Lambda m^2 \,\mathrm{d}\mu$. Then there exist $\breve{m}, \breve{\chi} \in \mathscr{C}^\alpha$ such that

$$\mathbb{E}\big[m^2 - \sigma^2 | T^{-1}\mathscr{B}\big] = \breve{m} + \breve{\chi} \circ T - \breve{\chi}, \tag{1.10}$$

with $\breve{m} \in \ker P \cap \mathscr{C}_0^\alpha$. Calling $\breve{v} := \mathbb{E}[m^2 - \sigma^2 | T^{-1}\mathscr{B}]$, we have furthermore that there exist constants $C', C'' > 0$ (independent of $\breve{v}$) such that

$$\|\breve{\chi}\|_\alpha \le C'\|\breve{v}\|_\alpha, \qquad \|\breve{m}\|_\alpha \le C''\|\breve{v}\|_\alpha.$$

Proof. The function m defined in (1.5) is Hölder thanks to Proposition 1.4, hence $m^2 \in \mathscr{C}^\alpha$ because it is a product of Hölder functions, and then $m^2 - \sigma^2 \in \mathscr{C}_0^\alpha$. Therefore (A) and (B) imply that $UP(m^2 - \sigma^2) \in \mathscr{C}_0^\alpha$, which is the left-hand side of (1.10) thanks to Lemma 1.2(b). Hence we can apply Proposition 1.4. $\square$

We use the decomposition (1.10) to write the Birkhoff sum $\breve{v}_n$ as done in (1.8) for v_n:

$$\breve{v}_n = \sum_{j=0}^{n-1} \breve{v} \circ T^j = \sum_{j=0}^{n-1} \breve{m} \circ T^j + \breve{\chi} \circ T^n - \breve{\chi} = \breve{m}_n + \breve{\chi} \circ T^n - \breve{\chi}.$$

1.3 Martingales and Inequalities

Henceforth we show how the terms of both decompositions in the previous section are controlled by the Hölder norm of v. Applying Burkholder's inequality [4], we will see in Corollary 1.11 that Birkhoff sums of the martingale parts grows as $\sqrt{n}$ in every p norm with $p \ge 2$. This is the first step leading to the eventual rate of convergence for the WIP.

1.3.1 Background

Definition 1.7. A sequence of random variables $\{S_n : n \in \mathbb{N}^+\}$ is called a *martingale* with respect to a filtration $\{\mathscr{F}_n : n \in \mathbb{N}^+\}$ if for every $n \in \mathbb{N}^+$:

(i) $S_n \in L^1(\Lambda)$;

(ii) S_n is $\mathscr{F}_n$-measurable;

(iii) $\mathbb{E}\big[S_{n+1} | \mathscr{F}_n\big] = S_n$.

Proposition 1.8. Consider a sequence of random variables $\{d_k\}_{k\in\mathbb{N}^+}$; requiring $S_n = \sum_{k=1}^n d_k$ to be a martingale wrt a filtration is equivalent to saying that

(a) $\mathbb{E}|d_k| < \infty$ for every $k \in \mathbb{N}^+$;

(b) d_k is $\mathscr{F}_k$-measurable for every $k \in \mathbb{N}^+$;

(c) $\mathbb{E}\big[d_{k+1}|\mathscr{F}_k\big] = 0$ for every $k \in \mathbb{N}^+$.

Proof. **(a)⇒(i)**: By the triangle inequality (a) implies $\mathbb{E}|S_n| < \infty$ for every $n \in \mathbb{N}^+$.
(b)⇒(ii): Assuming (b) we have that S_n is $\mathscr{F}_n$-measurable because it is a sum of $\mathscr{F}_n$-measurable random variables.
(i),(ii)⇒(a),(b): On the contrary, for $k = 1$ the result is trivial for $S_1 = d_1$; for $k \geq 2$ since S_k is integrable and $\mathscr{F}_k$-measurable for every k, so is $d_k = S_k - S_{k-1}$, because it is a sum of two integrable and $\mathscr{F}_k$-measurable random variables, proving (a) and (b).
(c)⇔(iii): We have for every $n \in \mathbb{N}^+$

$$S_n = \mathbb{E}\big[S_n|\mathscr{F}_n\big] + \mathbb{E}\big[d_{n+1}|\mathscr{F}_n\big] = S_n + \mathbb{E}\big[d_{n+1}|\mathscr{F}_n\big] = \mathbb{E}\big[S_{n+1}|\mathscr{F}_n\big]$$

if and only if $\mathbb{E}\big[d_{n+1}|\mathscr{F}_n\big] = 0$. □

Such $\{d_k\}_{k\in\mathbb{N}^+}$ is sometimes called a *martingale difference sequence* (briefly MDS).

Our work does not focus directly on a MDS but, because of technical reasons, we deal with a *martingale difference array* (briefly MDA). The latter is a collection of random variables $\{\xi_{n,k}\}_{0\leq k\leq n}$ together with a family of σ-algebras $\{\mathcal{G}_{n,k}\}_{0\leq k\leq n}$, increasing for every n, such that for every $n \in \mathbb{N}^+$ and $1 \leq k \leq n$:

- $\xi_{n,k} \in L^1(\Lambda)$;
- $\xi_{n,k}$ is $\mathcal{G}_{n,k}$-measurable;
- $\mathbb{E}[\xi_{n,k}|\mathcal{G}_{n,k-1}] = 0$.

Proposition 1.9. Let $m \in \ker P \cap L^p$ with $p \geq 2$. Call $\sigma^2 = \int_\Lambda m^2 \,\mathrm{d}\mu$ and define

$$\xi_{n,k} := \frac{1}{\sigma\sqrt{n}} m \circ T^{n-k} \qquad \mathcal{G}_{n,k} := T^{-(n-k)}\mathscr{B} \tag{1.11}$$

for $n \in \mathbb{N}$ and $0 \leq k \leq n$, where $\mathscr{B}$ is the Borel σ-algebra of Λ.
Then $\xi_{n,k}$ is a MDA wrt $\mathcal{G}_{n,k}$.

Proof. First of all we see that the σ-algebras are increasing: measurability of T implies

$$T^{-n}\mathscr{B} \subseteq T^{-n+1}\mathscr{B} \subseteq \cdots \subseteq T^{-1}\mathscr{B} \subseteq \mathscr{B}.$$

We have trivially that $\xi_{n,k}$ is integrable and $\mathcal{G}_{n,k}$-measurable for every n and k. Finally switching T with the conditional expectation by Lemma 1.3 we get

$$\sigma\sqrt{n}\,\mathbb{E}\big[\xi_{n,k+1}|\mathcal{G}_{n,k}\big] = \mathbb{E}\big[m \circ T^{n-(k+1)}|T^{n-k}\mathscr{B}\big] = \mathbb{E}\big[m|T^{-1}\mathscr{B}\big] \circ T^{n-(k+1)} = 0,$$

because $\mathbb{E}\big[m|T^{-1}\mathscr{B}\big]= UP(m)$ by Lemma 1.2(b), and $m \in \ker P$. $\square$

1.3.2 Inequalities

We state some inequalities that apply to $v \in \mathscr{C}_0^\alpha$ and its related martingale-coboundary decompositions (1.3), (1.10). Denote $v_n, \breve{v}_n, m_n, \breve{m}_n$ as the corresponding Birkhoff sums, for example $v_n := \sum_{j=0}^{n-1} v \circ T^j$.

Theorem 1.10 (Burkholder's inequality [4])**.** *Let $p \geq 2$ and suppose that $S_n = \sum_{k=1}^n d_k$ is a martingale with $d_k \in L^p$ for $k \geq 1$. Then there is a universal constant $C_p > 0$ such that*

$$\big\|\max_{1\leq k\leq n} |S_k|\big\|_p \leq C_p\big(\max_{1\leq k\leq n} \|d_k\|_p\big)\sqrt{n}$$

for all $n \geq 1$.

Corollary 1.11. *Let $p \geq 2$. There exists $C_p' > 0$ such that*

$$\bigg\|\max_{0\leq k\leq n} |m_k|\bigg\|_p \leq C_p' \|m\|_p \sqrt{n}$$

for every $m \in \ker P \cap L^p$ and $n \in \mathbb{N}^+$.

Proof. By Proposition 1.9 the array $\xi_{n,k} = m \circ T^{n-k}$ is an MDA. It follows that for fixed $n \in \mathbb{N}^+$ we can define a MDS setting $\xi_{n,k} = 0$ for every $k > n$. This generates a martingale to which Burkholder's inequality can be applied:

$$\bigg\|\max_{0\leq k\leq n}\bigg|\sum_{j=0}^{n} \xi_{n,k}\bigg|\bigg\|_p \leq C_p\Big(\max_{1\leq k\leq n}\big\|\xi_{n,k}\big\|_p\Big)\sqrt{n}. \tag{1.12}$$

We note that:

$$m_k = \sum_{j=0}^{k-1} m \circ T^j = \sum_{j=0}^{n} m \circ T^{n-j} - \sum_{j=0}^{n-k} m \circ T^{(n-k)-j} = \sum_{j=0}^{n} \xi_{n,j} - \sum_{j=0}^{n-k} \xi_{n-k,j}$$

for every $k \in \{0, \ldots, n\}$, which yields

$$\max_{0\leq k\leq n} |m_k| \leq 2 \max_{0\leq k\leq n}\bigg|\sum_{j=0}^{n-k} \xi_{n,j}\bigg| = 2 \max_{0\leq k\leq n}\bigg|\sum_{j=0}^{n} \xi_{n,j}\bigg|.$$

Applying p-norms on both sides we see that

$$\big\|\max_{0\leq k\leq n} |m_k|\big\|_p \leq 2\bigg\|\max_{0\leq k\leq n}\bigg|\sum_{j=0}^{n} \xi_{n,k}\bigg|\bigg\|_p \tag{1.13}$$

which is the left-hand side of (1.12).

