LOCAL LIMIT THEOREM FOR NONUNIFORMLY PARTIALLY HYPERBOLIC SKEW-PRODUCTS AND FAREY SEQUENCES

SÉBASTIEN GOUËZEL

Abstract

We study skew-products of the form $(x, \omega) \mapsto (Tx, \omega + \phi(x))$, where T is a nonuniformly expanding map on a space X, preserving a (possibly singular) probability measure $\tilde{\mu}$, and $\phi : X \to \mathbb{S}^1$ is a C^1 function. Under mild assumptions on $\tilde{\mu}$ and ϕ, we prove that such a map is exponentially mixing and satisfies both the central limit and local limit theorems. These results apply to a random walk related to the Farey sequence, thereby answering a question of Guivarc'h and Raugi [GR, Section 5.3].

Contents

1. Results . . . 193
2. Tools on transfer operators . . . 209
3. Exponential mixing . . . 212
4. Strategy and tools for the local limit theorem . . . 225
5. Proof of the local limit theorem . . . 235
6. Proofs for Farey sequences . . . 257
Appendix. Contraction properties of transfer operators . . . 267
References . . . 282

1. Results

Let $\mathscr{T}$ be a transformation on a compact manifold. If $\mathscr{T}$ is uniformly expanding or hyperbolic, the transfer operator associated to $\mathscr{T}$ admits a spectral gap on a well-chosen Banach space, which makes it possible to prove virtually any limit theorem (e.g., the local limit theorem) by using Nagaev's method (see, e.g., [GH], [HH]). This article is devoted to the proof of the local limit theorem for transformations of the form $\mathscr{T} : (x, \omega) \mapsto (Tx, \omega + \phi(x))$, where T is a nonuniformly expanding transformation on a compact manifold X, and $\phi : X \to \mathbb{S}^1$ is a C^1 function. This transformation $\mathscr{T}$ is an isometry in the fibers $\mathbb{S}^1$, which prevents us from obtaining a spectral gap.

DUKE MATHEMATICAL JOURNAL
Vol. 147, No. 2, © 2009 DOI 10.1215/00127094-2009-011
Received 22 March 2007. Revision received 19 June 2008.
2000 *Mathematics Subject Classification.* Primary 37A25, 37A50, 37D30; Secondary 37A30, 37D25.

Limit theorems have been obtained (in the more general setting of partially hyperbolic transformations) by Dolgopyat in [D3] when T is uniformly hyperbolic, and for a measure which is absolutely continuous with respect to Lebesgue measure in the unstable direction. However, he uses elementary arguments (moment methods) which cannot be used to get the local limit theorem. To the best of our knowledge, the only partially hyperbolic transformations for which a local limit theorem is proved in the literature are the Anosov flows, found in [W] (the specific algebraic structure of flows makes it possible to reduce the problem to the study of Axiom A maps, which are uniformly hyperbolic). With the techniques of [Ts], it is probably possible to obtain it also for skew-products over uniformly expanding maps for an absolutely continuous measure. Unfortunately, the main motivating example of our study, described in Section 1.1, is nonuniformly hyperbolic, and its invariant measure is singular. Hence, we need to introduce a new technique, essentially based on renewal theory.

The qualitative theory of skew-products as above has been studied by Brin and Pesin [BP]. We need more quantitative results, and to obtain them we use tools that are mainly due to Dolgopyat [D1], [D2]. These techniques of Dolgopyat have already proved very powerful in a variety of contexts (see [PS], [An], [St], [N], [BV1], [BV2], [AGY]), and the present article is yet another illustration of their usefulness.

1.1. Farey sequences

Before we give the precise definition of the systems to which our results apply, let us describe an interesting example that is, in fact, the main motivation for this article. The following discussion is essentially taken from [CG].

If p/q and p'/q' are two irreducible rational numbers in $[0, 1]$, they are *adjacent* if $|pq'-p'q| = 1$. We can then construct their median $p''/q'' = (p+p')/(q+q')$, which lies between p/q and p'/q' and is adjacent to any of them. Let $\mathscr{F}_0 = \{0/1, 1/1\}$, and define inductively $\mathscr{F}_n$ by enumerating the elements of $\mathscr{F}_{n-1}$ in increasing order (which gives a sequence of adjacent rational numbers) and by inserting the successive medians. For example, $\mathscr{F}_1 = \{0/1, 1/2, 1/1\}$ and $\mathscr{F}_2 = \{0/1, 1/3, 1/2, 2/3, 1/1\}$. The set $\mathscr{F}_n$ has cardinality $2^n + 1$. Let also $\mathscr{F}_n^* = \mathscr{F}_n - \{0\}$; it has cardinality 2^n. Any rational number of $(0, 1]$ belongs to $\mathscr{F}_n^*$ for any large enough n. Let $\mu_n = 2^{-n}\sum_{x\in\mathscr{F}_n^*}\delta_x$; this sequence of measures converges exponentially fast to a measure μ, in the following sense: for any $\alpha > 0$, there exist $C > 0$ and $\theta < 1$ such that, for any function $f : [0, 1] \to \mathbb{C}$ which is Hölder-continuous of exponent α,

$$\left|\int f \,\mathrm{d}\mu_n - \int f \,\mathrm{d}\mu\right| \leq C\theta^n \|f\|_{C^\alpha}. \tag{1.1}$$

The measure μ is called *Minkowski's measure*. It has full support in $[0, 1]$ and is totally singular with respect to Lebesgue measure. It is the Stieltjes measure associated to Minkowski's ? function.

To prove the exponential convergence (1.1), it is more convenient to reformulate everything in terms of a random walk on a homogeneous space for the group $\mathrm{SL}(2,\mathbb{R})$. Consider the two matrices $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$ in $\mathrm{SL}(2,\mathbb{R})$; we are interested in the random walk in $\mathrm{SL}(2,\mathbb{R})$ going from a matrix M to AM or BM with probability $1/2$. The Markov operator of this random walk is simply the convolution with the measure $\nu = (\delta_A + \delta_B)/2$. Since $\mathrm{SL}(2,\mathbb{R})$ is noncompact, this random walk does not have interesting recurrence properties. However, whenever X is a homogeneous space for the action of $\mathrm{SL}(2,\mathbb{R})$, the previous random walk induces a random walk on X, going from x to $A \cdot x$ or $B \cdot x$ with probability $1/2$, which can be recurrent.

For instance, let $X = \mathbb{P}(\mathbb{R}^2)$. Since $\mathrm{SL}(2,\mathbb{R})$ acts linearly on $\mathbb{R}^2$, it also acts on the compact space X. The actions of the matrices A and B on $\mathbb{R}^2$ leave invariant the cone $\mathbf{C} = \{(x, y) \mid 0 \le x \le y\}$, and its projectivization $\mathbb{P}(\mathbf{C})$ is the unique closed subset of $\mathbb{P}(\mathbb{R}^2)$ which is invariant and minimal for the action of the semigroup Σ generated by A and B. Let us identify $\mathbb{P}(\mathbf{C})$ with the interval $[0, 1]$ by intersecting $\mathbf{C}$ with the line $y = 1$; we obtain an action of Σ on $[0, 1]$. The actions of the matrices A and B are given by the transformations

$$h_A(x) = \frac{x}{1+x}, \qquad h_B(x) = \frac{1}{2-x}. \tag{1.2}$$

It can easily be checked inductively that

$$\mathscr{F}_n^* = \Big\{ M_n \cdots M_1 \cdot 1 \;\Big|\; M_i \in \{A, B\} \text{ for } i = 1, \dots, n \Big\}. \tag{1.3}$$

In particular, we have $\mu_n = \nu^n \star \delta_1$. The measure μ is the unique stationary measure for the random walk given by ν (i.e., such that $\nu \star \mu = \mu$). The exponential convergence (1.1) is then proved by showing that the Markov operator associated to the random walk has a spectral gap when it acts on the space of Hölder-continuous functions.

In [CG] (see also [GR]), Conze and Guivarc'h have considered the same random walk, but on homogeneous spaces that are larger than $\mathbb{P}(\mathbb{R}^2)$. More precisely, let us fix $r > 1$, and consider the quotient X of $\mathbb{R}^2 - \{0\}$ by the subgroup H_r of homotheties of ratio $\pm r^n$, $n \in \mathbb{Z}$. This is a compact space, endowed with an action of $\mathrm{SL}(2,\mathbb{R})$. In particular, the semigroup Σ acts on $\bar{\mathbf{C}} = \mathbf{C}/H_r$, which is a compact extension (with fiber $\mathbb{S}^1$) of $\mathbb{P}(\mathbf{C})$. Let us identify $\bar{\mathbf{C}}$ with $[0, 1] \times \mathbb{R}/(\log r)\mathbb{Z}$ by $(x, y) \mapsto (x/y, \log y + (\log r)\mathbb{Z})$. With this identification, a matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ acts by

$$M \cdot (x, \omega) = \Big(\frac{ax+b}{cx+d}, \omega + \log(cx+d) \Big). \tag{1.4}$$

Hence, the random walk given by ν on $\bar{\mathbf{C}}$ jumps from (x, ω) to $\bar{h}_A(x, \omega) := (h_A(x), \omega + \log(1+x))$ or $\bar{h}_B(x, \omega) := (h_B(x), \omega + \log(2-x))$ with probability $1/2$. Moreover, if the random walk starts from $(1, 0) \in \bar{\mathbf{C}}$, then the formula (1.4) shows that, at time n, this random walk reaches the points of the form $(p/q, \log q)$ with $p/q \in \mathscr{F}_n^*$, with equal probability 2^{-n}: the random walk in $\bar{\mathbf{C}}$ given by ν and starting from the point $(1,0)$ describes the rational numbers obtained by the Farey process, as well as the logarithm of their denominators, modulo $\log r$. In other words, let $\bar{\mathscr{F}}_n^* = \{(p/q, \log q) \mid p/q \in \mathscr{F}_n^*\} \subset [0,1] \times \mathbb{R}/(\log r)\mathbb{Z}$, then the measure $\bar{\mu}_n := \nu^n \star \delta_{(1,0)}$ is the average of the Dirac masses at the points of $\bar{\mathscr{F}}_n^*$. By general results on random walks on compact extensions, Conze and Guivarc'h [CG] and Guivarc'h and Raugi [GR] proved that $\bar{\mu}_n$ converges weakly to $\mu \otimes \mathrm{Leb}$, where Leb denotes the normalized Lebesgue measure on $\mathbb{R}/(\log r)\mathbb{Z}$. This is an equirepartition result of the denominators modulo $\log r$.

In this article, we are interested in more precise results for this random walk. First of all, we prove that the previous convergence is exponentially fast.

THEOREM 1.1
For any $\alpha > 0$, there exist $C > 0$ and $\theta < 1$ such that for any function $f : \bar{\mathbf{C}} \to \mathbb{C}$ which is Hölder-continuous of exponent α,

$$\left| \int f \,\mathrm{d}\bar{\mu}_n - \int f \,\mathrm{d}(\mu \otimes \mathrm{Leb}) \right| \leq C\theta^n \|f\|_{C^\alpha}. \tag{1.5}$$

We also obtain limit theorems for this random walk. In particular, we prove that it satisfies the local limit theorem. This answers a question raised by Guivarc'h and Raugi in [GR, Section 5.3]

THEOREM 1.2
Let $\psi : \bar{\mathbf{C}} \to \mathbb{R}$ be a C^6 function. Assume that there does not exist a continuous function $f : \bar{\mathbf{C}} \to \mathbb{R}$ such that $\psi \circ \bar{h}_M = f \circ \bar{h}_M - f$ for $M = A$ and B. Then the Markov chain X_n on $\bar{\mathbf{C}}$, starting from $(1,0)$ and whose transition probability is given by ν, satisfies a nondegenerate central limit theorem for the function ψ; that is, there exists $\sigma^2 > 0$ such that, for any $a \in \mathbb{R}$,

$$P\Big(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \psi(X_k) < a\Big) \to \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{a} e^{-t^2/(2\sigma^2)} \,\mathrm{d}t. \tag{1.6}$$

Assume additionally that there do not exist constants $a > 0$, $\lambda > 0$, and a continuous function $f : \bar{\mathbf{C}} \to \mathbb{R}/\lambda\mathbb{Z}$ such that $\psi \circ \bar{h}_M = f \circ \bar{h}_M - f + a \mod \lambda\mathbb{Z}$ for $M = A$ and B. Then ψ satisfies the local limit theorem: for any compact subinterval I of $\mathbb{R}$ and any real sequence k_n such that $k_n/\sqrt{n} \to \kappa \in \mathbb{R}$, then

$$\sqrt{n}\, P\Big(\sum_{k=1}^{n} \psi(X_k) \in I + k_n\Big) \to \mathrm{Leb}(I) \frac{e^{-\kappa^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}. \tag{1.7}$$

This result, as well as Theorem 1.1, in fact holds for any starting point of the random walk; there is nothing specific about $(1, 0)$. Note that aperiodicity conditions on ψ are clearly necessary to get the theorem. For $\kappa = 0$, the local limit theorem can be reformulated as follows. Consider a random walk on $\bar{\mathbf{C}} \times \mathbb{R}$ whose transition probability is $Q((x, \omega, z) \to (x', \omega', z')) = P((x, \omega) \to (x', \omega'))1_{z'=z+\psi(x,\omega)}$. The local limit theorem simply means that the measure $\sqrt{n}Q^n\delta_{(1,0,0)}$ converges weakly to an explicit multiple of the measure $\mu \otimes \mathrm{Leb}_{\mathbb{R}/(\log r)\mathbb{Z}} \otimes \mathrm{Leb}_{\mathbb{R}}$.

Let us now describe the relationship between the previous random walk and skew-products. Let T be the transformation on the interval $[0, 1]$ given by

$$T(x) = \frac{x}{1-x} \quad \text{if } x < \frac{1}{2}, \qquad T(x) = 2 - \frac{1}{x} \quad \text{if } x \geq \frac{1}{2}. \tag{1.8}$$

Then h_A and h_B are the inverse branches of the transformation T. The Markov operator corresponding to the random walk on $[0, 1]$ is therefore the adjoint (for the measure μ) of the composition by T (i.e., the transfer operator associated to T). The transformation T is topologically conjugate to the transformation $x \mapsto 2x$ on $[0, 1]$, and μ is simply the maximal entropy measure of T (i.e., the pullback of Lebesgue measure under this conjugacy). Note that T is not uniformly expanding, since it has neutral fixed points at zero and 1. We can then define a transformation $\mathscr{T}$ on $[0, 1] \times \mathbb{R}/(\log r)\mathbb{Z}$ whose inverse branches are $\bar{h}_A$ and $\bar{h}_B$ by

$$\mathscr{T}(x, \omega) = \big(Tx, \omega + \phi(x)\big), \tag{1.9}$$

where $\phi(x) = \log(1-x)$ if $x < 1/2$, and $\phi(x) = \log(x)$ if $x \geq 1/2$. By construction, the Markov operator corresponding to the random walk on $\bar{\mathbf{C}}$ is the transfer operator associated to $\mathscr{T}$ (for the measure $\mu \otimes \mathrm{Leb}$).

We can reformulate the previous theorems in the general setting of this article: we are going to study transformations of the form $(x, \omega) \mapsto (Tx, \omega + \phi(x))$, where T is a nonuniformly expanding transformation of a manifold X, and ϕ is a C^1 function from X to the circle $\mathbb{S}^1$. Hence, to integrate the study of Farey sequences in our general setting, it is important not to demand uniform expansion and to be able to deal with measures that are singular with respect to Lebesgue measure. These two constraints justify the forthcoming definitions, but they bring along a certain number of technical difficulties.

1.2. Definition of nonuniformly partially hyperbolic skew-products

Definition 1.3

Let Z be a Riemannian manifold, endowed with a finite measure ν. An open subset O of Z is said to have the *weak Federer property* (for the measure ν) if it satisfies the following property. We work on O, with the induced metric, and the geodesic distance it defines. For any $C > 1$, there exist $D = D(O, C) > 1$ and $\eta_0 = \eta_0(O, C) > 0$ such that, for any $\eta < \eta_0$, there exist disjoint balls $B(x_1, C\eta), \dots, B(x_k, C\eta)$ that are compactly included in O and sets $A_1, \dots, A_k$ contained, respectively, in $B(x_1, DC\eta), \dots, B(x_k, DC\eta)$ whose union covers a full measure subset of O and such that, for any $x'_i \in B(x_i, (C-1)\eta)$, we have $\nu(B(x'_i, \eta)) \geq \nu(A_i)/D$.

A family of open subsets $(O_n)_{n\in\mathbb{N}}$ of Z is said to uniformly have the weak Federer property (for the measure ν) if each set O_n has the weak Federer property and if, furthermore, for any $C > 1$, $\sup_{n\in\mathbb{N}} D(O_n, C) < \infty$.

