Variations on a central limit theorem in infinite ergodic theory

Damien THOMINE

Abstract

In a previous article the author proved a distributional convergence for the Birkhoff sums of functions of null average defined over a dynamical system with an infinite, invariant, ergodic measure, akin to a central limit theorem. Here we extend this result to larger classes of observables, with milder smoothness conditions, and to larger classes of dynamical systems, which may not be mixing anymore. A special emphasis is given to continuous time systems: semi-flows, flows, and $\mathbb{Z}^d$-extensions of flows. The later generalization is applied to the geodesic flow on $\mathbb{Z}^d$-periodic manifolds of negative sectional curvature.

Given a dynamical system (Ω, T, μ) with an ergodic, T-invariant, infinite measure μ and a real-valued function f on Ω, what is the asymptotic behavior of the Birkhoff sums $\sum_{k=0}^{n-1} f \circ T^k$? This problem has received an extensive answer when f is integrable and has a non-zero integral, provided that the system behaves nicely, for instance if it can be written as a tower over a transformation with good mixing properties [1, Lemma 3.7.4]. The theory is much less developed if f has zero integral, in which case these results are not accurate. The first dent into this problem was made by R.L. Dobrushin in 1955 [9] for the simple random walk on $\mathbb{Z}$ and functions with finite support on $\mathbb{Z}$. This was progressively extended, until two articles by E. Csáki and A. Földes [7] [8] which dealt with random walks on $\mathbb{Z}$ or $\mathbb{Z}^2$ whose transition kernel has finite variance, and more general random walks on $\mathbb{Z}$.

In [16], the author extended these results to discrete time dynamical systems which induce a Gibbs-Markov map on some subset. The method was adapted from the works of E. Csáki and A. Földes, and can be described as exploiting *asymptotic independence via coupling methods*. The limit theorems therein cover random walks, but also hyperbolic maps with indifferent periodic points, provided that f satisfies some smoothness and integrability conditions.

With this article we will extend these theorems in different directions. In Section 3, the smoothness condition used in [16] is relaxed; the martingale methods we develop are then applied in Section 4 to observables which take their values in a Hilbert space. Section 5 replaces the hypothesis of mixing which was made in [16] with ergodicity. Finally, in Section 6 we study continuous time systems of varying complexity: suspension flows over Gibbs-Markov maps endowed with an infinite measure, $\mathbb{Z}^d$-extensions of suspension flows over Gibbs-Markov maps endowed with a finite measure, and natural extensions thereof. The article ends on a study of the geodesic flow on periodic manifolds of negative curvature (Proposition 6.15).

Before we start to prove these new results, let us recall the setting, and the strategy used in [16].

1 Setting

1.1 Gibbs-Markov maps

The main limit theorems of this paper shall be established for towers or suspension flows over Gibbs-Markov maps. We recall here some basic definitions and properties; a more in-depth introduction can be found for instance in [1, Chapter 4.7].

Definition 1.1 (Gibbs-Markov maps).

Let $(\Omega, d, \mathcal{B})$ be a measurable, metric, bounded Polish space, endowed with a probability measure μ. A non-singular, measurable map $T : \Omega \mapsto \Omega$ is said to be a Markov map if μ is T-invariant and if there exists a countable partition π of Ω in sets of positive measure such that:

- *for all a in π, the image of a by T is a union of elements of π (up to a set of null measure);*
- *for all a in π, the map T is an isomorphism from a onto its image;*
- *the completion for μ of the σ-algebra $\bigvee_{n \in \mathbb{N}} T^{-n}\pi$ is $\mathcal{B}$.*

The full data defining a Markov map is $(\Omega, \pi, d, \mathcal{B}, \mu, T)$, but we shall often omit some of the objects if there is no ambiguity. A Markov map is said to be Gibbs-Markov if it also has the following properties:

- *$\inf_{a \in \pi} \mu(Ta) > 0$ (large image property);*
- *it is locally uniformly expanding: there exists $\lambda > 1$ such that, for all a in π and x, y in a, we have $d(Tx, Ty) \geq \lambda d(x, y)$;*
- *it has a Lipschitz distortion: there exists a constant C such that, for all a in π, for almost every x and y in a:*

$$\left| \frac{\mathrm{d}\mu}{\mathrm{d}\mu \circ T}(x) - \frac{\mathrm{d}\mu}{\mathrm{d}\mu \circ T}(y) \right| \leq C d(Tx, Ty) \frac{\mathrm{d}\mu}{\mathrm{d}\mu \circ T}(x). \tag{1.1}$$

A Gibbs-Markov map is said to be mixing if, for any Borel sets A and B,

$$\lim_{n \to +\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

A function f on Ω is said to be a coboundary if there exists a measurable function g such that $f = g \circ T - g$.

For any points x and y, let us denote by $s(x, y)$ the time of separation of x and y for the partition π and the transformation T, i.e., the smallest time $n \geq 0$ at which the points $T^n x$ and $T^n y$ do not belong to the same element of the partition π. Then, for any $\kappa > 1$, one can define a metric d_κ on Ω by $d_\kappa(x, y) := \kappa^{-s(x,y)}$. The dynamical system $(\Omega, d_\kappa, \mathcal{B}, \mu, T)$ is also Gibbs-Markov if κ belongs to $(1, \lambda]$. With this canonical choice of a distance, a Gibbs-Markov map is entirely defined by the data $(\Omega, \pi, \kappa, T, \mu)$.

Let h be in $(0, 1)$. If κ is chosen close enough to 1, then any h-Hölder function (for the initial metric d) is Lipschitz for the metric d_κ. Hence, any result stated for Lipschitz functions actually holds for Hölder functions.

We shall denote by $g(x) := \frac{\mathrm{d}\mu}{\mathrm{d}\mu \circ T}(x)$ the inverse of the Jacobian of T at point x, and by $g^{(n)}(x) := g(x) \cdots g(T^{n-1}x)$ the inverse of the Jacobian of T^n. Thus, the bounded distortion property reads $|g(x) - g(y)| \leq C d(Tx, Ty) g(x)$ for all a in π and almost every x and y in a.

For every function f in $\mathbb{L}^1(\Omega,\mu)$, we put $\mathcal{L}f(x) := \sum_{Ty=x} g(y)f(y)$. Its iterates can be written in terms of $g^{(n)}$ by $\mathcal{L}^n f(x) = \sum_{T^n y=x} g^{(n)}(y)f(y)$. A cylinder is a set $\overline{a} = [a_0, \cdots, a_{n-1}]$ such that a_i belongs to π for all $0 \leq i \leq n-1$ and $\overline{a} = \bigcap_{i=0}^{n-1} T^{-i}a_i$. A Gibbs-Markov map then satisfies a stronger distortion property:

Lemma 1.2 (Distortion Lemma)**.**
Let (Ω, d, μ, T) be a Gibbs-Markov map. There exists a constant C such that, for almost every x and y in Ω, for all $n \leq s(x,y)$,

$$\left|g^{(n)}(x) - g^{(n)}(y)\right| \leq Cd(T^n x, T^n y) g^{(n)}(x), \tag{1.2}$$

and, for all cylinder of $\overline{a}$ of length n and all x in $\overline{a}$:

$$g^{(n)}(x) \leq C\mu(\overline{a}). \tag{1.3}$$

For any subset ω of Ω, we denote by $|\cdot|_{\mathrm{Lip}_d(\omega)}$ the Lipschitz semi-norm on ω: for any function from ω to a metric space (E, d'), it is defined by:

$$|f|_{\mathrm{Lip}_d(\omega)} := \inf\left\{C > 0 :\ \forall x \in \omega,\ \forall y \in \omega,\ d'(f(x), f(y)) \leq Cd(x,y)\right\}.$$

For any metric space E, we define an application:

$$D_{\pi,d} := \left\{ \begin{array}{lll} \mathcal{C}(\Omega, E) & \to & \mathcal{C}(\Omega, \mathbb{R}_+ \cup \{+\infty\}) \\ f & \mapsto & \sum_{a\in\pi} |f|_{\mathrm{Lip}_d(a)} 1_a \end{array} \right. .$$

For instance, $\|D_{\pi,d}(f)\|_\infty = \sup_{a\in\pi} |f|_{\mathrm{Lip}_d(a)}$ and $\mathbb{E}_\mu(D_{\pi,d}(f)) = \sum_{a\in\pi} \mu(a) |f|_{\mathrm{Lip}_d(a)}$. If there is no ambiguity on the partition nor on the metric, we may denote the semi-norm by $|\cdot|_{\mathrm{Lip}(\omega)}$ and the application by D. For any function f, the function $D(f)$ plays the role of the absolute value of a derivative. As customary in analysis, constraints on the regularity of f will be expressed as constraints on the integrability of $D(f)$.

Let Lip^∞ be the set of functions f from Ω to $\mathbb{R}$ such that $\|f\|_{\mathrm{Lip}^\infty} := \|f\|_{\mathbb{L}^\infty} + \|D(f)\|_\infty$ is finite. The transfer operator of a Gibbs-Markov system acting on Lip^∞ has a spectral gap, which entails many important results. We recall here a couple of them. The first is the exponential decay of correlations for observables in Lip^∞:

Proposition 1.3 (Exponential decay of correlations)**.**
If (Ω, d, μ, T) is a mixing Gibbs-Markov map, then there exist $C > 0$ and $\rho \in (0,1)$ such that, for all $f \in \mathrm{Lip}^\infty$ and $g \in \mathbb{L}^1$ and for every integer n:

$$\left| \int f \cdot g \circ T^n \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \cdot \int g \, \mathrm{d}\mu \right| \leq C\rho^n \|f\|_{\mathrm{Lip}^\infty} \|g\|_{\mathbb{L}^1}. \tag{1.4}$$

The second proposition is a central limit theorem for smooth observables, which can be proved either by spectral perturbation of the transfer operator [10, Theorem 3.7] or martingale methods:

Proposition 1.4 (Central limit theorem)**.**
Let (Ω, d, μ, T) be a mixing Gibbs-Markov map. Let $f \in \mathbb{L}^2(\Omega)$ be such that $\int_\Omega f \, \mathrm{d}\mu = 0$ and $\mathbb{E}(D(f)) < +\infty$. Then:

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k \to \sigma(f)\mathcal{N},$$

where the convergence is in distribution, $\mathcal{N}$ is a standard Gaussian random variable, and:

$$\sigma(f)^2 = \int_\Omega f^2 \, \mathrm{d}\mu + 2\sum_{n=1}^{+\infty} \int_\Omega f \cdot f \circ T^n \, \mathrm{d}\mu.$$

With stronger assumptions, Burkholder-Rosenthal's inequality also holds: if we assume that $f \in \mathbb{L}^p$ for some $p > 2$, on top of the hypotheses of Proposition 1.4, then there exists a constant C such that, for all n:

$$\left\| \sum_{k=0}^{n-1} f \circ T^k \right\|_{\mathbb{L}^p} \leq C n^{\frac{1}{2}}.$$

1.2 Stopping times

Let $(\Omega, \pi, \lambda, T, \mu)$ be a Gibbs-Markov map. For some applications, most notably $\mathbb{Z}^d$-extensions of Gibbs-Markov maps, we will be interested not in the application T, but in some of its iterates T^φ, where φ is a function on Ω taking its values in non-negative integers. The good notion which ensures that T^φ is "nice" is the one of stopping time. When applied to Markov maps, this type of acceleration is sometimes called Schweiger's jump transformation (see [14] for the initial construction, and [1, Section 4.6] for an alternative presentation of the Gibbs-Markov case). We recall its definition, and describe the properties of the accelerated system T^φ.

Definition 1.5 (Stopping time).
Let $(\Omega, (\mathcal{F}_n)_{n \geq 0})$ be a measurable space with a filtration. A function $\varphi : \Omega \to \mathbb{N} \cup \{+\infty\}$ is a stopping time if $\{\varphi \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$. The σ-algebra at the stopping time φ is the σ-algebra:

$$\mathcal{F}_\varphi := \{A \in \mathcal{B} : \ A \cap \{\varphi \leq n\} \in \mathcal{F}_n \ \forall n \geq 0\}.$$

The natural filtration for the Gibbs-Markov map is:

$$\mathcal{F}_n := \sigma\left(\bigvee_{k=0}^{n-1} T^{-k}\pi\right).$$

For this filtration, $\mathcal{F}_0$ is trivial, so for any stopping time φ, either $\varphi = 0$ almost surely, or $\varphi > 0$ almost surely. Since we want to work with the application T^φ, we should also require that $\varphi < +\infty$ almost surely.

If φ is almost surely finite, we will denote by π_φ the countable partition of Ω such that $\mathcal{F}_\varphi = \sigma(\pi_\varphi)$, by s_φ the separation time for points in Ω for the partition π_φ and the transformation T^φ, and by D_φ the application $D_{\pi_\varphi, d_\varphi}$. We can state a first result.

Lemma 1.6.
Let $(\Omega, \pi, \lambda, T, \mu)$ be a Gibbs-Markov map, and let φ be a stopping time. Assume that φ is almost surely positive and finite, and that T^φ preserves the measure μ. Then $(\Omega, \pi_\varphi, \lambda, T^\varphi, \mu)$ is a Gibbs-Markov map.

Proof.
Since $\varphi \geq 1$ and π_φ is finer than π, the partition $\bigvee_{k=0}^{n-1} (T^\varphi)^{-k} \pi_\varphi$ is finer than π_n, and the partition π_φ together with the transformation T^φ generate the Borel σ-algebra on Ω.

Let $\overline{a}$ be in π_φ. Then there exist $n \geq 1$ and a sequence $(a_0, \cdots, a_{n-1})$ of elements of π such that $\overline{a} = [a_0, \cdots, a_{n-1}]$ and $\varphi = n$ on $\overline{a}$. By induction, we see that $T^k \overline{a} = [a_k, \cdots, a_{n-1}]$ for all $k < n$, and that T^k is an isomorphism from $\overline{a}$ to $[a_k, \cdots, a_{n-1}]$. Hence, $T^{n-1}\overline{a} = a_{n-1} \in \pi$. By applying T one more time, we see that $T^n \overline{a} = T^\varphi \overline{a}$ is a union of elements of π, that T^n is an isomorphism from $\overline{a}$ onto its image, and that $(\Omega, \pi_\varphi, T^\varphi, \mu)$ has the large image property.

The new distance on Ω is $d_\varphi := \lambda^{-s_\varphi}$, so tautologically $d_\varphi(T^\varphi x, T^\varphi y) = \lambda d_\varphi(x, y)$ whenever x and y are in the same element of π_φ. Finally, the Lipschitz distortion property is an application of the distortion lemma (Lemma 1.2), together with the fact that $d_\varphi \geq d$. $\square$

1.3 Assumptions on the observables

In this article we often work with a pair of real-valued functions (X, φ) defined on a space Ω on which a Gibbs-Markov map acts. The function X stands for a well-behaved observable with good integrability, so that we can apply, for instance, a central limit theorem. The function φ stands for a first return time to a set of finite measure; since this work is about ergodic theory with infinite measure, the function φ will not be integrable in general. The standard assumption will bear on the behavior of the tails of φ, or in other words on the asymptotic decay of $\mu(\{\varphi > x\})$. We will also need a regularity condition on φ, which is trivially satisfied for discrete time dynamical systems (then, φ is constant on each element of the partition π).

The standard regularity condition we demand for X and φ was already used in Proposition 1.4. It is:

$$\mathbb{E}_\mu(D(X)) + \mathbb{E}_\mu(D(\varphi)) < +\infty. \tag{1.5}$$

This condition implies some kind of exponential decay of correlations, although the wording of such a decorrelation property is not obvious: since φ need not be integrable, the cross-correlation coefficients are not defined. A weaker condition is used in Section 3.

As for integrability, X will usually be assumed to be in $\mathbb{L}^p(\Omega, \mu)$ for some $p > 2$. The assumption on φ is that its tails are regularly varying. We introduce now the relevant definitions. A measurable function $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is said to have regular variation of index $\beta \in \mathbb{R}$ at infinity if, for all positive x:

$$\lim_{y \to +\infty} \frac{\psi(xy)}{\psi(y)} = x^\beta.$$

In addition, if β is nonnegative and ψ is a nondecreasing, unbounded and càglàd (left-continuous with right limits) function with regular variation of index β, we define its generalized inverse ψ^* by:

$$\psi^*(x) = \sup \{t \geq 0 : \ \psi(t) \leq x\}. \tag{1.6}$$

In this article we work with random variables with regularly varying tails. We will routinely use the following conditions on φ:

$$\mathbb{P}(\varphi \geq x) = O(1/\psi(x)), \tag{1.7}$$

where the function ψ is nondecreasing, unbounded, càglàd, and has regular variation of index $\beta \in [0, 1]$ at infinity.

We will sometimes need a stronger assumption on the tail of φ, such as:

$$\forall x > 0, \ \mathbb{P}(\varphi \geq x) = 1/\psi(x), \tag{1.8}$$

where the function ψ has regular variation of index $\beta \in [0, 1]$ at infinity (by construction, such a function ψ is automatically nondecreasing, unbounded, and càglàd).

2 Two limit theorems

We present the main results of [16]. The first of them is expressed in terms of coupling. Let us recall the definition of a coupling, and a few useful facts. Let X and Y be two random variables taking their values respectively in some Polish spaces $\mathcal{X}$ and $\mathcal{Y}$. We call a coupling between X and Y a random variable taking its value in $\mathcal{X} \times \mathcal{Y}$ whose first marginal (its projection onto $\mathcal{X}$) has the same law as X, and whose second marginal (its projection onto $\mathcal{Y}$) has the same law as Y.

Let (Ω, μ, T) be a conservative and ergodic dynamical system, where Ω is a Polish space and μ is an infinite, nonnegative T-invariant measure. For any Borel set A such that $\mu(A) > 0$, we denote by φ_A the first return time in A, i.e., $\varphi_A(x) := \inf\{i > 0 : T^i x \in A\}$ for all x in Ω, and we put $T_A := T_{|A}^{\varphi_A}$. Since the transformation T is conservative and ergodic, the return time is finite almost everywhere in A (see e.g. [1, Proposition 1.2.2]), and $(A, \mu_{|A}, T_A)$ is an ergodic dynamical system.

Definition 2.1.

We say that a dynamical system (Ω, μ, T) induces a Gibbs-Markov map on a Borel set A if $0 < \mu(A) < +\infty$ and $(A, \pi, d, \mu(\cdot|A), T_A)$, for some metric d on A and with respect to some partition π, is a Gibbs-Markov map, and if φ_A is constant on each set of the partition π.

Notice that, if the set A is given, we can rescale μ so that $\mu_{|A}$ is a probability measure. From now on, when we restrict such a system to its induced system on some Borel set A, we shall always assume that $\mu(A) = 1$. We denote by ψ the inverse of the tail of the random variable φ_A under $\mu_{|A}$, that is, for all $x \geq 0$:

$$\psi(x) := \frac{1}{\mu_{|A}(\varphi_A \geq x)}.$$

Let A be a Borel subset of Ω with positive and finite measure. For any measurable real-valued function f on Ω, we denote by X_f the function on A defined by:

$$X_f(x) := \sum_{i=0}^{\varphi_A(x)-1} f(T^i x).$$

If f is integrable, then X_f is also integrable.

If f is a coboundary for T, then one can check that X_f is a coboundary for T_A. Conversely, if X_f is a coboundary for T_A, then f is a coboundary for T. This is because if we have $X_f = \tilde{g} \circ T_A - \tilde{g}$ and we put:

$$g(x) := \tilde{g}(T^{\varphi_A(x)} x) - \sum_{k=0}^{\varphi_A(x)-1} f(T^k x),$$

then one can check that $f = g \circ T - g$.

Let us introduce the Mittag-Leffler distributions, which naturally appear when one deals with the distributional limit of local times. They are specified by their moment-generating functions.

Definition 2.2 (Mittag-Leffler distribution)**.**

Let β be in $[0, 1]$. The Mittag-Leffler distribution of order β is the distribution on $[0, +\infty)$ such that, for any random variable Y_β with this distribution, for all z in $\mathbb{C}$ (or all z in the open unit disc of $\mathbb{C}$ if $\beta = 0$):

$$\mathbb{E}(e^{zY_\beta}) = \sum_{n=0}^{+\infty} \frac{\Gamma(1+\beta)^n}{\Gamma(1+n\beta)} z^n. \tag{2.1}$$

We will also denote by $\operatorname{sinc}(x) := \sin(x)/x$ the cardinal sine. Then, the two main theorems of [16] are the following.

Theorem 2.3.

Let (Ω, d, μ, T) be a mixing Gibbs-Markov map. Let X and φ be measurable functions from Ω to $\mathbb{R}$ and to $\mathbb{R}_+$ respectively, satisfying the condition (1.5). We put $(X_i, \varphi_i) := (X \circ T^i, \varphi \circ T^i)$. Let $(\tilde{X}_i)_{i \in \mathbb{N}}$ and $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$ be copies of the processes $(X_i)_{i \in \mathbb{N}}$ and $(\varphi_i)_{i \in \mathbb{N}}$ respectively, such that $(\tilde{X}_i)_{i \in \mathbb{N}}$ and $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$ are mutually independent.