On the other side, $\left\|\xi_{n,k}\right\|_p = \|m\|_p$ by T-invariance; therefore, applying (1.12) to (1.13) gives

$$\left\|\max_{0\le k\le n}|m_k|\right\|_p \le 2C_p\|m\|_p\sqrt{n}.$$

The same procedure can be applied for every $n \in \mathbb{N}^+$ with the same universal constant C_p from Burkholder's inequality. Hence we can conclude taking $C_p' = 2C_p$.

□

Proposition 1.12. Let $p \ge 2$. Then there exists $K_p > 0$ such that

$$\left\|\max_{0\le k\le n}|v_k|\right\|_p \le K_p\|v\|_\alpha\sqrt{n}$$

for every $v \in \mathscr{C}_0^\alpha$ and $n \in \mathbb{N}^+$.

Proof. By decomposition (1.3) and the triangle inequality we have that

$$\left\|\max_{0\le k\le n}|v_k|\right\|_p = \left\|\max_{0\le k\le n}\left|m_k+\chi\circ T^k-\chi\right|\right\|_p \le \left\|\max_{0\le k\le n}|m_k|\right\|_p + \left\|\max_{0\le k\le n}\left|\chi\circ T^k-\chi\right|\right\|_p. \tag{1.14}$$

Since $\chi \in L^\infty(\Lambda)$, then so is $\max_{0\le k\le n}|\chi\circ T^k|$, and therefore

$$\left\|\max_{0\le k\le n}\left|\chi\circ T^k-\chi\right|\right\|_p \le \left\|\max_{0\le k\le n}\left|\chi\circ T^k\right|\right\|_\infty + \|\chi\|_\infty \le 2\|\chi\|_\infty \tag{1.15}$$

Using the fact that $m \in \ker P$ we can apply Corollary 1.11 to (1.14), together with (1.15). Since $\|\chi\|_\infty \le \|\chi\|_\alpha$ and $\|m\|_p \le \|m\|_\alpha$, we conclude by Proposition 1.4 that

$$\left\|\max_{0\le k\le n}|v_k|\right\|_p \le C_p\|m\|_p\sqrt{n}+2\|\chi\|_\infty \le C_pC''\|v\|_\alpha\sqrt{n}+2C'\|v\|_\alpha = \left(C_pC''+\frac{2C'}{\sqrt{n}}\right)\|v\|_\alpha\sqrt{n}.$$

The statement is hence proved taking $K_p = C_pC'' + 2C'$.

□

Proposition 1.13. Let $p \ge 2$. There exists $\breve{K}_p > 0$ such that

$$\left\|\max_{0\le k\le n}\left|\sum_{j=0}^{k-1}\mathbb{E}\left[m^2-\sigma^2|T^{-1}\mathscr{B}\right]\circ T^j\right|\right\|_p \le \breve{K}_p\|v\|_\alpha^2\sqrt{n},$$

for every $v \in \mathscr{C}_0^\alpha$ and $n \in \mathbb{N}^+$, where $m = m(v)$ is defined in (1.5) and $\sigma^2 = \int_\Lambda m^2\,\mathrm{d}\mu$.

Proof. We notice that Corollary 1.11 can be applied to the martingale part $\breve{m} \in \ker P$. By Proposition Proposition 1.12, taking $\breve{v} = \mathbb{E}\left[m^2-\sigma^2|T^{-1}\mathscr{B}\right]$ in place of v we find some $K_p > 0$ such that

$$\left\|\max_{0\le k\le n}|\breve{v}_k|\right\|_p \le K_p\|\breve{v}\|_\alpha\sqrt{n}. \tag{1.16}$$

We are left to estimate the Hölder norm of the conditional expectation $\breve{v}$ which can be written as $UP(m^2-\sigma^2)$ from Lemma 1.2(b). Using together assumptions (A) and (B) (because $m^2-\sigma^2 \in \mathscr{C}_0^\alpha$) we get

$$\|\breve{v}\|_\alpha = \big\|UP\big(m^2-\sigma^2\big)\big\|_\alpha \le K\big\|P\big(m^2-\sigma^2\big)\big\|_\alpha \le KC\gamma\big\|m^2-\sigma^2\big\|_\alpha, \tag{1.17}$$

for constants $K, C > 0$ and $\gamma \in (0,1)$. Propositions A.1 and A.2 give us that $\|m^2\|_\infty \le \|m\|_\infty^2$ and $\mathcal{H}(m^2) \le 2\|m\|_\infty \mathcal{H}(m)$; therefore by the triangle inequality

$$\begin{aligned} \|m^2-\sigma^2\|_\alpha &\le \|m^2\|_\alpha + \|\sigma^2\|_\alpha = \|m^2\|_\infty + \mathcal{H}(m^2) + \int_\Lambda m^2 \,\mathrm{d}\mu \\ &\le \|m\|_\infty^2 + 2\|m\|_\infty \mathcal{H}(m) + \|m\|_\infty^2 = 2\|m\|_\infty\big(\|m\|_\infty + \mathcal{H}(m)\big) \\ &\le 2\|m\|_\alpha^2 \le 2C''\|v\|_\alpha^2, \end{aligned} \tag{1.18}$$

using the second inequality of Proposition 1.4. Finally, applying (1.17) and (1.18) to (1.16) we get the statement with $\breve{K}_p = 2K_pKCC''\gamma$. □

For a $v \in \mathscr{C}_0^\alpha$ we take its martingale part $m \in \ker P \cap \mathscr{C}_0^\alpha$ found in (1.5) and obtain a MDA $\xi_{n,k}$ like in (1.11). Our aim is to use Proposition 1.13 to estimate the second moments of $\xi_{n,k}$. Let us define the array $V_{n,k}$ for $n \in \mathbb{N}$ and $0 \le k \le n$,

$$V_{n,k} \coloneqq \sum_{j=1}^{k} \mathbb{E}\big[\xi_{n,j}^2 | \mathcal{G}_{n,j-1}\big], \tag{1.19}$$

where $\mathcal{G}_{n,k} \coloneqq T^{-(n-k)}\mathscr{B}$. By Lemma 1.3 we can rewrite $V_{n,k}$ as

$$V_{n,k} = \frac{1}{\sigma^2 n}\sum_{j=1}^{k} \mathbb{E}\big[m^2 \circ T^{n-j} | T^{-(n-j+1)}\mathscr{B}\big] = \frac{1}{\sigma^2 n}\sum_{j=1}^{k} \mathbb{E}\big[m^2 | T^{-1}\mathscr{B}\big] \circ T^{n-j}. \tag{1.20}$$

Proposition 1.14. Fix $p \ge 2$. Then there exists $C_p > 0$ such that

$$\Big\|\max_{0\le k\le n}\Big|V_{n,k} - \frac{k}{n}\Big|\Big\|_p \le \frac{C_p\|v\|_\alpha^2}{\sigma^2\sqrt{n}},$$

for every $v \in \mathscr{C}_0^\alpha$ and $n \in \mathbb{N}^+$, where $V_{n,k}$ is the array defined in (1.19).

Proof. Denote with $\breve{v} = \mathbb{E}[m^2 - \sigma^2 | T^{-1}\mathscr{B}]$ and with $\breve{v}_n$ its Birkhoff sum. By formula (1.20) and letting j start from 0 we get

$$V_{n,k} - \frac{k}{n} = \frac{1}{\sigma^2 n}\sum_{j=0}^{k-1} \breve{v} \circ T^{n-j-1} = \frac{1}{\sigma^2 n}\big(\breve{v}_n - \breve{v}_{n-k}\big).$$

Therefore $|V_{n,k} - \frac{k}{n}| \le \sigma^{-2}n^{-1}(|\breve{v}_n| + |\breve{v}_{n-k}|)$. Passing to the max, applying p-norms and using Proposition 1.13, we get

$$\left\| \max_{0\le k\le n} \left| V_{n,k} - \frac{k}{n} \right| \right\|_p \le \frac{2}{\sigma^2 n} \left\| \max_{0\le k\le n} |\breve{v}_k| \right\|_p \le \frac{2\breve{K}_p \|v\|_\alpha^2}{\sigma^2 \sqrt{n}}.$$

We conclude taking $C_p = 2\breve{K}_p$.

□

1.4 Rates of Convergence

The following section regards the convergence rate in the Prokhorov metric of a sequence of probability measures on $\mathscr{C}[0,1]$ to a Wiener measure. The sequence of laws is defined by pushforward measures generated by deterministic processes $\{W_n\}_{n\in\mathbb{N}^+}$, which are obtained by iteration of a dynamical system T. We always require that the Koopman and transfer operators of T satisfy the assumptions (A) and (B) displayed in section 1.2. The methods and techniques adopted follow the paper of Antoniou and Melbourne [2], which proved a similar rate for non-uniformly expanding dynamical systems. This general setting doesn't always satisfy our working assumptions, especially the preservation of Hölder property by the transfer operator P. Indeed our advantage is that the various term of the martingale-coboundary decomposition (1.3) are in $\mathscr{C}^\alpha$, and so they lie in every L^p because of boundedness of Λ.