This is a technical covering condition. It is a kind of weakening of the classical doubling condition having the following advantages. On the one hand, it is satisfied in many examples (particularly for Farey sequences, where the doubling condition does not hold). On the other hand, it is sufficient to carry out the forthcoming proofs (essentially, it is the technical condition that is required for Dolgopyat-type arguments to work). The main point of the definition is that D can be chosen independently of η: in some sense, the weak Federer property is a covering lemma with built-in uniformity.

The following definition describes the class of applications T to which the results of this article apply. It is large enough to contain the map (1.8), as we see later on.

Definition 1.4

Let T be a nonsingular transformation on a Riemannian compact manifold X (possibly with boundary), endowed with a Borel measure μ. Let Y be a connected open subset of X, with finite measure and finite diameter for the induced metric. We say that T is a *nonuniformly expanding transformation of base Y, with exponential tails and the uniform weak Federer property* if the following properties are satisfied.

(1) There exist a finite or countable partition (modulo 0) $(W_l)_{l\in\Lambda}$ of Y and times $(r_l)_{l\in\Lambda}$ such that, for all $l \in \Lambda$, the restriction of T^{r_l} to W_l is a diffeomorphism between W_l and Y, satisfying $\kappa \|v\| \leq \|DT^{r_l}(x)v\| \leq C_l \|v\|$ for any $x \in W_l$ and for any tangent vector v at x, for some constants $\kappa > 1$ (independent of l) and C_l. We denote by $T_Y : Y \to Y$ the map which is equal to T^{r_l} on each set W_l.

(2) Let $\mathscr{H} = \mathscr{H}_1$ denote the set of inverse branches of T_Y and, more generally, let $\mathscr{H}_n$ denote the set of inverse branches of T_Y^n. Let $J(x)$ be the inverse of the Jacobian of T_Y at x with respect to μ. We assume that there exists a constant $C > 0$ such that, for any inverse branch $h \in \mathscr{H}$, $\|D((\log J) \circ h)\| \leq C$.

(3) There exists a constant C such that, for any l, if $h_l : Y \to W_l$ denotes the corresponding inverse branch of T_Y, for any $k \leq r_l$, then $\left\|T^k \circ h_l\right\|_{C^1(Y)} \leq C$.

(4) Let $r : Y \to \mathbb{N}$ be the function that is equal to r_l on W_l. Then there exists $\sigma_0 > 0$ such that $\int_Y e^{\sigma_0 r} \,\mathrm{d}\mu < \infty$.

(5) Let μ_Y denote the probability measure induced by μ on Y. Then the sets $h(Y)$, for $h \in \bigcup_{n\in\mathbb{N}} \mathscr{H}_n$, uniformly have the weak Federer property (with respect to μ_Y).

In this article, we only consider transformations T of this type. Hence, we simply say that T is nonuniformly expanding with base Y.

The first four conditions above roughly mean that T is nonuniformly expanding and that an induced map T_Y (which is not necessarily a first-return map) is uniformly expanding and Markov, with exponential tails. This kind of assumption is described in [Y1], [Y2], and is often called a *Young tower structure* in the literature. The fifth condition is a covering condition. It is probably not very natural to require it uniformly over the inverse branches of the iterates of T_Y, but it is satisfied in all the examples that we consider here.

Under the first two assumptions, it is a folklore result that T_Y preserves a probability measure that is equivalent to μ_Y, whose density is C^1 and bounded away from zero and ∞. Without loss of generality, we may replace μ_Y by this measure (which does not change the assumptions), and we therefore always assume that μ_Y is invariant under T_Y (and has mass 1). Inducing from μ_Y (and using the fourth assumption) and then renormalizing, we obtain a probability measure $\tilde{\mu}$ on X which is invariant under T and ergodic. However, the restriction of $\tilde{\mu}$ to Y is, in general, not proportional to μ_Y, when the return times r_l are not first-return times.

The measure $\tilde{\mu}$ is always ergodic for T, but sometimes not for its iterates: in general, there exist a divisor d of $\gcd\{r_l \mid l \in \Lambda\}$ and open sets $(O_i)_{i\in\mathbb{Z}/d\mathbb{Z}}$ such that T maps O_i to O_{i+1}, and the restriction of T^d to each O_i is mixing. For the sake of simplicity, we only consider transformations T that are mixing (i.e., for which $d = 1$). However, the results we give have their counterparts in the general case, since they can be applied to T^d on each set O_i. Note that the mixing of T is equivalent to the ergodicity of all the iterates T^n and is implied by the equality $\gcd\{r_l\} = 1$.

Remark 1.5

Under the first four assumptions of Definition 1.4, and if T is mixing for the probability measure $\tilde{\mu}$, then T is exponentially mixing (for Hölder-continuous functions). This has been proved by Young in [Y1, Theorem 2] (in a slightly different setting) using a spectral gap argument and again in [Y2, Theorem 3] using coupling. We do not use these results of Young. Indeed, our arguments yield yet another proof of this exponential mixing, through operator renewal theory (see, in particular, Corollary 3.5). This proof is not new; it is already implicit in [S] and explicit in [G4, Remark 2.3.7].

In a similar setting (the study of expanding semiflows), Ruelle shows in [Ru] that a suspension over an expanding map cannot be exponentially mixing if the roof function is locally constant. Therefore, it is not surprising that this case should be excluded from our study, since, among other results, we prove exponential mixing.

Definition 1.6
Let T be a nonuniformly expanding transformation of base Y on a manifold X. Let $\phi : X \to \mathbb{R}$ be a C^1 function. Denote by ϕ_Y the induced function on Y given by $\phi_Y(x) = \sum_{i=0}^{r(x)-1} \phi(T^i x)$. We say that ϕ is *cohomologous to a locally constant function* if there exists a C^1 function $f : Y \to \mathbb{R}$ such that the function $\phi_Y - f + f \circ T_Y$ is constant on each set W_l, $l \in \Lambda$.

If ϕ is not cohomologous to a locally constant function, we define a map $\mathscr{T} : X \times \mathbb{S}^1 \to X \times \mathbb{S}^1$ by $\mathscr{T}(x, \omega) = (Tx, \omega + \phi(x))$. It preserves the probability measure $\tilde{\mu} \otimes \mathrm{Leb}$ (in this article, the Lebesgue measure on the circle $\mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$, denoted by Leb or $\mathrm{d}\omega$, is always normalized of mass 1). The transformation $\mathscr{T}$ is nonuniformly partially hyperbolic in the following sense: in each fiber $\mathbb{S}^1$, $\mathscr{T}$ is an isometry while it is expanding in the direction of X. Hence, we would like to talk of partial hyperbolicity. However, since the expansion of T is not uniform, T can have neutral fixed points or even critical points. Hence, there may exist points where the expansion in the X direction does not dominate what is happening in the fiber. Therefore, the partial hyperbolicity is asymptotic rather than instantaneous.

1.3. Limit theorems for nonuniformly partially hyperbolic skew-products
Let T be a nonuniformly expanding map with base Y, preserving the probability measure $\tilde{\mu}$, and mixing. Assume that μ_Y has full support in Y. Let $\phi : X \to \mathbb{R}$ be a C^1 function which is not cohomologous to a locally constant function. We consider the skew-product $\mathscr{T}(x, \omega) = (Tx, \omega + \phi(x))$.

THEOREM 1.7
For any $\alpha > 0$, there exist $\bar{\theta} < 1$ and $C > 0$ such that, for all functions f, g from $X \times \mathbb{S}^1$ to $\mathbb{C}$, respectively, bounded and Hölder-continuous with exponent α, and for all $n \in \mathbb{N}$, we have

$$\left| \int f \circ \mathscr{T}^n \cdot g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) - \Big(\int f \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \Big) \Big(\int g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \Big) \right| \\ \leq C \bar{\theta}^n \, \|f\|_{L^\infty} \, \|g\|_{C^\alpha} . \tag{1.10}$$

We then are interested in limit theorems for the transformation $\mathscr{T}$. Let $\psi : X\times\mathbb{S}^1 \to \mathbb{R}$ be a Hölder-continuous function, such that $\int \psi \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) = 0$. Let

$$\sigma^2 = \int \psi^2 \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) + 2 \sum_{k=1}^{\infty} \int \psi \cdot \psi \circ \mathscr{T}^k \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}). \tag{1.11}$$

This quantity is well defined by Theorem 1.7.

PROPOSITION 1.8
We have $\sigma^2 \geq 0$. Moreover, $\sigma^2 = 0$ if and only if there exists a measurable function $f : X \times \mathbb{S}^1 \to \mathbb{R}$ such that $\psi = f - f \circ \mathscr{T}$ almost everywhere. In this case, the function f has a version that is continuous on $Y \times \mathbb{S}^1$, and it belongs to $L^p(X \times \mathbb{S}^1)$ for all $p < \infty$.

Let us denote by $S_n\psi$ the Birkhoff sums $\sum_{i=0}^{n-1} \psi \circ \mathscr{T}^i$. When σ^2 is nonzero (i.e., ψ is not a coboundary), then ψ satisfies the central limit theorem, as follows.

THEOREM 1.9
Let ψ be a Hölder-continuous function on $X \times \mathbb{S}^1$ with zero average, such that $\sigma^2 > 0$. Then $S_n\psi/\sqrt{n}$ satisfies the central limit theorem; that is, $S_n\psi/\sqrt{n}$ converges in distribution (for the probability measure $\tilde{\mu}\otimes\mathrm{Leb}$) towards the Gaussian distribution $\mathscr{N}(0, \sigma^2)$.

Let us say that ψ is *periodic* if there exist $a > 0$, $\lambda > 0$, and $f : X \times \mathbb{S}^1 \to \mathbb{R}/\lambda\mathbb{Z}$ measurable such that $\psi = f - f \circ \mathscr{T} + a \mod \lambda$ almost everywhere. Otherwise, we say that ψ is *aperiodic* (this implies, in particular, that ψ is not a coboundary and hence that $\sigma^2 > 0$).

PROPOSITION 1.10
If ψ is a periodic C^6 function, there exist $a > 0$, $\lambda > 0$, and $f : X \times \mathbb{S}^1 \to \mathbb{R}/\lambda\mathbb{Z}$ measurable such that $\psi = f - f \circ \mathscr{T} + a \mod \lambda$ almost everywhere, and f is continuous on $Y \times \mathbb{S}^1$.

This proposition makes it possible to check in practice whether a function ψ is periodic or not using periodic points: let x be a point in Y, fixed under an iterate $\mathscr{T}^n$ of $\mathscr{T}$ such that $\mathscr{T}$ is continuous on a neighborhood of the orbit of x. If $\psi = f - f \circ \mathscr{T} + a \mod \lambda$ almost everywhere, with f continuous on Y, then f is continuous (modulo λ) along the orbit of x, and we obtain $\psi = f - f \circ \mathscr{T} + a \mod \lambda$ along this orbit. In particular, $S_n\psi(x) = an \mod \lambda$; this restricts the possible values of a and λ. Using three periodic orbits of $\mathscr{T}$, we may reach a contradiction, showing that no values of a and λ are possible, and proving that ψ is aperiodic.

The notion of periodicity is interesting, since it gives the only obstruction to the local limit theorem as follows.

THEOREM 1.11
Let ψ be a C^6 function on $X \times \mathbb{S}^1$ with vanishing average, aperiodic (which implies $\sigma^2 > 0$). Then the Birkhoff sums $S_n\psi$ satisfy the local limit theorem in the following sense: for any compact interval I and any real sequence k_n such that $k_n/\sqrt{n} \to \kappa \in \mathbb{R}$, we have, when $n \to \infty$, the following:

$$\sqrt{n}\,(\tilde{\mu} \otimes \mathrm{Leb})\big\{(x,\omega) \in X \times \mathbb{S}^1 \;\big|\; S_n\psi(x,\omega) \in I + k_n\big\} \to \mathrm{Leb}(I)\frac{e^{-\kappa^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}}. \quad (1.12)$$

We also obtain numerous other limit theorems (such as the Berry-Esseen theorem on the speed of $1/\sqrt{n}$ in the central limit theorem, the renewal theorem, and so on). Instead of giving precise statements, we give the key estimate that implies all of them by showing that the Birkhoff sums $S_n\psi$ essentially behave like a sum of independent identically distributed random variables.

THEOREM 1.12
Let ψ be a C^6 function with zero average, such that $\sigma^2 > 0$. There exist $\tau_0 > 0$, $C > 0$, $c > 0$, and $\bar{\theta} < 1$ such that, for all functions f, g from $X \times \mathbb{S}^1$ to $\mathbb{C}$, respectively, bounded and C^6, for any $n \in \mathbb{N}$, and for any $t \in [-\tau_0, \tau_0]$, we have

$$\begin{aligned} \Big| \int e^{itS_n\psi} \cdot f \circ \mathscr{T}^n \cdot g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \\ - \Big(1 - \frac{\sigma^2 t^2}{2}\Big)^n \Big(\int f \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\Big)\Big(\int g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\Big)\Big| \\ \leq C\big(\bar{\theta}^n + |t|(1-ct^2)^n\big) \,\|f\|_{L^\infty} \,\|g\|_{C^6}. \end{aligned} \quad (1.13)$$

Moreover, if ψ is aperiodic, for all $t_0 > \tau_0$ there exist $C > 0$ and $\bar{\theta} < 1$ such that, for all $|t| \in [\tau_0, t_0]$, we have

$$\Big| \int e^{itS_n\psi} \cdot f \circ \mathscr{T}^n \cdot g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\Big| \leq C\bar{\theta}^n \,\|f\|_{L^\infty} \,\|g\|_{C^6}. \quad (1.14)$$

Taking $f = g = 1$, we obtain that the characteristic function of $e^{itS_n\psi}$ essentially behaves like $(1 - \sigma^2 t^2/2)^n$, which makes it possible to prove Theorem 1.9 for C^6 functions and to prove Theorem 1.11, as well as numerous limit theorems, by mimicking the classical methods in probability theory for sums of independent identically distributed random variables. It should just be checked that the additional error term

$\bar{\theta}^n + |t|(1 - ct^2)^n$ does not spoil the arguments. This has already been done in [G2, pages 30–33]. We do not give further details on these classical arguments in the following.

Note that, taking $t = 0$, Theorem 1.12 implies Theorem 1.7 (for $\alpha = 6$, but this easily implies the general case by a regularization argument). However, the proof of Theorem 1.7 is considerably easier than the proof of Theorem 1.12; hence, we give proof of Theorem 1.7 with full details, which also allows us to introduce, in a simple setting, some tools used later in more sophisticated versions.

Remark 1.13
Propositions 1.8 and 1.10 give automatic regularity for solutions of the cohomological equation, with a loss of regularity (arbitrarily small in Proposition 1.8, of 6 derivatives in Proposition 1.10). The loss of 6 derivatives is probably not optimal but, with the method of proof we use, some loss seems to be unavoidable.

In general, the continuity of f on $Y \times \mathbb{S}^1$ cannot be extended to a continuity on the whole space (e.g., think of a map T with discontinuities). Nevertheless, using the specificities of T, it is often possible to obtain the continuity of f on larger sets.

Remark 1.14
Theorem 1.9 is first proved for C^6 functions by using Theorem 1.12 and then is extended to Hölder-continuous functions by an approximation argument. This argument does not apply for the local limit theorem, which explains our stronger regularity assumption in Theorem 1.11.

Remark 1.15
We require that μ_Y have full support in Y. For some interesting maps (e.g., maps on Cantor sets; see [N]), this condition is not satisfied. The full support condition is used only to get Dolgopyat-like contraction in the proof of Lemma A.8, and it can be dispensed with under a stronger condition on ϕ. Indeed, if there exist two sequences $h_1, h_2, \dots$ and $h'_1, h'_2, \dots$ of elements of $\mathscr{H}$ and a point x in the support of μ_Y such that the series $\sum_{n=1}^{\infty} D(\phi_Y \circ h_n \cdots h_1)(x)$ and $\sum_{n=1}^{\infty} D(\phi_Y \circ h'_n \cdots h'_1)(x)$ converge and are not equal, then the proof of this lemma goes through (note that this condition is very similar to the nonlocal integrability property (NLI) in [N]). When μ_Y has full support, this condition is equivalent to ϕ not being cohomologous to a locally constant function, as shown in the proof of Lemma A.8.

1.4. Examples
In the examples, if T and ϕ are given, and one wants to apply the previous results, then one should first check that T is nonuniformly expanding of base Y for some Y and

then prove that ϕ is not cohomologous to a locally constant function. The first issue depends strongly on the map T (see the following list of examples), but the second one is generally easy to check by using periodic orbits, as follows.

Assume—this is the case in all our examples—that every inverse branch $h \in \mathscr{H}$ of T_Y has a unique fixed point x_h. Let f be a C^1 function on Y. If $\phi_Y - f + f \circ T_Y$ is constant on each set $h(Y)$, it has to be equal to $\phi_Y(x_h)$ there. Consequently, the function g, equal to $\phi_Y - \phi_Y(x_h)$ on each set $h(Y)$, is cohomologous to zero. In particular, if one can find a periodic orbit of T_Y along which the Birkhoff sum of g is nonzero, then this is a contradiction, and ϕ cannot be cohomologous to a locally constant function. This can easily be checked in practice: for example, we use this argument in the specific case of Farey sequences.