Assume that X belongs to $\mathbb{L}^p(\Omega,\mu)$ for some $p>2$. Assume that φ satisfies the condition (1.7) *for some $\beta\in[0,1)$ and some auxiliary function ψ. Then there exist $r\in(0,1)$ and a coupling between $(X_i,\varphi_i)_{i\in\mathbb{N}}$ and $(\tilde{X}_i,\tilde{\varphi}_i)_{i\in\mathbb{N}}$ such that, almost surely, for all large enough integer N,*

$$\left|\sum_{i=0}^{N-1} X_i - \sum_{i=0}^{N-1} \tilde{X}_i\right| \leq N^{\frac{r}{2}},$$
$$\left|\sum_{i=0}^{N-1} \varphi_i - \sum_{i=0}^{N-1} \tilde{\varphi}_i\right| \leq \psi^*(N^r).$$

Theorem 2.4.

Let (Ω,μ,T) be a dynamical system which induces a mixing Gibbs-Markov map on a Borel set A. Assume that the function ψ associated with φ_A through Equation (1.8) *is regularly varying with index $\beta\in[0,1)$.*

Let f be in $\mathbb{L}^1(\Omega,\mu)$. Assume that $\int_\Omega f \,\mathrm{d}\mu=0$, that the random variable $X_{|f|}$ belongs to $\mathbb{L}^p(A,\mu_{|A})$ for some $p>2$ and that $\mathbb{E}_{\mu_{|A}}(D(X_f))$ is finite. Then, for any probability measure $\nu\ll\mu$:

$$\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(N)}}\sum_{i=0}^{N-1} f\circ T^i \to \sigma(f)\sqrt{Y_\beta}\mathcal{N}, \tag{2.2}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω,ν) to $\mathbb{R}$, where Y_β and $\mathcal{N}$ are independent, Y_β is a standard Mittag-Leffler distribution of order β and $\mathcal{N}$ is a standard Gaussian random variable, and where:

$$\sigma(f)^2 = \int_A X_f^2 \,\mathrm{d}\mu + 2\sum_{i=1}^{+\infty}\int_A X_f\cdot X_f\circ T_A^i \,\mathrm{d}\mu. \tag{2.3}$$

Moreover, $\sigma(f)=0$ if and only if f is a coboundary.

This article generalizes these two theorems in multiple directions. We will explain how these two theorems can be proved, so that we can control the influence of any change in the hypotheses.

2.1 Proof of the asymptotic independence

Here is a sketch of the proof of Theorem 2.3 as done in [16]. The techniques were initially designed by E. Csáki and A. Földes and were applied to random walks in [7] and [8]. The argument works well with independent excursions, but does not apply directly to Gibbs-Markov systems. The author had to find a way to get back to a situation where independence is satisfied, hence the notion of *piecewise i.i.d. processes*. We recall the relevant definitions.

We first choose two parameters q and ε in $(0,1)$. Then, we cut the set of non-negative integers into boxes of polynomial size in the following way. For any positive integers n and k with $k<2^{(1-q)n}$, we define $I_{n,k}:=\{i\in\mathbb{N} : 2^n+k2^{qn}\leq i<2^n+(k+1)2^{qn}-2^{q\varepsilon n}\}$ and $J_{n,k}:=\{i\in\mathbb{N} : 2^n+(k+1)2^{qn}-2^{q\varepsilon n}\leq i<2^n+(k+1)2^{qn}\}$. By convention, if $k\geq 2^{(1-q)n}$, we put $I_{n,k}=J_{n,k}=\varnothing$. We put $I:=\bigcup_{(n,k)\in\mathbb{N}^2} I_{n,k}$ and $J:=\bigcup_{(n,k)\in\mathbb{N}^2} J_{n,k}$. The set J is the set of gaps between the boxes; it will be large enough that the processes defined on two distinct boxes $I_{n,k}$ and $I_{n',k'}$ with $(n,k)\neq(n',k')$ are almost independent.

Definition 2.5 (Piecewise i.i.d. processes)**.**

Let $\mathbb{B}$ be a Banach space. We say that a sequence of $\mathbb{B}$-valued random variables $(Y_i)_{i\in\mathbb{N}}$ is a piecewise i.i.d. process with parameters q and ε if:

- *the $(Y_i)_{i\in I}$ are identically distributed;*
- *if i belongs to J, then $Y_i = 0$;*
- *for all n, the $\mathbb{B}^{|I_{n,k}|}$-valued random variables $((Y_i)_{i\in I_{n,k}})_{0\le k<2^{(1-q)n}}$ are independent and identically distributed.*

The proof of Theorem 2.3 proceeds in three steps. First, we prove that, under suitable conditions on X and φ, we can couple the process $(X_i, \varphi_i)_{i\in I}$ with a piecewise i.i.d. process with an almost surely bounded difference. This is Lemma 3.2 in [16]:

Lemma 2.6.
Let (Ω, μ, T) be a mixing Gibbs-Markov map. Let Y be a function from Ω to a Banach space $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$, and let q and ε be in $(0,1)$. We put $Y_i := Y \circ T^i$. Let (Y_i^*) be a piecewise i.i.d. process with parameters q and ε such that, for all integers n and k, the sequences $(Y_i)_{i\in I_{n,k}}$ and $(Y_i^*)_{i\in I_{n,k}}$ have the same law.

Assume furthermore that $\mathbb{E}(D(Y))$ is finite. Then there exists a coupling between $(Y_i)_{i\in\mathbb{N}}$ and $(Y_i^*)_{i\in\mathbb{N}}$ such that, almost surely,

$$\sum_{i\in I} \|Y_i - Y_i^*\|_{\mathbb{B}} < +\infty.$$

Then, we check that throwing away the part of the process which are in the gaps makes only a small difference, provided that the gaps are small enough. These are Lemma 3.3 and Lemma 3.4 in [16], respectively:

Lemma 2.7.
Assume that φ satisfies the condition (1.7). Let $r \in (1-(1-\varepsilon)q, 1)$. Almost surely, for all large enough integer N,

$$\sum_{\substack{i\le N\\ i\in J}} \varphi_i \le \psi^*(N^r). \tag{2.4}$$

Lemma 2.8.
Assume that $X \in \mathbb{L}^p(\Omega, \mu)$ for some $p > 2$. Let $r \in (1-(1-\varepsilon)q, 1)$. Almost surely, for all large enough integer N,

$$\left|\sum_{\substack{i\le N\\ i\in J}} X_i\right| \le N^{\frac{r}{2}}. \tag{2.5}$$

At last, we prove the asymptotic independence for the piecewise i.i.d. process. This is Proposition 3.5 in [16]:

Proposition 2.9.
Let $(X_i^*, \varphi_i^*)_{i\in\mathbb{N}}$ be a piecewise i.i.d. process with parameters q and ε. Let $(\overline{X}_i)_{i\in\mathbb{N}}$ and $(\overline{\varphi}_i)_{i\in\mathbb{N}}$ be two independent processes, such that $(\overline{X}_i)_{i\in\mathbb{N}}$ and $(X_i^*)_{i\in\mathbb{N}}$ have the same law, and so do $(\overline{\varphi}_i)_{i\in\mathbb{N}}$ and $(\varphi_i^*)_{i\in\mathbb{N}}$.

Assume that there exist $p > 2$ and a constant C such that, for all n,

$$\mathbb{E}\left(\sup_{i<2^{qn}} \left|\sum_{\ell=2^n}^{2^n+i} X_\ell^*\right|^p\right) \le C 2^{\frac{pqn}{2}}. \tag{2.6}$$

If φ satisfies the condition (1.7), then there exist $r \in (0,1)$ and a coupling between $(X_i^*, \varphi_i^*)_{i\in\mathbb{N}}$ and $(\overline{X}_i, \overline{\varphi}_i)_{i\in\mathbb{N}}$ such that, almost surely, for all large enough integer N,

$$\left|\sum_{i=0}^{N-1} X_i^* - \sum_{i=0}^{N-1} \overline{X}_i\right| \leq N^{\frac{r}{2}}, \tag{2.7}$$

$$\left|\sum_{i=0}^{N-1} \varphi_i^* - \sum_{i=0}^{N-1} \overline{\varphi}_i\right| \leq \psi^*(N^r).$$

To deduce Theorem 2.3 from these lemmas and propositions, we start from the process $(X_i, \varphi_i) = (X \circ T^i, \varphi \circ T^i)$. With Lemma 2.6, we couple it with a piecewise i.i.d. process (X_i^*, φ_i^*). With Proposition 2.9, we couple (X_i^*, φ_i^*) with a piecewise i.i.d. process $(\overline{X}_i, \overline{\varphi}_i)$, where $(\overline{X}_i)$ and $(\overline{\varphi}_i)$ are independent. Finally, using Lemma 2.6 again, we couple $(\overline{X}_i)$ with $(\tilde{X}_i) \simeq (X \circ T^i)$, and $(\overline{\varphi}_i)$ with $(\tilde{\varphi}_i) \simeq (\varphi \circ T^i)$. The results copied here ensure that, with well-chosen coupling, at each step the processes stay close one to the other.

Since we want to expand our initial result, we will check that each of these propositions remains true under weaker conditions. The proof of Lemma 2.6 is very technical, and we will have to redo some non-trivial computations. The proof of Lemma 2.7 will stay unchanged. The proof of Lemma 2.8 relies on two main ingredients; the first is a decorrelation as in Lemma 2.6 (except this time we work with the blocks $J_{n,k}$ and throw away the blocks $I_{n,k}$), and the second is a consequence of Burkholder-Rosenthal's inequality, i.e., we need to prove that there exists a constant C (depending on X) such that, for all N,

$$\mathbb{E}\left(\sup_{k\leq N}\left|\sum_{i=0}^{k-1} X \circ T^i\right|^p\right) \leq CN^{\frac{p}{2}}.$$

Finally, Proposition 2.9 remains the same, but the same consequence of Burkholder-Rosenthal's inequality is needed to check that its hypotheses are satisfied.

In a nutshell, to get a version of Theorem 2.3 under weaker conditions, there are two non-trivial ingredients: the proof of Lemma 2.6, and Burkholder-Rosenthal's inequality.

2.2 Proof of the generalized CLT

We now explain how Theorem 2.3 can be used so as to get Theorem 2.4 as in [16]. In the following, (Ω, μ, T) is an ergodic dynamical system, where μ is an infinite, σ-finite T-invariant measure. Assume that there is a measurable subset $A \subset \Omega$ such that $\mu(A) = 1$ and the induced system $(A, \pi, d, \mu_{|A}, T_A)$ is a mixing Gibbs-Markov map (the former condition can be achieved by normalizing the measure μ).

We denote by $\xi_N(x) := \sum_{i=1}^N 1_A(T^i x)$ the local time in A at time N, and by $\tau_N(x) := \sum_{i=0}^{N-1} \varphi_A(T_A^i x)$ the sequence of return times in A (which is also the generalized inverse of $(\xi_N)_{N\geq 0}$).

In this setting, the function φ_A is constant on each element of the partition π (so that $D(\varphi_A) \equiv 0$), and not integrable. We assume that φ_A has nice tails, i.e., that it satisfies Condition (1.8) for some $\beta \in [0,1)$ and some auxiliary function ψ.

We assume that $X_{|f|}$ belongs to $\mathbb{L}^p(A, \mu_{|A})$ for some $p > 2$, that $\mathbb{E}(D(X_f))$ is finite, and that $\int_\Omega f \, \mathrm{d}\mu = 0$.

For any x in A and any large enough (depending on x) integer N, we have:

$$\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(N)}} \sum_{i=0}^{N-1} f(T^i x) = \sqrt{\frac{\xi_N(x)}{\operatorname{sinc}(\beta\pi)\psi(N)}} \frac{1}{\sqrt{\xi_N}} \sum_{i=0}^{\xi_N(x)-1} X_f(T_A^i x) + \frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(N)}} \sum_{i=\tau_{\xi_N}}^{N-1} f(T^i x).$$

The last term will prove to be negligible under our assumptions. By [16, Proposition 4.4],

$$\frac{\xi_N}{\operatorname{sinc}(\beta\pi)\psi(N)} \to Y_\beta,$$

where the convergence is in distribution on $(A, \mu_{|A})$ and Y_β is a normalized Mittag-Leffler random variable with parameter β. The assumption that $(A, \pi, d, \mu_{|A}, T_A)$ be mixing is not even necessary: ergodicity is enough here.

Since X_f belongs to $\mathbb{L}^p(A, \mu_{|A})$, has a null integral and $\mathbb{E}(D(X_f))$ is finite, we have a central limit theorem 1.4:

$$\frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X_f \circ T_A^i \to \sigma(f)\mathcal{N},$$

where the convergence is in distribution on $(A, \mu_{|A})$, where $\mathcal{N}$ is a normalized Gaussian random variable and where the standard deviation $\sigma(f)$ is given by Equation (2.3).

If (ξ_N) and $(\sum_{i=0}^{N-1} X_f \circ T_A^i)$ were to be independent, by [16, Lemma 4.1] we would have:

$$\sqrt{\frac{\xi_N(x)}{\operatorname{sinc}(\beta\pi)\psi(N)}} \frac{1}{\sqrt{\xi_N}} \sum_{i=0}^{\xi_N(x)-1} X_f(T_A^i x) \to \sigma(f)\sqrt{Y_\beta}\mathcal{N},$$

with the same conventions as above. However, these processes are far from being independent. The critical point here is that, by Theorem 2.3, the processes (τ_N) and $(\sum_{i=0}^{N-1} X_f \circ T_A^i)$ are asymptotically independent. We can couple them closely with two processes $(\tilde{\tau}_N)$ and $(\sum_{i=0}^{N-1} \tilde{X}_{f,i})$ which are actually independent, and for which the convergence in distribution holds. Then, we need to check that the distance between $(\sum_{i=0}^{N-1} X_f \circ T_A^i, \tau_N)$ and $(\sum_{i=0}^{N-1} \tilde{X}_{f,i}, \tilde{\tau}_N)$ is not too large that the convergence in distribution fails for the initial process. This is the role of the two following lemmas, respectively Lemma 4.4 and Lemma 4.5 in [16].

Lemma 2.10.

Let (Ω, μ, T) be a dynamical system which induces a mixing Gibbs-Markov map on a Borel set A. Assume that the random variable φ_A on $(A, \mu_{|A})$ fulfills the condition (1.8). *Let $r > 0$ and $r^* > r$. Then, $\mu_{|A}$-almost surely, for all large enough integer N,*

$$\sup_{k \le N} \left(\xi_{k+\psi^*(\psi(N)^r)} - \xi_k \right) \le \psi(N)^{r^*}. \tag{2.8}$$

Lemma 2.11.

Let $(A, d, \mu_{|A}, T_A)$ be a mixing Gibbs-Markov map. Let $X \in \mathbb{L}^p(A, \mu_{|A})$, with $p > 2$, be such that $\int_A X \, \mathrm{d}\mu_{|A} = 0$ and $\mathbb{E}_{\mu_{|A}}(D(X)) < +\infty$.

For all $r \in [0, 1]$ and all $r^* > 2/p + (1 - 2/p)r$, almost surely, for all large enough integer N,

$$\sup_{k \le N} \sup_{i \le N^r} \left| \sum_{j=k}^{k+i-1} X \circ T_A^j \right| \le N^{\frac{r^*}{2}}. \tag{2.9}$$

Lemma 2.11 is a consequence of Burkholder-Rosenthal's inequality. Lemma 2.10 is trickier to prove, and we use a coupling technique as in the proof of Lemma 2.6.

Finally, once we get the convergence in distribution according to the invariant distribution on the induced system in A, we extend it to any absolutely continuous starting distribution using [17, Corollary 1].

In a nutshell, to get a version of Theorem 2.4 under weaker conditions, provided that Theorem 2.3 has already been proved, the critical components are some decorrelation results for the proof of Lemma 2.10, Burkholder-Rosenthal's inequality, a central limit theorem, and the convergence in distribution of the local times (see Proposition 4.4 in [16]). We shall also use in Section 6 a generalization of [16, Lemma 4.1] (which, roughly, states the asymptotic independence of the processes (A_{B_n}) and (B_n) given the independence of (A_n) and (B_n)) to continuous times, but the changes to apply are extremely straightforward. Finally, [17, Corollary 1], which extends the convergence in distribution from one initial probability measure to every absolutely continuous probability measure, works well as long as the dynamical systems are indexed by a discrete time, but breaks down when the time is continuous. We will have to replace it in Section 6.

3 Weakening the smoothness condition

The goal of this section is to prove the results stated in Theorem 2.3 and Theorem 2.4 under a milder regularity condition. In Theorem 2.3, we require the real-valued random variables X and φ defined on (Ω, d, μ, T) to satisfy the condition (1.5):

$$\mathbb{E}(D(X)) + \mathbb{E}(D(\varphi)) < +\infty.$$

Let θ be in $(0, 1]$. One can check that the following condition yields the same conclusions:

$$\mathbb{E}(D(X)^\theta) + \mathbb{E}(D(\varphi)^\theta) < +\infty.$$

This weaker condition still guarantees an exponential decay of correlations, although one has to let $\mathcal{L}$ act on different Banach spaces to prove it, and the decorrelation will be slower. However, we can go further and choose a criterion which only guarantees a polynomial decay of correlations. We shall assume that, for some parameter $\theta > 0$, the functions X and φ satisfy:

$$\mathbb{E}(\ln_+(D(X))^{1+\theta}) + \mathbb{E}(\ln_+(D(\varphi))^{1+\theta}) < +\infty, \tag{3.1}$$

where $\ln_+$ is the positive part of the logarithm. The price we pay for this is that the regularity condition on X and φ on the one hand, and the integrability condition on X on the other hand, are no longer independent. Namely, we will prove in Theorem 3.5 that the asymptotic independence (Theorem 2.3) and the generalized central limit theorem (Theorem 2.4) hold under (3.1) if $\theta > 1/(p-1)$.

3.1 Preliminary results

We will begin by getting the central limit theorem and Burkholder-Rosenthal's inequality under the new regularity condition. As we can expect only a polynomial decay of correlation, spectral methods can no longer be used directly; we use martingale methods instead. Our first result is a bound on the speed at which the norm of $\mathcal{L}^i X$ decays.

Lemma 3.1.

Let (Ω, d, μ, T) be a Gibbs-Markov map. Let $X \in \mathbb{L}^p(\Omega)$, with $p > 1$, be such that $\int_\Omega X \, \mathrm{d}\mu = 0$. Assume that, for some $\theta > 0$, we have $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$. Then $\left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p} = O(N^{-\frac{(p-1)(1+\theta)}{p}})$.

In particular, if $\theta > 1/(p-1)$ then $\sum_{N\in\mathbb{N}} \left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p}$ is finite.

Proof.

The first part of the proof is a light version of the proof of Lemma 2.6 (Lemma 3.2 in [16]), where we construct a coupling which satisfies some bounds in probability. The second part yields our claim.

First step: Construction of a nice coupling

For any point x in Ω and any $N \geq 0$, we define:

$$\tilde{\mu}_x^{(N)} := \sum_{\{y : T^N y = x\}} g^{(N)}(y)\delta_y.$$

For any measurable and integrable function X defined on Ω, we have $\mathcal{L}^n X(x) = \int_\Omega X \, \mathrm{d}\tilde{\mu}_x^{(N)}$. Let $M \geq 0$. By the second step of the proof of Lemma 3.2 in [16], there exist constants $C \geq 1$ and $\rho \in (0,1)$ such that, for any cylinder $\overline{a}$ in π^M,

$$\left|\tilde{\mu}_x^{(N+M)}(\overline{a}) - \mu(\overline{a})\right| = \left|\mathcal{L}^{N+M} 1_{\overline{a}}(x) - \int_\Omega 1_{\overline{a}} \, \mathrm{d}\mu\right| \leq C\mu(\overline{a})\rho^N. \tag{3.2}$$

Let X be as in Lemma 3.1. Let x be in Ω; let $X_x^{(N+M)}$ be the random variable X defined on the probability space $(\Omega, \tilde{\mu}_x^{(N+M)})$, and let X^* be a random variable distributed as X on (Ω, μ). We now construct a coupling between $X_x^{(N+M)}$ and X^* such that these two random variables are close with high probability.

Let $\mathcal{X}^*$ be a random variable with values in Ω and whose law is μ, and $\mathcal{X}_x^{(N+M)}$ be a random variable with values in Ω and whose law is $\tilde{\mu}_x^{(N+M)}$, so that $X^* = X(\mathcal{X}^*)$ and $X_x^{(N+M)} = X(\mathcal{X}_x^{(N+M)})$ in distribution. By Equation (3.2), there exists a coupling between $\mathcal{X}_x^{(N+M)}$ and $\mathcal{X}^*$ such that, with probability at least $1 - C\rho^N$, those two random variable take their values in the same cylinder of length M. This induces a coupling between $X_x^{(N+M)}$ and X^*.

A cylinder $\overline{a} = [a_0, \cdots, a_{M-1}]$ of length M has a diameter of at most $\mathrm{Diam}(\Omega)\lambda^{-M}$, so that $X(\overline{a})$ has a diameter of at most $|X|_{\mathrm{Lip}(a_0)} \mathrm{Diam}(\Omega)\lambda^{-M}$. This gives us a control on the difference between $X_x^{(N+M)}$ and X_x^* as long as $\mathcal{X}_x^{(N+M)}$ and $\mathcal{X}^*$ land in the same cylinder of length M.