Let $v \in \mathscr{C}_0^\alpha$ with Birkhoff sums v_n for $n \in \mathbb{N}^+$. We define $W_n \in \mathscr{C}[0,1]$ as the function such that

$$W_n(t) = \frac{v_{nt}}{\sqrt{n}} \qquad \text{for } t = \frac{j}{n},\ \ 0 \le j \le n, \tag{1.21}$$

with linear interpolation between the nodes j/n. We show that the law induced by W_n (starting from $\mu \in \mathscr{M}_1(\Lambda)$) on the space of continuous functions, converges weakly to a Wiener measure induced by a Brownian motion with variance $\sigma^2 = \int_\Lambda m^2 \,\mathrm{d}\mu$, where m is the martingale part of v defined in (1.5). We estimate the rate of convergence through a distance defined on the space of probability measures $\mathscr{M}_1\big(\mathscr{C}[0,1]\big)$.

Definition 1.15. Given X and Y random variables taking values into a metric space (S, d_S), we define the *Prokhorov distance* π_1 on the space of probability measures $\mathscr{M}_1(S)$:

$$\pi_1(X,Y) = \pi_1\big(\mathscr{L}(X), \mathscr{L}(Y)\big) := \inf\big\{\varepsilon > 0 \mid \mathbb{P}\big(X \in A\big) \le \mathbb{P}\big(Y \in A^\varepsilon\big) + \varepsilon \text{ for all closed } A \subset S\big\},$$

where A^ε is the ε-neighbourhood of A,

$$A^\varepsilon := \bigcup_{f\in A} \{g \in S \colon d_S(f,g) < \varepsilon\}.$$

We take $S = \mathscr{C}[0,1]$ with d_S the uniform distance given by the infinity norm.

Theorem 1.16 ([2, Theorem 2.2])**.** *Suppose $v \in \mathscr{C}_0^\alpha$ and consider $W_n \in \mathscr{C}[0,1]$ as defined in (1.21). Then for every $\gamma > 0$ there is a constant $C > 0$ such that*

$$\pi_1(W_n, W) \leq Cn^{-(\frac{1}{4}-\gamma)},$$

for all $n \in \mathbb{N}^+$.

1.4.1 Preliminaries

Before proving Theorem 1.16, we need first to show the same rate of convergence for a sequence of processes $\{X_n \in \mathscr{C}[0,1] \colon n \in \mathbb{N}^+\}$ defined similarly to W_n. Fix $v \in \mathscr{C}_0^\alpha$ and consider the related MDAs $\xi_{n,k}$ found in (1.11) and $V_{n,k}$ found in (1.19). For $n \in \mathbb{N}^+$, we partition the interval $[0,1]$ into $n+1$ random nodes

$$t_k \coloneqq \frac{V_{n,k}}{V_{n,n}}, \qquad 0 = t_0 < t_1 < \cdots < t_n = 1,$$

using the notation $V_{n,0} \coloneqq 0$. We define the process $X_n \colon \Lambda \to \mathscr{C}[0,1]$,

$$X_n(t) = \sum_{j=1}^{k} \xi_{n,j} = \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{k} m \circ T^{n-j}, \qquad \text{for } t = \frac{V_{n,k}}{V_{n,n}}, \tag{1.22}$$

with linear interpolation in between. We see that $X_n(0) = 0$ and $X_n(1) = \sum_{j=1}^n \xi_{n,j}$.

The central tool from probability theory to estimate the Prokhorov distance between X_n and B, a standard Brownian motion on $\mathscr{C}[0,1]$, is the following.

Theorem 1.17 (Kubilius [11, Theorem 1])**.** *Let X_n be the process defined in (1.22) by using $\xi_{n,k}$ of (1.11) and take $\delta \in [0, \frac{3}{4}] \cup \{1\}$. There is a constant $C > 0$ such that*

$$\pi_1(X_n, B) \leq C\lambda |\log \lambda|,$$

where $\lambda = \lambda_1 + \lambda_2$ and

$$\lambda_1 = \inf_{0 \leq \varepsilon \leq 1} \left\{ \varepsilon^{\frac{1}{2}} + \mathbb{E}\left[\sum_{j=1}^{n} |\xi_{n,j}|^{2+2\delta} \mathbb{1}_{\{|\xi_{n,j}| > \varepsilon\}} \right]^{\frac{1}{3+2\delta}} \right\};$$
$$\lambda_2 = \inf_{0 \leq \varepsilon \leq 1} \left\{ \varepsilon + \mathbb{P}\left(|V_{n,n} - 1| > \varepsilon^2 \right) \right\}.$$

Notation. Henceforth we will use the notation $a_n \ll b_n$ for two positive sequences if there exists a constant $C > 0$ such that $a_n \leq C b_n$ for every $n \in \mathbb{N}^+$.

Lemma 1.18 ([2, Lemma 4.3]). *Let $v \in \mathscr{C}_0^\alpha$ and let X be the process defined in (1.22). Then for every $\gamma > 0$ there exists $C > 0$ such that*

$$\pi(X_n, B) \leq C n^{-(\frac{1}{4}-\gamma)},$$

for every $n \in \mathbb{N}^+$.

Proof. Fix $\gamma > 0$ and choose $p \geq 4$ such that $\frac{1}{2p} \leq \gamma$. We are going to show that there exists $C > 0$ such that $\pi(X_n, B) \leq Cn^{-r}$, with $r = \frac{p-2}{4p}$. For this purpose, we consider λ_1 and λ_2 defined in Theorem 1.17 as function of n and prove that

$$\lambda_1 \ll \|v\|_\alpha^{r'} n^{-r_1} \qquad\qquad \lambda_2 \ll \|v\|_\alpha^{r''} n^{-r_2}, \tag{1.23}$$

for some $r', r'' > 0$, $r_1 = \frac{p-2}{4p-6}$, and $r_2 = \frac{p}{4p+2}$.

Assuming (1.23) true we have that $\lambda_2 \ll n^{-r_1}$ (since $r_2 \geq r_1$), and hence $\lambda = \lambda_1 + \lambda_2 \ll n^{-r_1}$. Applying Theorem 1.17 we get that

$$\pi_1(X_n, B) \ll \lambda\,|\log \lambda| \ll n^{-r_1} |\log n^{-r_1}| \ll n^{-r_1} \log n, \tag{1.24}$$

because the function $x \mapsto x|\log x|$ is increasing in $(0, e^{-1}]$ and for n large enough

$$0 < \lambda \leq \text{const } n^{-r_1} \leq e^{-1}.$$

Because $r_1 > r$, we can choose $K > 0$ such that $\log n \leq Kn^{r_1 - r}$ for every $n \in \mathbb{N}^+$: therefore $n^{-r_1} \log n \leq Kn^{-r}$ for every $n \in \mathbb{N}^+$ and the statement of the theorem is proven.

Estimate for $\boldsymbol{\lambda_1}$ – Fix $\delta = 1$. By T-invariance we evaluate

$$\mathbb{E}_\mu\left[\sum_{j=1}^n |\xi_{n,j}|^4 \, \mathbb{1}_{\{|\xi_{n,j}| > \varepsilon\}}\right] = \frac{1}{\sigma^4\, n} \mathbb{E}_\mu\left[|m|^4 \mathbb{1}_{\{|m| \geq \varepsilon\sigma\sqrt{n}\}}\right], \tag{1.25}$$

for any $\varepsilon \in [0, 1]$. We know that the martingale part $m \in \mathscr{C}_0^\alpha \subset L^p$ and $|m|^4 \in L^{p/4}$. Since $\frac{p}{4}$ and $\frac{p}{p-4}$ are conjugate exponents, we can apply Hölder inequality and obtain

$$\begin{aligned} \mathbb{E}_\mu\left[|m|^4 \mathbb{1}_{\{|m| \geq \varepsilon\sigma\sqrt{n}\}}\right] &\leq \|m\|_p^4 \mu\big(|m| \geq \varepsilon\sigma\sqrt{n}\big)^{\frac{p-4}{p}} \\ &\leq \|m\|_p^4 \left(\frac{\|m\|_p^p}{\varepsilon^p \sigma^p n^{p/2}}\right)^{\frac{p-4}{p}} = \frac{\|m\|_p^p}{\big(\varepsilon\sigma\sqrt{n}\big)^{p-4}} \end{aligned} \tag{1.26}$$

where we have applied Markov's Inequality $\mu[|m| > a] = \mu[|m|^p > a^p] \leq \frac{\mathbb{E}|m|^p}{a^p}$, with

$a = \varepsilon\sigma\sqrt{n} > 0$. Combining (1.25) and (1.26) we obtain

$$\mathbb{E}_\mu\left[\sum_{j=1}^{n} |\xi_{n,j}|^4 \, \mathbb{1}_{\{|\xi_{n,j}|>\varepsilon\}}\right] \leq \frac{\|m\|_p^p}{\sigma^p \; n^{\frac{p}{2}-1}\varepsilon^{p-4}}$$

that used in the definition of λ_1 yields

$$\lambda_1 \leq \inf_{0\leq\varepsilon\leq 1}\left\{\varepsilon^{\frac{1}{2}} + \left[\frac{\|m\|_p^p}{\sigma^p \; n^{\frac{p}{2}-1}\varepsilon^{p-4}}\right]^{\frac{1}{5}}\right\}.$$