If $1 \leq k \leq \infty$, the previous argument shows, moreover, that in the space of C^k functions on X, the set of functions ϕ which are cohomologous to a locally constant function is contained in a closed vector subspace of infinite codimension. Hence, the theorems of Section 1.3 can be applied for most (in a very strong sense) functions ϕ.

Let us now describe different classes of maps T which satisfy Definition 1.4.

Nonuniformly expanding maps and Lebesgue measure

Let T be a C^2 map on a compact Riemannian manifold X (possibly with boundary). We assume that T is nonuniformly expanding in the following sense (see [ABV], [ALP], [G3]). Let S be a closed subset of X with zero Lebesgue measure (corresponding to the singularities of T), possibly empty, and containing the boundary of X. We assume that T is a local diffeomorphism on $X - S$, nondegenerate close to S: there exist $B > 1$ and $\beta > 0$ such that, for any $x \in X - S$ and any nonzero tangent vector v at x,

$$\frac{1}{B} d(x, S)^{\beta} \leq \frac{\|DT(x)v\|}{\|v\|} \leq B d(x, S)^{-\beta}. \tag{1.15}$$

Assume also that, for any $x, y \in X$ with $d(x, y) < d(x, S)/2$,

$$\left|\log \|DT(x)^{-1}\| - \log \|DT(y)^{-1}\|\right| \leq B \frac{d(x, y)}{d(x, S)^{\beta}} \tag{1.16}$$

and

$$\left|\log |\det DT(x)^{-1}| - \log |\det DT(y)^{-1}|\right| \leq B \frac{d(x, y)}{d(x, S)^{\beta}}. \tag{1.17}$$

For $\delta > 0$, let $d_\delta(x, S) = d(x, S)$ if $d(x, S) < \delta$, and let $d_\delta(x, S) = 1$ otherwise. Let $\delta : (0, \varepsilon_0) \to \mathbb{R}_+$ be a positive function, and let $\kappa > 0$. Assume that, for any $\varepsilon < \varepsilon_0$, there exist $C > 0$ and $\theta < 1$ such that, for any $N \in \mathbb{N}$,

$$\mathrm{Leb}\left\{x \in X \;\middle|\; \exists n \geq N, \frac{1}{n}\sum_{k=0}^{n-1} \log \left\| DT(T^k x)^{-1} \right\|^{-1} < \kappa\right.$$

or

$$\left.\frac{1}{n}\sum_{k=0}^{n-1} - \log d_{\delta(\varepsilon)}(T^k x, S) > \varepsilon\right\} \leq C\theta^N.$$

This assumption means that the points that do not see the expansion or are too close to the singularities after time N have an exponentially small measure.

As examples of such applications, let us first mention uniformly expanding maps, of course, but also multimodal maps with infinitely many branches (see [AP]), which have thereby infinitely many critical points, as well as small perturbations of uniformly expanding maps (such perturbations can have saddle fixed points; see [Al, Section 6]).

PROPOSITION 1.16
Under these assumptions, there exists a subset Y of X such that T is nonuniformly expanding of base Y for Lebesgue measure.

Proof
This theorem is essentially proved in [G3, Theorem 4.1]. More precisely, this theorem constructs a subset Y of X and a partition of Y such that the first four properties of Definition 1.4 are satisfied. The set Y is an open set with piecewise C^1 boundary, and each inverse branch h can be extended to a neighborhood of Y.

If the boundary of Y were C^1 (and not merely piecewise C^1), each set $h(Y)$ would also be an open set with C^1 boundary, and the uniform weak Federer property would directly result from the good doubling properties of Lebesgue measure. However, if the boundary of Y is only piecewise C^1, the images of the boundary components by an inverse branch h could meet with smaller and smaller angles, which could prevent the uniform weak Federer property from holding.

Therefore, we have to modify slightly the construction in [G3] to obtain a set Y with C^1 boundary. In that article, one starts from a partition U_i of X (into sets with piecewise C^1 boundary), and one subdivides each set U_i into subsets V_j that are sent by some iterate of T on one of the sets U_k. The set Y is then one of the U_i's, and the desired partition of Y is obtained by inducing from the V_j's (see [G3, Section 4] for details).

To obtain a smooth Y, we also start from a partition U_i, but we decompose U_i as $U_i^1 \cup U_i^2$, where U_i^1 is a ball inside U_i and U_i^2 is its complement. Applying the construction of [G3] separately to each set U_i^1 and U_i^2, we subdivide them into sets

V_j that are sent by some iterate of T to some U_k. We finish the construction by taking for Y one of the sets U_i^1 and inducing on it. □

To apply the results of Section 1.3, one needs an additional mixing assumption, which is satisfied as soon as all the iterates of T are topologically transitive on the attractor $\bigcap_{n\geq 0} T^n(X)$ (see [G3]).

Multimodal maps of Collet-Eckmann type

Let T be a multimodal map on a compact interval I. If the derivative of T^n along the postcritical orbits grow exponentially fast, and T is not renormalizable (which prevents periodicity problems), [BLV] shows that there exists a unique absolutely continuous invariant probability measure $\tilde{\mu}$, and that T is exponentially mixing for this measure.

To prove this result, the authors show that there exist an interval Y and a subpartition W_l of Y satisfying the first four properties of Definition 1.4 for Lebesgue measure. Since the sets $h(Y)$ (for $h \in \bigcup_{n\in\mathbb{N}} \mathscr{H}_n$) are all intervals, the uniform weak Federer property is also trivially satisfied by Lebesgue measure.

Gibbs measures in dimension 1

If T is a C^2 uniformly expanding map on a compact connected manifold X, and if $u : X \to \mathbb{R}$ is a C^1 function, there exists a unique invariant probability measure μ that maximizes the quantity $h_\nu(T) + \int u \, \mathrm{d}\nu$ over all invariant probability measures ν. This is the so-called *Gibbs measure* associated to the potential u.

In general, it is unlikely that such a Gibbs measure satisfies the weak Federer property (unless μ is equivalent to Lebesgue measure, which corresponds to potentials u which are cohomologous to $-\log\det(DT)$). Indeed, the proof of the weak Federer property in the previous examples relies in an essential way on the good doubling properties of Lebesgue measure.

However, in dimension 1 (i.e., if T is a circle map), the iterates of T are conformal, which implies that μ satisfies the weak Federer property, and our results apply. Proofs of the Federer property in this setting have been given by Dolgopyat and by Pollicott, but with small imprecisions, so we give a full proof in Proposition 6.2 (as a very simple consequence of the methods we develop to treat the Farey sequence). Note that the same results also apply in higher dimension for conformal uniformly expanding maps (since uniformly expanding maps always admit Markov partitions).

Farey sequences

The results of Section 1.3 also apply to the map (1.9), which generates the Farey sequence. However, the proof requires more work, since checking the weak Federer property is not trivial. Moreover, the most interesting results stated in Theorem 1.2

are pointwise results (for a random walk starting from $(1, 0)$), while the statements of Section 1.3 are on average results. To prove the pointwise statements, we therefore need to use more technical results, established during the course of the proof of Theorems 1.7 and 1.12. As a consequence, the results of Section 1.1 are proved at the end of the article in Section 6.

1.5. Method of proof and contents of the article

In general, to prove exponential mixing and a local limit theorem, it is very comfortable to have a *spectral gap property* for a transfer operator (the spectral perturbation methods then yield the desired results quite automatically). The spectral gap is in general a consequence of some expansion or contraction properties. However, in our setting, the map $\mathscr{T}$ is an isometry in the fibers, and a spectral gap seems therefore difficult to obtain. Note that [Ts] manages to construct a space with a spectral gap for such maps, but under strong assumptions: the map T should be uniformly expanding, and $\tilde{\mu}$ should be absolutely continuous with respect to Lebesgue measure. These properties are unfortunately not satisfied in our setting, and we thus have to work without a spectral gap (on the space $X \times \mathbb{S}^1$).

In [D1], [D2], Dolgopyat developed techniques that he used to prove the exponential decay of correlations for maps $\mathscr{T}$, as above, if T is uniformly expanding. His main idea is to work in Fourier coordinates, to see that each frequency is left invariant by the transfer operator associated to $\mathscr{T}$ and to obtain explicit bounds on the mixing speed in each frequency (by using oscillatory integrals, which give explicit compensations). The gain is not uniform with respect to the frequency (which accounts for the lack of spectral gap), but the estimates are nevertheless sufficiently good to obtain exponential mixing.

In an essential way, we use Dolgopyat's ideas in this article as a technical tool. This tool applies to uniformly expanding maps, which is not the case of our map T; we therefore need to induce on the set Y to get uniform expansion. To obtain information on the initial map, we then make use of (elementary) ideas of generating series and renewal theory.

The real difficulty of the article lies in the local limit theorem, since a spectral gap property seems more or less necessary to any known proof of the local limit theorem, while Dolgopyat's arguments do not give such a spectral gap. If we try to work at the level of frequencies, just like for the exponential mixing, we quickly run into the following additional difficulty: if f is a function of frequency k (i.e., $f(x, \omega) = u(x)e^{ik\omega}$), then $e^{it\psi} f$ is no longer a function of frequency k. In other words, the multiplication by $e^{it\psi}$—which is at the heart of the proof of the local limit theorem for the function ψ—mixes the different frequencies together. Hence, even though Dolgopyat's techniques give a good control at high frequencies, this control is instantaneously ruined by the multiplication by $e^{it\psi}$, which can go back into low frequencies, where no control is available.

The central idea for the proof of the local limit theorem is to induce at the same time in x and in k: we consider some kind of random walk on the space $X \times \mathbb{Z}$ (where the $\mathbb{Z}$ factor corresponds to the space of frequencies), and we induce on a subset $Y \times [-K, K]$, where K is large enough so that what happens outside of this set can be controlled by Dolgopyat's tools. The main interest of this process is that the induced operator on $Y \times [-K, K]$ has a spectral gap and can be studied very precisely. Using techniques of operator renewal theory in [S], [G2], we then use this information to obtain a global control on $X \times \mathbb{Z}$, finally yielding Theorem 1.12.

Remark 1.17

The next natural question is to study maps of the form $\mathscr{T}' : (x, \omega, \omega') \mapsto (Tx, \omega + \phi(x), \omega' + \psi(x, \omega))$, where T and ϕ are as above. If ψ is aperiodic, then Theorem 1.12 shows that the correlations of functions of the form $u(x, \omega)e^{ik\omega'}$ (where u is C^6 and $k \in \mathbb{Z}$) tend to zero. Since the linear combinations of such functions are dense in L^2, this implies that $\mathscr{T}'$ is mixing. It is even Bernoulli, by the following argument. First, T (or rather its natural extension) is Bernoulli since it is mixing and nonuniformly hyperbolic (see, e.g., [OW]). Since $\mathscr{T}$ is a mixing isometric extension of T, it is also Bernoulli by [R]. The same argument applied to $\mathscr{T}$ then implies that $\mathscr{T}'$ is Bernoulli.

However, to prove further results on $\mathscr{T}'$ such as exponential mixing or the local limit theorem (probably under stronger assumptions on ψ) seems out of reach by current techniques. More precisely, we use Dolgopyat's techniques (which give precise explicit estimates for the map T) to study the map $\mathscr{T}$ (and obtain, by an abstract compactness argument, nonexplicit estimates for $\mathscr{T}$). To go one step further and study precisely $\mathscr{T}'$, we would need *explicit* estimates for $\mathscr{T}$ (i.e., in (1.14), we would need to control $\bar{\theta}$ and C in terms of t_0), which seems considerably more difficult.

The article is organized as follows. In Section 2, we state a theorem on transfer operators giving all the technical estimates we need further on (with contraction in the classical sense or in Dolgopyat norms). This technical theorem is proved in the appendix. In Section 3, it is used to prove Theorem 1.7. The proof is a baby version of the proof of the local limit theorem, introducing some tools on renewal operators that are used further on. In Section 4, we describe in detail the strategy of the proof of the local limit theorem and give two technical results essential to its proof. The proof itself is given in Section 5. Finally, Section 6 is devoted to the proof of the results on Farey sequences as stated in Section 1.1.

In the remainder of this article, we fix once and for all a map T that is nonuniformly expanding of base Y and mixing, together with a function ϕ that is not cohomologous to a locally constant function.

2. Tools on transfer operators

Let us first give two general tools on transfer operators. If T_0 is a probability-preserving map, its transfer operator $\hat{T}_0$ is defined by duality by $\int \hat{T}_0 u \cdot v = \int u \cdot v \circ T_0$. It acts on integrable functions. It turns out that the spectral properties of the transfer operator (acting on a suitable Banach subspace of L^1) give a very powerful tool to study the dynamical properties of T_0. The following theorem of Hennion is especially important for us.

THEOREM 2.1

Let $\mathscr{M}$ be a linear operator acting continuously on a Banach space $\mathscr{B}$ endowed with a norm $\|\cdot\|$. Assume that there exists a seminorm $|\cdot|$ on $\mathscr{B}$ such that the unit ball of $\|\cdot\|$ is relatively compact for $|\cdot|$. Assume also that there exist $n > 0$, $C > 0$, and $\sigma > 0$ such that, for any $x \in \mathscr{B}$,

$$\|\mathscr{M}^n x\| \leq \sigma^n \|x\| + C|x|. \tag{2.1}$$

Then the essential spectral radius of $\mathscr{M}$ is at most σ; that is, for any $r > \sigma$, the intersection of the spectrum of $\mathscr{M}$ with $\{|z| > r\}$ consists in finitely many eigenvalues of finite multiplicity.

Proof

Hennion proves this theorem in [H, Corollary 1], assuming additionally that $\mathscr{M}$ is also continuous for the seminorm $|\cdot|$; [BGK, Lemma 2.2] shows that this assumption can be dispensed with. □

We also need the following general lemma on transfer operators.

LEMMA 2.2

Let T_0 be an ergodic transformation of a probability space, with corresponding transfer operator $\hat{T}_0$. Let g be a nonzero integrable function, let f be a measurable function with modulus at most 1, and let $\lambda \in \mathbb{C}$ with $|\lambda| \geq 1$. We assume that $\lambda g = \hat{T}_0(fg)$. Then $|\lambda| = 1$, $|f| = 1$ almost everywhere, and $\lambda g \circ T = fg$ almost everywhere.

Proof

We have $|\lambda||g| \leq \hat{T}_0|g|$. Integrating this equation yields $|\lambda| \, \|g\|_{L^1} \leq \|g\|_{L^1}$, which implies $|\lambda| = 1$. Moreover, the function $\hat{T}_0|g| - |g|$ is nonnegative and has zero integral, hence it vanishes almost everywhere. Since $\hat{T}_0|g| = |g|$, the measure with density $|g|$ is invariant. By ergodicity, $|g|$ is almost everywhere constant (and this constant is nonzero). The equation $\lambda g = \hat{T}_0(fg)$ becomes $\hat{T}_0(\lambda^{-1} fg/g \circ T_0) = 1$.

Therefore,

$$1 = \int \lambda^{-1} f \frac{g}{g \circ T_0} \leq \int \left| \lambda^{-1} f \frac{g}{g \circ T_0} \right| \leq 1. \tag{2.2}$$

This shows that the function $\lambda^{-1} f \frac{g}{g\circ T}$ has to be equal to 1 almost everywhere. □

We now turn to the specific situation of skew-products. For $k \in \mathbb{Z}$ and $v \in C^1(Y)$, we define an operator

$$\mathscr{L}_k v(x) = \sum_{h \in \mathscr{H}} e^{-ik\phi_Y(hx)} J(hx) v(hx). \tag{2.3}$$

These operators are the building blocks of the transfer operator associated to $\mathscr{T}$. In particular, let $\mathscr{L} = \mathscr{L}_0$ (this is the transfer operator associated to T_Y). For $x \in Y$ and $n \in \mathbb{N}$, let us also write $S_n^Y \phi_Y(x) = \sum_{i=0}^{n-1} \phi_Y(T_Y^i x)$.

For $n \in \mathbb{N}$ and $x \in Y$, let $r^{(n)}(x) = \sum_{i=0}^{n-1} r(T_Y^i x)$. This is simply the Birkhoff sum $S_n^Y r(x)$, but we favor the more compact notation $r^{(n)}$ since we think of $r^{(n)}$ as a return time for the iterated map T_Y^n, and not as a Birkhoff sum.