Let γ be a random variable taking its values in π such that $\mathbb{P}(\gamma = a) = \mu(a)$. Then, for any $\delta > 0$:

$$\begin{aligned}
\mathbb{P}\left(\left|X_x^{(N+M)} - X^*\right| > \delta\right) &\leq C\rho^N + \sum_{a \in \pi} \mu(a) 1\left(|X|_{\mathrm{Lip}(a)} \mathrm{Diam}(\Omega)\lambda^{-M} > \delta\right) \\
&= C\rho^N + \mathbb{P}_\gamma\left(|X|_{\mathrm{Lip}(\gamma)} \mathrm{Diam}(\Omega)\lambda^{-M} > \delta\right).
\end{aligned} \tag{3.3}$$

Now, we use a logarithmic Markov's inequality. If $\lambda^M \mathrm{Diam}(\Omega)^{-1}\delta \geq 1$, then:

$$\begin{aligned}
\mathbb{P}\left(\left|X_x^{(N+M)} - X^*\right| > \delta\right) &\leq C\rho^N + \mathbb{P}_\gamma\left(\ln\left(1 + |X|_{\mathrm{Lip}(\gamma)}\right)^{1+\theta} > \ln\left(1 + \lambda^M \mathrm{Diam}(\Omega)^{-1}\delta\right)^{1+\theta}\right) \\
&\leq C\rho^N + \frac{\mathbb{E}\left(\ln\left(1 + |X|_{\mathrm{Lip}(\gamma)}\right)^{1+\theta}\right)}{\ln\left(1 + \lambda^M \mathrm{Diam}(\Omega)^{-1}\delta\right)^{1+\theta}} \\
&\leq C\rho^N + \frac{\mathbb{E}\left(\left(1 + \ln_+\left(|X|_{\mathrm{Lip}(\gamma)}\right)\right)^{1+\theta}\right)}{\ln\left(\lambda^M \mathrm{Diam}(\Omega)^{-1}\delta\right)^{1+\theta}} \\
&\leq C\rho^N + \frac{2^\theta\left(1 + \mathbb{E}(\ln_+(D(X))^{1+\theta})\right)}{\ln\left(\lambda^M \mathrm{Diam}(\Omega)^{-1}\delta\right)^{1+\theta}}.
\end{aligned}$$

Let us take, for instance, $\delta = \mathrm{Diam}(\Omega)\lambda^{-\frac{M}{2}}$. Then we have, for some constant $C \geq 0$:

$$\mathbb{P}\left(\left|X_x^{(N+M)} - X^*\right| > \mathrm{Diam}(\Omega)\lambda^{-\frac{M}{2}}\right) \leq C\rho^N + CM^{-(1+\theta)}. \tag{3.4}$$

Second step: $\mathbb{L}^p$ bound

Let x be in Ω. Let $N \geq 0$. Let X^* and $X_x^{(N)}$ be as in the first part of the proof, and coupled in such a way that $\mathbb{P}_\mu\left(\left|X_x^{(N)} - X^*\right| > \mathrm{Diam}(\Omega)\lambda^{-\frac{N}{4}}\right) \leq C\rho^{\frac{N}{2}} + CN^{-(1+\theta)}$.

Let p^* be such that $1/p + 1/p^* = 1$. We denote by O the set on which $\left|X^* - X_x^{(N)}\right| > \mathrm{Diam}(\Omega)\lambda^{-\frac{N}{4}}$. Since $\mathbb{E}(X^*) = 0$:

$$\begin{aligned}
\left|\mathcal{L}^N X(x)\right| &= \left|\mathbb{E}(X^*) - \mathbb{E}(X_x^{(N)})\right| \\
&\leq \mathbb{E}\left(\left|X^* - X_x^{(N)}\right|\right) \\
&\leq \mathrm{Diam}(\Omega)\lambda^{-\frac{N}{4}} + \left\|1_O(X^* - X_x^{(N)})\right\|_{\mathbb{L}^1} \\
&\leq CN^{-\frac{(1+\theta)}{p^*}}\left(1 + \|X\|_{\mathbb{L}^p} + \left\|X_x^{(N)}\right\|_{\mathbb{L}^p}\right),
\end{aligned}$$

where the last inequality is an application of Hölder's inequality. Moreover, notice that $\left\|X_x^{(N)}\right\|_{\mathbb{L}^p} = \left(\mathcal{L}^N|X|^p\right)^{\frac{1}{p}}(x)$. Let us take the $\mathbb{L}^p$ norm on both sides. We recall the fact that the operator $\mathcal{L}$ is a contraction when acting on $\mathbb{L}^1$.

$$\begin{aligned}
\left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p} &\leq CN^{-\frac{(1+\theta)}{p^*}}\left(1 + \|X\|_{\mathbb{L}^p} + \left\|\left\|X_x^{(N)}\right\|_{\mathbb{L}^p}\right\|_{\mathbb{L}^p}\right) \\
&= CN^{-\frac{(1+\theta)}{p^*}}\left(1 + \|X\|_{\mathbb{L}^p} + \left\|\left(\mathcal{L}^N|X|^p\right)^{\frac{1}{p}}\right\|_{\mathbb{L}^p}\right) \\
&= CN^{-\frac{(1+\theta)}{p^*}}\left(1 + \|X\|_{\mathbb{L}^p} + \left\|\left(\mathcal{L}^N|X|^p\right)\right\|_{\mathbb{L}^1}^{\frac{1}{p}}\right) \\
&\leq CN^{-\frac{(1+\theta)}{p^*}}\left(1 + \|X\|_{\mathbb{L}^p} + \left\||X|^p\right\|_{\mathbb{L}^1}^{\frac{1}{p}}\right) \\
&= CN^{-\frac{(1+\theta)}{p^*}}\left(1 + 2\|X\|_{\mathbb{L}^p}\right) \\
&= O(N^{-\frac{(p-1)(1+\theta)}{p}}).
\end{aligned}$$

□

Thanks to martingale methods developed by Gordin, this is enough to get a central limit theorem and Burkholder-Rosenthal's inequality, as we will explain now. Let $(\Omega, d, \mathcal{B}, \mu, T)$ be a Gibbs-Markov map. Let X be a function on Ω satisfying the assumptions of Lemma 3.1. We define $C(X) := \sum_{N=1}^{+\infty} \mathcal{L}^N X$ and $\tilde{X} := X + C(X) - C(X) \circ T$. Then, $\tilde{X}$ and X are cohomologous, and by Lemma 3.1 the function $\tilde{X}$ belongs to $\mathbb{L}^p$. Moreover, a short computation shows that $\mathcal{L}\tilde{X} = 0$.

The variables X and $\tilde{X}$ differ only by a coboundary, whose influence on the partial sums will be in general negligible. For all $\varepsilon > 0$, the random variables $N^{-\varepsilon}C(X)$ and $N^{-\varepsilon}C(X) \circ T^N$ converge in $\mathbb{L}^p$ to 0, and $N^{-\varepsilon}C(X)$ converges almost surely to 0 as N goes to $+\infty$. Since $C(X)$ belongs to $\mathbb{L}^p$, the function $|C(X)|^p$ belongs to $\mathbb{L}^1$, and by Birkhoff's theorem, $N^{-1}|C(X)|^p \circ T^n$ converges almost surely to 0. Hence, $N^{-1/p}C(X) \circ T^n$ converges almost surely to 0. These facts mean that the asymptotic behavior of $N^{-\varepsilon}\sum_{i=0}^{N-1} X \circ T^i$ and $N^{-\varepsilon}\sum_{i=0}^{N-1} \tilde{X} \circ T^i$ is the same for all $\varepsilon > 0$ if we look for $\mathbb{L}^p$ bounds and convergence, and for all $\varepsilon \geq 1/p$ if we look for almost sure convergence.

Now, we define, for all $N \in \mathbb{N}$, the σ-algebra $\mathcal{G}_N := T^{-N}\mathcal{B}$. The sequence $(\mathcal{G}_N)_{N \geq 0}$ is a decreasing filtration. For all $i > N \geq 0$, we have $\mathbb{E}(\tilde{X} \circ T^N | \mathcal{G}_i) = (\mathcal{L}^{i-N}\tilde{X}) \circ T^i = 0$. Hence, $(\tilde{X} \circ T^N)_{N \geq 0}$ is a sequence of reverse martingale differences for the filtration $(\mathcal{G}_N)_{N \geq 0}$. The central limit theorem follows:

Theorem 3.2.

Let (Ω, d, μ, T) be a Gibbs-Markov map. Let $X \in \mathbb{L}^p(\Omega)$, with $p \geq 2$, be such that $\int_\Omega X \, \mathrm{d}\mu = 0$. Assume that, for some $\theta > 1/(p-1)$, we have $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$. Then:

$$\lim_{N \to +\infty} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X \circ T^i = \mathcal{N}(0, \sigma^2),$$

where the convergence is in distribution on (Ω, μ), and where $\mathcal{N}(0, \sigma^2)$ is a centered Gaussian random variable of variance:

$$\sigma^2 = \int_\Omega X^2 \, \mathrm{d}\mu + 2 \sum_{N=1}^{+\infty} \int_\Omega X \cdot X \circ T^N \, \mathrm{d}\mu.$$

The variance σ^2 is zero if and only X is a coboundary.

We now prove a version (or rather, a consequence) of Burkholder-Rosenthal's inequality.

Theorem 3.3 (Burkholder-Rosenthal's inequality).
Let (Ω, d, μ, T) be a Gibbs-Markov map. Let $X \in \mathbb{L}^p$, with $p \geq 2$, be a function satisfying the assumptions of Theorem 3.2. Then there exists a constant C such that, for all $N \geq 0$,

$$\left\| \sum_{i=0}^{N-1} X \circ T^i \right\|_{\mathbb{L}^p} \leq C N^{\frac{1}{2}}. \tag{3.5}$$

Proof.

We have put $C(X) = \sum_{N=1}^{+\infty} \mathcal{L}^N X$. Since $C(X) - C(X) \circ T$ is a coboundary,

$$\lim_{N \to +\infty} N^{-\frac{1}{2}} \left\| \sum_{i=0}^{N-1} (C(X) - C(X) \circ T) \circ T^N \right\|_{\mathbb{L}^p} = 0.$$

Hence, we only have to prove this theorem for $\tilde{X} = X + C(X) - C(X) \circ T$. For any given positive N, we apply Burkholder-Rosenthal's inequality (Theorem 21.1 in [6]) to the reverse martingale differences $\left(\tilde{X} \circ T^i\right)_{0 \leq i \leq N-1}$. There exists a constant C_p, depending only on p, such that, for all $N \geq 0$,

$$\left\| \sum_{i=0}^{N-1} \tilde{X} \circ T^i \right\|_{\mathbb{L}^p} \leq C_p \left(\left(\mathbb{E} \left(\sum_{i=0}^{N-1} \mathbb{E}(\tilde{X}^2 \circ T^i | \mathcal{G}_{i+1}) \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} + \left(\mathbb{E} \left(\sum_{i=0}^{N-1} \left| \tilde{X} \circ T^i \right|^p \right) \right)^{\frac{1}{p}} \right).$$

We first use the identity $\mathbb{E}(\tilde{X}^2 \circ T^i | \mathcal{G}_{i+1}) = \left(\mathcal{L}\tilde{X}^2\right) \circ T^{i+1}$. This yields, for all $N > 0$:

$$\left(\mathbb{E} \left(\sum_{i=0}^{N-1} \mathbb{E}(\tilde{X}^2 \circ T^i | \mathcal{G}_{i+1}) \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} = \left\| \sum_{i=0}^{N-1} \left(\mathcal{L}\tilde{X}^2\right) \circ T^i \right\|_{\mathbb{L}^{\frac{p}{2}}}^{\frac{1}{2}} \leq \left\| \tilde{X} \right\|_{\mathbb{L}^p} N^{\frac{1}{2}}.$$

We deal now with the last member of the inequality. One has:

$$\left(\mathbb{E} \left(\sum_{i=0}^{N-1} |\tilde{X} \circ T^i|^p \right) \right)^{\frac{1}{p}} = \left(N \mathbb{E}(\left|\tilde{X}\right|^p) \right)^{\frac{1}{p}} = \left\| \tilde{X} \right\|_{\mathbb{L}^p} N^{\frac{1}{p}}.$$ □

The process $(X \circ T^i)$ is stationary. By an inequality by Serfling [15, Corollary B1], together with the bound (3.5), under the conditions of Theorem 3.3 and if additionally $p > 2$, there exists a constant C such that, for all N,

$$\left\| \sup_{k \leq N} \sum_{i=0}^{k-1} X \circ T^i \right\|_{\mathbb{L}^p} \leq C N^{\frac{1}{2}}.$$

This is actually this inequality we use, for instance to check the hypotheses of Proposition 2.9.

3.2 Adaptation of the Csáki-Földes argument

In this subsection we shall get generalizations of the main results of [16], Theorems 2.3 and 2.4, in a low regularity setting. The central limit theorem and Burkholder-Rosenthal's inequality will be instrumental, so the condition (3.1) with $\theta > 1/(p-1)$ will appear. Other inequalities involving the parameters q and ε chosen to construct a piecewise i.i.d. process will have to be satisfied, and in the end we will check that one can still find such parameters under those constraints. As stated in Subsection 2.1, the only significant point left to prove is the equivalent of Lemma 2.6.

Lemma 3.4.
In Lemma 2.6, the condition (1.5) *on* Y *can be replaced by the condition* (3.1), *provided that:*

$$q(1 + \varepsilon\theta) > 1. \tag{3.6}$$

Proof.
We have to adapt some arguments of the original proof of Lemma 2.6. Using the same notations as in the proof of Lemma 3.1, given any non-negative integers n and k and any x in Ω, we define two finite sequences of random variables. The sequence $(Y_{i,y})_{i \in I_{n,k}}$ is the process $(Y \circ T^i x)_{i \in I_{n,k}}$ when the distribution of x is $\tilde{\mu}_x^{2^{qn}}$, and the sequence $(Y_{i,y}^*)_{i \in I_{n,k}}$ is the process $(Y \circ T^i x)_{i \in I_{n,k}}$ when the distribution of x is μ. In [16, Equation (3.3)], the author proved that, under the condition (1.5) on Y, there exists a coupling between $(Y_{i,y})_{i \in I_{n,k}}$ and $(Y_{i,y}^*)_{i \in I_{n,k}}$ such that:

$$\mathbb{P}\left(\sum_{i \in I_{n,k}} \left\| Y_{i,y} - Y_{i,y}^* \right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n - 2}} \right) \leq C \max\left\{\rho, \lambda^{-1}\right\}^{2^{q\varepsilon n - 2}}, \tag{3.7}$$

which was enough to ensure that:

$$\sum_{n \in \mathbb{N}} 2^{(1-q)n} \mathbb{P}\left(\sum_{i \in I_{n,k}} \left\| Y_{i,y} - Y_{i,y}^* \right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n - 2}} \right) < +\infty. \tag{3.8}$$

This last property is the key to prove Lemma 3.4. Here, the bound we will get will not be as good as Equation (3.7), but still enough to prove that Equation (3.8) holds. We use the same coupling between $(Y_{i,y})_{i \in I_{n,k}}$ and $(Y_{i,y}^*)_{i \in I_{n,k}}$ as the one constructed in the fourth step of the proof of [16, Lemma 3.2]. One this coupling is done, by [16, Equation (3.5)], for any $\tau \in (\lambda^{-1}, 1)$, for some positive constants C and C', for any $\delta > 0$,

$$\mathbb{P}\left(\sum_{i \in I_{n,k}} \left\| Y_{i,y} - Y_{i,y}^* \right\|_{\mathbb{B}} > \delta \right) \leq C \rho^{2^{q\varepsilon n - 1}} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P}_\gamma \left(|Y|_{\mathrm{Lip}(\gamma)} > C' (\lambda\tau)^{|I_{n,k}| - i} \lambda^{2^{q\varepsilon n - 1}} \delta \right),$$

where γ is a random variable taking values in π such that $\mathbb{P}(\gamma = a) = \mu(a)$.

This bound is very similar to Equation (3.3), and is proved in the same way. As in the proof of Lemma 3.1, we use a logarithmic Markov's inequality. If $C'\lambda^{2^{q\varepsilon n-1}}\delta > 1$, then:

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y_{i,y}^*\right\|_{\mathbb{B}} > \delta \right) \\
&\qquad \leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P}_\gamma \left(\ln\left(1 + |Y|_{\mathrm{Lip}(\gamma)}\right)^{1+\theta} > \ln\left(1 + C'\lambda^{2^{q\varepsilon n-1}} (\lambda\tau)^{|I_{n,k}|-i}\delta\right)^{1+\theta} \right) \\
&\qquad \leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \frac{\mathbb{E}_\gamma \left(\ln\left(1 + |Y|_{\mathrm{Lip}(\gamma)}\right)^{1+\theta} \right)}{\ln\left(C'\lambda^{2^{q\varepsilon n-1}} (\lambda\tau)^{|I_{n,k}|-i}\delta\right)^{1+\theta}} \\
&\qquad \leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \frac{2\left(1 + \mathbb{E}\left((\ln_+ D(Y))^{1+\theta}\right)\right)}{\left(\ln\left(C'\lambda^{2^{q\varepsilon n-1}}\delta\right) + (|I_{n,k}| - i)\ln(\lambda\tau)\right)^{1+\theta}} \\
&\qquad \leq C\rho^{2^{q\varepsilon n-1}} + \frac{C''}{\ln\left(C'\lambda^{2^{q\varepsilon n-1}}\delta\right)^{\theta}}.
\end{aligned}$$

Let us take $\delta = \lambda^{-2^{q\varepsilon n-2}}$, which is a valid choice for all large enough n. Then we get:

$$\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y_{i,y}^*\right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}} \right) = O(2^{-q\varepsilon\theta n}),$$

which in turn implies that Equation (3.8) holds as long as $1-q-q\varepsilon\theta < 0$, which is condition (3.6). □

All the other ingredients we used to prove Theorems 2.3 and 2.4 only need minor modifications to apply to the low regularity setting, either because the regularity of the observables has nothing to do with them (Lemma 2.7), or because some intermediate result (for instance Burkholder-Rosenthal's inequality) has already been proved.

Theorem 3.5.

In Theorem 2.3, the condition (1.5) *on X and φ can be replaced by the condition* (3.1)*, provided that:*

$$\theta > \frac{1}{p-1}.$$

In Theorem 2.4, the condition (1.5) *on X_f can be replaced by the condition* (3.1)*, provided that:*

$$\theta > \frac{1}{p-1}.$$

Proof.

We begin by proving our new version of Theorem 2.3. The conclusion of Subsection 2.1 is that the only non-trivial modifications are Burkholder-Rosenthal's inequality (Theorem 3.3) and the existence of a nice coupling (Lemma 3.4). As long as their hypotheses are simultaneously satisfied, the conclusion of Theorem 2.3 holds.

We need to choose the parameters q and ε in such a way that Condition (3.6) is satisfied. If we choose the parameter q in $((1+\theta)^{-1}, 1)$, the inequality $q(1+\theta) > 1$ is satisfied. Then, if ε is chosen close enough to 1, the inequality $q(1+\varepsilon\theta) > 1$ is also satisfied. Hence, there is a coherent way to choose the parameters q and ε, which finishes the proof of the first part of Theorem 3.5.

The proof of our new version of Theorem 2.4 given that we have proved this version of Theorem 2.3 is in Section 4 of [16]. The smoothness of X_f is only relevant in proving the central limit theorem (Theorem 3.2) and Burkholder-Rosenthal's inequality for the sequence $\left(\sum_{i<N} X_f \circ T_A^i\right)$, which we have got under our milder regularity condition. $\square$

4 Hilbert space-valued observables

Now that we have replaced spectral methods by martingale methods, we can pick a low-hanging fruit: the case of observables X and f which take their values not in the real line, but in a Hilbert space. This is mostly a matter of finding the relevant literature, and it will be mildly useful in order to remove the assumption of mixing in the next section.

For any measured space (Ω, μ) and any Hilbert space $\mathcal{H}$, we shall denote by $\mathbb{L}^p(\Omega, \mu; \mathcal{H})$ (or simply $\mathbb{L}^p(\Omega; \mathcal{H})$ if there is no ambiguity for the choice of the measure) the space of $\mathcal{H}$-valued Bochner-integrable functions on Ω with a finite moment of order p, endowed with the norm:

$$\|X\|_{\mathbb{L}^p(\Omega;\mathcal{H})} := \left(\int_\Omega \|X\|_{\mathcal{H}}^p \, \mathrm{d}\mu\right)^{\frac{1}{p}}.$$

As in the previous section, we start by proving the central limit theorem and Burkholder-Rosenthal's inequality. Since we use the same martingale methods as in Section 3, we will state our results under the same regularity conditions.

The reader can check first that any proof we wrote, in this article or in [16], and which does not rely either on the central limit theorem or on Burkholder-Rosenthal's inequality can be very easily adapted to fit our new setting. Some results are already designed to work in any Banach space (e.g. Lemma 2.6), or do not involve the observable X in any way (e.g. Lemma 2.7). Otherwise, whenever an inequality involves the absolute value of X, take the Hilbert norm instead; whenever the $\mathbb{L}^p$ norm of X appears, take the $\mathbb{L}^p(\Omega; \mathcal{H})$ norm instead. In particular, if we adapt Lemma 3.1, we get:

Lemma 4.1.