We have that $\varepsilon^{\frac{5}{2}} = \frac{\|m\|_p^p}{\sigma^p \; n^{p/2-1}\varepsilon^{p-4}}$ when

$$\varepsilon = \varepsilon_* := \left(\frac{\|m\|_p}{\sigma n^{(p-2)/2p}}\right)^{\frac{2p}{2p-3}} = \left(\frac{\|m\|_p}{\sigma}\right)^{\frac{2p}{2p-3}} n^{(p-2)/(2p-3)} = \left(\frac{\|m\|_p}{\sigma}\right)^{\frac{2p}{2p-3}} n^{-2r_1}.$$

Note that $\varepsilon_* \in [0,1]$ for n large enough. Therefore, by the second inequality of Proposition 1.4 we get

$$\lambda_1 \ll \varepsilon_*^{\frac{1}{2}} = \left(\frac{\|m\|_p}{\sigma}\right)^{\frac{p}{2p-3}} n^{-r_1} \ll \|v\|_\alpha^{p/(2p-3)} n^{-r_1}.$$

Call $r' := \frac{p}{2p-3} > 0$ to conclude the first inequality of (1.23).

Estimate for $\boldsymbol{\lambda_2}$ – By Markov's inequality and then Proposition 1.14,

$$\mu\left(|V_{n,n}-1| > \varepsilon^2\right) = \mu\left(|V_{n,n}-1|^p > \varepsilon^{2p}\right) \leq \frac{\left\|V_{n,n}-1\right\|_p^p}{\varepsilon^{2p}} \ll \left(\frac{\|v\|_\alpha^2}{\varepsilon^2\sqrt{n}}\right)^p.$$

Therefore by definition of λ_2 we can write

$$\lambda_2 \ll \inf_{0\leq\varepsilon\leq 1}\left\{\varepsilon + \frac{\|v\|_\alpha^{2p}}{\varepsilon^{2p} \; n^{\frac{p}{2}}}\right\}.$$

We have that $\varepsilon = \frac{\|v\|_\alpha^{2p}}{\varepsilon^{2p} \; n^{p/2}}$ when

$$\varepsilon = \varepsilon^* := \left(\frac{\|v\|_\alpha^2}{\sqrt{n}}\right)^{\frac{p}{2p+1}} = \|v\|_\alpha^{2p/(2p+1)} n^{-r_2}.$$

Note that $\varepsilon^* \in [0,1]$ for n large enough. In conclusion,

$$\lambda_2 \ll \varepsilon^* = \|v\|_\alpha^{2p/(2p+1)} n^{-r_2}.$$

Call $r'' := \frac{2p}{2p+1} > 0$ to prove the second inequality of (1.23). □

The process X_n in (1.22) is determined on the nodes $t_k = \frac{V_{n,k}}{V_{n,n}} \in [0,1]$, which position depend on Λ. The array $V_{n,j} \geq 0$ defined in (1.19) is non-decreasing in j; therefore, for fixed $k \in \{0, \dots, n\}$ we have that

$$V_{n,k} = \frac{V_{n,k}}{V_{n,n}} V_{n,n} \le tV_{n,n} \le V_{n,k+1},$$

for every $t \in [t_k, t_{k+1}]$. On the other hand, given $t \in [0,1]$, the integer $k = k_t \in \{0, \dots, n\}$ satisfies the relation

$$V_{n,k} \le tV_{n,n} \le V_{n,k+1}. \tag{1.27}$$

The following proposition, found in [2, Proposition 4.4] gives us a control over the growth of k as a function of n and t.

Proposition 1.19. Let k be defined as above. For every $p \ge 2$ there exists $C > 0$ such that

$$\Big\| \sup_{t\in[0,1]} |k - \lfloor nt \rfloor| \Big\|_p \le C\sqrt{n},$$

where we have used the floor function $\lfloor nt \rfloor := \sup\{s \in \mathbb{N} : s \le nt\}$.

Proof. Fix $p \ge 2$. Writing $\widetilde{V}_{n,k} := nV_{n,k} - k$, inequality (1.27) give us that

$$n\widetilde{V}_{n,k} + k \le t\widetilde{V}_{n,n} + nt \le \widetilde{V}_{n,k+1} + k + 1,$$

for every $t \in [0,1]$ and $n \in \mathbb{N}^+$. Applying the latter we get

$$\begin{aligned} k - nt &\le t\widetilde{V}_{n,n} - \widetilde{V}_{n,k} \le 2\max_{j\le n} |\widetilde{V}_{n,j}| \\ k - nt &\ge t\widetilde{V}_{n,n} - \widetilde{V}_{n,+k1} - 1 \ge -2\max_{j\le n} |\widetilde{V}_{n,j}| - 1. \end{aligned}$$

These inequalities imply that

$$|k - \lfloor nt \rfloor| \le |k - nt| + 1 \le 2\max_{j\le n+1} |\widetilde{V}_{n,j}| + 2 = 2n \max_{j\le n+1} \big| V_{n,j} - \frac{j}{n} \big| + 2.$$

Hence by Proposition 1.14

$$\Big\| \sup_{t\in[0,1]} |k - \lfloor nt \rfloor| \Big\|_p \le 2n \Big\| \max_{j\le n+1} \big| V_{n,j} - \frac{j}{n} \big| \Big\|_p + 2 \ll n\|v\|_\alpha^2 n^{-\frac{1}{2}} + 1 \ll n^{\frac{1}{2}}.$$

□

1.4.2 Passing to W_n

The strategy to complete the proof of Theorem 1.16 is as follows: first of all, we consider Proposition 1.20 from probability theory to estimate the Prokhorov distance between two random paths; then we use it and Proposition 1.21 to prove Lemma 1.22; finally, using the two lemmas together, we conclude the proof of the main theorem.

Proposition 1.20 ([2, Proposition 4.5]). Let $Y, Y' \in \mathscr{C}[0,1]$ be random paths defined on a common probability space. Let $\varepsilon_0, \varepsilon_1 > 0$ and $q \geq 1$. Then

(a) If $\mathbb{P}\big(\|Y - Y'\|_\infty \geq \varepsilon_0\big) \leq \varepsilon_1$, then $\pi_1(Y, Y') \leq \max\{\varepsilon_0, \varepsilon_1\}$.

(b) If $\big\|\|Y - Y'\|_\infty\big\|_q \leq \varepsilon_0$, then $\pi_1(Y, Y') \leq \varepsilon_0^{q/(q+1)}$.

Proof. (a) By Definition 1.15 it is sufficient to prove that for every $A \subset [0,1]$ closed

$$\mathbb{P}\big(Y \in A\big) \leq \mathbb{P}\big(Y' \in A^{\varepsilon_0 \vee \varepsilon_1}\big) + \varepsilon_0 \vee \varepsilon_1,$$

using the notation $\varepsilon_0 \vee \varepsilon_1 = \max\{\varepsilon_0, \varepsilon_1\}$. By the hypothesis we have

$$\begin{aligned}\mathbb{P}\big(Y \in A\big) - \varepsilon_0 \vee \varepsilon_1 &\leq \mathbb{P}\big(Y \in A\big) - \mathbb{P}\big(\|Y - Y'\|_\infty \geq \varepsilon_0\big) \leq \mathbb{P}\big(Y \in A, \|Y - Y'\|_\infty < \varepsilon_0\big) \\ &\leq \mathbb{P}\big(Y \in A, Y' \in A^{\varepsilon_0}\big) \leq \mathbb{P}\big(Y' \in A^{\varepsilon_0}\big) \leq \mathbb{P}\big(Y' \in A^{\varepsilon_0 \vee \varepsilon_1}\big),\end{aligned}$$

where we have used the fact $\mathbb{P}(E) - \mathbb{P}(F^c) \leq \mathbb{P}(E \cap F)$ for every pair of events in $\mathscr{C}[0,1]$.

(b) By Markov's inequality and the assumption we know that for every $\delta > 0$

$$\mathbb{P}\big(\|Y - Y'\|_\infty \geq \delta\big) = \mathbb{P}\big(\|Y - Y'\|_\infty^q \geq \delta^q\big) \leq \frac{\big\|\|Y - Y'\|_\infty\big\|_q^q}{\delta^q} \leq \frac{\varepsilon_0^q}{\delta^q}.$$

In particular, if δ is such that $\delta = (\varepsilon_0/\delta)^q$, i.e. $\delta = \varepsilon_0^{q/(q+1)}$, then we have that

$$\mathbb{P}\big(\|Y - Y'\|_\infty \geq \delta\big) \leq \delta = \varepsilon_0^{q/(q+1)},$$

and the results follows from (a).