For $n \in \mathbb{N}$, $A > 0$, and $\varepsilon > 0$, we denote by $\mathscr{C}_n^{A,\varepsilon}$ the set of functions v from Y to $\mathbb{C}$ which are C^1 on each set $h(Y)$ for $h \in \mathscr{H}_n$, and such that the quantity

$$\|v\|_{\mathscr{C}_n^{A,\varepsilon}} = \sup_{h \in \mathscr{H}_n} \sup_{x \in Y} \max(|v(hx)|, \|D(v \circ h)(x)\| / A)/e^{\varepsilon r^{(n)}(hx)} \tag{2.4}$$

is finite. These are the functions we work with. They can be unbounded, but their explosion speed is controlled by the return time. Typically, if one starts from a smooth function on X and induces, then the resulting function is unbounded, but it belongs to $\mathscr{C}_1^{A,\varepsilon}$ for any A, ε. In particular, for any $A > 0$ and any $\varepsilon > 0$, we have $\sup_{n\in\mathbb{N}} \left\| S_n^Y \phi_Y \right\|_{\mathscr{C}_n^{A,\varepsilon}} < \infty$. Note that the set of functions $\mathscr{C}_n^{A,\varepsilon}$ does not depend on A, but the corresponding norm does.

Let $k \in \mathbb{Z}$, and let $C_0 > 1$. We denote by $\mathscr{E}_k(C_0)$ the set of pairs (u, v) of functions from Y to $\mathbb{C}$ such that u takes only nonnegative values, $|v| \leq u$, and $\max(\|Dv\|, \|Du\|) \leq C_0 \max(1, |k|) u$. This set is a cone (i.e., it is stable under addition and multiplication by nonnegative real numbers). We also write $\|v\|_{D_k(C_0)}$ (or simply $\|v\|_{D_k}$) for the infimum of the quantities $\|u\|_{L^4}$ over all functions u such that $(u, v) \in \mathscr{E}_k(C_0)$. Since $\mathscr{E}_k(C_0)$ is a cone, this is a norm, satisfying $\|v\|_{L^4} \leq \|v\|_{D_k} \leq \|v\|_{C^1}$. The D_k norm has been (implicitly) used by Dolgopyat, and it is very useful since it enjoys good contraction properties for the action of the transfer operator $\mathscr{L}_k$.

We freely use the following trivial inequalities: if $|k| \leq |\ell|$, then $\|v\|_{D_\ell} \leq \|v\|_{D_k}$. Moreover, for any k, we have $\|v\|_{D_k} \leq \|v\|_{C^1}$. Finally, we have $\|v\|_{\mathscr{C}_n^{A,\varepsilon'}} \leq \|v\|_{\mathscr{C}_n^{A,\varepsilon}}$ as soon as $\varepsilon' \geq \varepsilon$.

The theorem we use is the following. Recall that T is a fixed nonuniformly expanding transformation of base Y and that ϕ is a C^1 function that is not cohomologous to a locally constant function, also fixed once and for all.

THEOREM 2.3

There exist $N > 0$, $C_0 > 1$, $\varepsilon > 0$, and $\theta \in (2^{-1/(1010N)}, 1)$ such that, for any $M \geq 1$, the following properties hold.

(Classical contraction) *For any $A \geq 1$, there exists a constant $C(A)$ such that, for any $\psi \in \mathscr{C}_{MN}^{A,4\varepsilon}$ and for any $v \in C^1(Y)$,*

$$\|\mathscr{L}^{MN}(\psi v)\|_{C^1} \leq \theta^{100MN}\Big(\sup_{x\in Y}|\psi(x)|/e^{4\varepsilon r^{(MN)}(x)}\Big)\,\|v\|_{C^1} + C(A)\,\|\psi\|_{\mathscr{C}_{MN}^{A,4\varepsilon}}\,\|v\|_{C^0}\,. \tag{2.5}$$

Moreover, there exists $C > 0$ satisfying the following property. Let $A \geq 1$, let $\psi_1, \ldots, \psi_n \in \mathscr{C}_{MN}^{A,4\varepsilon}$, and let $v \in C^1(Y)$. Write $v^0 = v$, and write $v^i = \mathscr{L}^{MN}(\psi_i v^{i-1})$. Then

$$\|v^n\|_{C^1} \leq CA\Big(\prod_{i=1}^{n}\|\psi_i\|_{\mathscr{C}_{MN}^{A,4\varepsilon}}\Big)\Big(\theta^{100MNn}\,\|v\|_{C^1} + \theta^{-MNn}\,\|v\|_{L^2}\Big). \tag{2.6}$$

(Dolgopyat's contraction) *For any $A \geq 1$, there exists $K = K(A, M)$ such that, for any $|k| \geq K$, for any C^1 function $v : Y \to \mathbb{C}$, and for any function $\psi \in \mathscr{C}_{MN}^{A,4\varepsilon}$,*

$$\|\mathscr{L}_k^{MN}(\psi v)\|_{D_k} \leq \theta^{100MN}\,\|\psi\|_{\mathscr{C}_{MN}^{A,4\varepsilon}}\,\|v\|_{D_{2^M k}}\,. \tag{2.7}$$

Moreover, for any $|\ell| \geq |k| \geq K$, we also have

$$\|\mathscr{L}_k^{MN}(\psi v)\|_{D_\ell} \leq \theta^{-MN}\,\|\psi\|_{\mathscr{C}_{MN}^{A,4\varepsilon}}\,\|v\|_{D_{2^M \ell}}\,. \tag{2.8}$$

The first half of the theorem is really classical (it is a consequence of the usual contraction of transfer operators on spaces of Lipschitz or C^1 functions), the second half is less classical but should not be surprising to a reader familiar with Dolgopyat's techniques. However, this result contains additional technical difficulties with respect to the same kind of results in the literature. Indeed, the functions in $\mathscr{C}_{MN}^{A,\varepsilon}$ are usually unbounded and have unbounded derivatives. Moreover, the application of Dolgopyat's arguments is problematic since the function ϕ_Y is also unbounded with unbounded derivative. As a consequence, the proof of this theorem is quite unpleasant, even though it does not need additional conceptual ideas, only technical ones. Therefore, the proof of Theorem 2.3 is postponed to the appendix.

In the rest of this article (except the appendix), N, C_0, ε, and θ are fixed once and for all and denote the constants given by Theorem 2.3.

Remark 2.4

Note that the bounds with $\|\psi\|_{\mathscr{C}^{A,4\varepsilon}_{MN}}$ imply the same bounds with $\|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}}$. Most of the time, we only need this weaker version (the inequalities with 4ε simply give a small additional margin, which is useful from time to time). In some applications of Theorem 2.3, the function ψ does not play any role (i.e., the theorem with $\psi = 1$ would suffice). However, to study the local limit theorem, the relevant operators incorporate the iteration of the map *and* the multiplication by a characteristic function. The presence of the function ψ in the above statements makes it possible to study such an operator using Theorem 2.3.

Remark 2.5

Concerning the precise formulation of Theorem 2.3, let us make two additional remarks that are apparently technical but are in fact extremely important for the forthcoming proofs.

(1) The theorem for $M = 1$ is sufficient to obtain the exponential mixing (and to prove the theorem for $M = 1$, we only need the weak Federer property of Y and no uniformity on the inverse branches). However, to prove the local limit theorem, we need to take larger and larger M's; since θ is independent of M, the gain θ^{100MN} enables us to control some terms that are polynomially growing with M. The uniformity in M in Theorem 2.3 is therefore crucial.

(2) Since $\|v\|_{D_{2M_k}} \leq \|v\|_{D_k}$, the inequality (2.7) is stronger than

$$\|\mathscr{L}_k^{MN}(\psi v)\|_{D_k} \leq \theta^{100MN} \|\psi\|_{\mathscr{C}^{A,4\varepsilon}_{MN}} \|v\|_{D_k}. \tag{2.9}$$

The inequality (2.9) would be sufficient to prove the exponential mixing. However, to prove the local limit theorem, we jump from one frequency to another, and the additional gain in the index given by (2.7) is crucial (especially in the proof of Lemma 4.3).

3. Exponential mixing

3.1. A model for $\mathscr{T}$

For $n \in \mathbb{N}$, we define an artificial transformation, which models the dynamics of $\mathscr{T}$, as follows. Let $X^{(n)} = \{(x,i) \mid x \in Y, i < r^{(n)}(x)\}$, define a map $U^{(n)}$ (or simply U if n is implicit) on $X^{(n)}$ by $U(x,i) = (x, i+1)$ if $i + 1 < r^{(n)}(x)$, and let $U(x, r^{(n)}(x) - 1) = (T_Y^n(x), 0)$. Let $\pi^{(n)} : X^{(n)} \to X$ be given by $\pi^{(n)}(x,i) = T^i(x)$; we obtain $\pi^{(n)} \circ U = T \circ \pi^{(n)}$. We endow each set $h(Y) \times \{i\}$, for $h \in \mathscr{H}_n$ and $i < r^{(n)} \circ h$, with the restriction of the measure μ_Y to $h(Y)$. This yields a measure $\mu^{(n)}$ that is invariant under U and whose restriction to $Y \times \{0\}$ is equal to μ_Y. Strictly speaking, the map U is not defined everywhere since some points of Y do not come back to Y. However, it is defined $\mu^{(n)}$ almost everywhere, which is sufficient for our needs. The measure $\pi_*^{(n)}\mu^{(n)}$ is absolutely continuous with respect to $\tilde\mu$ and invariant; hence, these measures are proportional by ergodicity. In particular, setting $\tilde\mu^{(n)} = \mu^{(n)}/\mu^{(n)}(X^{(n)})$, we have $\pi_*^{(n)}\tilde\mu^{(n)} = \tilde\mu$.

We also endow $X^{(n)}$ with a metric, as follows. The set Y is canonically embedded in $X^{(n)}$ by $y \mapsto (y, 0)$; we endow the image of this embedding by the metric of Y. Let $h \in \mathscr{H}_n$, and let $0 < i < r^{(n)} \circ h$ (this function is constant on Y). The map $U^{r^{(n)}\circ h - i}$ is a bijection between $h(Y) \times \{i\}$ and $Y \times \{0\}$; we choose the metric on $h(Y) \times \{i\}$ so that this map is an isometry.

With this choice of the metric, the map U is very expanding on the points of the form $(y, 0)$ (it expands the metric by at least κ^n), and it is a local isometry on the points (y, i) with $i > 0$. Since T satisfies the third property of Definition 1.4, the map $\pi^{(n)}$ is almost a contraction for each n: there exists a constant C such that

$$\|D\pi^{(n)}(x) \cdot v\| \leq C\,\|v\| \tag{3.1}$$

for any $x \in X^{(n)}$ and v tangent at x. If $u : X \to \mathbb{C}$ is a C^1 function, the function $u \circ \pi^{(n)}$ is then also C^1 on $X^{(n)}$, and $\|u \circ \pi^{(n)}\|_{C^1} \leq C\,\|u\|_{C^1}$.

We finally define a map $\mathscr{U} = \mathscr{U}^{(n)}$ on $X^{(n)} \times \mathbb{S}^1$ by $\mathscr{U}(x, \omega) = (Ux, \omega + \phi \circ \pi^{(n)}(x))$. If we define $\tilde\pi^{(n)} : X^{(n)} \times \mathbb{S}^1 \to X \times \mathbb{S}^1$ as $\pi^{(n)} \times \mathrm{Id}$, then $\mathscr{U}$ is a model for $\mathscr{T}$ since $\tilde\pi^{(n)} \circ \mathscr{U} = \mathscr{T} \circ \tilde\pi^{(n)}$. To study the properties of $\mathscr{T}$, it is therefore sufficient to understand $\mathscr{U}^{(n)}$ (for any conveniently chosen n). Abusing notations, we simply write ϕ on $X^{(n)}$ instead of $\phi \circ \pi^{(n)}$. We also identify Y with $Y \times \{0\} \subset X^{(n)}$.

The map U is not always mixing for the measure $\tilde\mu^{(n)}$; setting

$$d = d^{(n)} = \gcd\left\{r^{(n)}(x) \,\middle|\, x \in Y\right\}, \tag{3.2}$$

then U is mixing if and only if $d = 1$. If $d > 1$, let us write, for $k \in \mathbb{Z}/d\mathbb{Z}$, $\tilde\mu_k^{(n)}$ for the probability measure induced by $\tilde\mu^{(n)}$ on the set $\{(x, i) \mid i = k \mod d\}$. Then each measure $\tilde\mu_k^{(n)}$ is invariant under U^d and mixing. The measure $\pi_*^{(n)}\tilde\mu_k^{(n)}$ is absolutely continuous with respect to $\tilde\mu$ and invariant under T^d. Since T^d is ergodic (because T is mixing), this yields $\pi_*^{(n)}\tilde\mu_k^{(n)} = \tilde\mu$.

3.2. The transfer operator associated to $\mathscr{U}^{(N)}$

In the rest of this section, we work on $X^{(N)}$, where N is given by Theorem 2.3 (and fixed once and for all). This theorem makes it possible to study the transfer operator $\hat{\mathscr{U}}$ associated to the map $\mathscr{U} = \mathscr{U}^{(N)}$. Our goal in this section is to use this information to prove Theorem 1.7.

To keep the arguments as transparent as possible, we assume (until the end of the proof and without repeating it each time) that $d^{(N)} = \gcd\{r^{(N)}(x)\}$ is equal to 1. At the end of the proof, we indicate the modifications to be done in the general case.

Let us write a function v on $X^{(N)} \times \mathbb{S}^1$ as $v(x,\omega) = \sum_{k\in\mathbb{Z}} v_k(x)e^{ik\omega}$, that is,

$$v_k(x) = \int v(x,\omega)e^{-ik\omega}\,\mathrm{d}\omega, \tag{3.3}$$

where $\mathrm{d}\omega$ denotes the normalized Lebesgue measure on $\mathbb{S}^1$. If $\hat{\mathscr{U}}$ is the transfer operator associated to $\mathscr{U}$, and if $\mathscr{J}$ is the inverse of the Jacobian of U for $\mu^{(N)}$, then

$$\begin{aligned}\hat{\mathscr{U}}v(x,\omega) &= \sum_{\mathscr{U}(x',\omega')=(x,\omega)} \mathscr{J}(x')v(x',\omega') = \sum_{U(x')=x} \mathscr{J}(x')v\big(x',\omega-\phi(x')\big)\\ &= \sum_{k\in\mathbb{Z}}\sum_{Ux'=x} \mathscr{J}(x')v_k(x')e^{ik(\omega-\phi(x'))}.\end{aligned}$$

In the same way, if $\mathscr{J}^{(n)}$ denotes the Jacobian of U^n, then

$$\hat{\mathscr{U}}^n v(x,\omega) = \sum_{k\in\mathbb{Z}}\sum_{U^n x'=x} \mathscr{J}^{(n)}(x')v_k(x')e^{ik(\omega-S_n\phi(x'))}. \tag{3.4}$$

Hence, the operator $\hat{\mathscr{U}}^n$ is diagonal, acting on the kth frequency by the operator

$$\mathscr{M}_k^n v(x) = \sum_{U^n x'=x} \mathscr{J}^{(n)}(x')v(x')e^{-ikS_n\phi(x')}. \tag{3.5}$$

We understand separately the action of $\mathscr{M}_k$ for each k. Using the induction process, we are able to understand this operator for points x, x' belonging to the base Y of $X^{(N)}$. We then use this information to reconstruct the whole operator $\mathscr{M}_k$. To do so, let us define the following operators:

$$R_{n,k}v(x) = \sum_{\substack{U^n x'=x\\ x'\in Y, Ux',\dots,U^{n-1}x'\notin Y, U^n x'\in Y}} \mathscr{J}^{(n)}(x')v(x')e^{-ikS_n\phi(x')}, \tag{3.6}$$

$$T_{n,k}v(x) = \sum_{\substack{U^n x'=x\\ x'\in Y, U^n x'\in Y}} \mathscr{J}^{(n)}(x')v(x')e^{-ikS_n\phi(x')}, \tag{3.7}$$

$$A_{n,k}v(x) = \sum_{\substack{U^n x'=x\\ x'\in Y, Ux',\dots,U^n x'\notin Y}} \mathscr{J}^{(n)}(x')v(x')e^{-ikS_n\phi(x')}, \tag{3.8}$$

$$B_{n,k}v(x) = \sum_{\substack{U^n x'=x\\ x',\dots,U^{n-1}x'\notin Y, U^n x'\in Y}} \mathscr{J}^{(n)}(x')v(x')e^{-ikS_n\phi(x')}, \tag{3.9}$$

$$C_{n,k}v(x) = \sum_{\substack{U^n x'=x\\ x',\dots,U^n x'\notin Y}} \mathscr{J}^{(n)}(x')v(x')e^{-ikS_n\phi(x')}. \tag{3.10}$$

The main interest of these definitions is the following. First, cutting an orbit according to the first and last time it belongs to Y, we get

$$\mathscr{M}_k^n = C_{n,k} + \sum_{a+i+b=n} A_{a,k} T_{i,k} B_{b,k}. \tag{3.11}$$

Moreover, considering all the times an orbit belongs to Y, we obtain

$$T_{n,k} = \sum_{p=1}^{\infty} \sum_{j_1+\cdots+j_p=n} R_{j_1,k} \cdots R_{j_p,k}. \tag{3.12}$$

Finally, for $z \in \mathbb{C}$ with modulus at most e^{ε}, we have

$$\sum_{n>0} z^n R_{n,k} v = \mathscr{L}_k^N (z^{r^{(N)}} v). \tag{3.13}$$

The restriction $|z| < e^{\varepsilon}$ ensures that this operator is well defined by Theorem 2.3. More precisely, we even have the following.