Let (Ω, μ, T) be a Gibbs-Markov map. Let $\mathcal{H}$ be a Hilbert space. Let $X \in \mathbb{L}^p(\Omega; \mathcal{H})$, with $p \geq 2$, be a function such that $\int_\Omega X \, \mathrm{d}\mu = 0$ and which satisfies the regularity condition (3.1) *for some $\theta > 1/(p-1)$.*

Then $\sum_{N \in \mathbb{N}} \left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p(\Omega;\mathcal{H})}$ is finite.

All the discussion of Subsection 3.1, where we explained how to add a coboundary to X and how to define a filtration to get a reverse martingale array, is still valid. If we want to prove new versions of Theorem 2.3 and Theorem 2.4, we only have to prove a central limit theorem and Burkholder-Rosenthal's inequality.

Theorem 4.2 (Central limit theorem in Hilbert spaces)**.**

Let (Ω, μ, T) be a Gibbs-Markov map. Let $\mathcal{H}$ be a Hilbert space. Let X be a function which satisfies all the hypotheses of Lemma 4.1. Then there exists a centered Gaussian variable $\mathcal{N}(0, S)$ on $\mathcal{H}$ with covariance operator S such that:

$$\lim_{N \to +\infty} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X \circ T^i = \mathcal{N}(0, S),$$

where the convergence is in law on (Ω, μ), and where the covariance operator is such that, for any two vectors u and v in $\mathcal{H}$,

$$(Su, v) = \int_\Omega (X, u) (X, v) \, \mathrm{d}\mu + \sum_{i=1}^{+\infty} \int_\Omega \left(X \circ T^i, u\right) (X, v) \, \mathrm{d}\mu + \sum_{i=1}^{+\infty} \int_\Omega (X, u) \left(X \circ T^i, v\right) \, \mathrm{d}\mu.$$

The operator S is degenerate if and only if (X, u) is a coboundary for some non-zero vector u.

Proof.

In order to prove this theorem, we shall use a central limit theorem for Hilbert space-valued martingales, namely Theorem 5.1 in [11]. We have to construct a martingale array, and then check the three conditions of this theorem.

We put, as in Subsection 3.1, $C(X) := \sum_{i=1}^{+\infty} \mathcal{L}^i X$ and $\tilde{X} := X + C(X) - C(X) \circ T$. Then, for each $i \in \mathbb{N}$, we write $\tilde{X}_i := \tilde{X} \circ T^i$. As we have seen, for each integer N and with the natural filtration $(\mathcal{G}_N)_{N \geq 0}$, the sequence $(\tilde{X}_{N-i})_{1 \leq i \leq N}$ is a martingale. Thus, if we define $\tilde{X}_{N,i} := N^{-\frac{1}{2}} \tilde{X}_{N-i}$, the process $(\tilde{X}_{N,i})_{1 \leq i \leq N}$ is a martingale array. Now we check the three conditions to apply Jakubowski's theorem.

To begin with, for any $N > 0$ and $x > 0$:

$$\mathbb{P}\left(\sup_{1 \leq i \leq N} \left\| \tilde{X}_{N,i} \right\|_{\mathcal{H}} > x \right) \leq N \mathbb{P}\left(\left\| \tilde{X} \right\|_{\mathcal{H}} > x N^{\frac{1}{2}} \right);$$

then, we use this estimate to bound the expectation:

$$\begin{aligned}
\mathbb{E}\left(\sup_{1 \leq i \leq N} \left\| \tilde{X}_{N,i} \right\|_{\mathcal{H}}^2 \right) &= \int_0^{+\infty} \mathbb{P}\left(\sup_{1 \leq i \leq N} \left\| \tilde{X}_{N,i} \right\|_{\mathcal{H}} > x \right) \, \mathrm{d}x \\
&\leq \int_0^{+\infty} \min\left\{1, N \mathbb{P}\left(\left\| \tilde{X} \right\|_{\mathcal{H}} > x N^{\frac{1}{2}} \right) \right\} \, \mathrm{d}x \\
&= \frac{1}{\sqrt{N}} \int_0^{+\infty} \min\left\{1, N \mathbb{P}\left(\left\| \tilde{X} \right\|_{\mathcal{H}} > x \right) \right\} \, \mathrm{d}x.
\end{aligned}$$

Let $\varepsilon > 0$, and let $M > 0$ be such that $\mathbb{P}\left(\left\| \tilde{X} \right\|_{\mathcal{H}} > x \right) \leq \varepsilon x^{-2}$ for all $x \geq M$. Then:

$$\begin{aligned}
\mathbb{E}\left(\sup_{1 \leq i \leq N} \left\| \tilde{X}_{N,i} \right\|_{\mathcal{H}}^2 \right) &\leq \frac{M}{\sqrt{N}} + \frac{1}{\sqrt{N}} \int_M^{+\infty} \min\left\{1, \varepsilon N x^{-2}\right\} \, \mathrm{d}x \\
&= \frac{\max\{M, \sqrt{\varepsilon N}\}}{\sqrt{N}} + \frac{1}{\sqrt{N}} \int_{\max\{M, \sqrt{\varepsilon N}\}}^{+\infty} \varepsilon N x^{-2} \, \mathrm{d}x \\
&\leq \frac{\max\{M, \sqrt{\varepsilon N}\}}{\sqrt{N}} + \frac{\varepsilon \sqrt{N}}{\max\{M, \sqrt{\varepsilon N}\}},
\end{aligned}$$

and $\limsup_{N \to +\infty} \mathbb{E}\left(\sup_{1 \leq i \leq N} \left\| \tilde{X}_{N,i} \right\|_{\mathcal{H}}^2 \right) \leq 2\sqrt{\varepsilon}$. Since this is true for all $\varepsilon > 0$, the first condition is satisfied.

Then, for any two vectors u and v in $\mathcal{H}$,

$$\sum_{i=1}^{N} \left(\tilde{X}_{N,i}, u \right) \left(\tilde{X}_{N,i}, v \right) = \frac{1}{N} \sum_{i=1}^{N} \left[\left(\tilde{X}, u \right) \left(\tilde{X}, v \right) \right] \circ T^i,$$

which by Birkhoff's ergodic theorem converges almost surely to a constant as N goes to $+\infty$.

Now, let $(e_j)_{j\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}$, let $M\in\mathbb{N}$ and $\varepsilon>0$. We compute:

$$\mathbb{P}\left(\sum_{1\leq i\leq N}\sum_{j\geq M}\left(\tilde{X}_{N,i},e_j\right)^2>\varepsilon\right)=\mathbb{P}\left(\frac{1}{N}\sum_{1\leq i\leq N}\left(\sum_{j\geq M}\left(\tilde{X},e_j\right)^2\right)\circ T^i>\varepsilon\right)$$
$$\underset{N\to+\infty}{\to} 1\left[\sum_{j\geq M}\mathbb{E}\left(\left(\tilde{X},e_j\right)^2\right)>\varepsilon\right]\underset{M\to+\infty}{\to} 0.$$

Hence, all the hypotheses of the central limit theorem are satisfied, and $N^{-\frac{1}{2}}\sum_{i=0}^{N-1}\tilde{X}\circ T^i$ converges in distribution to a Gaussian random variable. Since the addition of a coboundary does not change such a limit, the partial sums $N^{-\frac{1}{2}}\sum_{i=0}^{N-1}X\circ T^i$ have the same limit in distribution. All we have left to find is the expression of the covariance operator S. The same theorem by Jakubowski states that, for any two vectors u and u in $\mathcal{H}$,

$$(Su,v)=\lim_{N\to+\infty}\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}\tilde{X}\circ T^i,u\right)\left(\sum_{i=0}^{N-1}\tilde{X}\circ T^i,v\right)\right].$$

We first prove that one can replace $\tilde{X}$ by X in this formula. Indeed,

$$\begin{aligned}\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}\tilde{X}\circ T^i,u\right)\left(\sum_{i=0}^{N-1}\tilde{X}\circ T^i,v\right)\right]&=\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}X\circ T^i,u\right)\left(\sum_{i=0}^{N-1}X\circ T^i,v\right)\right]\\&\quad+\mathbb{E}\left[\left(C(X)-C(X)\circ T^N,u\right)\left(\frac{1}{N}\sum_{i=0}^{N-1}\tilde{X}\circ T^i,v\right)\right]\\&\quad+\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=0}^{N-1}X\circ T^i,u\right)\left(C(X)-C(X)\circ T^N,v\right)\right],\end{aligned}$$

and since both $N^{-1}\sum_{i=0}^{N-1}(X,u)\circ T^i$ and $N^{-1}\sum_{i=0}^{N-1}(\tilde{X},v)\circ T^i$ converge to 0 in $\mathbb{L}^2$, the last two lines vanish as N goes to $+\infty$. Hence:

$$\begin{aligned}(Su,v)&=\lim_{N\to+\infty}\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}X\circ T^i,u\right)\left(\sum_{i=0}^{N-1}X\circ T^i,v\right)\right]\\&=\lim_{N\to+\infty}\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\mathbb{E}\left[\left(X\circ T^i,u\right)\left(X\circ T^j,v\right)\right]\\&=\lim_{N\to+\infty}\left[\mathbb{E}\left((X,u)\,(X,v)\right)+\sum_{i=1}^{N-1}\left(1-\frac{i}{N}\right)\mathbb{E}\left(\left(X\circ T^i,u\right)(X,v)\right)\right.\\&\qquad\left.+\sum_{i=1}^{N-1}\left(1-\frac{i}{N}\right)\mathbb{E}\left((X,u)\left(X\circ T^i,v\right)\right)\right]\\&=\mathbb{E}\left[(X,u)\,(X,v)\right]+\sum_{i=1}^{+\infty}\mathbb{E}\left[\left(X\circ T^i,u\right)(X,v)\right]+\sum_{i=1}^{+\infty}\mathbb{E}\left[(X,u)\left(X\circ T^i,v\right)\right],\end{aligned}$$

where the last equality comes from the dominated convergence theorem. The fact that the operator S is degenerate if and only if (X,u) is a coboundary for some non-zero vector u is similar to the 1-dimensional case, for which it is a classical result. □

Burkholder-Rosenthal's inequality for Hilbert space-valued martingales is [12, Theorem 8.33]. Since the central limit theorem and Burkholder-Rosenthal's inequality were the only missing ingredients, new versions of Theorem 2.3 and Theorem 2.4 follow immediately:

Proposition 4.3.

Let (Ω, d, μ, T) be a mixing Gibbs-Markov map. Let $\mathcal{H}$ be a Hilbert space. Let X and φ be measurable functions from Ω to $\mathcal{H}$ and to $\mathbb{R}_+$ respectively, satisfying the condition (3.1) for some parameter θ. We put $(X_i, \varphi_i) := (X \circ T^i, \varphi \circ T^i)$. Let $(\tilde{X}_i)_{i \in \mathbb{N}}$ and $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$ be copies of the processes $(X_i)_{i \in \mathbb{N}}$ and $(\varphi_i)_{i \in \mathbb{N}}$ respectively, such that $(\tilde{X}_i)_{i \in \mathbb{N}}$ and $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$ are mutually independent.

Assume that X belongs to $\mathbb{L}^p(\Omega; \mathcal{H})$ for some $p > 2$; that φ satisfies the condition (1.7) for some $\beta \in [0, 1)$ and some auxiliary function ψ; and that $\theta > 1/(p-1)$.

Then there exist $r \in (0, 1)$ and a coupling between $(X_i, \varphi_i)_{i \in \mathbb{N}}$ and $(\tilde{X}_i, \tilde{\varphi}_i)_{i \in \mathbb{N}}$ such that, almost surely, for all large enough integer N,

$$\left\| \sum_{i=0}^{N-1} X_i - \sum_{i=0}^{N-1} \tilde{X}_i \right\|_{\mathcal{H}} \leq N^{\frac{r}{2}},$$

$$\left| \sum_{i=0}^{N-1} \varphi_i - \sum_{i=0}^{N-1} \tilde{\varphi}_i \right| \leq \psi^*(N^r).$$

Proposition 4.4.

Let (Ω, μ, T) be a dynamical system which induces a mixing Gibbs-Markov map on a Borel set A. Assume that the function ψ associated with φ_A is regularly varying with index $\beta \in [0, 1)$.

Let f be in $\mathbb{L}^1(\Omega, \mu; \mathcal{H})$. Assume that $\int_\Omega f \, \mathrm{d}\mu = 0$, that the random variable $X_{|f|}$ belongs to $\mathbb{L}^p(A, \mu_{|A}; \mathcal{H})$ for some $p > 2$ and satisfies the regularity condition (3.1) for some $\theta > 1/(p-1)$.

Then, for any probability measure $\nu \ll \mu$:

$$\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i \to \sqrt{Y_\beta} \mathcal{N}(0, S), \tag{4.1}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω, ν) to $\mathbb{R}$, where Y_β and $\mathcal{N}(0, S)$ are independent, Y_β is a standard Mittag-Leffler distribution of order β and $\mathcal{N}(0, S)$ is a Gaussian random variable on $\mathcal{H}$ with covariance operator S, and where, for any two vectors u and v in $\mathcal{H}$,

$$(Su, v) = \int_A (X_f, u)(X_f, v) \, \mathrm{d}\mu + \sum_{i=1}^{+\infty} \int_A \left(X_f \circ T_A^i, u\right)(X_f, v) \, \mathrm{d}\mu + \sum_{i=1}^{+\infty} \int_A (X_f, u)\left(X_f \circ T_A^i, v\right) \, \mathrm{d}\mu.$$

Moreover, S is degenerate if and only if (f, u) is a coboundary for some non-zero vector u.

5 Non-mixing maps

The goal of this section is to weaken the hypothesis of mixing of Theorems 2.3 and 2.4, so as to prove the following:

Proposition 5.1.

In Theorem 2.3 and Theorem 2.4, one can replace "mixing" by "ergodic", provided that the variance in Equation (2.3) is replaced by:

$$\sigma(f)^2 = \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_A^i \right)^2 \, \mathrm{d}\mu_{|A}.$$

For the sake of brevity, we use the strong regularity condition (1.5), even though the weaker condition (3.1) works as well.

We shall use the fact that, when an ergodic Gibbs-Markov map is not mixing, some of its iterates will be. More precisely: let (Ω, d, π, T, μ) be an ergodic Gibbs-Markov map. We can assume without loss of generality that there exists a parameter $\lambda > 1$ such that $d(x,y) = \lambda^{-s(x,y)}$, where $s(x,y)$ is the time of separation of x and y for the partition π and the transformation T. Then there exist an integer $M \geq 1$ and a partition $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}}$ of Ω into M subsets such that:

- $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}}$ is coarser than π;
- the subsets A_k are T^M-invariant modulo μ;
- $T(A_k) = A_{k+1}$ for all k in $\mathbb{Z}/M\mathbb{Z}$;
- the dynamical systems (A_k, d, T^M) are topologically mixing.

We use the acceleration process described in Subsection 1.2, with the stopping time constant and equal to M. The map T^M is an iterate of T, and as such preserves μ. We get a Gibbs-Markov map $(\Omega, d_M, \pi_M, T^M, \mu)$ which is not ergodic, but has exactly M ergodic components $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}}$. Moreover, these components have the same measure, and the system is mixing on each of them. We will use the notation $\mu_k := M\mu_{|A_k}$, so that $T_*^\ell \mu_k = \mu_{k+\ell}$ for all non-negative integer ℓ. We will also write $D = D_{\pi,d}$ and $D_M = D_{\pi_M, d_M}$.

This is enough for the consequences of Theorem 2.3 to hold for the sequence of processes $(X_{Mi+\ell}, \varphi_{Mi+\ell})_{i \geq 0}$. To get a control on the initial process, we will work with the vector-valued processes $((X_{Mi+\ell}, \varphi_{Mi+\ell})_{0 \leq \ell < M})_{i \geq 0}$ defined on A_k, and apply a version of Theorem 2.3. We now describe how the integrability and regularity of observables behave with respect to the iteration, and how nice we can expect $(X_{Mi+\ell})_{0 \leq \ell < M}$ and $(\varphi_{Mi+\ell})_{0 \leq \ell < M}$ to be.

Let f be a function defined on Ω. Let k be in $\mathbb{Z}/M\mathbb{Z}$ and $0 \leq \ell < M$. If f belongs to $\mathbb{L}^p(\Omega, \mu)$, then:

$$\left\| f \circ T^\ell \right\|_{\mathbb{L}^p(A_k, \mu_k)} = \left(\int_{A_k} |f \circ T^\ell|^p M \, \mathrm{d}\mu_{|A_k} \right)^{\frac{1}{p}} = M^{\frac{1}{p}} \left(\int_{A_{k+\ell}} |f|^p \, \mathrm{d}\mu_{|A_{k+\ell}} \right)^{\frac{1}{p}} \leq M^{\frac{1}{p}} \left\| f \right\|_{\mathbb{L}^p(\Omega,\mu)},$$

so that each of the functions $f \circ T^\ell$ belongs to $\mathbb{L}^p(A_k, \mu_k)$. Similarly, if f satisfies the condition (1.7) with some auxiliary function ψ, then, for all $x > 0$:

$$\mathbb{P}_{\mu_k}(f \circ T^\ell > x) = \mathbb{P}_{\mu_{k+\ell}}(f > x) \leq M \mathbb{P}_\mu(f > x) \leq \frac{M}{\psi(x)},$$

so that each of the functions $f \circ T^\ell$ satisfies the condition (1.7) with the auxiliary function ψ/M.

Assume now that f is Lipschitz on each element of the partition π. Let a_M be in π_M, with $a_M \subset A_k$, and let x and y be in a_M. Let a be the element of the partition π such that $T^\ell a_k \subset a$. Then:

$$|f(T^\ell x) - f(T^\ell y)| \leq |f|_{\mathrm{Lip}_d(a)} d(T^\ell x, T^\ell y) = \lambda^{\ell - M} |f|_{\mathrm{Lip}_d(a)} d(T^M x, T^M y) \leq \lambda^{\ell - M + 1} |f|_{\mathrm{Lip}_d(a)} d_M(x,y),$$

so that $|f \circ T^\ell|_{\mathrm{Lip}_{d_M}(a_M)} \leq \lambda^{\ell - M + 1} |f|_{\mathrm{Lip}(a)}$. Integrating over all a_k such that $T^\ell a_k \subset a$, and then over all a in $A_{k+\ell} \cap \pi$, we get:

$$\mathbb{E}_{\mu_k} \left(D_M(f \circ T^\ell) \right) \leq \lambda^{\ell - M + 1} \mathbb{E}_\mu \left(D(f) | A_{k+\ell} \right) \leq M\lambda \mathbb{E}_\mu \left(D(f) \right).$$

Hence, if the regularity condition (1.5) is satisfied by f on (Ω, π, d, T, μ), then it is satisfied by $f \circ T^\ell$ on $(A_k, \pi_M, d_M, T^M, \mu_k)$ for all $0 \leq \ell < M$. The same is true for any weakened regularity condition such as condition (3.1).

Now that these fundamental properties are laid down, let us go for the proof of Proposition 5.1. The reader may check that Theorem 2.3 can be trivially generalized to account for multiple occurrences of the heavy-tailed observable:

Proposition 5.2.
Let (Ω, d, μ, T) be a mixing Gibbs-Markov map. Let M be a positive integer. Let $X^{(0)}, \cdots, X^{(M-1)}$ and $\varphi^{(0)}, \cdots, \varphi^{(M-1)}$ be measurable functions from Ω to $\mathbb{R}$ and to $\mathbb{R}_+$ respectively, each one satisfying the condition (1.5). We put $(X_i^{(\ell)}, \varphi_i^{(\ell)}) := (X^{(\ell)} \circ T^i, \varphi^{(\ell)} \circ T^i)$. Let $((\tilde{X}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ and $((\tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ be copies of the processes $((X_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ and $((\varphi_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ respectively, such that $((\tilde{X}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ and $((\tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ are mutually independent.

Assume that there exists $p > 2$ such that each $X^{(\ell)}$ belongs to $\mathbb{L}^p(\Omega, \mu)$. Assume that there exists $\beta \in [0, 1)$ such that each $\varphi^{(\ell)}$ satisfies the condition (1.7) with the same auxiliary function ψ. Then there exist $r \in (0, 1)$ and a coupling between $((X_i^{(\ell)}, \varphi_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ and $((\tilde{X}_i^{(\ell)}, \tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ such that, almost surely, for all large enough integer N,

$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} X_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{X}_i^{(\ell)} \right| \leq N^{\frac{r}{2}},$$
$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} \varphi_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{\varphi}_i^{(\ell)} \right| \leq \psi^*(N^r).$$

This result for the $X^{(\ell)}$ side is a direct consequence of Proposition 4.3 and the equivalence of norms, although it can be proved in a more elementary way. For the $\varphi^{(\ell)}$ side, the proof requires minor tweaks of Lemma 2.7 and Proposition 2.9.

Proof of Proposition 5.1.