□

Proposition 1.21 ([2, Proposition 4.6]). Define for $v \in \mathscr{C}^\alpha$ and $n \in \mathbb{N}^+$

$$Z_n := \max_{0 \leq i, \ell \leq \lfloor \sqrt{n} \rfloor} \left| \sum_{j=0}^{\ell} v \circ T^j \right| \circ T^{i\lfloor \sqrt{n} \rfloor}.$$

Fix $a, b \in \mathbb{N}$ such that $a \leq b \leq n$. For every $p \geq 2$ there exists $C_p > 0$ such that

$$\left| \sum_{j=a}^{b-1} v \circ T^j \right| \leq Z_n \left(\frac{b-a}{\sqrt{n} - 1} + 3 \right) \qquad \text{and} \qquad \|Z_n\|_p \leq C_p \|v\|_\alpha \, n^{\frac{1}{4} + \frac{1}{2p}},$$

for all $v \in \mathscr{C}^\alpha$, $n \in \mathbb{N}^+$.

Proof. Fix a, b as above and $n \in \mathbb{N}^+$. Choose $\ell_1, \ell_2 \in \mathbb{N}$ such that $0 \leq \ell_1 \leq \ell_2 \leq \lfloor \sqrt{n} \rfloor$, greatest such that $\ell_1 \lfloor \sqrt{n} \rfloor \leq a$ and $\ell_2 \lfloor \sqrt{n} \rfloor \leq b$. It follows that $(\ell_1 + 1)\lfloor \sqrt{n} \rfloor > a$ which yields $-\ell_1 < -a/\lfloor\sqrt{n}\rfloor + 1$. Then

$$\ell_2 - \ell_1 \le \frac{b-a}{\lfloor\sqrt{n}\rfloor} + 1 \le \frac{b-a}{\sqrt{n}-1} + 1.$$

It follows that

$$\left|\sum_{j=a}^{b-1} v \circ T^j\right| \le (\ell_2 - \ell_1 + 2)\, Z_n \le \left(\frac{b-a}{\sqrt{n}-1} + 3\right) Z_n$$

that proves the first inequality.

Using the definition of Z_n and the fact that T is measure-preserving

$$\begin{aligned}\int_\Lambda |Z_n|^p \, d\mu \le \sum_{i=0}^{\lfloor\sqrt{n}\rfloor} \int_\Lambda \Big(\max_{0\le\ell\le\lfloor\sqrt{n}\rfloor} |v_\ell|\Big)^p \circ T^{i\lfloor\sqrt{n}\rfloor}\, \mathrm{d}\mu = \sum_{i=0}^{\lfloor\sqrt{n}\rfloor} \int_\Lambda \max_{0\le\ell\le\lfloor\sqrt{n}\rfloor} |v_\ell|^p \, \mathrm{d}\mu \\ \le \sqrt{n} \int_\Lambda \max_{0\le\ell\le\lfloor\sqrt{n}\rfloor} |v_\ell|^p \, \mathrm{d}\mu.\end{aligned}$$

Raising to the power of $\frac{1}{p}$ we obtain

$$\|Z_n\|_{p)} \le n^{\frac{1}{2p}} \Big\| \max_{0\le\ell\le\lfloor\sqrt{n}\rfloor} |v_\ell| \Big\|_p \ll n^{\frac{1}{2p}} \lfloor\sqrt{n}\rfloor^{\frac{1}{2}} \|v\|_\alpha \le n^{\frac{1}{4}+\frac{1}{2p}} \|v\|_\alpha,$$

applying finally Proposition 1.12. □

Let us define the linear operator

$$g\colon \mathscr{C}[0,1] \to \mathscr{C}[0,1], \qquad (gu)(t) := u(1) - u(1-t). \tag{1.28}$$

Lemma 1.22 ([2, Lemma 4.7])**.** *Fix $v \in \mathscr{C}_0^\alpha$ and consider σ^2 defined in (1.9). Let W_n be as in (1.21) and X_n in (1.22). For every $\gamma > 0$ there exists $C > 0$ such that*

$$\pi_1(g \circ W_n, \sigma X_n) \le C n^{-(\frac{1}{4}-\gamma)}.$$

Proof. Fix $\gamma > 0$ and choose $p \ge 2$ such that $\frac{1}{2p} \le \gamma$. We are going to show that there exists $C > 0$ such that $\pi_1(g \circ W_n, \sigma X_n) \le C n^{-r}$, with $r = \frac{p-2}{4p}$.

For every $t \in [0,1]$ we have that

$$\begin{aligned} g \circ W_n(t) - \sigma X_n(t) &= \frac{1}{\sqrt{n}}\left(\sum_{j=0}^{n-1} v \circ T^j - \sum_{j=0}^{\lfloor n(1-t)\rfloor} v \circ T^j - \sum_{j=1}^{k} m \circ T^{n-j}\right) + E_n(t) \\ &= \frac{1}{\sqrt{n}}\left(\sum_{j=n-\lfloor nt\rfloor}^{n-1} v \circ T^j - \sum_{j=1}^{k} m \circ T^{n-j}\right) + E_n(t), \end{aligned}$$

where

$$E_n(t) = \big(X_n(t) - X_n(t_k)\big) + \big(W_n(1-t) - W_n(\lfloor n(1-t)\rfloor/n)\big)$$

is the sum of interpolation errors of X_n and W_n. This term is negligible, because $\|E_n\|_\infty \le 2n^{-\frac{1}{2}}\|v\|_\infty$ for every $n \in \mathbb{N}^+$. Since by formula (1.8) we can write

$$\sum_{j=1}^{k} m \circ T^{n-j} = m_n - m_{n-k} = v_n - v_{n-k} - \chi \circ T^n + \chi \circ T^{n-k},$$

we obtain

$$\begin{aligned}
\big|g \circ W_n(t) - \sigma X_n(t)\big| &\le \frac{1}{\sqrt{n}}\bigg|\sum_{j=n-\lfloor nt\rfloor}^{n-1} v \circ T^j - \sum_{j=n-k}^{n-1} v \circ T^j - \chi \circ T^n + \chi \circ T^{n-k}\bigg| + \|E_n\|_\infty \\
&\le |V_n'(t)| + \frac{1}{\sqrt{n}}\Big(\big|\chi \circ T^{n-k} - \chi \circ T^n\big| + \frac{2}{\sqrt{n}}\|v\|_\alpha,
\end{aligned}$$

defining the piecewise constant process $V_n'(t) = \frac{1}{\sqrt{n}}\sum_{j=n-\lfloor nt\rfloor}^{n-k+1} v \circ T^j$.

Recall that we have k as function of t. We have by Proposition 1.4

$$\Big\|\frac{1}{\sqrt{n}}\sup_{t\in[0,1]}\big|\chi \circ T^{n-k} - \chi \circ T^n\big|\Big\|_{p-1} \le \frac{2}{\sqrt{n}}\|\chi\|_\alpha \ll \frac{\|v\|_\alpha}{\sqrt{n}}$$

We reason now using the first inequality of Proposition 1.21 and with Cauchy-Schwarz,

$$\begin{aligned}
\Big\|\sup_{t\in[0,1]}\big|V_n'(t)\big|\Big\|_{p-1} &\le n^{-\frac{1}{2}}\Big\|Z_n\big((n^{\frac{1}{2}}-1)^{-1}\sup_{t\in[0,1]}|\lfloor nt\rfloor - k| + 3\big)\Big\|_{p-1} \\
&\le n^{-\frac{1}{2}}\|Z_n\|_{2(p-1)}\ \Big\|(n^{\frac{1}{2}}-1)^{-1}\sup_{t\in[0,1]}|\lfloor nt\rfloor - k| + 3\Big\|_{2(p-1)} \\
&\ll n^{-\frac{1}{2}}\|Z_n\|_{2(p-1)}\Big(n^{-\frac{1}{2}}\big\|\sup_{t\in[0,1]}|\lfloor nt\rfloor - k|\big\|_{2(p-1)} + 3\Big).
\end{aligned}$$

Using now Proposition 1.19, followed by the second inequality of Proposition 1.21 we get

$$\Big\|\sup_{t\in[0,1]}\big|V_n'(t)\big|\Big\|_{p-1} \ll n^{-\frac{1}{2}}\|Z_n\|_{2(p-1)} \ll n^{-\frac{1}{4}\frac{p-2}{p-1}}. \tag{1.29}$$

Combining now these estimates we obtain

$$\Big\|\sup_{t\in[0,1]}\big|g \circ W_n(t) - \sigma X_n(t)\big|\Big\|_{p-1} \ll n^{-\frac{1}{4}\frac{p-2}{p-1}}.$$

We conclude by Proposition 1.20(b)

$$\pi_1(g \circ W_n, \sigma X_n) \ll n^{-\frac{1}{4}\frac{p-2}{p-1}\frac{p-1}{p}} = n^{-r}.$$

□

Observation (about the function g)**.** We have just proved something very similar to what stated in Theorem 1.16, but with an auxiliary function $g : \mathscr{C}[0,1] \to \mathscr{C}[0,1]$ defined in (1.28). Let W be a Wiener process with variance σ^2 found in (1.9) and let B be a standard Brownian motion. We get that W **is invariant under** g, i.e.

$$(gW)(t) = W_1 - W_{1-t} =_d W = \sigma B. \tag{1.30}$$

This is because

- $(gW)(0) = 0$;
- For fixed $0 \le s \le t \le 1$, we have $(gW)(t) - (gW)(s) = W_{1-s} - W_{1-t} =_d \mathcal{N}(0, t-s)$
- For fixed $k \in \mathbb{N}^+$ and $0 \le t_1 < \cdots < t_k \le 1$, the increments

  $$\big(W_{1-t_1} - W_{1-t_2}\big), \big(W_{1-t_2} - W_{1-t_3}\big), \dots, \big(W_{1-t_{k-1}} - W_{1-t_k}\big)$$

  are independent.