LEMMA 3.1
There exists $C > 0$ such that, for any $n \in \mathbb{N}$, for any $k \in \mathbb{Z}$, we have

$$\|R_{n,k} v\|_{C^1(Y)} \leq C \max(1, |k|) e^{-2\varepsilon n} \|v\|_{C^1(Y)}. \tag{3.14}$$

Proof
Let $\psi_{n,k}(x) = e^{-ikS_N^Y \phi_Y(x)}$ if $r^{(N)}(x) = n$, and 0 otherwise, so that $R_{n,k} v = \mathscr{L}^N(\psi_{n,k} v)$. We show that $\|\psi_{n,k}\|_{\mathscr{C}_N^{1,4\varepsilon}} \leq C \max(1, |k|) e^{-2\varepsilon n}$, which concludes the proof by (2.6).

We have $|\psi_{n,k}(x)| \leq e^{-2n\varepsilon} e^{2\varepsilon r^{(N)}(x)}$. Moreover, if $h \in \mathscr{H}_N$ satisfies $r^{(N)} \circ h = n$, we have

$$\|D(\psi_{n,k} \circ h)(x)\| \leq C|k| r^{(N)}(hx) \leq C|k| e^{2\varepsilon r^{(N)}(hx)} \leq C|k| e^{-2\varepsilon n} e^{4\varepsilon r^{(N)}(hx)}. \tag{3.15}$$

This proves the lemma. □

3.3. Study of the operators $T_{n,k}$
In Equation (3.11), the complicated part in the expression of $\mathscr{M}_k^n$ comes from $T_{i,k}$, since the other operators are more or less explicit. This section is devoted to the study of the operators $T_{i,k}$, by using (3.12).

LEMMA 3.2
There exist $C > 0$ and $\bar{\theta} < 1$ such that, for any $k \in \mathbb{Z} - \{0\}$, for any $n \in \mathbb{N}$, and for any $v \in C^1(Y)$, we have $\|T_{n,k} v\|_{C^1} \leq C k^2 \bar{\theta}^n \|v\|_{C^1}$.

Proof
For $k \in \mathbb{Z}$ and $|z| \leq e^{\varepsilon}$, let us write $\mathscr{L}_{k,z} v = \mathscr{L}_k^N(z^{r^{(N)}} v) = \mathscr{L}^N(e^{-ikS_N^Y \phi_Y} z^{r^{(N)}} v)$. Since $\mathscr{L}_{k,z} = \sum z^j R_{j,k}$ by (3.13), Lemma 3.1 shows that this operator acts continuously on $C^1(Y)$ and that $z \mapsto \mathscr{L}_{k,z}$ is holomorphic on the disk $\{|z| \leq e^{\varepsilon}\}$. Formally, we can rewrite (3.12) as $\sum T_{n,k} z^n = \left(I - \sum R_{j,k} z^j\right)^{-1} = (I - \mathscr{L}_{k,z})^{-1}$. Hence, for any path γ in $\mathbb{C}$ around zero bounding a domain on which $I - \mathscr{L}_{k,z}$ is invertible for any z, we have, for any $n \in \mathbb{N}$,

$$T_{n,k} = \frac{1}{2i\pi} \int_{\gamma} z^{-n-1} (I - \mathscr{L}_{k,z})^{-1} \, \mathrm{d}z. \tag{3.16}$$

We use this equation as well as the information on $\mathscr{L}_{k,z}$ to estimate $T_{n,k}$.

Step 1. Fix $A_0 = 1$, and let $K_0 = K(A_0, 1)$ be given by the second half of Theorem 2.3 for this value of A. We first prove the lemma for $|k| \geq K_0$. Let us fix such a k.

Let $|z| \leq e^{\varepsilon}$. The function $z^{r^{(N)}}$ belongs to $\mathscr{C}_N^{A_0,\varepsilon}$ and its norm is bounded by 1. For $n \in \mathbb{N}$, we can iterate n times (2.7) (or rather (2.9)) (for $M = 1$) to obtain

$$\|\mathscr{L}_{k,z}^n v\|_{L^4} \leq \|\mathscr{L}_{k,z}^n v\|_{D_k} \leq \theta^{100Nn} \|v\|_{D_k} \leq \theta^{100Nn} \|v\|_{C^1}. \tag{3.17}$$

We then use (2.6). Note that the function $\psi(x) = e^{-ikS_N^Y \phi_Y(x)} z^{r^{(N)}(x)}$ is bounded by $e^{\varepsilon r^{(N)}(x)}$, and that, for $h \in \mathscr{H}_N$, we have

$$\begin{aligned} \|D(\psi \circ h)(x)\| &\leq |k| \|D(S_N^Y \phi_Y \circ h)(x)\| e^{\varepsilon r^{(N)}(x)} \\ &\leq C|k| r^{(N)}(x) e^{\varepsilon r^{(N)}(x)} \leq C'|k| e^{2\varepsilon r^{(N)}(x)}. \end{aligned}$$

Letting $A = C'|k|$, we have proved that $\psi \in \mathscr{C}_N^{A,2\varepsilon}$ and that $\|\psi\|_{\mathscr{C}_N^{A,2\varepsilon}} \leq 1$. Applying (2.6) for n iterates, we obtain, for any C^1 function w,

$$\|\mathscr{L}_{k,z}^n w\|_{C^1} \leq C|k| (\theta^{100Nn} \|w\|_{C^1} + \theta^{-Nn} \|w\|_{L^2}). \tag{3.18}$$

Applying this equation to $w = \mathscr{L}_{k,z}^n v$ and using (3.17), we get

$$\|\mathscr{L}_{k,z}^{2n} v\|_{C^1} \leq C|k| (\theta^{100Nn} \|\mathscr{L}_{k,z}^n v\|_{C^1} + \theta^{-Nn} \theta^{100Nn} \|v\|_{C^1}). \tag{3.19}$$

Applying once again (3.18) but this time to v, we finally get $\|\mathscr{L}_{k,z}^{2n} v\|_{C^1} \leq C|k|^2 \theta^{99Nn} \|v\|_{C^1}$. We can argue, in the same way for odd times, to finally obtain the existence of C such that, for any $n \in \mathbb{N}$, $v \in C^1(Y)$, $|k| \geq K_0$, and $|z| \leq e^{\varepsilon}$,

$$\|\mathscr{L}_{k,z}^n v\|_{C^1} \leq Ck^2 \theta^{40Nn} \|v\|_{C^1}. \tag{3.20}$$

In particular, this shows that the operator $I - \mathscr{L}_{k,z}$ is invertible on $C^1(Y)$ and that its inverse $\sum \mathscr{L}_{k,z}^n$ has a norm that is bounded by $Ck^2/(1 - \theta^{40N})$.

We can then use equation (3.16) by taking for γ a circle of radius e^{ε}. We obtain

$$\|T_{n,k}\| \leq Ck^2 \int_{\gamma} |z|^{-n} \leq Ck^2 e^{-n\varepsilon}. \tag{3.21}$$

This concludes the proof for $|k| \geq K_0$.

Step 2. Consider now $|k| < K_0$, $k \neq 0$. We show that, for any z with $|z| \leq 1$, the operator $I - \mathscr{L}_{k,z}$ is invertible on $C^1(Y)$. Since the invertible operators form an open set, this implies the existence of $\varepsilon(k)$ such that, for $|z| \leq e^{\varepsilon(k)}$, $I - \mathscr{L}_{k,z}$ is invertible on $C^1(Y)$. Using a path γ that is a circle of radius $e^{\varepsilon(k)}$, we can then conclude as above (without explicit control, but since there are only finitely many values of k to deal with, this is not a problem).

Thus, consider z with $|z| \leq 1$. The inequality (3.18) still holds (its proof does not use $|k| \geq K_0$). Therefore, there exists $C > 0$ such that, for any $n \in \mathbb{N}$, $\|\mathscr{L}_{k,z}^n v\|_{C^1} \leq C\theta^{100Nn} \|v\|_{C^1} + C(n) \|v\|_{L^2}$. Since the injection of $C^1(Y)$ in $L^2(Y)$ is compact, this is a Lasota-Yorke inequality. Hennion's Theorem 2.1 therefore shows that the essential spectral radius of $\mathscr{L}_{k,z}$ is less than 1. If $I - \mathscr{L}_{k,z}$ is not invertible, there must therefore exist $v \in C^1(Y)$ nonzero such that $\mathscr{L}_{k,z} v = v$ (i.e., $\mathscr{L}^N(e^{-ikS_N^Y \phi_Y} z^{r^{(N)}} v) = v$). The operator $\mathscr{L}^N$ is the transfer operator associated to the map T_Y^N, which is ergodic on Y. Lemma 2.2 applies and shows on the one hand that $|z|^{r^{(N)}}$ is almost everywhere equal to 1 (hence $|z| = 1$) and on the other hand that $v \circ T_Y^N = z^{r^{(N)}} e^{-ikS_N^Y \phi_Y} v$ almost everywhere. Raising this equation to the power K_0, we obtain that v^{K_0} is invariant under the operator $\mathscr{L}_{kK_0, z^{K_0}}$. But we have already proved that $I - \mathscr{L}_{kK_0, z^{K_0}}$ is invertible on $C^1(Y)$. As a consequence, $v^{K_0} = 0$, and $v = 0$, which is a contradiction. This concludes the proof for $|k| \in [1, K_0)$. □

To obtain an estimate on $T_{n,0}$, we must also take into account the fact that $I - \mathscr{L}_{0,1}$ is not invertible (its kernel corresponds to constant functions), which adds a residue in the integral calculus of the previous proof. In the following definition, we introduce a tool that makes the computation of this residue possible. We write $\mathbb{D}$ for the open unit disk in $\mathbb{C}$, and $\overline{\mathbb{D}}$ for its closure.

Definition 3.3

Let $\mathscr{B}$ be a Banach space, and let R_j be operators acting on $\mathscr{B}$ for $j > 0$. We say that they form a *renewal sequence of operators with exponential decay* if the following conditions hold.

(1) There exist $\delta > 0$ and $C > 0$ such that $\|R_j\| \leq Ce^{-\delta j}$. We can thus define an operator $R(z) = \sum R_j z^j$ for $|z| < e^{\delta}$.

(2) For any $z \in \overline{\mathbb{D}} - \{1\}$, the operator $I - R(z)$ is invertible on $\mathscr{B}$.

(3) The operator $R(1)$ has a simple isolated eigenvalue at 1. Let $P = P(1)$ be the corresponding spectral projection, and let $R'(1) = \sum jR_j$. We assume that there exists $\mu > 0$ such that $PR'(1)P = \mu P$.

PROPOSITION 3.4
Let R_j be a renewal sequence of operators with exponential decay on a Banach space $\mathscr{B}$. Let us define an operator T_n by $T_n = \sum_{p=1}^{\infty} \sum_{j_1+\cdots+j_p=n} R_{j_1} \cdots R_{j_p}$. Then there exist $C > 0$ and $\bar{\theta} < 1$ such that, for any $n \in \mathbb{N}$, $\|T_n - P/\mu\| \leq C\bar{\theta}^n$.

Proof
For z close to 1, the operator $R(z)$ is close to $R(1)$. Hence, it has an eigenvalue $\lambda(z)$ close to 1, with a corresponding spectral projection $P(z)$ (and all these quantities depend holomorphically on z). Let us compute the derivative $\lambda'(1)$.

We denote with a prime the derivative with respect to z. For any $x \in \mathscr{B}$, $R(z)P(z)x = \lambda(z)P(z)x$. Differentiating with respect to z and then multiplying on the left by $P(z)$, we get (omitting the variable z)

$$PR'Px + PRP'x = \lambda'Px + \lambda PP'x. \tag{3.22}$$

Moreover, $PRP' = P^2RP' = PRPP' = \lambda PP'$. After simplification, we obtain $PR'Px = \lambda'Px$. For $z = 1$, $PR'P = \mu P$. Choosing x such that $Px \neq 0$, we finally get

$$\lambda'(1) = \mu \neq 0. \tag{3.23}$$

In particular, on a small enough disk O around 1, the function $z \mapsto \lambda(z)$ is injective, and takes the value 1 only for $z = 1$.

The operators $I - R(z)$ are invertible for $z \in \overline{\mathbb{D}} - O$, hence also for z in a neighborhood of this compact set. We can therefore choose a path γ around zero going along an arc of a circle of radius greater than 1 and the inner part of ∂O. It satisfies the equation

$$T_n = \frac{1}{2i\pi} \int_{\gamma} z^{-n-1} \big(I - R(z)\big)^{-1} \, \mathrm{d}z. \tag{3.24}$$

We modify γ into a new path $\tilde{\gamma}$ that runs along the same arc of a circle of radius greater than 1 and the outer part of ∂O. To obtain an analogue of (3.24), we need to add the residue of $z^{-n-1}(I - R(z))^{-1}$ inside O. We have $(I - R(z))^{-1} = (1 - \lambda(z))^{-1}P(z) + Q(z)$, where $Q(z)$ is holomorphic inside O (hence without residue). The only pole is thus at 1, and we get

$$T_n = \frac{1}{2i\pi} \int_{\tilde{\gamma}} z^{-n-1} \big(I - R(z)\big)^{-1} \, \mathrm{d}z + \frac{1}{\lambda'(1)} P. \tag{3.25}$$

On $\tilde{\gamma}$, $|z| \geq e^{\delta'}$ for some $\delta' > 0$. As $\|(I - R(z))^{-1}\|$ is uniformly bounded along $\tilde{\gamma}$, the integral term is therefore $O(e^{-n\delta'})$. The remaining term gives the conclusion of the proposition. $\square$

We can now come back to the study of the transfer operator associated to $\mathscr{U}$, and more precisely to the operators $T_{n,0}$, which have not yet been estimated.

COROLLARY 3.5
For any C^1 function v on Y, let $Pv = \int v \, \mathrm{d}\mu_Y$. Then there exist $C > 0$ and $\bar{\theta} < 1$ such that, for any $n \in \mathbb{N}$ and any $v \in C^1(Y)$, we have

$$\left\| T_{n,0}v - \frac{1}{\mu^{(N)}(X^{(N)})} Pv \right\|_{C^1} \leq C\bar{\theta}^n \, \|v\|_{C^1} \, . \tag{3.26}$$

Proof
We use the fact that the Markov transformations T_Y and U are mixing. Since these transformations are topologically mixing (by the equality $\gcd\{r^{(N)}(x)\} = 1$ for U), the mixing in measure results, for example, from [A, Theorem 4.4.7].

Let us show that $R_{n,0}$ is a renewal sequence of operators with exponential decay on the Banach space $\mathscr{B} = C^1(Y)$. The exponential decay of $\|R_{n,0}\|$ is given by Lemma 3.1. Let $\mathscr{L}_{0,z}v = \mathscr{L}^N(z^{r^{(N)}}v) = \sum z^n R_{n,0} = R(z)$.

Let us check that $I - R(z) = I - \mathscr{L}_{0,z}$ is invertible for $z \in \overline{\mathbb{D}} - \{1\}$. As in the proof of Lemma 3.2, the operators $\mathscr{L}_{0,z}$ (for $|z| \leq 1$) have an essential spectral radius less than 1 on C^1. If $I - \mathscr{L}_{0,z}$ were not invertible, there would exist a nonzero C^1 function v such that $\mathscr{L}_{0,z}v = v$. Lemma 2.2 implies that $|z| = 1$ and $v \circ T_Y^N = z^{r^{(N)}}v$. Let us extend v to the whole space $X^{(N)}$ by setting $v(x,i) = z^i v(x,0)$. Thus, the function v is bounded (and therefore integrable) and satisfies $v \circ U = zU$. This is a contradiction since U is mixing.

For $z = 1$, $R(1) = \mathscr{L}_{0,1}$ simply is the transfer operator associated to T_Y^N. It has a simple eigenvalue at 1 (the corresponding spectral projection being P), and no other eigenvalue of modulus 1. Let us compute $PR'(1)P$. We have

$$PR_{n,0}Pu = \mu_Y\{r^{(N)} = n\}Pu. \tag{3.27}$$

As a consequence, Kac's formula gives $PR'(1)P = \left(\sum n\mu_Y\{r^{(N)} = n\}\right)P = \mu^{(N)}(X^{(N)})P$.

We can then apply Proposition 3.4 and get the conclusion of the corollary. $\square$

3.4. The exponential mixing

The estimates on $T_{n,k}$ given in Section 3.3 enable us to describe $\mathscr{M}_k^n$ for any k and then to describe the full transfer operator $\hat{\mathscr{U}}$.