Proof of Theorem 2.3 without mixing

Let (Ω, d, T, μ) be an ergodic Gibbs-Markov map. Let X and φ be measurable functions from Ω to $\mathbb{R}$ and to $\mathbb{R}_+$ respectively, both satisfying the condition (1.5), such that X belongs to $\mathbb{L}^p(\Omega, \mu)$ for some $p > 2$ and φ satisfies the condition (1.7) for some $\beta \in [0, 1)$ and some auxiliary function ψ. Let $M \geq 1$ and $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}} \subset \pi$ be as studied in the beginning of this section. Let k be in $\mathbb{Z}/M\mathbb{Z}$.

For all $0 \leq \ell < M$, let us put $X^{(\ell)} := X \circ T^\ell$ and $\varphi^{(\ell)} := \varphi \circ T^\ell$. All these observables are well defined almost everywhere on A_k; moreover, each function $X^{(\ell)}$ belongs to $\mathbb{L}^p(A_k, \mu_k)$ and satisfies the condition (1.5), and each function $\varphi^{(\ell)}$ satisfies the conditions (1.7) (with the same parameter $\beta \in [0, 1)$ and the auxiliary function ψ/M) and (1.5). Hence, we can apply Proposition 5.2 to T^M on A_k. Notice that, for all $x \geq 0$, we have $(\psi/M)^*(x) = \psi^*(Mx)$.

Let $((\tilde{X}_i^{(\ell)}, \tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ be as in Proposition 5.2 and already coupled with $((X_i^{(\ell)}, \varphi_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ in such a way that, μ_k-almost surely, for all large enough integer N:

$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} X_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{X}_i^{(\ell)} \right| \leq N^{\frac{r}{2}},$$
$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} \varphi_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{\varphi}_i^{(\ell)} \right| \leq \psi^*(MN^r).$$

We put $\overline{X}_{Mi+\ell} := \tilde{X}_i^{(\ell)}$, and define a process $(\overline{\varphi}_i)$ in the same way. Then the processes $(\overline{X}_i)$ and $(\overline{\varphi}_i)$ are independent, and distributed respectively as $(X \circ T^i)$ and $(\varphi \circ T^i)$, when the starting point is chosen under the distribution μ_k. In addition, μ_k-almost surely, for all large enough integer $N = N_0 M + N_1$ (with $N_1 < M$),

$$\left|\sum_{i=0}^{N-1} (X \circ T^i - \overline{X}_i)\right| \leq \sum_{\ell=0}^{N_1-1} \left|\sum_{i=0}^{N_0+1} \left(X_i^{(\ell)} - \tilde{X}_i^{(\ell)}\right)\right| + \sum_{\ell=N_1}^{M-1} \left|\sum_{i=0}^{N_0} \left(X_i^{(\ell)} - \tilde{X}_i^{(\ell)}\right)\right| \leq 2\left(\frac{N-1}{M}+1\right)^{\frac{r}{2}},$$

and a similar bound holds for the difference between $\sum_{i=0}^{N-1} \varphi \circ T^i$ and $\sum_{i=0}^{N-1} \overline{\varphi}_i$. Up to an arbitrarily small increase in r, if N is large enough, the coupling between $(X \circ T^i, \varphi \circ T^i)$ and $(\overline{X}_i, \overline{\varphi}_i)$ fulfills the conclusions of Theorem 2.3 when the starting point is chosen according to the distribution μ_k.

To finish the proof, notice that μ is the average of the μ_k:

$$\mu = \frac{1}{M} \sum_{k=0}^{M-1} \mu_k.$$

Since for each k there is a coupling between $(X \circ T^i, \varphi \circ T^i)$ under the distribution μ_k and some suitable process, we have a canonical coupling between $(X \circ T^i, \varphi \circ T^i)$ under the distribution μ and a process $(\hat{X}_i, \hat{\varphi}_i)$, by taking the average of the couplings. The process $(\hat{X}_i, \hat{\varphi}_i)$, together with the coupling we constructed, satisfies all the conclusions of Theorem 2.3.

Proof of Theorem 2.4 without mixing

As we explained in Subsection 2.2, the points about which we need to be careful are Burkholder-Rosenthal's inequality, the central limit theorem, the convergence in distribution of the local times, and the proof of Lemma 2.10.

Burkholder-Rosenthal's inequality for the sequence $(X_f \circ T^i)$ comes either from the multidimensional version of this inequality we used in Section 4, or from the fact that the inequality holds along each of the subsequences $(X_f \circ T^{Mi+k})$, where $0 \leq k < M$.

Once we have induced the system on any $A_k \subset A$ such that T^M is a mixing Gibbs-Markov map on A_k, a central limit theorem holds:

$$\lim_{N \to +\infty} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X_f \circ T_A^i = \sigma(f)\mathcal{N},$$

where the convergence is in distribution on any (A_k, μ_k), where $\mathcal{N}$ is a normalized Gaussian random variable, and where:

$$\begin{aligned}
\sigma(f)^2 &= \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_A^i\right)^2 \mathrm{d}\mu \\
&= \frac{1}{M}\left[\int_A \left(\sum_{k=0}^{M-1} X_f \circ T_A^k\right)^2 \mathrm{d}\mu + 2\sum_{i=1}^{+\infty} \int_A \left(\sum_{k=0}^{M-1} X_f \circ T_A^k\right)\left(\sum_{k=Mi}^{Mi+M-1} X_f \circ T_A^k\right) \mathrm{d}\mu\right].
\end{aligned}$$

The convergence in distribution of the local times stays unchanged, as [16, Proposition 4.1] only requires ergodicity, and not mixing.

The only point which is not essentially trivial is the adaptation of Lemma 2.10 (we refer to [16, Lemma 4.4] for its proof). This lemma means roughly that the process $(\xi_N)_{n \geq 0}$ does not increase too fast; to prove it, we show instead that its inverse $(\tau_N) = \sum_{i=0}^{N-1} \varphi_A \circ T^i$ increases quickly enough.

More specifically, at some point we prove that the sequence $(\mathbb{P}(\xi_{\psi^*(n^r)} > n^{r^*}))_{n\geq 0}$ is summable if $r < r^*$. This requires some amount of independence for the sequence $(\varphi_A \circ T^i)$. In the mixing case, we work with the subsequence $(\varphi_A \circ T^{n^{qr^*}i})$ for some well-chosen q. In the ergodic, non-mixing case, our best tool is to induce the transformation on some A_k. This leads us to work with the subsequences $((\sum_{\ell=0}^{M-1} \varphi_A \circ T^\ell) \circ T^{Mn^{qr^*}i})$. Most of the estimates in the proof will change by a constant factor depending on M, but a small change of q will compensate for them. The summability of the sequence $(\mathbb{P}(\xi_{\psi^*(n^r)} > n^{r^*}))_{n\geq 0}$ follows, and with it versions of Lemma 2.10 and Theorem 2.4 which do not require mixing. □

6 Semi-flows and flows

Our ultimate goal in this article is to get a version of Theorem 2.4 for the geodesic flow on $\mathbb{Z}$ or $\mathbb{Z}^2$-periodic manifolds of negative curvature. There are three directions in which we would like to adapt our results:

- instead of working with discrete time dynamical systems (towers over a Gibbs-Markov system), we want results in a continuous time setting (suspension flows over a Gibbs-Markov system);
- we want results for $\mathbb{Z}$ or $\mathbb{Z}^2$-extensions of a Gibbs-Markov map (sometimes called "random walks driven by a Gibbs-Markov map"), where the data is not directly given in terms of a first return time but in terms of a transition kernel;
- we want results for invertible systems, and more precisely for natural extensions of Gibbs-Markov maps.

The geodesic flow on periodic manifolds mixes those three features: it can be seen for instance as a suspension flow over a $\mathbb{Z}^d$-extension of the natural extension of a Gibbs-Markov map (and the terms "suspension flow", "$\mathbb{Z}^d$-extension" and "natural extension" in this sentence commute).

We will not state a version of Theorem 2.4 for each combination of "suspension flow", "$\mathbb{Z}^d$-extension" and "natural extension". Instead, we will go to the most complex case by the following path: first we extend Theorem 2.4 to suspension flows (Subsection 6.1), then to $\mathbb{Z}^d$-extensions of suspension flows (Subsection 6.2), and finally to the geodesic flow (Subsection 6.3). Our results can be downgraded to accommodate, for instance, $\mathbb{Z}^d$-extensions of Gibbs-Markov maps.

6.1 Suspension flows

In discrete time we studied non-invertible transformations; continuous time non-invertible transformations are semi-flows, that is, actions of the additive semi-group $\mathbb{R}_+$. In discrete time, given an ergodic transformation (Ω, μ, T) and a measurable subset of non-zero measure A, we could easily define an induced dynamical system $(A, \mu_{|A}, T_A)$ and a first return time φ_A. We could also see (Ω, μ, T) as a tower of base $(A, \mu_{|A}, T_A)$ and height function φ_A. For semi-flows, it will be more convenient to adopt this second point of view.

Definition 6.1 (Gibbs-Markov semi-flow).

Let $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ be an ergodic Gibbs-Markov map. Let φ_A be a measurable function defined on A and such that $\varphi_A > 0$ almost surely. The Gibbs-Markov semi-flow of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ and height φ_A is the dynamical system $(\Omega, \mu, (g_t)_{t\geq 0})$, where:

- *Ω is the quotient of the topological space $A \times \mathbb{R}_+$ by the equivalence relation $(x, \varphi_A(x) + t) \sim (T_A(x), t)$ for all (x, t) in $A \times \mathbb{R}_+$;*

- *the semi-flow $(g_t)_{t\geq 0}$ acts on $A \times \mathbb{R}_+$ by translation on the second coordinate, and on Ω by its canonical projection;*

- $\mu = \mu_A \otimes \mathrm{Leb}$ *on $A \times \mathbb{R}_+$ and is defined canonically on Ω by restriction to a fundamental domain.*

Gibbs-Markov semi-flows arise naturally when one studies the geodesic flow on $\mathbb{Z}$ or $\mathbb{Z}^2$ covers of compact manifolds of negative curvature, or $\mathbb{Z}$ or $\mathbb{Z}^2$ periodic hyperbolic billiards. Since (A, μ_A, T_A) is ergodic, so is $(\Omega, \mu, (g_t)_{t\geq 0})$. Up to a subset of measure 0, we can identify A with the subset $A \times \{0\}$ of Ω, and extend μ_A to Ω by putting $\mu_A(B) = \mu_A(x \in A : \ x \times \{0\} \in B)$ for any measurable subset $B \subset \Omega$.

For almost any point x in A and any integer $N \geq 0$, the N-th return time is $\tau_N(x) := \sum_{k=0}^{N-1} \varphi_A \circ T_A^k(x)$. For almost any point x in A and any non-negative t, the local time in A at time t is $\xi_t(x) := \mathrm{Card}\{s \in (0,t] : \ g_s(x) \in A\}$, so that $\xi_{\tau_N} = N$ for all N and $t \leq \tau_{\xi_t}$ for all t. All these functions can be extended to Ω.

For any real-valued function f on Ω such that $\int_0^{\varphi_A(x)} |f(x,t)| \, \mathrm{d}t$ is finite for almost every x in A, we put $X_f(x) := \int_0^{\varphi_A(x)} f(x,t) \, \mathrm{d}t$. If f belongs to $\mathbb{L}^1(\Omega, \mu)$, then X_f is well-defined on A and belongs to $\mathbb{L}^1(A, \mu_A)$.

Our goal in this section is to extend Theorems 2.3 and 2.4 to ergodic Gibbs-Markov semi-flows. Fortunately, Proposition 5.1 already extends Theorem 2.3 in a satisfactory manner. Our work will consist in adapting Theorem 2.4, of which we give the following variant:

Proposition 6.2.

Let $(\Omega, \mu, (g_t)_{t\geq 0})$ be an ergodic Gibbs-Markov semi-flow of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ and of height function φ_A. Let f be a real-valued function from Ω to $\mathbb{R}$. Assume that:

- *f is measurable and $X_{|f|}$ belongs to $\mathbb{L}^p(\Omega, \mu_A)$ for some $p > 2$;*

- $\int_\Omega f \, \mathrm{d}\mu = 0$;

- *the function $\psi(t) = \mu_A(\varphi_A \geq t)^{-1}$ is regularly varying with index $\beta \in [0, 1)$;*

- *$\mathbb{E}(D(X_f))$ and $\mathbb{E}(D(\varphi_A))$ are finite.*

Then, for any probability measure ν absolutely continuous with respect to either μ or μ_A:

$$\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma(f)\sqrt{Y_\beta}\mathcal{N}, \tag{6.1}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω, ν) to $\mathbb{R}$, where Y_β and $\mathcal{N}$ are independent, Y_β is a standard Mittag-Leffler distribution of order β and $\mathcal{N}$ is a standard Gaussian random variable, and:

$$\sigma(f)^2 = \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_A^i \right)^2 \mathrm{d}\mu_A. \tag{6.2}$$

Moreover, $\sigma(f) = 0$ if and only if f is a coboundary.

In order prove the convergence in distribution of the local time, it will be convenient to relate the properties of the local time of a Gibbs-Markov semi-flow with those of a suitably discretized version. This is done with the next two lemmas.

Lemma 6.3.

Let $a : \mathbb{R}_+ \to \mathbb{R}_+$ be a function with regular variation. Let $(\xi_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$ be two stochastic processes taking their values in $\mathbb{R}_+$. Assume that there exists a positive constant C such that $\lim_{t\to+\infty} X_t = C$ in distribution, that there exists a random variable Y such that $\lim_{t\to+\infty} a(t)^{-1}\xi_t = Y$ in distribution, and that $(\xi_t)_{t\geq 0}$ is non-decreasing. Then, in distribution,

$$\lim_{t\to+\infty} a(Ct)^{-1}\xi_{tX_t} = Y.$$

Proof.

Without loss of generality, we can assume that $C = 1$. Let $\varepsilon \in (0,1)$. For all large enough t, we have $\mathbb{P}(X_t \notin (1-\varepsilon, 1+\varepsilon)) \leq \varepsilon$. Since (ξ_t) is non-decreasing, this inequality yields:

$$\mathbb{P}\left(\xi_{tX_t} \notin \left(\xi_{(1-\varepsilon)t}, \xi_{(1+\varepsilon)t}\right)\right) \leq \varepsilon$$

Since this inequality holds for all $\varepsilon > 0$ and the sequence $(a(t)^{-1}\xi_{(1+\varepsilon)t})$ is tight, the sequence $(a(t)^{-1}\xi_{tX_t})$ is also tight, and thus any subsequence of $(a(t)^{-1}\xi_{tX_t})$ has limit points in $\mathcal{P}(\mathbb{R}_+)$.

Let β be the index of the regular variation of a. For all $\varepsilon \in (0,1)$ and $\rho \geq 0$, for all large enough t, we have:

$$\begin{aligned}
\mathbb{E}\left(e^{-\rho(1+\varepsilon)^\beta Y}\right) - \varepsilon &= \lim_{t\to+\infty} \mathbb{E}\left(e^{-\rho\frac{\xi_{(1+\varepsilon)t}}{a(t)}}\right) - \varepsilon \\
&\leq \liminf_{t\to+\infty} \mathbb{E}\left(e^{-\rho\frac{\xi_{tX_t}}{a(t)}}\right) \\
&\leq \limsup_{t\to+\infty} \mathbb{E}\left(e^{-\rho\frac{\xi_{tX_t}}{a(t)}}\right) \\
&\leq \lim_{t\to+\infty} \mathbb{E}\left(e^{-\rho\frac{\xi_{(1-\varepsilon)t}}{a(t)}}\right) + \varepsilon = \mathbb{E}\left(e^{-\rho(1-\varepsilon)^\beta Y}\right) + \varepsilon.
\end{aligned}$$

Since this sequence of inequalities holds for all $\varepsilon > 0$, for all $\rho \geq 0$,

$$\lim_{t\to+\infty} \mathbb{E}\left(e^{-\rho\frac{\xi_{tX_t}}{a(t)}}\right) = \mathbb{E}\left(e^{-\rho Y}\right).$$

Let us take any subsequence $(a(t_n)^{-1}\xi_{t_n X_{t_n}})_{n\geq 0}$ of $(a(t)^{-1}\xi_{tX_t})_{t\geq 0}$ with $t_n \to +\infty$, and any limit point Y^* of this subsequence for the weak topology in $\mathcal{P}(\mathbb{R}_+)$. Then, by the dominated convergence theorem, for all $\rho \geq 0$ we have $\mathbb{E}\left(e^{-\rho Y}\right) = \mathbb{E}\left(e^{-\rho Y^*}\right)$. The Laplace transform from $\mathbb{P}(\mathbb{R}_+)$ to $\mathcal{C}(\mathbb{R}, \mathbb{R})$ is injective, so $Y = Y^*$ in distribution. Since this holds for any subsequence, we have proved at last the convergence in distribution we claimed: $a(t)^{-1}\xi_{tX_t} \to Y$. □

Lemma 6.4.

Let $(\Omega, \mu, (g_t)_{t\geq 0})$ be an ergodic Gibbs-Markov semi-flow of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ and of height function φ_A, where $\mathbb{E}(D(\varphi_A))$ is finite. Let $\psi(t) = \mu_A(\varphi_A \geq t)^{-1}$.

If ψ is regularly varying with index $\beta \in [0,1)$, then:

$$\lim_{t\to+\infty} \frac{1}{\operatorname{sinc}(\beta\pi)\psi(t)}\xi_t = Y_\beta, \tag{6.3}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (A, μ_A) to $\mathbb{R}$ and where Y_β is a standard Mittag-Leffler random variable of order β.

Proof.

If φ_A only takes integer values and is constant on each element of π, this is [16, Proposition 4.2]. In the general case, let us define a function $\overline{\varphi}_A$ which, for all a in π, is constant on a and takes there the value $\lceil \sup_a \varphi_A \rceil > 0$. Let $\Delta := \overline{\varphi}_A - \varphi_A \geq 0$. We define $(\overline{\xi}_t)$, $(\overline{\tau}_N)$ and $\overline{\psi}$ for the Gibbs-Markov semi-flow of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ and of height function $\overline{\varphi}_A$ in the obvious way.

We have $\Delta \leq 1 + \mathrm{Diam}(A) D(\varphi_A)$. Since $\overline{\varphi}_A \geq \varphi_A$, we have $\overline{\psi} \leq \psi$. On the other hand, for all $t \geq 0$ and $\varepsilon \in (0,1)$,

$$\begin{aligned}\mu_A(\overline{\varphi}_A \geq t) &\leq \mu_A(\varphi_A \geq (1-\varepsilon)t) + \mu_A(\Delta \geq \varepsilon t) \\ &\leq \mu_A(\varphi_A \geq (1-\varepsilon)t) + \frac{1 + \mathrm{Diam}(A)\mathbb{E}(D(\varphi_A))}{\varepsilon t}.\end{aligned}$$

Hence, for all $\varepsilon > 0$, we have $\psi \geq \overline{\psi} \geq (1-\varepsilon)^\beta (1+o(1))\psi$, which finally yields $\overline{\psi} \sim \psi$.

Let $t \geq 0$, and put $N := \xi_t$. Then $t \in [\tau_N, \tau_{N+1})$. Moreover, $\overline{\tau}_n = \tau_n + \sum_{k=0}^{n-1} \Delta \circ T_A^k$ for all $n \geq 1$, so that $t + \sum_{k=0}^{N-1} \Delta \circ T_A^k \in [\overline{\tau}_N, \overline{\tau}_{N+1})$. Hence, for all $t \geq 0$:

$$\xi_t = \overline{\xi}_{t + \sum_{k=0}^{\xi_t - 1} \Delta \circ T_A^k} = \overline{\xi}_{t\left(1 + \frac{1}{t}\sum_{k=0}^{\xi_t - 1} \Delta \circ T_A^k\right)}.$$

By Birkhoff's ergodic theorem, almost surely,

$$\lim_{t \to +\infty} \frac{1}{\xi_t} \sum_{k=0}^{\xi_t - 1} \Delta \circ T_A^k = \mathbb{E}(\Delta).$$

By Birkhoff's ergodic theorem again, since μ is infinite, $N^{-1}\tau_N$ converges to $+\infty$ almost surely. Taking the generalized inverse to the sequence (τ_N), one gets that $t^{-1}\xi_t$ converges to 0 almost surely. Hence, $1 + t^{-1} \sum_{k=0}^{\xi_t - 1} \Delta \circ T_A^k$ converges in distribution to 1. Lemma 6.4 follows from [16, Proposition 4.2] applied to $(\overline{\xi}_t)$, and from Lemma 6.3. □

With Lemma 6.4, we are ready to prove Proposition 6.2.

Proof of Proposition 6.2.

The first and second steps of the proof of Theorem 1.11 in [16] (adapted to take into account the fact that we work with an ergodic system as in Subsection 5, and not necessarily a mixing one) can be adapted straightforwardly. Any result they refer to (e.g. Lemma 2.11, Lemma 2.10...) has a direct translation in the continuous-time setting, where the first return time φ_A takes values in $\mathbb{R}_+$, where the local time (ξ_t) is defined on $\mathbb{R}_+$, etc.

Thus, under the assumptions of the proposition:

$$\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma(f)\sqrt{Y_\beta}\mathcal{N},$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω, μ_A) to $\mathbb{R}$, where Y_β and $\mathcal{N}$ are independent, Y_β is a standard Mittag-Leffler distribution of order β and $\mathcal{N}$ is a standard Gaussian random variable, and $\sigma(f)$ is given by the formula (6.2).