We see now that for every $u, v \in \mathscr{C}[0,1]$ and $t \in [0,1]$,

$$|(gu)(t) - (gv)(t)| \le |u(1) - v(1)| + |u(1-t) - v(1-t)| \le 2\|u - v\|_\infty.$$

Passing to the $\sup_{t\in[0,1]}$ we get that g **is Lipschitz** with $Lip(g) \le 2$. Whitt shows in [14, Theorem 3.2] that the rate of convergence in the Prokhorov metric is preserved under the action of Lipschitz maps. As a matter of fact, for every pair of random paths $X, Y \in \mathscr{C}[0,1]$,

$$\pi_1\big(g(X), g(Y)\big) \le Lip(g)\ \pi_1(X, Y). \tag{1.31}$$

If now we consider a function $u \in \mathscr{C}[0,1]$ such that $u(0) = 0$, then

$$g\big(gu\big)(t) = g\big(u(1) - u(1-t)\big) = \big(u(1) - u(0)\big) - \big(u(1) - u(1-1+t)\big) = u(t), \tag{1.32}$$

and therefore $\boldsymbol{g \circ g = Id}$ on the subspace $\{f \in \mathscr{C}[0,1] : f(0)_{=}0\}$.

Proof (of Theorem 1.16). Fix $\gamma > 0$. We are going to show that $\pi_1\big(W_n, W\big) \ll n^{-(\frac{1}{4}-\gamma)}$.

By (1.30)and the triangle inequality we get for every $n \in \mathbb{N}^+$,

$$\pi_1\big(g(W_n), g(W)\big) = \pi_1\big(g(W_n), \sigma B\big) \le \pi_1\big(g(W_n), \sigma X_n\big) + \pi_1\big(\sigma X_n, \sigma B\big) \ll n^{-(\frac{1}{4}-\gamma)},$$

combining Lemma 1.18 and Lemma 1.22. Since $W(0) = W_n(0) = 0$, we can use (1.32) and finally, by (1.31) (because $Lip(g) \leq 2$),

$$\pi_1\big(W_n, W\big) = \pi_1\big(g(g \circ W_n), g(g \circ W)\big) \leq 2\pi_1\big(g \circ W_n, g \circ W\big) \ll n^{-(\frac{1}{4}-\gamma)}.$$

□

Chapter 2

Extension to the Continuum

The previous Chapter 1 showed how, starting with a discrete dynamical system $T\colon \Lambda \to \Lambda$, it is possible to prove a rate of convergence for the Weak Invariance Principle. In order to find a better estimate and prove a similar result for d-dimensions, it is reasonable to extend the previous setting to continuous time, eventually to use estimates from Courbot [6, 7] for continuous-time martingales.

To start moving in this direction, we have to substitute the discrete transformation T with a family $\{T_t\}_{t\geq 0}$ which preserves the probability measure on Λ, acting as a semigroup in t, mimicking the behaviour of iterations T^n for $n \in \mathbb{N}$. Still, this has to be thought as the evolution of a dynamical system $T_t\colon \Lambda \to \Lambda$, for which we can define Koopman and transfer operators.

2.1 Semigroups of Operators

Suppose $(\Lambda, \mathscr{B}, \mu)$ a probability space. Consider a family of measure-preserving transformation $\{T_t\colon \Lambda \to \Lambda | t \geq 0\}$ and suppose that $T_0 = Id$ and $T_t \circ T_s = T_{t+s}$ for every $t, s \geq 0$. The latter says that T_t is a *semigroup* wrt maps composition. We deduce that $T_t^{-1}\mathscr{B} \subseteq T_s^{-1}\mathscr{B}$ for every $0 \leq s \leq t$: as a matter of fact

$$T_t^{-1}\mathscr{B} = \big(T_s \circ T_{t-s}\big)^{-1}\mathscr{B} = T_{t-s}^{-1}\big(T_s^{-1}\mathscr{B}\big) \subseteq T_s^{-1}\mathscr{B},$$

by measurability of T_{t-s}. Let now $\{U_t\colon L^1(\Lambda) \to L^1(\Lambda) | t \geq 0\}$ and $\{P_t\colon L^1(\Lambda) \to L^1(\Lambda) | t \geq 0\}$ be the families of Koopman and transfer operators for T_t, i.e.

$$U_t v := v \circ T_t \qquad\qquad \int_\Lambda P_t v \; w \,\mathrm{d}\mu = \int_\Lambda v \; U_t w \,\mathrm{d}\mu$$

for every $v \in L^1(\Lambda)$, $w \in L^\infty(\Lambda)$, and $t \geq 0$.

Proposition 2.1. The families U_t and P_t are semigroups.

Proof. For every $v \in L^1(\Lambda)$ we have for U_t

$$U_0 v = v \circ T_0 = v \circ Id = v, \qquad U_t(U_s v) = v \circ T_s \circ T_t = v \circ T_{t+s} = U_{t+s} v.$$

For every $v \in L^1(\Lambda)$ and $w \in L^\infty(\Lambda)$ we have for P_t

$$\int_\Lambda P_0 v \; w \, \mathrm{d}\mu = \int_\Lambda v \; U_0 w \, \mathrm{d}\mu = \int_\Lambda vw \, \mathrm{d}\mu,$$

$$\int_\Lambda P_t(P_s v) \; w \, \mathrm{d}\mu = \int_\Lambda P_s v \; U_t w \, \mathrm{d}\mu = \int_\Lambda v \; U_s(U_t w) \, \mathrm{d}\mu = \int_\Lambda v \; U_{t+s} w \, \mathrm{d}\mu = \int_\Lambda P_{t+s} v \; w \, \mathrm{d}\mu.$$

□

Suppose now Λ a bounded metric space with a Borel probability measure μ . The notation $(\mathscr{C}^\alpha, \|\cdot\|_\alpha)$ with $\alpha \in (0,1]$ indicates as usual the Banach space of Hölder functions on Λ and $\mathscr{C}_0^\alpha$ denotes its subspace with mean 0. Let us assume some working hypotheses that can be thought as continuous versions of (A) and (B) made in section 1.2:

(A') We have $U_t \mathscr{C}^\alpha \subseteq \mathscr{C}^\alpha$ and there exists $K \geq 1$ such that $\|U_t v\|_\alpha \leq K \|v\|_\alpha$ for every $v \in \mathscr{C}^\alpha$ and $t \in [0,1]$.

(B') We have that $P_t \mathscr{C}^\alpha \subseteq \mathscr{C}^\alpha$ and there exist $C, a > 0$ such that for every $v \in \mathscr{C}_0^\alpha(\Lambda)$ and $t \geq 0$

$$\|P_t v\|_\alpha \leq C e^{-at} \|v\|_\alpha.$$

Similarly with what was explained before Proposition 1.4, assumption (A') is implied by a uniform Lipschitz condition of transformations $\{T_t\}_{t \in [0,1]}$. This is a reasonable equivalent with the discrete dynamical system of section 1.2. Furthermore we have by induction

Proposition 2.2. Assuming (A'), then for every $t \in [j-1, j]$, $j \in \mathbb{N}^+$ and $v \in \mathscr{C}^\alpha$

$$\|U_t v\|_\alpha \leq K^j \|v\|_\alpha.$$

Proof. Let us reason by induction on j.

$\boldsymbol{j = 1}$ – assumption (A');

$\boldsymbol{j \mapsto j+1}$ – Take $t \in [j, j+1]$, then

$$\|U_t v\|_\alpha = \|U_{t-1} U_1 v\|_\alpha \leq K^j \|U_1 v\|_\alpha \leq K^{j+1} \|v\|_\alpha,$$

because $t - 1 \in [j-1, j]$.

□

Remark 2.3. Since $K \geq 1$, this implies $\|U_t v\|_\alpha \leq K^j \|v\|_\alpha$ for every $t \in [0, j]$. The same could be restated using the ceiling function $\lceil t \rceil \coloneqq \inf\{s \in \mathbb{N} : s \geq t\}$ as

$$\|U_t v\|_\alpha \leq K^{\lceil t \rceil} \|v\|_\alpha,$$

for every $t \geq 0$.