For $x \in X^{(N)}$, denote by $h(x)$ its height in the tower (i.e., if $x = (y, i)$ with $y \in Y$ and $i < r^{(N)}(x)$, let $h(x) = i$). We write $C^{5,1}(X^{(N)} \times \mathbb{S}^1)$ for the set of functions $v : X^{(N)} \times \mathbb{S}^1 \to \mathbb{C}$ such that $\partial^i v / \partial \omega^i$ is C^1 for $0 \leq i \leq 5$ with its canonical norm.

THEOREM 3.6
There exist constants $C > 0$ and $\bar{\theta} < 1$ such that, for any $C^{5,1}$ function $v : X^{(N)} \times \mathbb{S}^1 \to \mathbb{C}$, for any $n \in \mathbb{N}$, and for any $(x, \omega) \in X^{(N)} \times \mathbb{S}^1$ with $h(x) \leq n/2$, we have

$$\left| \hat{\mathscr{U}}^n v(x, \omega) - \int v \, \mathrm{d}(\tilde{\mu}^{(N)} \otimes \mathrm{Leb}) \right| \leq C \bar{\theta}^n \|v\|_{C^{5,1}} . \tag{3.28}$$

For the proof, we need information on the operators $T_{i,k}$, but we also need to describe precisely the operators $B_{i,k}$ (defined in (3.9)).

LEMMA 3.7
There exist $\bar{\theta} < 1$ and $C > 0$ such that, for any $k \in \mathbb{Z}$, $v \in C^1(X^{(N)})$, and $n \in \mathbb{N}$,

$$\|B_{n,k} v\|_{C^1} \leq C(1 + |k|) \bar{\theta}^n \|v\|_{C^1} . \tag{3.29}$$

Moreover,

$$\left| \int_{X^{(N)}} v \, \mathrm{d}\mu^{(N)} - \sum_{j=0}^{n} \int_Y B_{j,0} v \, \mathrm{d}\mu^{(N)} \right| \leq C \bar{\theta}^n \|v\|_{C^1} . \tag{3.30}$$

Proof
For $y \in Y$, let $v_n(y) = 0$ if $r^{(N)}(y) \leq n$, and let

$$v_n(y) = v\big(y, r^{(N)}(y) - n\big) \exp\Big(-ik \sum_{j = r^{(N)}(y) - n}^{r^{(N)}(y) - 1} \phi(y, j) \Big) \tag{3.31}$$

otherwise. For $x \in Y$, we then have $B_{n,k} v(x) = \mathscr{L}^N v_n(x)$ since $B_{n,k} v(x)$ takes into account the values of v on the set Z_n of points that enter Y after exactly n iterations, that is, points of the form $(y, r^{(N)}(y) - n)$ with $r^{(N)}(y) > n$.

Let us check that the function v_n belongs to $\mathscr{C}_N^{1,\varepsilon}$. First, since v_n vanishes for $r^{(N)} \leq n$, we have

$$|v_n(x)| \leq 1_{r^{(N)}(x) > n} \|v\|_{C^0} \leq e^{-\varepsilon n} e^{\varepsilon r^{(N)}(x)} \|v\|_{C^0} . \tag{3.32}$$

Moreover, if $h \in \mathscr{H}_N$, then

$$\|D(v_n \circ h)(x)\| \leq 1_{r^{(N)} \circ h > n} (\|v\|_{C^1} + kn \|v\|_{C^0}) \leq C(1 + |k|) n e^{-\varepsilon n} e^{\varepsilon r^{(N)}(hx)} \|v\|_{C^1} . \tag{3.33}$$

Hence, v_n belongs to $\mathscr{C}_N^{1,\varepsilon}$, and its norm is bounded by $C(1+|k|)\bar{\theta}^n \|v\|_{C^1}$. Applying (2.6), this yields (3.29).

For (3.30), note that $\sum_{j=0}^{\infty} \int_Y B_{j,0} v = \int v$ since $\int_Y B_{j,0} v$ is the integral of v on Z_j. Therefore,

$$\left| \int v - \sum_{j=0}^{n} \int_Y B_{j,0} v \right| \leq \sum_{j=n+1}^{\infty} \left| \int_Y B_{j,0} v \right| \leq \sum_{j=n+1}^{\infty} \|B_{j,0} v\|_{C^1} \leq C\bar{\theta}^n \|v\|_{C^1} \quad (3.34)$$

by (3.29). $\square$

COROLLARY 3.8
There exist $C > 0$ and $\bar{\theta} < 1$ such that, for any $k \in \mathbb{Z}$, any $n \in \mathbb{N}$, any $x \in X^{(N)}$ with $h(x) \leq n/2$, and any $v \in C^1(X^{(N)})$, we have

$$\left| \mathscr{M}_k^n v(x) - 1_{k=0} \int v \, \mathrm{d}\tilde{\mu}^{(N)} \right| \leq C(1 + |k|^3)\bar{\theta}^n \|v\|_{C^1} . \quad (3.35)$$

Proof
Assume first that $x \in Y$. Then (3.11) simply becomes

$$\mathscr{M}_k^n v(x) = \sum_{i=0}^{n} T_{n-i,k} B_{i,k} v(x). \quad (3.36)$$

If $k \neq 0$, then

$$\|T_{n-i,k} B_{i,k} v\|_{C^1} \leq Ck^2 \bar{\theta}^{n-i} \|B_{i,k} v\|_{C^1} \leq C|k|^3 \bar{\theta}^{n-i} \bar{\theta}^i \|v\|_{C^1} , \quad (3.37)$$

by Lemmas 3.2 and 3.7. Summing over i, we obtain the desired bound.

If $k = 0$, Corollary 3.5 gives an additional term

$$\begin{aligned} \sum_{i=0}^{n} P B_{i,0} v/\mu^{(N)}(X^{(N)}) &= \sum_{i=0}^{n} \int_Y B_{i,0} v \, \mathrm{d}\mu^{(N)}/\mu^{(N)}(X^{(N)}) \\ &= \int v \, \mathrm{d}\mu^{(N)}/\mu^{(N)}(X^{(N)}) + O(\bar{\theta}^n) = \int v \, \mathrm{d}\tilde{\mu}^{(N)} + O(\bar{\theta}^n) \end{aligned}$$

by (3.30). This proves (3.35) for $x \in Y$.

If x has height $j \in (0, n/2]$, let us write $x = U^j(x')$, so that

$$\mathscr{M}_k^n u(x) = e^{-ikS_j\phi(x')} \mathscr{M}_k^{n-j} u(x'). \quad (3.38)$$

The estimate for x' gives the desired conclusion (after replacing $\bar{\theta}$ with $\bar{\theta}^{1/2}$). $\square$

Proof of Theorem 3.6
Let $v : X^{(N)} \times \mathbb{S}^1 \to \mathbb{R}$ be a $C^{5,1}$ function. We decompose it as $v(x,\omega) = \sum_{k\in\mathbb{Z}} v_k(x) e^{ik\omega}$. Then

$$\hat{\mathscr{U}}^n v(x,\omega) = \sum_{k\in\mathbb{Z}} \mathscr{M}_k^n v_k(x) \cdot e^{ik\omega} \tag{3.39}$$

by (3.4). Therefore, if $h(x) \le n/2$, Corollary 3.8 gives

$$\begin{aligned}\left|\hat{\mathscr{U}}^n v(x,\omega) - \int v \,\mathrm{d}(\tilde{\mu}^{(N)} \otimes \mathrm{Leb})\right| &\le \left|\mathscr{M}_0^n v_0(x) - \int v_0 \,\mathrm{d}\tilde{\mu}^{(N)}\right| + \sum_{k\neq 0} \left|\mathscr{M}_k^n v_k(x)\right| \\ &\le C \sum_{k\in\mathbb{Z}} (1+|k|^3)\bar{\theta}^n \|v_k\|_{C^1}.\end{aligned}$$

With 5 integrations by parts with respect to ω, we show that $\|v_k\|_{C^1} \le C \|v\|_{C^{5,1}} / (1+|k|^5)$. This implies the theorem after summation. □

Proof of Theorem 1.7 (under the assumption that $d^{(N)} = 1$)
Let us first show that, on $X^{(N)} \times \mathbb{S}^1$,

$$\left\|\hat{\mathscr{U}}^n v - \int v \,\mathrm{d}(\tilde{\mu}^{(N)} \otimes \mathrm{Leb})\right\|_{L^1} \le C\bar{\theta}^n \|v\|_{C^{5,1}} \tag{3.40}$$

for some constants $C > 0$ and $\bar{\theta} < 1$. To do this, we decompose $X^{(N)}$ as $\{x \mid h(x) > n/2\}$ and $\{x \mid h(x) \le n/2\}$. The first set has an exponentially small measure, its contribution is therefore exponentially small. If x belongs to the second set, $\left|\hat{\mathscr{U}}^n v(x,\omega) - \int v\right| \le C\bar{\theta}^n \|v\|_{C^{5,1}}$ by Theorem 3.6. This proves (3.40).

This implies that, for any functions $v \in C^{5,1}$ and $u \in L^\infty$,

$$\begin{aligned}&\left|\int u \circ \mathscr{U}^n \cdot v \,\mathrm{d}(\tilde{\mu}^{(N)} \otimes \mathrm{Leb}) - \left(\int u \,\mathrm{d}(\tilde{\mu}^{(N)} \otimes \mathrm{Leb})\right)\left(\int v \,\mathrm{d}(\tilde{\mu}^{(N)} \otimes \mathrm{Leb})\right)\right| \\ &\quad \le C\bar{\theta}^n \|u\|_{L^\infty} \|v\|_{C^{5,1}}.\end{aligned} \tag{3.41}$$

Now take $f \in L^\infty(X \times \mathbb{S}^1)$, and take $g \in C^6(X \times \mathbb{S}^1)$. The functions $u = f \circ \tilde{\pi}^{(N)}$ and $v = g \circ \tilde{\pi}^{(N)}$ are defined on $X^{(N)} \times \mathbb{S}^1$, respectively, bounded and in $C^{5,1}$. Moreover, (3.1) shows that $\|v\|_{C^{5,1}} \le C \|g\|_{C^6}$. Since $\pi_*^{(N)} \tilde{\mu}^{(N)} = \tilde{\mu}$, (3.41) implies that

$$\begin{aligned}&\left|\int f \circ \mathscr{T}^n \cdot g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) - \left(\int f \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\right)\left(\int g \,\mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\right)\right| \\ &\quad \le C\bar{\theta}^n \|f\|_{L^\infty} \|g\|_{C^6}.\end{aligned} \tag{3.42}$$

Let $n \in \mathbb{N}$, and let $f \in L^\infty$. The linear operator

$$g \mapsto \int f \circ \mathscr{T}^n \cdot g \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) - \Big(\int f \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \Big) \Big(\int g \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \Big) \quad (3.43)$$

is then bounded by $2 \| f \|_{L^\infty}$ in C^0 norm, and by $C\bar{\theta}^n \| f \|_{L^\infty}$ in C^6 norm. For any noninteger $\alpha \in (0, 6)$, interpolation theory on the compact manifold $X \times \mathbb{S}^1$ (possibly with boundary) shows that there exists a constant C_α such that any operator that is bounded by A in C^0 norm and by B in C^6 norm is then bounded by $C_\alpha A^{1-\alpha/6} B^{\alpha/6}$ in C^α norm (see [T, page 200]). As a consequence, we get

$$\Big| \int f \circ \mathscr{T}^n \cdot g \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) - \Big(\int f \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \Big) \Big(\int g \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \Big) \Big| \\ \leq C_\alpha \| f \|_{L^\infty} 2^{1-\alpha/6} (C\bar{\theta}^n)^{\alpha/6} \| g \|_{C^\alpha} .$$

This concludes the proof of the theorem for noninteger α. The general case follows readily. The interpolation argument can also be replaced by an elementary (but less synthetic) convolution argument. The idea of using interpolation theory in this kind of setting was suggested by Dinh and Sibony in [DS]. □

Proof of Theorem 1.7 in the general case
If $d = d^{(N)} > 1$, the transformation U is not mixing, and the arguments used above (especially in the proof of Corollary 3.5) do not apply anymore.

However, they can be applied to the transformation U^d and its invariant measure $\tilde{\mu}_0^{(N)}$ (defined in Section 3.1). As $\pi_*^{(N)} \tilde{\mu}_0^{(N)} = \tilde{\mu}$, this implies Theorem 1.7 for times n of the form kd. To deduce the general case, one writes $n = kd + r$ with $0 \leq r < d$ and applies the theorem to the time kd and to the functions $f \circ \mathscr{T}^r$ and g (which are, respectively, bounded and C^α). □

3.5. Proof of one implication in Proposition 1.8

PROPOSITION 3.9
Let $\psi : X \times \mathbb{S}^1 \to \mathbb{R}$ be a Hölder-continuous function of zero average, and define σ^2 by (1.11). Then $\sigma^2 \geq 0$. Moreover, if $\sigma^2 = 0$, there exists a measurable function $f : X \times \mathbb{S}^1$, continuous on $Y \times \mathbb{S}^1$, belonging to L^p for any $p < \infty$ such that $\psi = f - f \circ \mathscr{T}$ almost everywhere.

This is one of the implications in Proposition 1.8. Theorem 1.9 is required for the other half, hence its proof is postponed to Section 5.6.

Proof
We have

$$\int_{X\times\mathbb{S}^1}\Big(\sum_{i=0}^{n-1}\psi\circ\mathscr{T}^i\Big)^2 = n\int\psi^2 + 2\sum_{i=0}^{n-1}(n-i)\int\psi\cdot\psi\circ\mathscr{T}^i. \tag{3.44}$$

Since $\sum_{i>0} i\left|\int\psi\cdot\psi\circ\mathscr{T}^i\right| < \infty$ by Theorem 1.7, this yields

$$\int_{X\times\mathbb{S}^1}\Big(\sum_{i=0}^{n-1}\psi\circ\mathscr{T}^i\Big)^2 = n\sigma^2 + O(1). \tag{3.45}$$

As a consequence, $\sigma^2 \geq 0$. Moreover, if $\sigma^2 = 0$, the Birkhoff sums of ψ are uniformly bounded in L^2. By [K, Remark 1], there exists an L^2 function f with zero average such that $\psi = f - f\circ\mathscr{T}$ almost everywhere. We have to prove that f is continuous on $Y\times\mathbb{S}^1$ and belongs to every L^p, $p<\infty$.

Theorem 3.6 implies that there exist $\bar\theta < 1$ and $C > 0$ such that, for any C^6 function $v : X\times\mathbb{S}^1 \to \mathbb{C}$, for any $n\in\mathbb{N}$ and for any $x\in X^{(N)}$ with $h(x)\leq n/2$, we have

$$\left|\hat{\mathscr{U}}^n(v\circ\tilde\pi^{(N)})(x,\omega) - \int v\right| \leq C\bar\theta^n \|v\|_{C^6}. \tag{3.46}$$

Since $|\hat{\mathscr{U}}^n(v\circ\tilde\pi^{(N)})(x,\omega) - \int v| \leq 2\|v\|_{C^0}$, interpolation theory as above implies that, for any $\alpha > 0$, there exist $C_\alpha > 0$ and $\bar\theta_\alpha < 1$ such that, for any $x\in X^{(N)}$ with $h(x)\leq n/2$, we have

$$\left|\hat{\mathscr{U}}^n(v\circ\tilde\pi^{(N)})(x,\omega) - \int v\right| \leq C_\alpha\bar\theta_\alpha^n \|v\|_{C^\alpha}. \tag{3.47}$$

As ψ belongs to C^α and has vanishing integral, we can therefore define a function g on $X^{(N)}\times\mathbb{S}^1$ by

$$g(x,\omega) = -\sum_{n=1}^{\infty}\hat{\mathscr{U}}^n(\psi\circ\tilde\pi^{(N)})(x,\omega). \tag{3.48}$$

This function is continuous on $Y\times\mathbb{S}^1$. Moreover, if $h(x) = H$, then the first $2H$ terms in the sum defining $g(x,\omega)$ are bounded by $\|\psi\|_{C^0}$, while the other ones are bounded by $C\bar\theta_\alpha^n$ by (3.47). Hence, $g(x,\omega)$ is bounded by $C(1+h(x))$. Therefore, g belongs to L^p for any $p<\infty$ (it even has an exponential moment), since $\mu^{(N)}\{h(x)\geq n\}$ decays exponentially with n. Moreover, by construction, $\hat{\mathscr{U}}g - g = \hat{\mathscr{U}}(\psi\circ\tilde\pi^{(N)})$.

We know that $\psi = f - f \circ \mathcal{T}$, where $f \in L^2$. As a consequence, $\psi \circ \tilde{\pi}^{(N)} = f \circ \tilde{\pi}^{(N)} - f \circ \tilde{\pi}^{(N)} \circ \mathcal{U}$, whence $\hat{\mathcal{U}}(\psi \circ \tilde{\pi}^{(N)}) = \hat{\mathcal{U}}(f \circ \tilde{\pi}^{(N)}) - f \circ \tilde{\pi}^{(N)}$. We get

$$g - f \circ \tilde{\pi}^{(N)} = \hat{\mathcal{U}}(g - f \circ \tilde{\pi}^{(N)}). \tag{3.49}$$

In particular, for any $n \in \mathbb{N}$, $g - f \circ \tilde{\pi}^{(N)} = \hat{\mathcal{U}}^n(g - f \circ \tilde{\pi}^{(N)})$.