We shall now extend this result to any starting distribution absolutely continuous with respect to either μ_A or μ. However, we cannot use [17, Corollary 1], which only deals with discrete time systems. Moreover, this result does not translate readily to the continuous time setting, due to a couple of obstructions. First, we know that the result holds for the starting distribution μ_A, but μ_A is not absolutely continuous with respect to μ. Moreover, since the first return time φ_A is not assumed to be bounded from below by a positive constant, it is not obvious that one can extend the convergence in distribution with starting distribution μ_A to a starting distribution absolutely continuous with respect to μ by looking at boxes $A \times [0, \varepsilon]$. The bad case which could happen *a priori* is that X_f is large where φ_A is close to 0, and that the underlying Gibbs-Markov dynamic can spend a lot of consecutive iterations where φ_A is small, making for instance $\int_A \int_0^\varepsilon |f|(x,t) \,\mathrm{d}t \,\mathrm{d}\mu_A$ infinite. This phenomenon would make ineffective the control on X_f on what happens during a small window of time. We will exclude this possibility by using Lemma 2.10.

We begin by proving the result for starting distributions absolutely continuous with respect to μ_A, then deduce the case of starting distributions absolutely continuous with respect to μ.

Probability measures absolutely continuous with respect to μ_A

Let f be a function satisfying the assumptions of the proposition we want to prove.

For now, assume that the base (A, μ_A, T_A) is not only ergodic, but mixing; it is then also exact, so that, for any function h in $\mathbb{L}^1(A, \mu_A)$, the sequence $(\mathcal{L}^n h)_{n \geq 0}$ converges to the constant function $\int_A h \,\mathrm{d}\mu_A$ in $\mathbb{L}^1$ norm [1, Theorem 1.3.3].

Let $\nu = h \,\mathrm{d}\mu_A$ be a probability measure absolutely continuous with respect to μ_A, and let $\varepsilon > 0$. Let n be such that $\|\mathcal{L}^n h - 1\|_{\mathbb{L}^1} \leq \varepsilon$.

We already know that, for any real-valued, uniformly continuous and bounded function G on $\mathbb{R}$,

$$\lim_{t \to +\infty} \mathbb{E}_{\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \,\mathrm{d}s\right)\right) = \mathbb{E}\left(G\left(\sigma(f)\sqrt{Y_\beta}\mathcal{N}\right)\right).$$

Since $\mathcal{L}^n h \,\mathrm{d}\mu_A$ and μ_A differ by at most ε in total variation norm,

$$\limsup_{t \to +\infty} \left|\mathbb{E}_{\mathcal{L}^n h d\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \,\mathrm{d}s\right)\right) - \mathbb{E}\left(G\left(\sigma(f)\sqrt{Y_\beta}\mathcal{N}\right)\right)\right| \leq \varepsilon \|G\|_\infty .$$

Moreover, $\mathcal{L}^n h \,\mathrm{d}\mu_A$ is the distribution of $T_A^n x$ given that x has distribution $h \,\mathrm{d}\mu_A$, so that, for all t:

$$\mathbb{E}_{\mathcal{L}^n h d\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \,\mathrm{d}s\right)\right) = \mathbb{E}_{h d\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}} \int_{\tau_n}^{t+\tau_n} f \circ g_s \,\mathrm{d}s\right)\right).$$

Obviously, $\psi(t)^{-1/2} \int_0^{\tau_n} f \circ g_s \,\mathrm{d}s$ converges to 0 almost surely. By Lemma 2.10, for all $r > 0$, almost surely, for all large enough t, we have $\xi_{t+\tau_n} - \xi_t \leq \psi(t)^r$. Looking at the second step of the proof of Theorem 1.11 in [16] and adapting it to the continuous time setting, we also learn that for all $\delta > 0$, almost surely, for all large enough t, we have $X_{|f|} \circ T_A^{\xi_t} \leq \psi(t)^{\frac{1}{p}+\delta}$. Hence, for all $r, \delta > 0$, almost surely, for all large enough t,

$$\left|\int_t^{t+\tau_n} f \circ g_s \,\mathrm{d}s\right| \leq \sum_{k=\xi_t}^{\xi_{t+\tau_n}} X_{|f|} \circ T_A^k \leq (\xi_{t+\tau_n} - \xi_t + 1)\psi(t+\tau_n)^{\frac{1}{p}+\delta} \leq (\psi(t)^r + 1)\psi(t+\tau_n)^{\frac{1}{p}+\delta}.$$

By taking r and δ small enough, we see that $\psi(t)^{-1/2} \int_t^{t+\tau_n} f \circ g_s \,\mathrm{d}s$ converges to 0 almost surely. Hence,

$$\limsup_{t \to +\infty} \left|\mathbb{E}_{h d\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \,\mathrm{d}s\right)\right) - \mathbb{E}\left(G\left(\sigma(f)\sqrt{Y_\beta}\mathcal{N}\right)\right)\right| \leq \varepsilon \|G\|_\infty .$$

Since this is true for all $\varepsilon > 0$, the convergence in distribution we claimed occurs when the starting point is chosen according to the measure ν, where ν is any distribution absolutely continuous with respect to μ_A, as long as the base of the Gibbs-Markov semi-flow is mixing.

For the ergodic but not mixing case, let M be the period of the system. Then for any probability density h in $\mathbb{L}^1(A, \mu_A)$, the densities $M^{-1} \sum_{k=0}^{M-1} \mathcal{L}^{n+k} h$ converge to 1 in $\mathbb{L}^1(A, \mu_A)$, and the same strategy works.

Probability measures absolutely continuous with respect to μ

Let ν be a probability measure absolutely continuous with respect to μ, and let ν_A be its projection onto A. By construction, ν_A is absolutely continuous with respect to μ_A, and we have a natural coupling between ν and ν_A. Moreover, if we denote by $(f_t)_{t \geq 0}$ the process $(f \circ g_t)_{t \geq 0}$ where the starting point is chosen according to ν, and $(\tilde{f}_t)_{t \geq 0}$ the same process but where the starting point is chosen according to ν_A, then, under this coupling, they differ only by a translation in time:

$$\left| \int_0^t f_s \, \mathrm{d}s - \int_0^t \tilde{f}_s \, \mathrm{d}s \right| \leq \int_0^{\varphi_A} \left| \tilde{f}_s \right| \, \mathrm{d}s + \int_t^{t+\varphi_A} \left| \tilde{f}_s \right| \, \mathrm{d}s.$$

Using the same method as above, we can show that $\psi(t)^{-1/2} \left| \int_0^t f_s \, \mathrm{d}s - \int_0^t \tilde{f}_s \, \mathrm{d}s \right|$ converges almost surely to 0. We only need to consider the asymptotic behavior of $(\tilde{f}_t)_{t \geq 0}$, which is already covered by the previous case since $\nu_A \ll \mu_A$.

Finally, if X_f is a coboundary, let $\tilde{u}$ be a real-valued function on A such that $X_f = \tilde{u} \circ T_A - \tilde{u}$. For all $t \in [0, \varphi_A)$, let us define:

$$u(x,t) := \tilde{u}(T_A x) - \int_t^{\varphi_A(x)} f \circ g_s(x, 0) \, \mathrm{d}s.$$

Then $\int_0^t f \circ g_s \, \mathrm{d}s = u \circ g_s - u$ for all $t \geq 0$, so f is a coboundary. Conversely, if f is a coboundary, there exists a measurable function u such that, for all t:

$$\int_0^t f \circ g_s \, \mathrm{d}s = u \circ g_t - u,$$

and thus $X_f = u \circ T_A - u$ is also a coboundary. □

6.2 $\mathbb{Z}^d$-extensions of semi-flows

An important class of dynamical systems endowed with an infinite measure comes from $\mathbb{Z}$ or $\mathbb{Z}^2$-extensions of Gibbs-Markov maps (or suspension flows over a Gibbs-Markov map). We will adapt our result to this setting, and will apply it further to the study of the geodesic flow on some $\mathbb{Z}$ or $\mathbb{Z}^2$ periodic manifolds with negative sectional curvature. One of the important features of those systems is that the data is usually given in terms of a step function and a step time, and not directly in terms of a first return time.

Definition 6.5 ($\mathbb{Z}^d$ extension of a Gibbs-Markov semi-flow).

Let $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$ be an ergodic Gibbs-Markov map. Let r be a real-valued, measurable function on A which is almost surely positive. Let d be a non-negative integer, and let F be a $\mathbb{Z}^d$-valued measurable function on A which is $\sigma(\pi)$-measurable (i.e. almost surely constant on each element of π).

The $\mathbb{Z}^d$ extension with base $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$, with step F and step time r is the dynamical system $(\Omega, \mu, (g_t)_{t \geq 0})$, where:

- *Ω is the quotient of the topological space $A \times \mathbb{Z}^d \times \mathbb{R}_+$ by the equivalence relation $(x, q, r(x)+t) \sim (T(x), q + F(x), t)$ for all (x, q, t) in $A \times \mathbb{Z}^d \times \mathbb{R}_+$;*

- *the semi-flow $(g_t)_{t\geq 0}$ acts on $A \times \mathbb{Z}^d \times \mathbb{R}_+$ by translation on the third coordinate, and on Ω by its canonical projection;*

- $\mu = \mu_A \otimes \mathrm{Leb} \otimes \mathrm{Leb}$ *on $A \times \mathbb{Z}^d \times \mathbb{R}_+$ and is defined canonically on Ω by restriction to a fundamental domain.*

Such an extension is said to be degenerate if F can be written as a sum of a coboundary and a function with values in a translate of a proper sublattice of $\mathbb{Z}^d$, and non-degenerate otherwise.

We are interested in recurrent extensions, which is quite restrictive: it excludes non-degenerate $\mathbb{Z}^d$-extensions for all $d \geq 3$. In order to apply Proposition 6.2, we describe such a process as a Gibbs-Markov semi-flow. The method is the same as with random walks: we look at the shift on the set of excursions starting from zero. Let us consider a recurrent extension with base $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$, step F and step time r. Without loss of generality, we can assume that $d_A(x, y) = \lambda^{-s(x,y)}$ with $\lambda > 1$ and where $s(x, y)$ is the separation time of x and y for the partition π and the transformation T.

We can identify A with $A_0 := A \times \{0\} \times \{0\}$, and define:

- the first return time φ in A_0 by $\varphi(x) := \inf\{t > 0 : \ g_t(x, 0, 0) \in A_0\}$ (this function can be extended to Ω);

- the discretized first return time $\overline{\varphi}$ in A_0 as the return time for the extension with the same base, the same step function but with a step time constant and equal to 1;

- the position at time t as the unique element $S_t \in \mathbb{Z}^d$ such that, for some $y \in A$ and $s < r(y)$, we have $g_t(x, 0, 0) \sim (y, S_t, s)$.

Since the semi-flow is assumed to be recurrent, φ and $\overline{\varphi}$ are finite almost surely. Since $\overline{\varphi}$ is the first return time to a subset of the initial dynamical system, by Kac's theorem, the application $T^{\overline{\varphi}}$ preserves the measure μ_A. We use the setting of Subsection 1.2 to describe an accelerated Gibbs-Markov map. For all $n \geq 0$, the event $\{\overline{\varphi} \leq n\}$ only depends on the value of $(F, F \circ T, \cdots, F \circ T^{n-1})$; as F is $\mathcal{F}_1$-measurable, this event is $\mathcal{F}_n$-measurable. Hence, $\overline{\varphi}$ is a stopping time which is almost surely positive and finite. We use Lemma 1.6 to define a new Gibbs-Markov map $(A_0, \pi_{\overline{\varphi}}, d_{\overline{\varphi}}, \mathcal{B}, \mu_A, T_{\overline{\varphi}})$. The $\mathbb{Z}^d$-extension is then measurably isomorphic to the Gibbs-Markov semi-flow of base $(A_0, \pi_{\overline{\varphi}}, d_{\overline{\varphi}}, \mathcal{B}, \mu_A, T_{\overline{\varphi}})$ and height φ. We will write $D = D_{\pi, d_A}$ and $D_{\overline{\varphi}} = D_{\pi_{\overline{\varphi}}, d_{\overline{\varphi}}}$.

We can use Section 6.1 to define the sequence of return times (τ_N) in A_0, the local time (ξ_t) in A_0, and so on. By putting $r \equiv 1$, we also define discretized versions of these quantities, which we distinguish with a bar: $(\overline{S}_t)$, $(\overline{\tau}_N)$, $(\overline{\xi}_t)$, $\overline{\psi}(t)$... They are useful to relate the properties of the initial $\mathbb{Z}^d$-extension with those of its discretized counterpart, which have been extensively studied.

First, we will relate some properties of the induced Gibbs-Markov semi-flow to those of the initial data (for instance, the functions r and F) so as the express the hypotheses of Proposition 6.2 in a way which is better suited to the setting of $\mathbb{Z}^d$-extensions. Then, we will relate the distributional properties of (ξ_t) to those of $(\overline{\xi}_t)$ to leverage the existing literature on the discretized version.

Lemma 6.6 (Smoothness of the first return time)**.**

For a recurrent $\mathbb{Z}^d$-extension with base $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$ and step time r, with $d_A(x, y) = \lambda^{-s(x,y)}$ and the choice of distance $d_{\overline{\varphi}}$ mentioned before,

$$\mathbb{E}\left(D_{\overline{\varphi}}(\varphi)\right) \leq \frac{\lambda}{\lambda - 1} \mathbb{E}\left(D(r)\right).$$

Proof.

Let a be in $\pi_{\overline{\varphi}}$, and let x, y be in a. For all $k < \overline{\varphi}(x)$, the points $T^k(x)$ and $T^k(y)$ are in the same element of π, so that $d_A(T_{\overline{\varphi}}(x), T_{\overline{\varphi}}(y)) = \lambda^{\overline{\varphi}(x)-k} d_A(T^k(x), T^k(y))$. For $0 \leq k < \overline{\varphi}(x)$, let us denote by $a_k(a)$ the element of π in which $T^k(x)$ and $T^k(y)$ lie. Then,

$$\begin{aligned}
|\varphi(x) - \varphi(y)| &= \left| \sum_{k=0}^{\overline{\varphi}(x)-1} (r(T^k(x)) - r(T^k(y))) \right| \\
&\leq \sum_{k=0}^{\overline{\varphi}(x)-1} |r|_{\mathrm{Lip}_{d_A}(a_k(a))} d_A(T^k(x), T^k(y)) \\
&\leq d_A(T_{\overline{\varphi}}(x), T_{\overline{\varphi}}(y)) \sum_{k=0}^{\overline{\varphi}(x)-1} \lambda^{k-\overline{\varphi}(x)} |r|_{\mathrm{Lip}_{d_A}(a_k(a))} \\
&\leq d_{\overline{\varphi}}(x, y) \sum_{k=0}^{\overline{\varphi}(x)-1} \lambda^{k-\overline{\varphi}(a)+1} |r|_{\mathrm{Lip}_{d_A}(a_k(a))}.
\end{aligned}$$

In short, we have the inequality:

$$D_{\overline{\varphi}}(\varphi) \leq \sum_{k=0}^{\overline{\varphi}-1} \lambda^{-k} D(r) \circ T^{\overline{\varphi}-1-k}. \tag{6.4}$$

At this point, we use the structure of the extension to leverage this inequality. Let M be a positive integer. Let $(q_0, \cdots, q_M)$ be in $(\mathbb{Z}^d)^M$. Assume that there exists some $0 \leq k < M$ such that $q_k = q_M$. Then the initial path $(q_0, \cdots, q_M)$ can be split in a unique way into two sub-paths $(q_0, \cdots, q_{M-N-1})$ and $(q_{M-N}, \cdots, q_M)$ such that $q_{M-N} = q_M$ and $q_k \neq q_M$ for all $M - N < k < M$. Conversely, given two such sub-paths, by concatenation, we can create a new path $(q_0, \cdots, q_M)$ with $q_{M-N} = q_M$. If we apply this observation to the path $(\overline{S}_0(x), \cdots, \overline{S}_M(x))$ for any $x \in A$, we see that the two following sets are identical:

- $\{x \in A : \ \exists\, 0 \leq k < M, \ \overline{S}_k(x) = \overline{S}_M(x)\}$;
- $\displaystyle\bigsqcup_{N=1}^{M} \{x \in A : \ T^{M-N}(x) \in \{\overline{\varphi} = N\}\}$.

The most interesting consequence for us will be the following inequality, which holds for all $M \geq 1$:

$$\sum_{N=1}^{M} 1_{\{\overline{\varphi}=N\}} \circ T^{M-N} \leq 1. \tag{6.5}$$

If we take the expectation on both sides of Equation (6.4), we get:

$$\begin{aligned}
\mathbb{E}\left(D_{\overline{\varphi}}(\varphi)\right) &\leq \mathbb{E}\left(\sum_{k=0}^{\overline{\varphi}-1} \lambda^{-k} D(r) \circ T^{\overline{\varphi}-1-k}\right) \\
&= \sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{E}\left(1_{\{\overline{\varphi}>k\}} \cdot D(r) \circ T^{\overline{\varphi}-1-k}\right).
\end{aligned}$$

Let us fix the integer $k \geq 0$. Then, because of Equation (6.5), for all $M > k$,

$$\begin{aligned}
\mathbb{E}\left(1_{\{\overline{\varphi} \leq M\}} 1_{\{\overline{\varphi} > k\}} \cdot D(r) \circ T^{\overline{\varphi}-1-k}\right) &= \sum_{N=k+1}^{M} \mathbb{E}\left(1_{\{\overline{\varphi} = N\}} \cdot D(r) \circ T^{N-1-k}\right) \\
&= \sum_{N=k+1}^{M} \mathbb{E}\left(1_{\{\overline{\varphi} = N\}} \circ T^{M-N} \cdot D(r) \circ T^{M-1-k}\right) \\
&\leq \mathbb{E}\left(D(r) \circ T^{M-1-k}\right) \\
&= \mathbb{E}\left(D(r)\right).
\end{aligned}$$

Since the system is recurrent, $\overline{\varphi}$ is almost surely finite, so that $\lim_{M \to +\infty} 1_{\{\overline{\varphi} \leq M\}} = 1$ almost surely. By the monotone convergence theorem,

$$\mathbb{E}\left(D_{\overline{\varphi}}(\varphi)\right) \leq \sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{E}\left(1_{\{\overline{\varphi} > k\}} \cdot D(r) \circ T^{\overline{\varphi}-1-k}\right) \leq \sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{E}\left(D(r)\right) = \frac{\lambda}{\lambda - 1} \mathbb{E}\left(D(r)\right). \qquad \square$$

Let f be in $\mathbb{L}^1(\Omega, \mu)$. We define on A (or, similarly, A_0):

- for all q in $\mathbb{Z}^d$, the function $X_{f,q}(x) := \int_0^{r(x)} f \circ g_s(x, q, 0) \, \mathrm{d}s$;
- $X_f(x) := \int_0^{\varphi(x)} f \circ g_s(x, 0, 0) \, \mathrm{d}s$.

Then $X_f(x) = \sum_{k=0}^{\overline{\varphi}(x)-1} X_{f, \overline{S}_k(x)}(T^k x)$. Using the same argument as in the proof of Lemma 6.6, we get:

Lemma 6.7 (Smoothness of the observables)**.**
For a recurrent $\mathbb{Z}^d$-extension with base $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$, where $d_A(x, y) = \lambda^{-s(x,y)}$, for all f in $\mathbb{L}^1(\Omega, \mu)$,

$$\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right) \leq \frac{\lambda}{\lambda - 1} \mathbb{E}\left(\sup_{q \in \mathbb{Z}^d} D\left(X_{f,q}\right)\right). \tag{6.6}$$

Proof.
The difference between this lemma and Lemma 6.6 is that the function $D\left(X_{f,q}\right)$ depends on q, so that its weighted sum along an excursion will depend on that excursion. Bounding each of these functions by $\sup_{q \in \mathbb{Z}^d} D\left(X_{f,q}\right)$, which does not depend on q anymore, solves this problem. $\square$

In particular, if $\mathbb{E}\left(D\left(X_{f,q}\right)\right)$ is finite for all q and $X_{f,q} \equiv 0$ for all but a finite number of positions q, then $\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right)$ is also finite.

Remark 6.8 (Alternative upper bounds)**.**
By manipulating more precisely the proof of Lemma 6.6, one can find other upper bounds for $\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right)$. For example, for all θ in $[0, 1]$,

$$\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right) \leq \sum_{q \in \mathbb{Z}^d} \left(\sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{P}(\overline{\varphi} > k \text{ and } \overline{S}_{\overline{\varphi}-1-k} = q)^{1-\theta}\right) \left\|D\left(X_{f,q}\right)\right\|_{\mathbb{L}^{1/\theta}}.$$

For $\theta = 1$, this bound is worse than the one we already proved. However, for $\theta = 0$, it becomes:

$$\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right) \leq \sum_{q \in \mathbb{Z}^d} \left(\sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{P}(\overline{\varphi} > k \text{ and } \overline{S}_{\overline{\varphi}-1-k} = q)\right) \left\|D\left(X_{f,q}\right)\right\|_{\infty}. \tag{6.7}$$

This bound is especially useful if π is a finite partition, for then the semi-norms $\|D(\cdot)\|_\infty$ and $\mathbb{E}(D(\cdot))$ are equivalent, and:

$$\mathbb{E}\left(\sup_{q\in\mathbb{Z}^d} D\left(X_{f,q}\right)\right) \simeq \sup_{q\in\mathbb{Z}^d} \left\|D\left(X_{f,q}\right)\right\|_\infty .$$

Hence, the bound (6.7) is improved from the bound (6.6) by the addition of the summable weights $\sum_{k=0}^{+\infty} \lambda^{-k}\mathbb{P}(\overline{\varphi} > k$ and $\overline{S}_{\overline{\varphi}-1-k} = q)$. This function is similar to the Green function for some kind of "reversed random walk" (this analogy can be made rigorous if the base of the extension is an ergodic Markov chain), and decreases exponentially fast in $|q|$ if the steps are bounded (which is the case if the partition is finite).