Remark 2.4. Lemma 1.2 is true for every $t \geq 0$; therefore, for every $0 \leq s \leq t$

$$\begin{aligned} P_t U_s &= P_{t-s} P_s U_s = P_{t-s}, \\ U_t P_s &= U_{t-s} U_s P_s = \mathbb{E}\left[\,\cdot\, | T_s^{-1} \mathscr{B}\right] \circ T_{t-s}. \end{aligned}$$

2.2 Bochner Integration

In order to extend discrete Birkhoff sums of $v \in \mathscr{C}^\alpha$ to the continuum, we need to define an integration wrt to t that will be used for the functions $U_t v, P_t v \colon [0, \infty) \to \mathscr{C}^\alpha$. Our choice is to work with *Bochner integral* for which we refer to [15, pag.132] and [12, Appendix E]. Given $T \geq 0$ we consider the interval $[0, T]$ endowed with Lebesgue measure. We call $f \colon [0, T] \to X$ a *simple function* taking values in the Banach space X if it can be expressed for every $t \in [0, T]$ as

$$f(t) = \sum_{j=1}^{k} f_j \mathbb{1}_{E_j}(t),$$

for $f_1, \ldots, f_k \in X$, and $E_1, \ldots, E_k$ a measurable partition of $[0, T]$. Its integral is the element of X defined as

$$X \ni \int_0^T f(t)\, \mathrm{d}t \coloneqq \sum_{j=1}^{k} \mu(E_j) f_j.$$

We say that a function $f \colon [0, T] \to X$ is *strongly measurable* if there exists a sequence $\{f_n\}_{n \in \mathbb{N}^+}$ of simple functions such that for almost every $t \in [0, T]$,

$$f_n(t) \xrightarrow{X} f(t),$$

for $n \to \infty$. Finally, a strong measurable function $f \colon [0, T] \to X$ is called *Bochner summable*, if there exists a sequence of simple functions $\{f_n\}_{n \in \mathbb{N}^+}$ such that

$$\int_0^T \big\| f_n(t) - f(t) \big\|\, \mathrm{d}t \xrightarrow{n \to \infty} 0,$$

where the latter is a standard Lebesgue integral. In this case we can define

$$X \ni \int_0^T f(t)\,\mathrm{d}r := \lim_{n\to\infty} \int_0^T f_n(t)\,\mathrm{d}r$$

which converge in X thanks to the summability condition.

Proposition 2.5. Given X a Banach space, $F\colon X \to X$ a bounded linear operator, $T \geq 0$ and a Bochner summable function $f\colon [0,T] \to X$, then

(1) $F \displaystyle\int_0^T f(t)\,\mathrm{d}t = \int_0^T (Ff)(t)\,\mathrm{d}t.$

(2) $\displaystyle\left\| \int_0^T f(t)\,\mathrm{d}t \right\| \leq \int_0^T \|f(t)\|\,\mathrm{d}t.$

2.3 A Continuous-Martingale Decomposition

This section provides a continuous-time equivalent to what found in Proposition 1.4. Given $v \in \mathscr{C}_0^\alpha$, we are after a martingale-coboundary decomposition with some modifications to fit the setting established in section 2.1. Our previous discrete-time transformation $T : \Lambda \to \Lambda$ has now become the family $\{T_t\}_{t\geq 0}$ that consequently requires a continuous equivalent of Birkhoff sums of v, too. The function v_t will be defined in terms of a time-integral that is interpreted the Bochner sense. Theorem 2.8 decomposes each v_t into a function m_t (connected to a martingale), plusd a coboundary term χ. Proposition 2.12 shows finally how for every $t \geq 0$ we can use m_t to construct a continuous-martingale array $M_{t,s}$.

In order to give sense to time-integrals $\int_0^t U_r v\,\mathrm{d}r$ and $\int_0^t P_r v\,\mathrm{d}r$ and apply properties (1) and (2) of Proposition 2.5, we suppose that the functions $U_s v, P_s v\colon [0,t] \to \mathscr{C}^\alpha$ are Bochner summable for every $v \in \mathscr{C}^\alpha$ and $t \geq 0$. We remark that most of this section is rigorous, but some calculations are formal due to lack of time and will be sorted out later. Similarly, some hypotheses could be dropped or adapted to future specific cases that we are going to consider.

Definition 2.6. Let $v \in \mathscr{C}^\alpha$ and let T_t be a semigroup of transformations as above. For every $t \geq 0$ we define the function $v_t\colon \Lambda \to \mathbb{R}$ as

$$v_t := \int_0^t v \circ T_s\,\mathrm{d}s = \int_0^t U_s v\,\mathrm{d}s. \tag{2.1}$$

Remark 2.7. Bochner summability of $s \mapsto U_s v$ guarantees that v_t is well defined for every $t \geq 0$ and that $v_t \in \mathscr{C}^\alpha$. Furthermore, if $v \in \mathscr{C}_0^\alpha$ we can reason by Fubini ($\|U_s v\|_\infty \leq \|v\|_\infty$)

and T_t-invariance to obtain $v_t \in \mathscr{C}_0^\alpha$ for every $t \geq 0$,

$$\int_\Lambda v_t \,\mathrm{d}\mu = \int_0^t \int_\Lambda v \,\mathrm{d}\mu \,\mathrm{d}s = 0.$$

Theorem 2.8. *Let $v \in \mathscr{C}_0^\alpha$, v_t defined in (2.1) and let T_t be a semigroup of measure-preserving transformations. There exist $\chi \in \mathscr{C}^\alpha$ and $m\colon [0,+\infty) \to \mathscr{C}^\alpha$ such that*

$$v_t = m_t + \chi \circ T_t - \chi, \tag{2.2}$$

for every $t \geq 0$. Furthermore $m_t - m_s \in \ker P_t \cap \mathscr{C}_0^\alpha$ for every $t \geq 0$ and $s \in [0,t]$.

Proof. For every $t \geq 0$ let us define the function $\chi_t\colon \Lambda \to \mathbb{R}$ as $\chi_t := \int_0^t P_s v \,\mathrm{d}s$ which is defined pointwise in Λ thanks to continuity assumed in (B'). By Bochner summability we get $\chi_t \in \mathscr{C}^\alpha$ for every $t \geq 0$. Furthermore,using Proposition 2.5(2), we get by (B') that there exists a constant $C > 0$ such that, given $0 \leq s \leq t$,

$$\begin{aligned}\|\chi_t - \chi_s\|_\alpha =& \left\| \int_s^t P_r v \,\mathrm{d}r \right\|_\alpha \leq \int_s^t \|P_r v\|_\alpha \,\mathrm{d}r \leq C\|v\|_\alpha \int_s^t e^{-ar} \,\mathrm{d}r \\ &= \frac{C}{a}\|v\|_\alpha \big(e^{-as} - e^{-at}\big) \xrightarrow{s,t\to\infty} 0.\end{aligned}$$

By completeness of $\mathscr{C}^\alpha$ there exsits $\chi = \lim_{t\to\infty} \chi_t := \int_0^\infty P_s v \,\mathrm{d}s \in \mathscr{C}^\alpha$.

Fix now $t \geq 0$ and define $m_t\colon \Lambda \to \mathbb{R}$ as

$$m_t := v_t - \chi \circ T_t + \chi. \tag{2.3}$$

Decomposition (2.2) follows. By (A') and Bochner summability we get that both $\chi \circ T_t$ and v_t belong to $\mathscr{C}^\alpha$, therefore m_t is Hölder as well. In addition, using T_t-invariance and Remark 2.7 we obtain for every $t \geq 0$,

$$\int_\Lambda m_t \,\mathrm{d}\mu = \int_\Lambda v_t \,\mathrm{d}\mu - \int_\Lambda \chi \circ T_t \,\mathrm{d}\mu + \int_\Lambda \chi \,\mathrm{d}\mu = 0,$$

because $v \in \mathscr{C}_0^\alpha$. Thereofre $m_t \in \mathscr{C}_0^\alpha$, too.

Fix now $s, t \in [0,\infty)$ such that $s \leq t$. Let us show that $m_t - m_s \in \ker P_t$. By formula (2.3), definition (2.1) and additivity of integrals

$$m_t - m_s = v_t - v_s - \chi \circ T_t + \chi \circ T_s = \int_s^t U_r v \,\mathrm{d}r - U_t\chi + U_s\chi.$$

Changing variable $r \mapsto r - s$ we get

$$\int_s^t U_r v \, \mathrm{d}r = \int_0^{t-s} U_{s+r} v \, \mathrm{d}r.$$

Therefore applying P_t to $m_t - m_s$ and using Remarks 2.4 and 2.5 (1)(2),

$$P_t(m_t - m_s) = P_t \int_0^{t-s} U_{s+r} v \, \mathrm{d}r - P_t U_t \chi + P_t U_s \chi = \int_0^{t-s} P_{t-s-r} v \, \mathrm{d}r - \chi + P_{t-s}\chi.$$

By Proposition 2.5(1) we have that for every $h \geq t - s$,

$$P_{t-s}\chi_h - \chi_h = \int_0^h P_{t-s+r} v \, \mathrm{d}r - \int_0^h P_r v \, \mathrm{d}r = \int_{t-s}^h P_r v \, \mathrm{d}r - \int_0^h P_r v \, \mathrm{d}r = -\chi_{t-s}.$$

Therefore $P_{t-s}\chi - \chi = \chi_{t-s}$ by continuity of P_{t-s}, and we can conclude

$$P_t(m_t - m_s) = \int_0^{t-s} P_{t-s-r} v \, \mathrm{d}r - \int_0^{t-s} P_r v \, \mathrm{d}r = 0,$$

by a change of variable $r \mapsto t - s - r$. □

Corollary 2.9. *There exists a constant $C' \geq 0$ such that*

$$\|\chi\|_\alpha \leq C' \|v\|_\alpha$$

for every $v \in \mathscr{C}_0^\alpha$ and $t \geq 0$.