Theorem 3.6 shows that, for any function $v \in C^{5,1}(X^{(N)} \times \mathbb{S}^1)$ with zero integral, $\hat{\mathcal{U}}^n v$ converges to zero in L^2. By density, this convergence holds for any function $v \in L^2$ with zero integral. In particular, $\hat{\mathcal{U}}^n(g - f \circ \tilde{\pi}^{(N)})$ converges to zero, hence $g - f \circ \tilde{\pi}^{(N)} = 0$. As g is continuous on $Y \times \mathbb{S}^1$ and belongs to all spaces L^p, $p < \infty$, this concludes the proof. □

4. Strategy and tools for the local limit theorem

4.1. Description of the strategy of the proof

Let us fix an integer M. We work with the transformation $U = U^{(MN)}$ on $X^{(MN)}$ (hence also with $\mathcal{U}^{(MN)}$ on $X^{(MN)} \times \mathbb{S}^1$).

Let $\psi : X \times \mathbb{S}^1 \to \mathbb{R}$ be a C^6 function with zero average. We also write ψ instead of $\psi \circ \tilde{\pi}^{(MN)}$ on $X^{(MN)} \times \mathbb{S}^1$. To prove the local limit theorem for ψ, we consider for $t \in \mathbb{R}$ the operator $\hat{\mathcal{U}}_t(v) := \hat{\mathcal{U}}(e^{it\psi} v)$. If we understand well the iterates of $\hat{\mathcal{U}}_t$, we deduce the asymptotic behavior of $\int e^{itS_n\psi}$, since this quantity is equal to $\int \hat{\mathcal{U}}_t^n(1)$.

Instead of working with functions on $X^{(MN)} \times \mathbb{S}^1$, we have seen in the proof of the exponential mixing that it is worthwhile to use Fourier series, and work on $X^{(MN)} \times \mathbb{Z}$. If v is a function and $(v_k)_{k\in\mathbb{Z}}$ denote its Fourier coefficients, then the Fourier coefficients of $e^{it\psi} v$ are given by

$$(e^{it\psi} v)_k = \sum_{a+b=k} (e^{it\psi})_a v_b. \tag{4.1}$$

Applying then the operator $\hat{\mathcal{U}}$ (which acts at the level of the kth frequency by the operator $\mathcal{M}_k$), we obtain

$$(\hat{\mathcal{U}}_t v)_k(x) = \sum_{l\in\mathbb{Z}} \sum_{Ux'=x} \mathcal{J}(x') e^{-ik\phi(x')} (e^{it\psi})_{k-l}(x') v_l(x'). \tag{4.2}$$

This is some kind of Markov operator on $X^{(MN)} \times \mathbb{Z}$, for the "transition probability"

$$\mathcal{K}^t_{(x,k)\to(x',l)} := 1_{Ux'=x} \mathcal{J}(x') e^{-ik\phi(x')} (e^{it\psi})_{k-l}(x'). \tag{4.3}$$

The equality $\sum_{(x',l)} \mathcal{K}^t_{(x,k)\to(x',l)} = 1$ does not hold, so this is not a real transition kernel, but we nevertheless use the intuition of random walks. In particular, let us write, for $n \in \mathbb{N}$,

$$\mathscr{K}^{t,n}_{(x,k)\to(x',l)} = \sum_{\substack{k_0=l,k_1,\dots,k_{n-1},k_n=k \\ x_0=x',x_1,\dots,x_{n-1},x_n=x}} \mathscr{K}^t_{(x_n,k_n)\to(x_{n-1},k_{n-1})} \cdots \mathscr{K}^t_{(x_2,k_2)\to(x_1,k_1)} \mathscr{K}^t_{(x_1,k_1)\to(x_0,k_0)}. \tag{4.4}$$

In this expression, we consider trajectories of the random walk $x_n, x_{n-1}, \dots, x_0$. It may seem unnatural to write things in that direction, but it is designed to give the good order when we express things in terms of transfer operators. Let $\hat{\mathscr{K}}^t$ be the operator with kernel $\mathscr{K}^t$, acting on bounded functions on $X^{(MN)} \times \mathbb{Z}$, by

$$\hat{\mathscr{K}}^t v(x,k) = \sum_{(x',l)} \mathscr{K}^t_{(x,k)\to(x',l)} v(x',l). \tag{4.5}$$

By construction, the powers $\hat{\mathscr{K}}^{t,n}$ of $\hat{\mathscr{K}}^t$ have kernels $\mathscr{K}^{t,n}$. Moreover, $\hat{\mathscr{U}}_t$ corresponds to the operator $\hat{\mathscr{K}}^t$ at the level of frequencies; that is, if v is a smooth function on $X^{(MN)} \times \mathbb{S}^1$ with Fourier coefficients $(v_k)_{k\in\mathbb{Z}}$, then

$$(\hat{\mathscr{U}}_t^n v)_k(x) = \sum_{(x',l)} \mathscr{K}^{t,n}_{(x,k)\to(x',l)} v_l(x'). \tag{4.6}$$

To see that this expression and these computations are correct, we should check that

$$\sup_{(x,k)\in X^{(MN)}\times\mathbb{Z}} \sum_{(x',l)} |\mathscr{K}^t_{(x,k)\to(x',l)}| < \infty, \tag{4.7}$$

which is always the case if ψ is C^2 in the direction of $\mathbb{S}^1$ (by two integrations by parts) and is always satisfied in the following. A priori, this does not prevent $\mathscr{K}^{t,n}_{(x,k)\to(x',l)}$ from blowing up exponentially fast with n. However, $\mathscr{K}^{t,n}_{(x,k)\to(x',l)}$ is also the kernel of the operator obtained by multiplying v with $e^{itS_n\psi}$ and then applying $\hat{\mathscr{U}}^n$. Therefore,

$$\mathscr{K}^{t,n}_{(x,k)\to(x',l)} = 1_{U^n x'=x} \mathscr{J}^{(n)}(x') e^{-ikS_n\phi(x')} (e^{itS_n\psi})_{k-l}(x'), \tag{4.8}$$

and this quantity is bounded by $\mathscr{J}^{(n)}(x') \le 1$. Note that (4.8) can also be checked directly from the formula (4.4), with several successive integrations.

We let different operators (with kernels related to $\mathscr{K}^{t,n}$) act on spaces of functions from $X^{(MN)} \times \mathbb{Z}$ to $\mathbb{C}$ (or $Y \times \mathbb{Z}$ to $\mathbb{C}$ if we only consider trajectories starting from $Y \times \mathbb{Z}$ or ending in $Y \times \mathbb{Z}$). If $\mathscr{B}$ is such a functional space, and if $v \in \mathscr{B}$, we sometimes write $v_k(x)$ instead of $v(x,k)$.

To understand the previous "random walk," we study its successive returns to the set $Y \times [-K, K]$, where K is large enough. Indeed, outside of this set, we have a strong contraction (by Theorem 2.3), and hence excursions can be controlled. Only what happens inside $Y \times [-K, K]$ can therefore be problematic, and we use there an abstract compactness argument. Let us denote by $\mathscr{K}^{t,n,\mathrm{exc}}_{(x,k)\to(x',l)}$ the "probability" of an excursion (i.e., of starting from $(x, k) \in Y \times [-K, K]$ and coming back to $(x', l) \in Y \times [-K, K]$ after a time exactly n without entering $Y \times [-K, K]$ in between). Formally, for $(x, k) \in Y \times [-K, K]$ and $(x', l) \in Y \times [-K, K]$, we have

$$\mathscr{K}^{t,n,\mathrm{exc}}_{(x,k)\to(x',l)} = \sum_{\substack{k_0=l,\ldots,k_n=k \\ x_0=x',x_1,\ldots,x_{n-1}\in X,x_n=x \\ (x_i,k_i)\notin Y\times[-K,K] \text{ for } 0<i<n}} \mathscr{K}^t_{(x_n,k_n)\to(x_{n-1},k_{n-1})} \cdots \mathscr{K}^t_{(x_2,k_2)\to(x_1,k_1)} \mathscr{K}^t_{(x_1,k_1)\to(x_0,k_0)}.$$

Let $\mathscr{B}_K = \bigoplus_{|k|\leq K} C^1(Y)$. An element of $\mathscr{B}_K$ can therefore be seen as a function v on $X \times \mathbb{Z}$ such that v_k is C^1 for $|k| \leq K$, and $v_k = 0$ for $|k| > K$. We define then an operator R^t_n on $\mathscr{B}_K$ by

$$(R^t_n v)_k(x) = \sum_{(x',l)} \mathscr{K}^{t,n,\mathrm{exc}}_{(x,k)\to(x',l)} v_l(x'). \tag{4.9}$$

For $x \in Y$ and $|k| \leq K$, let also $(T^t_n v)_k(x) = \sum_{(x',l)\in Y\times[-K,K]} \mathscr{K}^{t,n}_{(x,k)\to(x',l)} v_l(x')$, (i.e., we consider all the returns of the "random walk" to $Y \times [-K, K]$ and not only the first ones). This means that $T^t_n v = 1_{Y\times[-K,K]} \hat{\mathscr{K}}^{t,n}(1_{Y\times[-K,K]} v)$ for $v \in \mathscr{B}_K$. By construction,

$$T^t_n = \sum_{p=1}^{\infty} \sum_{j_1+\cdots+j_p=n} R^t_{j_1} \cdots R^t_{j_p}. \tag{4.10}$$

This is a renewal equation that we already met in the course of the proof of exponential mixing. The main difference is that, for the mixing, each frequency was left invariant by the transfer operator, which means we only had to consider random walks on $X^{(N)}$ and excursions outside Y. Here, since there is also some interaction between the frequencies, we have to localize spatially (i.e., on Y) but also on the space of frequencies since the estimates given by Theorem 2.3 are not uniform in k.

The proof consists in understanding precisely the R^t_n's, deducing from that good estimates on T^t_n's, and using these to reconstruct precisely enough $\hat{\mathscr{U}}^n_t$. We thus need two technical tools: on the one hand, a tool on perturbations of renewal sequences of operators (we want estimates that are precise both with respect to n and t); and on the other hand, good estimates on the excursions outside of $Y \times [-K, K]$.

Before going on, let us give another expression of $\mathscr{K}^{t,n,\mathrm{exc}}$ that is needed later on by considering the successive returns to $Y \times \mathbb{Z}$. Let us define a function $\psi_Y : Y \times \mathbb{S}^1 \to \mathbb{R}$ by

$$\psi_Y(x, \omega) = \sum_{i=0}^{r(x)-1} \psi\Big(T^i x, \omega + \sum_{j=0}^{i-1} \phi(T^j x)\Big). \tag{4.11}$$

It is the function induced by ψ and $\mathscr{T}$ on the set $Y \times \mathbb{S}^1$. Let us denote by $S_n^Y \psi_Y$ the Birkhoff sums of ψ_Y for the map induced by $\mathscr{T}$ on $Y \times \mathbb{S}^1$. For $x, x' \in Y$ and $k, l \in \mathbb{Z}$, let $\mathscr{K}^{t,Y}_{(x,k)\to(x',l)} = 1_{T^{MN}x'=x} J^{(MN)}(x') e^{-ikS^Y_{MN}\phi_Y(x')} (e^{itS^Y_{MN}\psi_Y})_{k-l}(x')$, which corresponds to the "probability" (for the above random walk) of the first return in $Y \times \mathbb{Z}$. Considering the successive returns to $Y \times (\mathbb{Z} - [-K, K])$, for $x, x' \in Y$ and $k, l \in [-K, K]$ we get

$$\mathscr{K}^{t,n,\mathrm{exc}}_{(x,k)\to(x',l)} = \sum_{p\geq 0} \sum_{\substack{k_0=l,k_1,\dots,k_{p-1}\notin[-K,K],k_p=k \\ x_0=x',x_1,\dots,x_{p-1}\in Y,x_p=x \\ \sum_{i=0}^{p-1} r^{(MN)}(x_i)=n}} \mathscr{K}^{t,Y}_{(x_p,k_p)\to(x_{p-1},k_{p-1})} \cdots \mathscr{K}^{t,Y}_{(x_1,k_1)\to(x_0,k_0)}. \tag{4.12}$$

4.2. Perturbed renewal sequences of operators

Definition 4.1

Let $\mathscr{B}$ be a Banach space, and let R^t_j be operators acting on $\mathscr{B}$ for $j > 0$ and $t \in [-t_0, t_0]$ for some $t_0 > 0$. These operators form a *perturbed sequence of renewal operators with exponential decay* if the following conditions hold.

(1) The operators R^0_j form a renewal sequence of operators with exponential decay. In particular, we write P and μ for the associated spectral projection and coefficient, as in Definition 3.3.

(2) There exist $\delta > 0$ and $a, C > 0$ such that, for all $t, t' \in [-t_0, t_0]$ with $|t - t'| \leq a$, for any $j > 0$, $\|R^t_j - R^{t'}_j\| \leq C|t - t'|e^{-\delta j}$.

(3) Let us write $R(z, t) = \sum z^j R^t_j$ for $|z| < e^\delta$. For (z, t) close to $(1, 0)$, the operator $R(z, t)$ is a small perturbation of $R(1, 0)$. Therefore, it has an eigenvalue $\lambda(z, t)$ close to 1. We assume that, for some $\alpha > 0$, $\lambda(1, t) = 1 - \alpha t^2 + O(|t|^3)$.

We say that this sequence is *aperiodic* if, for any $(z, t) \in (\overline{\mathbb{D}} \times [-t_0, t_0]) - \{(1, 0)\}$, the operator $I - R(z, t)$ is invertible on $\mathscr{B}$.

THEOREM 4.2

Let R^t_j be a perturbed sequence of renewal operators with exponential decay. Let

$$T^t_n = \sum_{p=1}^{\infty} \sum_{j_1+\dots+j_p=n} R^t_{j_1} \cdots R^t_{j_p}. \tag{4.13}$$

Then there exist $\tau_0 \in (0, t_0)$, $\bar{\theta} < 1$, and $c, C > 0$ such that, for $t \in [-\tau_0, \tau_0]$, for $n > 0$, we have

$$\left\| T^t_n - \frac{1}{\mu}\left(1 - \frac{\alpha t^2}{\mu}\right)^n P \right\| \leq C\bar{\theta}^n + C|t|(1 - ct^2)^n. \tag{4.14}$$

Moreover, if R_j^t is aperiodic, one also has, for $|t| \in [\tau_0, t_0]$ and $n > 0$,

$$\|T_n^t\| \leq C\bar{\theta}^n. \tag{4.15}$$

Proof
If γ is a path around zero in $\mathbb{C}$, close enough to zero,

$$T_j^t = \frac{1}{2i\pi} \int_\gamma z^{-j-1}\big(I - R(z,t)\big)^{-1} \,\mathrm{d}z. \tag{4.16}$$

By analyticity, this equality holds true for any path γ around zero bounding a domain on which $I - R(z,t)$ is invertible for any z.

Let us first show (4.15) in the aperiodic case. Let $t \neq 0$. The operators $I - R(z,t)$ are invertible for any $z \in \overline{\mathbb{D}}$. Since invertible operators form an open set, there exist an open neighborhood I_t of t and $\varepsilon_t > 0$ such that $I - R(z,t')$ is invertible for $t' \in I_t$ and $|z| \leq e^{\varepsilon_t}$. Taking for γ the circle of radius e^{ε_t}, we obtain $\|T_j^{t'}\| \leq C(t)e^{-j\varepsilon_t}$. If $\tau > 0$, the compact set $[-t_0, -\tau] \cup [\tau, t_0]$ can be covered by a finite number of the intervals I_t, and we get the following: there exist $\delta_\tau > 0$ and $C_\tau > 0$ such that, for any $|t| \in [\tau, t_0]$, for any $j > 0$, $\|T_j^t\| \leq C_\tau e^{-j\delta_\tau}$. This proves (4.15) if we can choose τ so that (4.14) is satisfied.

For (4.14), we work in a neighborhood of $(z,t) = (1,0)$. There exist an open disk O around 1 and $\tau_0 > 0$, such that, for $(z,t) \in O \times [-\tau_0, \tau_0]$, the operator $R(z,t)$ has a unique eigenvalue $\lambda(z,t)$ close to 1. Let us also denote by $P(z,t)$ the corresponding spectral projection. These functions depend holomorphically on z and in a Lipschitz way on t.

We saw in the proof of Proposition 3.4 that $\lambda'(1,0) = \mu \neq 0$. Reducing O if necessary, we can therefore assume that $z \mapsto \lambda(z,0)$ is injective on O (and takes the value 1 only at $z = 1$).