There is unfortunately no such simple condition on the integrability of the array of functions $(X_{f,q})_{q\in\mathbb{Z}^d}$ which guarantees that X_f belongs to $\mathbb{L}^p$ for some $p > 2$, or at least no condition with such a short proof. This problem is difficult even in the case of true random walks. We will only give the following criterion, which holds for all $p \geq 1$:

$$\left\|X_{|f|}\right\|_{\mathbb{L}^p}^p \leq \mathbb{E}\left(\left(\sum_{k=0}^{\overline{\varphi}-1} \left\|X_{|f|,\overline{S}_k}\right\|_{\mathbb{L}^\infty}\right)^p\right).$$

We shall also give a sufficient condition, which is alas very strong.

Lemma 6.9.

Let (Ω, μ, T) be a recurrent $\mathbb{Z}^d$-extension with an ergodic Gibbs-Markov base (A, μ_A, T_A). Let $f \in \mathbb{L}^1(\Omega, \mu)$. Let $1 \leq p^* < p \leq +\infty$.

If $X_{|f|,q} \equiv 0$ for all but finitely many q, and if $X_{|f|,q} \in \mathbb{L}^p(A, \mu_A)$ for all q, then $X_{|f|} \in \mathbb{L}^{p^*}(A, \mu_A)$.

Proof.

Let $1 \leq p^* < p \leq +\infty$ and f be as in the hypotheses of the lemma. Since $X_{|f|,q} \equiv 0$ for all but finitely many q, we can find a finite box $B \subset \mathbb{Z}^d$ such that $0 \in B$ and $X_{|f|,q} \equiv 0$ whenever $q \notin B$. We induce the whole system into $A \times \{0\}$ in two steps: first we induce from $A \times \mathbb{Z}^d$ into $A \times B$, and then from $A \times B$ into $A \times \{0\}$.

On $A \times B$, the integral of f along an excursion is only the integral on f during the first step of the excursion: after that, the process is outside of the box B, so $f \equiv 0$. The function induced on $A \times B$ is just given by $\tilde{X}_{|f|}(x, q) := X_{|f|,q}(x)$ for all $x \in A$ and $q \in B$.

By Lemma 1.6, up to a scaling of the measure, the dynamical system induced on $A \times B$ is Gibbs-Markov, the stopping time being given by the first return time in B. The subset $A \times \{0\} \subset A \times B$ is a union of elements of the Gibbs-Markov partition of $A \times B$. Let τ be the first return time to $A \times \{0\}$ for the Gibbs-Markov map on $A \times B$. Then $\mathbb{P}_{\mu_A}(\tau \geq n)$ is exponentially decaying.

The function $X_{|f|}$ is the sum of the $\tilde{X}_{|f|}$ along an excursion from $A \times \{0\}$ in $A \times B$, so that:

$$\begin{aligned}
\mathbb{E}\left(X_{|f|}^{p^*}\right) &\leq \mathbb{E}\left(\left(\sum_{k=0}^{\tau-1} (\max_{q\in B} X_{|f|,q}) \circ T_A^k\right)^{p^*}\right) \\
&= \sum_{n=1}^{+\infty} \mathbb{E}\left(1_{\tau=n}\left(\sum_{k=0}^{n-1} (\max_{q\in B} X_{|f|,q}) \circ T_A^k\right)^{p^*}\right) \\
&\leq \sum_{n=1}^{+\infty} n^{p^*}\mathbb{E}\left(1_{\tau=n}\sum_{k=0}^{n-1} (\max_{q\in B} X_{|f|,q})^{p^*} \circ T_A^k\right).
\end{aligned}$$

Let $r > 1$ be such that $rp^* < p$. By Hölder's inequality,

$$\begin{aligned}
\mathbb{E}\left(X_{|f|}^{p^*}\right) &\leq \sum_{n=1}^{+\infty} n^{p^*} \left\| \max_{q\in B} X_{|f|,q}^{p^*} \right\|_{\mathbb{L}^r} \left\| 1_{\tau=n} \right\|_{\mathbb{L}^{\frac{r}{r-1}}} \\
&\leq \sum_{n=1}^{+\infty} n^{p^*} \left\| \max_{q\in B} X_{|f|,q} \right\|_{\mathbb{L}^{rp^*}}^{p^*} \mathbb{P}\left(\tau = n\right)^{1-\frac{1}{r}} \\
&\leq \left\| \max_{q\in B} X_{|f|,q} \right\|_{\mathbb{L}^p}^{p^*} \sum_{n=1}^{+\infty} n^{p^*} \mathbb{P}\left(\tau = n\right)^{1-\frac{1}{r}}.
\end{aligned}$$

Since the tails of τ decay exponentially and $r > 1$, the sequence $n^{p^*}\mathbb{P}(\tau = n)^{1-\frac{1}{r}}$ is summable and $\left\|X_{|f|}\right\|_{\mathbb{L}^{p^*}}$ is finite. □

In particular, if $p > 2$, we can choose $p^* > 2$ in Lemma 6.9 and apply Proposition 6.2

Finally, we shall adapt the convergence in law of the local time from the discrete case to the continuous case.

Lemma 6.10.

Let us consider a recurrent $\mathbb{Z}^d$-extension with base $(\underline{A}, \pi, d_A, \mathcal{B}, \mu_A, T)$ and step time r, with $r \in \mathbb{L}^2(A, \mu_A)$ and such that $\mathbb{E}(D(r))$ is finite. Assume that $\overline{\psi}$ has regular variation of index $\beta \in [0, 1)$.

Then $\psi(t) \sim \overline{\psi}(t/\int r \, \mathrm{d}\mu)$, and in particular ψ has regular variation of index β.

Proof.

The $\mathbb{Z}^d$-extension of a Gibbs-Markov semi-flow and its discretized counterpart differ only by a random change of time. Our goal is to prove that this distortion is almost linear, in order to relate ψ and $\overline{\psi}$.

Since $\mathbb{E}(D(r))$ is finite and r is assumed to be in $\mathbb{L}^2(A, \mu_A)$, the variance of the Birkhoff sums can be controlled: there exists a constant K such that, for all positive integer n, the variance of $\left(n \int r \, \mathrm{d}\mu_A\right)^{-1} \sum_{k=0}^{n-1} r \circ T^k$ is at most Kn^{-1}. By Chebyshev's inequality, for any positive integer n,

$$\mathbb{P}_{\mu_A}\left(\frac{1}{n}\sum_{k=0}^{n-1} r\circ T^k \notin \left[(1-\varepsilon)\int_A r \, \mathrm{d}\mu_A, (1+\varepsilon)\int_A r \, \mathrm{d}\mu_A\right]\right) \leq \frac{K}{n\varepsilon^2} = \frac{o(1)}{\overline{\psi}(n)},$$

as $\beta < 1$. Then, for all $\varepsilon > 0$ and large enough t, we have the upper bound:

$$\begin{aligned}
\frac{1}{\psi(t)} &= \mathbb{P}_{\mu_A}\left(\varphi \geq t\right) \\
&\leq \mathbb{P}_{\mu_A}\left(\varphi \geq t \text{ and } \sum_{k=0}^{\lfloor((1+\varepsilon)\int r \, \mathrm{d}\mu_A)^{-1}t\rfloor - 1} r\circ T^k \leq t\right) \\
&\qquad + \mathbb{P}_{\mu_A}\left(\sum_{k=0}^{\lfloor((1+\varepsilon)\int r \, \mathrm{d}\mu_A)^{-1}t\rfloor - 1} r\circ T^k \geq t\right) \\
&\leq \mathbb{P}_{\mu_A}\left(\overline{\varphi} \geq \frac{t}{(1+2\varepsilon)\int r \, \mathrm{d}\mu_A}\right) + \frac{1}{\overline{\psi}\left(\frac{t}{\int r \, \mathrm{d}\mu_A}\right)} o(1) \\
&= (1+o(1))\frac{(1+2\varepsilon)^\beta}{\overline{\psi}\left(\frac{t}{\int r \, \mathrm{d}\mu_A}\right)}.
\end{aligned}$$

A lower bound on $\psi(t)^{-1}$ can be obtained in the same way:

$$\frac{1}{\psi(t)} \geq (1+o(1)) \frac{(1-2\varepsilon)^\beta}{\overline{\psi}\left(\frac{t}{\int r \, \mathrm{d}\mu_A}\right)}.$$

By letting t go to $+\infty$, these bounds yield for all $\varepsilon \in (0, 1/2)$:

$$\frac{1}{(1+2\varepsilon)^\beta} \leq \liminf_{t \to +\infty} \frac{\psi(t)}{\overline{\psi}\left(\frac{t}{\int r \, \mathrm{d}\mu_A}\right)} \leq \limsup_{t \to +\infty} \frac{\psi(t)}{\overline{\psi}\left(\frac{t}{\int r \, \mathrm{d}\mu_A}\right)} \leq \frac{1}{(1-2\varepsilon)^\beta}.$$

We get Lemma 6.10 by letting ε go to 0. □

The assumption in Lemma 6.10 that r belongs to $\mathbb{L}^2$ and not only to $\mathbb{L}^1$ is quite artificial. One could compute the asymptotic distribution of the local time directly, and get the asymptotics of the function ψ under the assumption that r belong to $\mathbb{L}^1$. However, since the setting and methods involved are quite different from what we work with in this article, we are content from now with just deducing the asymptotics of ψ from those of $\overline{\psi}$. The price to pay is that r has to belong to $\mathbb{L}^2$, which excludes for instance $\mathbb{Z}$ or $\mathbb{Z}^2$-periodic billiards with unbounded horizon.

Proposition 6.2, together with Lemmas 6.6 to 6.10, gives:

Theorem 6.11.
Let $(\Omega, \mu, (g_t)_{t \geq 0})$ be an ergodic $\mathbb{Z}^d$-extension of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$, of step time r and of step function F. Let f be a real-valued function from Ω to $\mathbb{R}$. Assume that:

- *r belong to $\mathbb{L}^2(A, \mu_A)$;*
- *the function $\overline{\psi}$ is regularly varying with index $\beta \in [0, 1)$;*
- *f is measurable and $X_{|f|}$ belongs to $\mathbb{L}^p(\Omega, \mu_A)$ for some $p > 2$;*
- *$\int_\Omega f \, \mathrm{d}\mu = 0$;*
- *$\mathbb{E}(D(r))$ and $\mathbb{E}(\sup_{q \in \mathbb{Z}^d} D(X_{f,q}))$ are finite.*

Then $\psi(t) \sim \overline{\psi}(t / \int r \, \mathrm{d}\mu)$, and, for any probability measure ν absolutely continuous with respect to either μ or μ_A:

$$\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma(f)\sqrt{Y_\beta}\mathcal{N},$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω, ν) to $\mathbb{R}$, where Y_β and $\mathcal{N}$ are independent, Y_β is a standard Mittag-Leffler random variable of parameter β and $\mathcal{N}$ is a standard Gaussian random variable, and where:

$$\sigma_{\overline{\varphi}}(f)^2 = \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_{\overline{\varphi}}^i \right)^2 \mathrm{d}\mu_A. \tag{6.8}$$

The literature on the local time of random walks is extensive, and includes discrete-time random walks for which the randomness is generated by a Gibbs-Markov map (see among others [2], [3], [18, Section 7.3]). Together with the Darling-Kac theorem [1, Theorem 3.6.4], this yields the 1- and 2-dimensional cases:

Corollary 6.12.
Let $(\Omega, \mu, (g_t)_{t\geq 0})$ be an ergodic $\mathbb{Z}^d$-extension of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$, of step time r and of step function F on $\mathbb{Z}$. Assume that it is non-degenerate, that r and F belong to $\mathbb{L}^2(A, \mu_A)$, and that $\mathbb{E}(D(r))$ is finite. Under these hypotheses, the quantity $\sigma(F)$ is positive and finite, where:

$$\sigma(F)^2 = \lim_{N\to+\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} F \circ T^i \right)^2 \, \mathrm{d}\mu_A.$$

Let f be a real-valued, measurable function from Ω to $\mathbb{R}$. Assume that $X_{|f|}$ belongs to $\mathbb{L}^p(A, \mu_A)$ for some $p > 2$, that $\int_\Omega f \, \mathrm{d}\mu = 0$, and that $\mathbb{E}(\sup_{q\in\mathbb{Z}^d} D(X_{f,q}))$ is finite. Then, for any probability measure ν absolutely continuous with respect to either μ or μ_A:

$$\left(\frac{\pi\sigma(F)^2}{2} \cdot \frac{\int_A r \, \mathrm{d}\mu_A}{t} \right)^{\frac{1}{4}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma_{\overline{\varphi}}(f) \sqrt{|\mathcal{N}'|} \mathcal{N},$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω, ν) to $\mathbb{R}$, where $\mathcal{N}$ and $\mathcal{N}'$ are independent, $\mathcal{N}$ is a standard Gaussian random variable, $\mathcal{N}'$ is a centered Gaussian random variable of variance $\pi/2$ (so that $\mathbb{E}(|\mathcal{N}'|) = 1$), and $\sigma_{\overline{\varphi}}(f)$ is given by Equation (6.8).

Corollary 6.13.
Let $(\Omega, \mu, (g_t)_{t\geq 0})$ be an ergodic $\mathbb{Z}^d$-extension of base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$, of step time r and of step function F on $\mathbb{Z}^2$. Assume that it is non-degenerate, that r and F belong to $\mathbb{L}^2(A, \mu_A)$, and that $\mathbb{E}(D(r))$ is finite. Under these hypotheses, the covariance matrix $\Sigma(F)$ is positive definite, where, for all u and v in $\mathbb{R}^2$:

$$(u, \Sigma(F)v) = \lim_{N\to+\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} F \circ T^i, u \right) \left(\sum_{i=0}^{N-1} F \circ T^i, v \right) \, \mathrm{d}\mu_A.$$

Let f be a real-valued, measurable function from Ω to $\mathbb{R}$. Assume that $X_{|f|}$ belongs to $\mathbb{L}^p(A, \mu_A)$ for some $p > 2$, that $\int_\Omega f \, \mathrm{d}\mu = 0$, and that $\mathbb{E}(\sup_{q\in\mathbb{Z}^d} D(X_{f,q}))$ is finite. Then, for any probability measure ν absolutely continuous with respect to either μ or μ_A:

$$\left(\frac{2\pi\sqrt{\det(\Sigma(F))}}{\ln(t)} \right)^{\frac{1}{2}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma_{\overline{\varphi}}(f) \sqrt{\mathcal{E}} \mathcal{N},$$

where the convergence is in distribution when the left-hand side is seen as a random variable from (Ω, ν) to $\mathbb{R}$, where $\mathcal{N}$ and $\mathcal{E}$ are independent, $\mathcal{N}$ is a standard Gaussian random variable, $\mathcal{E}$ is an exponential random variable of parameter 1 *(so that $\mathbb{E}(\mathcal{E}) = 1$), and $\sigma_{\overline{\varphi}}(f)$ is given by Equation* (6.8).

There are analogous results for $\mathbb{Z}$-extensions with a transition kernel which is in the basin of attraction of a symmetric Lévy stable distribution of index $\beta \in (1, 2]$. In all those situations, the integrability and smoothness conditions on the observable f hold as soon as $X_{|f|,q} \equiv 0$ for all but finitely vertices q, and for those q the function $X_{|f|,q}$ belongs to $\mathbb{L}^p(A, \mu_A)$ for some $p > 2$ and $\mathbb{E}(D(X_{f,q}))$ is finite.

6.3 Flows

Our last goal is to prove distributional convergence theorems not for semi-flows, but for flows. Since flows are invertible dynamical systems, we will work not with Gibbs-Markov maps, but with natural extensions thereof. In this subsection, we will denote by $(A_+, \pi, \lambda, \mathcal{B}_+, \mu_+, T_+)$ a Gibbs-Markov map and by $(A, \mathcal{B}, \mu, T)$ its natural extension. We also denote by $\mathcal{B}_+$ the σ-algebra of the future on A, which is the backward image of $\mathcal{B}_+$ by the canonical projection from A onto A_+. The partition π is also pulled backwards to a countable partition of A, that we also denote by π. The dynamical system (A, μ, T) is isomorphic to a subshift of $\pi^{\mathbb{Z}}$. For each $a \in \pi$, let us choose an element $x_a = (x_{a,n})_{n \in \mathbb{Z}} \in a$. Then, for each a in π and each $x = (x_n)_{n \in \mathbb{Z}}$ in a, let us put:

$$p_+(x)_n := \begin{cases} x_{a,n} & \text{if} \quad n \leq 0 \\ x_n & \text{if} \quad n \geq 0 \end{cases}.$$

The function $p_+ : A \to A$ corresponds to a projection onto a piece of unstable manifold by holonomy along the stable manifold. A function f is $\mathcal{B}_+$-measurable is and only if $f \circ p_+ = f$; we then says that the function only depends on the future. Such functions are convenient, as they factorize over the initial Gibbs-Markov map. In particular, any limit theorem we have for observables of (A_+, μ_+, T_+) can be applied to observables which depends only on the future of (A, μ, T).

To define a distance on the natural extension of a Gibbs-Markov map, we use the separation times:

$$\begin{aligned} s_+(x,y) &:= \inf\{n \geq 0 : \ T^n x \text{ and } T^n y \text{ do not belong to the same element of } \pi\}, \\ s_-(x,y) &:= \inf\{n \geq 0 : \ T^{-n} x \text{ and } T^{-n} y \text{ do not belong to the same element of } \pi\}. \end{aligned}$$

With these two separation times we define two pseudo-distances $d_+ := \lambda^{-s_+}$ and $d_- := \lambda^{-s_-}$, and a distance $d := d_+ + d_-$.

The notion of $\mathbb{Z}^d$-extension of a Gibbs-Markov semi-flow was defined in the beginning of Subsection 6.2. Given a real-valued, measurable function r on A which is almost surely positive, and a $\mathbb{Z}^d$-valued measurable function F on A which is $\sigma(\pi)$-measurable, we can define in the same way a $\mathbb{Z}^d$-extension of base (A, π, d, μ_A, T_A), with step F and step time r. By the next lemma, with mild assumptions on r, we can assume that r depends only on the future. While this is a classic result [5, Lemma 1.6], the precise expression of the coboundary used in the proof will be useful later on.

Lemma 6.14.
Let f be a measurable real-valued function on A. Assume that $\|D(f)\|_\infty$ is finite. Then there exists a function u which is bounded by $\lambda(\lambda-1)^{-1} \|D(f)\|_\infty$, uniformly $1/2$-Hölder, and such that the function $f_+ := f + u \circ T - u$ is $\mathcal{B}_+$-measurable.

Proof.
Let us put:

$$u := \sum_{n=0}^{+\infty} f \circ T^n - f \circ T^n \circ p_+.$$

Let $n \geq 0$. If x and y are in the same element of $\bigvee_{k=-n}^{+\infty} T^{-k}\pi$, then $T^k x$ and $T^k y$ are in the same element of π for all $k \geq -n$, so that $d(x,y) \leq \lambda^{-n}$ and $|r(x) - r(y)| \leq \lambda^{-n} \|D(f)\|_\infty$. Hence, $|f \circ T^n - f \circ T^n \circ p_+| \leq \|D(f)\|_\infty \lambda^{-n}$, and the function u is well-defined and bounded. We compute:

$$\begin{aligned} f + u \circ T - u &= f + \sum_{n=1}^{+\infty} \left(f \circ T^n - f \circ T^{n-1} \circ p_+ \circ T \right) - \sum_{n=0}^{+\infty} (f \circ T^n - f \circ T^n \circ p_+) \\ &= f \circ p_+ + \sum_{n=1}^{+\infty} f \circ T^n \circ p_+ - f \circ T^{n-1} \circ p_+ \circ T. \end{aligned}$$

This shows that the function $f_+ = f + u \circ T - u$ is $\mathcal{B}_+$-measurable.