Proof. By Proposition 2.5(2) applied to $\chi_t = \int_0^t P_s v \, \mathrm{d}s$ and assumption (B') we get

$$\|\chi_t\|_\alpha \leq \int_0^t \|P_r v\|_\alpha \, \mathrm{d}r \leq C\|v\|_\alpha \int_0^t e^{-ar} \, \mathrm{d}r = C \, \frac{1 - e^{-at}}{a} \, \|v\|_\alpha \leq C'\|v\|_\alpha$$

with $C' = C/a$. By continuity of the norm, we get for $t \to \infty$ the condition for χ.

□

Definition 2.10. A family of integrable random variables $\{X_t : t \geq 0\}$ together with a filtration $\{\mathcal{F}_t : t \geq 0\}$ is called a continuous-time martingale if for every $t, h \geq 0$

- X_t is $\mathcal{F}_t$-measurable;
- $\mathbb{E}[X_{t+h} | \mathcal{F}_t] = X_t$.

The same definition can be extended to martingale arrays of the type $\{X_s : 0 \leq s \leq t\}$ with filtrations $\{\mathcal{F}_s : 0 \leq s \leq t\}$ for $t \geq 0$.

Fix $v \in \mathscr{C}_0^\alpha$ and a semigroup of transformations T_t as above. Define for every $t \geq 0$ a continuous time array $\{M_{t,s}\}_{0\leq s\leq t}$ together with a filtration array $\{\mathcal{G}_{t,s}\}_{0\leq s\leq t}$:

$$M_{t,s} := \frac{m_t - m_{t-s}}{\sigma_t\sqrt{t}} \qquad\qquad \mathcal{G}_{t,s} := T_{t-s}^{-1}\mathscr{B}, \tag{2.4}$$

where $\sigma_t^2 := \int_\Lambda m_t^2 \,\mathrm{d}\mu$ and $\mathscr{B}$ is the Borel σ-algebra over Λ.

Remark 2.11. Consider $v \in \mathscr{C}^\alpha$ and for every $t \geq 0$ let v_t be defined in (2.1). We suppose that for every $0 \leq s \leq t$ the function $v_t - v_{t-s}$ is $T_{t-s}^{-1}\mathscr{B}$-measurable. As a matter of facts we obtain an integral summing backwards,

$$v_t - v_{t-s} = \int_{t-s}^{t} v \circ T_r \,\mathrm{d}r = \int_0^s v \circ T_{t-r} \,\mathrm{d}r,$$

using a change of variable $r \mapsto t - r$. This integral plays the role of continuous equivalent of $\sum_{j=1}^k v \circ T^{n-j}$ which satisfied $T^{-(n-k)}\mathscr{B}$-measurability for the iterated transformation T.

Proposition 2.12. For every $t \geq 0$, the array $\{M_{t,s}\}_{0\leq s\leq t}$ defined in (2.4) is a continuous-time martingale array.

Proof. For every $0 \leq s \leq t$, the $T_{t-s}^{-1}\mathscr{B}$-measurability of $M_{t,s}$ follows from Remark 2.11:

$$\sigma_t\sqrt{t}M_{t-s} = m_t - m_{t-s} = v_t - v_{t-s} + \big(\chi \circ T_s + \chi\big) \circ T_{t-s}.$$

Fix $t \geq 0$ and choose $r, s \in [0, t]$ such that $0 \leq r \leq s \leq t$. Using Remark 2.4 we have that

$$\sigma_t\sqrt{t}\,\mathbb{E}\left[M_{t,s} - M_{t,r}|\mathcal{G}_{t,r}\right] = \mathbb{E}\left[m_{t-r} - m_{t-s}|T_{t-r}^{-1}\mathscr{B}\right] = U_{t-r}P_{t-r}(m_{t-r} - m_{t-s}) = 0,$$

because $t - r \geq t - s$ and by Proposition 2.2 $(m_{t-r} - m_{t-s}) \in \ker P_{t-r}$. □

The martingale $M_{t,s}$ of (2.4) is the continuous equivalent of the one obtained by summation of $\xi_{n,k}$, the discrete MDA defined in Proposition 1.9. As a matter of fact, taking $v \in \mathscr{C}_0^\alpha$ and denoting with m its martingale part deriving from the discrete dynamical system $T\colon \Lambda \to \Lambda$ of section 1.2, we have for $k, n \in \mathbb{N}^+$, with $k \leq n$,

$$\frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{k} m \circ T^{n-j} = \frac{1}{\sigma\sqrt{n}}\left(\sum_{j=0}^{n-1} m \circ T^j - \sum_{j=0}^{n-k-1} m \circ T^j\right) = m_n - m_{n-k},$$

where m_n is the Birkhoff sum of m and $\sigma^2 = \int_\Lambda m^2$. We see that this is consistent with definition of $M_{t,s}$. As already stated in the introduction, this new continuous-time setting looks promising to eventually apply techniques of Courbot [6, 7], in order to improve and extend rate estimates of chapter 1 to vector-valued observables $v\colon \Lambda \to \mathbb{R}^d$.

Appendix A

Some Results

The following are technical results which are used in the proof of Proposition 1.13. Because of their generality and non-particular relevance with the topics discussed in chapter 1, they are presented in this appendix.

Proposition A.1. Let $\big(\Omega, \mathcal{A}, \mu\big)$ be a measure space.
If $f, g \in L^\infty(\Omega)$, then $fg \in L^\infty(\Omega)$ and $\|fg\|_\infty \leq \|f\|_\infty \|g\|_\infty$.

Proof. W have that $|f(x)g(x)| = |f(x)||g(x)| \leq \|f\|_\infty \|g\|_\infty$ for almost every $x \in \Omega$, then we conclude. □

Proposition A.2. Let (S, d_S) be a metric space, $\alpha \in (0, 1]$ and $f, g : S \to \mathbb{R}$ be bounded α-Hölder functions, with Hölder constants $\mathcal{H}(f)$ and $\mathcal{H}(g)$. Then fg is α-Hölder and

$$\mathcal{H}(fg) \leq \|f\|_\infty \mathcal{H}(g) + \|g\|_\infty \mathcal{H}(f).$$

Proof. For every $x \neq y \in S$ we have

$$\big|f(x)g(x) - f(y)g(y)\big| \leq \big|f(x)g(x) - f(x)g(y)\big| + \big|f(x)g(y) - f(y)g(y)\big|$$

$$\leq \|f\|_\infty \big|g(x) - g(y)\big| + \|g\|_\infty \big|f(x) - f(y)\big| \leq \|f\|_\infty \mathcal{H}(g)\ d_S(x, y)^\alpha + \|g\|_\infty \mathcal{H}(f)\ d_S(x, y)^\alpha$$

$$= \big(\|f\|_\infty \mathcal{H}(g) + \|g\|_\infty \mathcal{H}(f)\big)\ d_S(x, y)^\alpha.$$

This yields

$$\mathcal{H}(fg) = \sup_{s \neq t} \frac{|fg(x) - fg(y)|}{d_S(x, y)^\alpha} = \leq \|f\|_\infty \mathcal{H}(g) + \|g\|_\infty \mathcal{H}(f).$$

□

Bibliography

[1] Antoniou, M. (2018). *Rates of convergence for statistical limit laws in deterministic dynamical systems.* PhD, University of Warwick.

[2] Antoniou, M. and Melbourne, I. (2019). Rate of Convergence in the Weak Invariance Principle for Deterministic Systems. *Communications in Mathematical Physics*, 369(3):1147–1165.

[3] Billingsley, P. (1999). *Convergence of Probability Measures.* Wiley Series in Probability and Statistics. Wiley-Interscience, 2nd edition.

[4] Burkholder, D. L. (1973). Distribution Function Inequalities for Martingales. *The Annals of Probability*, 1(1):19–42.

[5] Coquet, F., Mémin, J., and Vostrikova, L. (1994). Rate of convergence in the functional central limit theorem for likelihood ratio processes. *Math. Methods Statist*, 3(2):89–113.

[6] Courbot, B. (1999). Rates of convergence in the functional CLT for martingales. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 328(6):509–513.

[7] Courbot, B. (2001). Rates of convergence in the functional CLT for multidimensional continuous time martingales. *Stochastic Processes and their Applications*, 91(1):57–76.

[8] Korepanov, A., Kosloff, Z., and Melbourne, I. (2018). Martingale–coboundary decomposition for families of dynamical systems. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 35(4):859–885.

[9] Korepanov, A., Kosloff, Z., and Melbourne, I. (2019). Explicit coupling argument for non-uniformly hyperbolic transformations. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 149(1):101–130.

[10] Kubilius, K. (1985). Rate of convergence in the functional central limit theorem for semimartingales. *Liet. Mat. Rinkinys*, 25:84–96.

[11] Kubilius, K. (1994). Rate of convergence in the invariance principle for martingale difference arrays. *Lithuanian Mathematical Journal*, 34(4):383–392.

[12] Lawrence, C. E. (2010). *Partial Differential Equations*, volume 19 of *Graduate Series in Mathematics*. American Mathematical Society, 2nd edition.

[13] Melbourne, I. (2015). Fast-Slow Skew Product Systems and Convergence to Stochastic Differential Equations. *Lecture notes for the LMS-CMI Research School Statistical Properties of Dynamical Systems at Loughborough.*

[14] Whitt, W. (1974). Preservation of rates of convergence under mappings. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 29(1):39–44.

[15] Yosida, K. (1995). *Functional Analysis.* Classics in Mathematics. Springer-Verlag, Berlin Heidelberg, 6th edition.