Now we prove that u is 1/2-Hölder on each element of π. Let a be in π, and let x, y be in a. If $d_+(x,y) = 0$, then $p_+(x) = p_+(y)$ and $|f \circ T^n(x) - f \circ T^n(y)| \leq \|D(f)\|_\infty \lambda^{-n} d_-(x,y)$, so that:

$$|u(x) - u(y)| \leq \frac{\lambda}{\lambda - 1} \|D(f)\|_\infty d_-(x,y).$$

If $d_-(x,y) = 0$, then we split the sum:

$$\begin{aligned}
|u(x) - u(y)| \leq & \sum_{n=0}^{\lfloor s_+(x,y)/2 \rfloor} |f \circ T^n \circ p_+(x) - f \circ T^n \circ p_+(y)| + |f \circ T^n(x) - f \circ T^n(y)| \\
& + \sum_{n=\lfloor s_+(x,y)/2 \rfloor + 1}^{+\infty} |f \circ T^n(x) - f \circ T^n \circ p_+(x)| + |f \circ T^n(y) - f \circ T^n \circ p_+(y)| \\
& \leq 2 \|D(f)\|_\infty \left(\frac{\lambda^{\lfloor s_+(x,y)/2 \rfloor - s_+(x,y) + 1}}{\lambda - 1} + \frac{\lambda^{-\lfloor s_+(x,y)/2 \rfloor - 1}}{\lambda - 1} \right) \\
& \leq \frac{4\lambda \|D(f)\|_\infty}{\lambda - 1} d_+(x,y)^{\frac{1}{2}}.
\end{aligned}$$

For general x and y in a, let z be the element of a such that the positive coordinates of z are the positive coordinates of x, and the negative coordinates of z are the negative coordinates of y. Then $d_+(x,z) = d_-(y,z) = 0$, and $d_-(x,z) = d_-(x,y)$, and $d_+(y,z) = d_+(x,y)$. Thus:

$$|u(x) - u(y)| \leq |u(x) - u(z)| + |u(z) - u(y)| \leq \frac{4\lambda \|D(f)\|_\infty}{\lambda - 1} \left(d_+(x,y)^{\frac{1}{2}} + d_-(x,y)^{\frac{1}{2}} \right). \qquad \square$$

Let us apply Lemma 6.14 to the step time r. Since we can take $\sqrt{\lambda}$ instead of λ as an expansion factor for the Gibbs-Markov map, without loss of generality, we can assume that the functions u and r_+ are actually Lipschitz. The function r_+ is not always positive, but if r is bounded below by a positive constant, then there is an integer N such that $\sum_{k=0}^{N-1} r_+ \circ T^k$ is positive and can be used as a new step time for the suspension flow. Hence, if r is bounded below by a positive constant and satisfies the assumptions of Lemma 6.14, we can assume without loss of generality that r is $\mathcal{B}_+$-measurable.

We finally tackle our first concrete example: the geodesic flow on periodic manifolds of negative sectional curvature. Other systems, such as billiards, can be studied in a similar fashion; however, there are significant adaptations to do. Periodic manifolds of negative curvature are in many respects the easiest examples to work with.

Let M be a compact, connected manifold with a Riemannian metric of negative sectional curvature; let μ_M be a Gibbs measure on T^1M. For $d = 1$ or 2, let $p : N \to M$ be a connected $\mathbb{Z}^d$-cover of M. Let $(g_t)_{t \in \mathbb{R}}$ be the geodesic flow on the unit tangent bundle T^1N of N, which is endowed with a measure μ_N by lifting μ_M. With these notations, we will prove:

Proposition 6.15.

Let f be a real-valued Hölder function on T^1N with compact support. Assume that $\int_{T^1N} f \, \mathrm{d}\mu_N = 0$. Let $\nu \ll \mu_N$. If $d = 1$, there exists a non-negative constant $K(f)$ such that:

$$\lim_{t \to +\infty} \frac{1}{t^{\frac{1}{4}}} \int_0^t f \circ g_s(x,v) \, \mathrm{d}s = K(f) \sqrt{|\mathcal{N}'|} \mathcal{N},$$

where the convergence is in distribution when (x,v) has distribution ν on T^1N, where $\mathcal{N}$ and $\mathcal{N}'$ are independent, $\mathcal{N}$ is a standard Gaussian random variable, and $\mathcal{N}'$ is a centered Gaussian random variable of variance $\pi/2$ (so that $\mathbb{E}(|\mathcal{N}'|) = 1$).

If $d = 2$, there exists a non-negative constant $K(f)$ such that:

$$\lim_{t \to +\infty} \frac{1}{\sqrt{\ln(t)}} \int_0^t f \circ g_s(x, v) \, \mathrm{d}s = K(f)\sqrt{\mathcal{E}}\mathcal{N},$$

where the convergence is in distribution when (x, v) has distribution ν on T^1N, where $\mathcal{N}$ and $\mathcal{E}$ are independent, $\mathcal{N}$ is a standard Gaussian random variable, and $\mathcal{E}$ is an exponential random variable of parameter 1.

In both cases, $K(f) = 0$ if and only if f is a coboundary.

The proof relies on an isomorphism between the geodesic flow on T^1M and a suspension flow over the natural extension of a Gibbs-Markov map with finite alphabet [4], [5, Theorem 3.12]. We recall the main features of this isomorphism. One can find finitely many submanifolds $A_1, \dots, A_p$ transverse to the geodesic flow and a function $r : \bigcup_i A_i \to \mathbb{R}_+^*$ such that:

1. each A_i has a box structure given by the strong stable manifold and the weak unstable manifold of T^1M (actually, each A_i can be constructed as a union of pieces of strong stable manifold which intersect a given piece of strong unstable manifold; the non-integrability of the foliation by strong stable and unstable manifolds prevents it from being a union of pieces of strong unstable manifold);

2. the function r is bounded from below and from above by some positive constants, and for all $x \in \bigcup_i A_i$, one has $r(x) = \inf\{t > 0 : g_t(x) \in \bigcup_i A_i\}$;

3. r is Hölder on each subset $\{x \in A_i : g_{r(x)}(x) \in A_j\}$;

4. since the foliation by strong stable leaves is invariant under the flow, r is constant on each piece of strong stable manifold in each subset $\{x \in A_i : g_{r(x)}(x) \in A_j\}$;

5. the image of any maximal piece of weak unstable manifold in some A_i by g_r is a union of maximal pieces of weak unstable manifold;

6. the backward image of any maximal piece of strong stable manifold in some A_i by g_r is a union of maximal pieces of strong stable manifold.

Let $A := \bigcup_i A_i$, and $\pi := \{A_1, \dots, A_p\}$, and $Tx := g_{r(x)}(x)$ for $x \in A$. Let μ_A be the Liouville measure on A, renormalized so as to be a probability measure. Then T preserves the measure μ_A.

We can refine the partition π into a partition π' so that r is Hölder on each element of π'. Any point $x \in A$ can be encoded by the sequence of transversals in π' to which $T^n x$ belongs for all $n \in \mathbb{Z}$. By the points (5) and (6), the dynamical system (A, μ_A, T) is a subshift of finite type of $(\pi')^{\mathbb{Z}}$. If we denote by d_A the distance on A inherited from the distance on T^1M, then the geodesic flow on T^1M is measurably isomorphic to the suspension flow of base (A, π', d_A, μ_A, T) and height function r. By (4), the height function only depends on the σ-algebra of the future.

Proof of Proposition 6.15.

Encoding of the geodesic flow on T^1N: We use the symbolic encoding of the geodesic flow on T^1M described above. First, let us prove that the geodesic flow on N is isomorphic to a $\mathbb{Z}^d$-extension of a suspension flow whose base is the natural extension of a Gibbs-Markov map. This is done as in [13]. Each transversal $a \in \pi$ has a box structure, so the preimage $p^{-1}a$ is homeomorphic to $a \times \mathbb{Z}^d$. For each a, we choose one of the pieces of $p^{-1}a$ and let it correspond to $a \times \{0\} \subset \pi \times \mathbb{Z}^d$. This defines an origin on $p^{-1}\pi$, in that it distinguishes a single isomorphism between $p^{-1}\pi$ and $\pi \times \mathbb{Z}^d$, and from there an isomorphism between $p^{-1}\pi'$ and $\pi' \times \mathbb{Z}^d$.

Hence, the geodesic flow on N is measurably isomorphic to the $\mathbb{Z}^d$-extension of base (A, π', d_A, μ_A, T) and height function r for some step function F. The function $x \mapsto g_{r(x)}(x)$ is continuous from any element of $p^{-1}\pi'$ to its image, and by connectedness is a single element of $p^{-1}\pi$. Hence, F is constant on any element of π'.

The system we work on is measurably isomorphic to $A \times [0, r] \times \mathbb{Z}^d$. We extend the function r to the whole space by $r(x, t, q) = r(x)$. We can define a projection p_+ onto the reference unstable leaves in each A_i, that we extend to $A \times [0, r] \times \mathbb{Z}^d$ by $p_+(x, t, q) = (p_+(x), t, q)$; this is well-defined because $r = r \circ p_+$ on A. For any measurable and integrable function f on $A \times [0, r] \times \mathbb{Z}^d$, we put:

$$Y_f(x, q) := \int_0^{r(x)} f(x, t, q) \, \mathrm{d}t.$$

We have $Y_f(x, q) = X_{f,q}(x)$ for all $x \in A$ and $q \in \mathbb{Z}^d$. Having a single function defined on $A \times \mathbb{Z}^d$, and not a family of functions defined on A, will soon be convenient. Finally, we define T' from $A \times \mathbb{Z}^d$ onto itself by $T'(x, q) = (Tx, q + F(x))$.

Properties of the projection of a function f onto $A \times \mathbb{Z}^d$: Given the observable f, we want to prove that Y_f is at least Hölder, and that up to the addition of a coboundary, it depends only on the future. We focus now on the regularity of Y_f.

Let d_N denote the distance on T^1N, and κ_- the smallest absolute value of a Lyapunov exponent of the geodesic flow on M. Then there exists a positive constant K with the following property. Let α be in $(0, 1]$, and let f be a α-Hölder function on T^1N with compact support. Let x and y be in the same element a of π, and let $q \in \mathbb{Z}^d$. If x and y are in the same strong stable leaf, then $d_N(g_t(x, q), g_t(y, q)) \leq K d_N(x, y) e^{-\kappa_- t}$ for all $t \geq 0$, whence:

$$|f \circ g_t(x, q) - f \circ g_t(y, q)| \leq K |f|_{\mathrm{Hol}_\alpha} e^{-\alpha \kappa_- t} d_N(x, y)^\alpha,$$

and:

$$|Y_f(x, q) - Y_f(y, q)| \leq \frac{K |f|_{\mathrm{Hol}_\alpha}}{\alpha \kappa_-} d_N(x, y)^\alpha.$$

If x and y are in the same small piece of strong unstable manifold, we can repeat the same process with minor modifications. However, a is made of pieces of weak unstable manifold. To circumvent this issue, choose a point x in the piece of strong unstable manifold included in a, and then take pieces of strong unstable manifold intersecting the piece of strong stable manifold going through x; this process generates a new transversal b. By choosing pieces of strong stable manifold of the right size, one can ensure that b has a box structure, and is the image of a under the geodesic flow stopped at non-constant times. More precisely, we can ensure that there exists a function $\delta : a \to \mathbb{R}$ which is Hölder continuous, such that $b = g_\delta(a)$ and $\delta = 0$ on the pieces of strong stable and strong unstable manifold going through x. We do this on each element $a \in \pi$, so as to define a Hölder function δ on the Poincaré section π and from there on $p^{-1}\pi$. Let α' be the Hölder exponent of δ.

Let a' be in π'. If x and y are in the same piece of weak unstable manifold in a', then $g_{r(x)}(x)$ and $g_{r(y)}(y)$ are in the same piece of weak unstable manifold in some element $a \in \pi$, and $g_{r(x)+\delta(T'x)}(x)$ and $g_{r(y)+\delta(T'y)}(y)$ are on the same piece of strong unstable manifold. By inverting the flow, we get for all $t \geq 0$:

$$d_N(g_{r(x)+\delta(T'x)-t}(x), g_{r(y)+\delta(T'y)-t}(y)) \leq K d_N(g_{r(x)+\delta(T'x)}(x), g_{r(y)+\delta(T'y)}(y)) e^{-\kappa_- t},$$

whence, using the fact that $L := r + \delta \circ T' - \delta$ is constant on the pieces of weak unstable manifold in a':

$$\begin{aligned}
|Y_f(x,q) \ -&Y_f(y,q)| \\
&= \left| \int_{\delta(T'x)}^{r(x)+\delta(T'x)} f(g_{r(x)+\delta(T'x)-t}(x,q,0)) \,\mathrm{d}t - \int_{\delta(T'y)}^{r(y)+\delta(T'y)} f(g_{r(y)+\delta(T'y)-t}(y,q,0)) \,\mathrm{d}t \right| \\
&\leq \int_{\max\{\delta(T'x),\delta(T'x)\}}^{L+\min\{\delta(x),\delta(y)\}} |f|_{\mathrm{Hol}_\alpha} d_N(g_{r(x)+\delta(T'x)-t}(x), g_{r(y)+\delta(T'y)-t}(y))^\alpha \,\mathrm{d}t \\
&\qquad + |\delta(x)-\delta(y)| \, \|f\|_\infty + |\delta(T'x)-\delta(T'y)| \, \|f\|_\infty \\
&\leq \frac{K^\alpha |f|_{\mathrm{Hol}_\alpha}}{\alpha\kappa_-} d_N(g_{r(x)+\delta(T'x)}(x), g_{r(y)+\delta(T'y)}(y))^\alpha \\
&\qquad + |\delta|_{\mathrm{Hol}_{\alpha'}(a)} \, \|f\|_\infty \, d_N(x,y)^{\alpha'} + |\delta|_{\mathrm{Hol}_{\alpha'}(Ta)} \, \|f\|_\infty \, d_N(T'x,T'y)^{\alpha'}.
\end{aligned}$$

If λ is the smallest dilation constant of the map $x \mapsto g_{r(x)}(x)$ on the unstable leaves of T^1M, there is a positive constant K' such that $K' d_N(x,y) \leq d(x,y) := \lambda^{-s_+(x,y)} + \lambda^{s_-(x,y)}$, and the subshift on A_+ is Gibbs-Markov for the distance λ^{-s_+}. Thus the function Y_f is bounded by $\|f\|_\infty r$ and $\min\{\alpha,\alpha'\}$-Hölder for the distance d on each subset $A \times \{q\}$, uniformly in q.

Towards the limit theorem: We now introduce a well-chosen coboundary which, once added to f, yields a function which only depends on the future. We will then be able to factor the dynamical system over the initial Gibbs-Markov map, and use our previous theorems. Using the same method as in Lemma 6.14, we consider the following functions u and f_+, defined on $A \times \mathbb{Z}^d$ and T^1N respectively:

$$u(x,q) := Y_{f-f\circ p_+} + \sum_{n=1}^{+\infty} \left(Y_f \circ T'^n(x,q) - Y_f \circ T'^n \circ p_+(x,q) \right),$$

$$f_+(x,t,q) := f \circ p_+(x,t,q) + \frac{1}{r(x)} \sum_{n=1}^{+\infty} \left(Y_f \circ T'^n \circ p_+(x,q) - Y_f \circ T'^{n-1} \circ p_+ \circ T'(x,q) \right).$$

We extend u to T^1N by putting, for $t < r(x)$:

$$u(x,t,q) := u(x,q) + \int_0^t (f_+ - f)(x,t,s) \,\mathrm{d}s.$$

Since r is Lipschitz and bounded away from zero, the functions u and f_+ are bounded and $\min\{\alpha,\alpha'\}/2$-Hölder, and $f_+ = f_+ \circ p_+$. The function $f_+ - f$ is a bounded coboundary for the geodesic flow, as is the function $Y_{f_+} - Y_f$ for the transformation T and the function $X_{f_+} - X_f$ for the transformation T_A.

The return time r and the new observable f_+ only depend on the future. Proposition 6.15 is a consequence of Corollaries 6.12 and 6.13 applied to f_+. We need to check that f_+ satisfies the assumptions of these corollaries. However, the criterion "$X_{|f_+|} \in \mathbb{L}^p$ for some $p > 2$ and $\int f_+ \,\mathrm{d}\mu = 0$" is too restrictive for our purposes, as it behaves badly when a coboundary is added to f. As in [16, Remark 4.6], it can be replaced by:

- $\sup_{0 \leq t \leq \varphi} \left| \int_0^t f_+ \circ g_s \,\mathrm{d}s \right| \in \mathbb{L}^p(A, \mu_A)$ for some $p > 2$;
- $\int_A X_{f_+} \,\mathrm{d}\mu_A = 0.$

The function f_+ is measurable. Up to a change in the metric on A, the function Y_{f_+} can be assumed to be Lipschitz on each element $a \times \{q\}$ of $p^{-1}\pi'$ uniformly in a and q, so that $\mathbb{E}(\sup_{q \in Z^d} D(X_{f_+,q}))$ is finite.

The section π is made of finitely many boxes, which implies that the step time r is bounded. This makes the family $X_{|f|,q}$ uniformly bounded. Moreover, $p^{-1}\pi$ is locally finite, so there are only finitely many elements of $p^{-1}\pi$ which are at a distance less than $\|r\|_\infty$ of the support of f, which we assumed to be compact. Hence, $X_{|f|,q} \equiv 0$ for all but finitely many q, and $X_{|f|,q}$ is bounded for all q. By Lemma 6.9, we know that $X_{|f|} \in \mathbb{L}^p$ for all finite p. Then:

$$\left\| \sup_{0 \leq t \leq \varphi} \left| \int_0^t f_+ \circ g_s \, \mathrm{d}s \right| \right\|_{\mathbb{L}^p} \leq \left\| \sup_{0 \leq t \leq \varphi} \left| \int_0^t f \circ g_s \, \mathrm{d}s \right| \right\|_{\mathbb{L}^p} + 2 \|u\|_{\mathbb{L}^\infty} \leq \left\| X_{|f|} \right\|_{\mathbb{L}^p} + 2 \|u\|_{\mathbb{L}^\infty} ,$$

so $\sup_{0 \leq t \leq \varphi} \left| \int_0^t f_+ \circ g_s \, \mathrm{d}s \right|$ is also in $\mathbb{L}^p$ for all $p > 2$. Finally, $\int_A X_{f_+} \, \mathrm{d}\mu_A = \int_A X_f \, \mathrm{d}\mu_A = \int_N f \, \mathrm{d}\mu_N = 0$. □

The periodic billiards are more difficult to handle, because the step time r for the underlying Gibbs-Markov map is unbounded. Hence, the argument which concludes that $X_{|f|,q} = 0$ for all but finitely many q if f has compact support fails. It is possible to write down a limit theorem which applies to functions supported by finitely many boxes $A \times [0, r] \times \{q\}$, but that would be inelegant and hard to apply to any given example.

References

[1] (MR1450400) J. Aaronson, *An introduction to infinite ergodic theory*, American Mathematical Society, 1997.

[2] (MR1890328) J. Aaronson and M. Denker, *A local limit theorem for stationary processes in the domain of attraction of a normal distribution*, *Asymptotic methods in probability and statistics with applications (St. Petersburg, 1998)*, 215-223, Birkhäuser Boston, 2001.

[3] (MR1840194) J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stochastics and Dynamics, **1** (2001), 193-237.

[4] (MR0339281) R. Bowen, *Symbolic dynamics for hyperbolic flows*, American Journal of Mathematics, **95** (1973), 429-460.

[5] (MR0442989) R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics, 1975.

[6] (MR0365692) D.L. Burkholder, *Distribution function inequalities for martingales*, The Annals of Probability, **1** (1973), 19-42.

[7] (MR1661493) E. Csáki and A. Földes, *On asymptotic independence and partial sums*, Asymptotic methods in probability and statistics, A volume in honour of Miklós Csörgõ, Elsevier (2000), 373-381.

[8] (MR1820506) E. Csáki and A. Földes, *Asymptotic independence and additive functionals*, Journal of Theoretical Probability, **13** (2000), 1123-1144.

[9] (MR0071662) R.L. Dobrushin, *Two limit theorems for the simplest random walk on a line* (Russian), Uspekhi Matematicheskikh Nauk, **10** (1955), 139-146.

[10] (MR2041223) S. Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel Journal of Mathematics, **139** (2004), 29-65.

[11] (MR0945679) A. Jakubowski, *Tightness criteria for random measures with application to the principle of conditioning in Hilbert spaces*, Probability and Mathematical Statistics, **9** (1988), 95-114.

[12] G. Pisier, *Martingales in Hilbert spaces (in connection with type and cotype)*, URL: http://people.math.jussieu.fr/~pisier/ihp-pisier.pdf (February 9, 2011).

[13] (MR1273266) M. Pollicott and R. Sharp, *Orbit counting for some discrete groups acting on simply connected manifolds with negative curvature*, Inventiones Mathematicae, **117** (1994), 275-302.

[14] F. Schweiger, *Numbertheoretical endomorphisms with σ-finite invariant measure*, Israel Journal of Mathematics, **21** (1975), 308-318.

[15] (MR0268938) R.J. Serfling, *Moment inequalities for the maximum cumulative sum*, The Annals of Mathematical Statistics, **41** (1970), 1227-1234.

[16] D. Thomine, *A generalized central limit theorem in infinite ergodic theory*, to appear in Probability Theory and Related Fields.

[17] (MR2359068) R. Zweimüller, *Mixing limit theorems for ergodic transformations*, Journal of Theoretical Probability, **20** (2007), 1059-1071.

[18] (MR2373265) R. Zweimüller, *Infinite measure preserving transformations with compact first regeneration*, Journal d'Analyse Mathématique , **103** (2007), 93-131.