When t converges to zero, the function $z \mapsto \lambda(z,t)$ converges uniformly to $z \mapsto \lambda(z,0)$ (with a speed $O(t)$). Since all of these functions are holomorphic, the derivatives converge uniformly with the same speed. In particular, $z \mapsto \lambda(z,t)$ takes the value 1 at a unique point $\rho(t)$ in O (if t is small enough) by Rouché's theorem. Moreover, $\rho(t) \to 1$ when $t \to 0$.

Let us establish an asymptotic expansion of $\rho(t)$. We have

$$\begin{aligned}
\lambda\big(\rho(t),t\big) - \lambda(1,t) &= \int_1^{\rho(t)} \lambda'(z,t)\,\mathrm{d}z \\
&= \int_1^{\rho(t)} \big(\lambda'(z,t) - \lambda'(1,0)\big)\,\mathrm{d}z + \lambda'(1,0)\big(\rho(t) - 1\big).
\end{aligned}$$

Moreover, $|\lambda'(z,t)-\lambda'(1,0)|\leq C(|z-1|+|t|)\leq C(|\rho(t)-1|+|t|)$. As $\lambda(\rho(t),t)-\lambda(1,t)=1-\lambda(1,t)=\alpha t^2+O(|t|^3)$, we obtain

$$\lambda'(1,0)\big(\rho(t)-1\big)=\alpha t^2+O(t^3)+O\big(|t||\rho(t)-1|\big)+O\big(|\rho(t)-1|^2\big). \quad (4.17)$$

As $\lambda'(1,0)=\mu\neq 0$, this yields $\rho(t)-1\sim\alpha t^2/\mu$. In particular, $\rho(t)-1=O(t^2)$. Putting this information back in the equation, we finally obtain

$$\rho(t)=1+\alpha t^2/\mu+O(t^3). \quad (4.18)$$

The operators $I-R(z,0)$ are invertible for $z\in\overline{\mathbb{D}}-O$. By continuity, $I-R(z,t)$ is invertible for any z in a neighborhood of this compact set, and t close enough to zero, say $t\in[-\tau_0,\tau_0]$. We can therefore choose a path γ around zero made of an arc of a circle of radius greater than 1 and the inner part of ∂O, satisfying (4.16) for $|t|\leq\tau_0$. We modify γ into a new path $\tilde\gamma$ by replacing the inner part of ∂O with its outer part. To obtain an analogue of (4.16), we should add the residue of $z^{-j-1}(I-R(z,t))^{-1}$ inside O. We have $(I-R(z,t))^{-1}=(1-\lambda(z,t))^{-1}P(z,t)+Q(z,t)$, where $Q(z,t)$ is holomorphic inside O (hence without residue). The only pole is located at $\rho(t)$, and we obtain

$$T_j^t=\frac{1}{2i\pi}\int_{\tilde\gamma}z^{-j-1}\big(I-R(z,t)\big)^{-1}\,\mathrm{d}z+\frac{1}{\lambda'(\rho(t),t)}P\big(\rho(t),t\big)\rho(t)^{-j-1}. \quad (4.19)$$

On $\tilde\gamma$, we have $|z|\geq e^{\delta_0}$ for some $\delta_0>0$. As $\|(I-R(z,t))^{-1}\|$ is uniformly bounded on $\tilde\gamma$, the integral term is $O(e^{-\delta_0 j})$. For the remaining term, we have $P(\rho(t),t)/\lambda'(\rho(t),t)=P(1,0)/\lambda'(1,0)+O(t)$. Making this substitution gives an error of $O(|t||\rho(t)|^{-j})=O(|t|(1-ct^2)^j)$ by (4.18). We get

$$\left\|T_j^t-\frac{1}{\mu}P\rho(t)^{-j-1}\right\|\leq Ce^{-j\delta_0}+C|t|(1-ct^2)^j. \quad (4.20)$$

Finally, if we replace $\rho(t)^{-j-1}$ with $(1-\alpha t^2/\mu)^j$, the error is bounded, thanks to (4.18), by

$$C(1-ct^2)^j\big((1+C|t|^3)^j-1\big)\leq C(1-ct^2)^j(1+C|t|^3)^j j|t|^3.$$

If t is small enough, $(1-ct^2)(1+C|t|^3)\leq(1-ct^2/2)$. Finally,

$$\begin{aligned} j|t|^3(1-ct^2/2)^j &\leq j|t|^3(1-ct^2/4)^j(1-ct^2/4)^j \\ &\leq |t|(1-ct^2/4)^j\cdot jt^2\exp(-cjt^2/4)\leq C|t|(1-ct^2/4)^j \end{aligned} \quad (4.21)$$

since the function $x\mapsto xe^{-cx/4}$ is bounded on $\mathbb{R}_+$. □

4.3. Estimates on the excursions

In this section, we fix an integer M, a constant $A > 1$, and a sequence $(\gamma_d)_{d\in\mathbb{Z}}$ with $\gamma_d \in (0, 1]$ and $\gamma_d = O(1/|d|^4)$ when $d \to \pm\infty$.

We then choose an integer K such that

$$\forall |d| > K/2, \quad \gamma_d \leq \frac{1}{(1+|d|)^{60/17}}, \tag{4.22}$$

and

$$K \geq K(A, M) \quad \text{given by Theorem 2.3}, \tag{4.23}$$

and

$$\forall n \geq 1, \quad 2^{Mn} - 1 - n/2 \geq 2^{Mn}/K. \tag{4.24}$$

Let $\underline{k} = (k_0, k_1, \ldots, k_j)$ be a sequence of integers. We say that this sequence is admissible if $|k_i| > K$ for any $i \in (0, j)$. We say that it is strongly admissible if, additionally, $|k_j| > K$. We denote by $d_i = k_i - k_{i-1}$ the successive differences.

LEMMA 4.3

Let $\underline{k} = (k_0, k_1, \ldots, k_{j_0})$ be a strongly admissible sequence. Let $\psi_1, \ldots, \psi_{j_0}$ be functions from Y to $\mathbb{C}$, and let $\varepsilon_1, \ldots, \varepsilon_{j_0}$ belong to $[0, 1]$. Assume that $\|\psi_i\|_{\mathscr{C}^{A,3\varepsilon}_{MN}} \leq \varepsilon_i \gamma_{d_i}$.

Let $v^0 : Y \to \mathbb{C}$. Define a sequence of functions v^i by induction, by $v^i = \mathscr{L}^{MN}_{k_i}(\psi_i v^{i-1})$. Then

$$\|v^{j_0}\|_{L^2} \leq \Big(\prod_{i=1}^{j_0} \varepsilon_i \gamma_{d_i}^{9/10}\Big) \theta^{100MNj_0} \|v^0\|_{C^1}. \tag{4.25}$$

Proof

We use the following virtual heights

$$\beta_i = \max\Big(|k_i|, \frac{|k_{i-1}|}{2^M} \cdots \frac{|k_0|}{2^{Mi}}\Big). \tag{4.26}$$

Their interest is that we are able to control by induction the Dolgopyat norms $\|v^i\|_{D_{\beta_i}}$. (This would not be possible for the norm D_{k_i} if the jumps d_i are too large.)

If $|k_i| \geq \beta_{i-1}/2^M$, we have $\beta_i = |k_i|$. Then, by Theorem 2.3 (and more precisely (2.7)),

$$\begin{aligned} \|v^i\|_{D_{\beta_i}} = \|\mathscr{L}_{k_i}^{MN}(\psi_i v^{i-1})\|_{D_{k_i}} &\leq \theta^{100MN} \|\psi_i\|_{\mathscr{C}_{MN}^{A,3\varepsilon}} \|v^{i-1}\|_{D_{2^M k_i}} \\ &\leq \theta^{100MN} \varepsilon_i \gamma_{d_i} \|v^{i-1}\|_{D_{\beta_{i-1}}}. \end{aligned}$$

Otherwise, $\beta_i = \beta_{i-1}/2^M > |k_i|$, and (using (2.8)) we have

$$\begin{aligned} \|v^i\|_{D_{\beta_i}} = \|\mathscr{L}_{k_i}^{MN}(\psi_i v^{i-1})\|_{D_{\beta_i}} &\leq \theta^{-MN} \|\psi_i\|_{\mathscr{C}_{MN}^{A,3\varepsilon}} \|v^{i-1}\|_{D_{2^M \beta_i}} \\ &\leq \theta^{-MN} \varepsilon_i \gamma_{d_i} \|v^{i-1}\|_{D_{\beta_{i-1}}}. \end{aligned} \tag{4.27}$$

In both cases, we have similar equations with a large gain or a small loss.

Let us show by induction on i that

$$\|v^i\|_{D_{\beta_i}} \leq \theta^{100MNi} \varepsilon_1 \cdots \varepsilon_i (\gamma_{d_1} \cdots \gamma_{d_i})^{9/10} \|v^0\|_{D_{k_0}}, \tag{4.28}$$

the result being clear for $i = 0$.

Assume that the result is proved up to $i-1$, and let us prove it for i. If $\beta_i = |k_i|$, then

$$\|v^i\|_{D_{\beta_i}} \leq \theta^{100MN} \varepsilon_i \gamma_{d_i} \|v^{i-1}\|_{D_{\beta_{i-1}}} \leq \theta^{100MN} \varepsilon_i (\gamma_{d_i})^{9/10} \|v^{i-1}\|_{D_{\beta_{i-1}}} \tag{4.29}$$

since $\gamma_d \leq 1$ for any $d \in \mathbb{Z}$. The inductive assumption concludes the proof.

If $\beta_i > |k_i|$, consider ι the last time before i for which $\beta_\iota = |k_\iota|$. Iterating (4.27) up to ι, we get

$$\|v^i\|_{D_{\beta_i}} \leq \varepsilon_i \cdots \varepsilon_{\iota+1} \gamma_{d_i} \cdots \gamma_{d_{\iota+1}} \theta^{-MN(i-\iota)} \|v^\iota\|_{D_{\beta_\iota}}. \tag{4.30}$$

Moreover, $\beta_i = \beta_\iota / 2^{M(i-\iota)}$, and $\beta_i > K$ since $\underline{k}$ is strongly admissible. Hence,

$$|d_{\iota+1}| + \cdots + |d_i| \geq |k_\iota - k_i| \geq (2^{M(i-\iota)} - 1)\beta_i \geq (2^{M(i-\iota)} - 1)K. \tag{4.31}$$

Write J for the set of indexes $a \in (\iota, i]$ for which $|d_a| > K/2$. Then $\sum_J |d_a| \geq (2^{M(i-\iota)} - 1 - (i-\iota)/2)K$. By (4.24), we therefore get $\sum_J |d_a| \geq 2^{M(i-\iota)}$. By (4.22), $\gamma_d \leq 1/(1+|d|)$ for any $|d| > K/2$. We obtain

$$\begin{aligned} (\gamma_{d_i} \cdots \gamma_{d_{\iota+1}})^{1/10} &\leq \prod_{a\in J} \gamma_{d_a}^{1/10} \leq \prod_{a \in J} \frac{1}{(1+|d_a|)^{1/10}} = \left(\frac{1}{\prod_{a\in J}(1+|d_a|)} \right)^{1/10} \\ &\leq \left(\frac{1}{\sum_{a\in J} |d_a|} \right)^{1/10} \leq 2^{-M(i-\iota)/10}. \end{aligned}$$

By Theorem 2.3, $\theta^{101N} \geq 2^{-1/10}$. As a consequence, $2^{-M(i-\iota)/10} \leq \theta^{101MN(i-\iota)}$. Hence, we obtain from (4.30)

$$\begin{aligned}\|v^i\|_{D_{\beta_i}} &\leq \theta^{-MN(i-\iota)}(\gamma_{d_i}\cdots\gamma_{d_{\iota+1}})^{1/10}\cdot\varepsilon_i\cdots\varepsilon_{\iota+1}(\gamma_{d_i}\cdots\gamma_{d_{\iota+1}})^{9/10}\,\|v^\iota\|_{D_{\beta_\iota}}\\ &\leq \theta^{100MN(i-\iota)}\cdot\varepsilon_i\cdots\varepsilon_{\iota+1}(\gamma_{d_i}\cdots\gamma_{d_{\iota+1}})^{9/10}\,\|v^\iota\|_{D_{\beta_\iota}}.\end{aligned}$$

Using the induction assumption at ι, we get (4.28) at i. This concludes the induction and the proof of (4.28).

From (4.28) at j_0, we obtain, in particular,

$$\|v^{j_0}\|_{L^2} \leq \theta^{100MNj_0}\varepsilon_1\cdots\varepsilon_{j_0}(\gamma_{d_1}\cdots\gamma_{d_{j_0}})^{9/10}\|v^0\|_{D_{k_0}}. \tag{4.32}$$

As $\|v^0\|_{D_{k_0}} \leq \|v^0\|_{C^1}$, this concludes the proof. □

LEMMA 4.4

There exists a constant C (depending on $M, A, \{\gamma_d\}, K$) satisfying the following property. Let $(k_0, k_1, \dots, k_j)$ be an admissible sequence. Let $\psi_1, \dots, \psi_j$ be functions from Y to $\mathbb{C}$, and let $\varepsilon_1, \dots, \varepsilon_j$ belong to $[0, 1]$. We assume that $\|\psi_i\|_{\mathscr{C}_{MN}^{A,3\varepsilon}} \leq \varepsilon_i\gamma_{d_i}$.

Let $v^0 : Y \to \mathbb{C}$. Define a sequence of functions v^i by induction, by $v^i = \mathscr{L}_{k_i}^{MN}(\psi_i v^{i-1})$. Then

$$\|v^j\|_{C^1} \leq C(1+k_0^2)\Big(\prod_{i=1}^{j}\varepsilon_i\gamma_{d_i}^{1/3}\Big)\theta^{30MNj}\|v^0\|_{C^1}. \tag{4.33}$$

Proof

We write $j_0 = j/2$ or $(j-1)/2$, depending on whether j is even or odd.

Let $\varphi_i = e^{-ik_i S_{MN}^Y\phi_Y}\psi_i$, so that $v^i = \mathscr{L}^{MN}(\varphi_i v^{i-1})$. We have $|\varphi_i(x)| \leq \varepsilon_i\gamma_{d_i}e^{3\varepsilon r^{(MN)}(x)}$ and, for $h \in \mathscr{H}_{MN}$,

$$\begin{aligned}\|D(\varphi_i\circ h)(x)\| &\leq \|D(\psi_i\circ h)(x)\| + |k_i|\|D(S_{MN}^Y\phi_Y\circ h)(x)\||\psi_i(hx)|\\ &\leq A\varepsilon_i\gamma_{d_i}e^{3\varepsilon r^{(MN)}(hx)} + C|k_i|r^{(MN)}(hx)\varepsilon_i\gamma_{d_i}e^{3\varepsilon r^{(MN)}(hx)}\\ &\leq C|k_i|\varepsilon_i\gamma_{d_i}e^{4\varepsilon r^{(MN)}(hx)}\end{aligned}$$

for some constant $C \geq 1$ depending only on M and A. Let $B = C\max|k_i|$; this shows that $\|\varphi_i\|_{\mathscr{C}_{MN}^{B,4\varepsilon}} \leq \varepsilon_i\gamma_{d_i}$.

We can apply (2.6) between the indexes 1 and j_0 to get

$$\begin{aligned}\|v^{j_0}\|_{C^1} &\leq C(\max|k_i|)\Big(\prod_{i=1}^{j_0}\varepsilon_i\gamma_{d_i}\Big)(\theta^{100MNj_0}\|v^0\|_{C^1}+\theta^{-MNj_0}\|v^0\|_{L^2})\\ &\leq C\theta^{-MNj_0}\Big(\prod_{i=1}^{j_0}\varepsilon_i\gamma_{d_i}\Big)(\max|k_i|)\|v^0\|_{C^1}.\end{aligned}$$

Applying (2.6) between the indexes j_0+1 and j, we obtain

$$\|v^j\|_{C^1}\leq C(\max|k_i|)\Big(\prod_{i=j_0+1}^{j}\varepsilon_i\gamma_{d_i}\Big)(\theta^{100MN(j-j_0)}\|v^{j_0}\|_{C^1}+\theta^{-MN(j-j_0)}\|v^{j_0}\|_{L^2}).$$

We use the bound on $\|v^{j_0}\|_{C^1}$ given by the previous equation and the bound on $\|v^{j_0}\|_{L^2}$ from Lemma 4.3 (if $j_0=0$, this lemma does not apply since the sequence (k_0) is not necessarily strongly admissible, but the estimate (4.25) is trivial in this case). Since $100MN(j-j_0)-MNj_0$ and $-MN(j-j_0)+100MNj_0$ are both at least $C+99MNj/2\geq C+40MNj$, we obtain

$$\begin{aligned}\|v^j\|_{C^1} &\leq C\Big(\prod_{i=1}^{j}\varepsilon_i\gamma_{d_i}\Big)\theta^{40MNj}(\max|k_i|)^2\|v^0\|_{C^1}\\ &\qquad+C\Big(\prod_{i=1}^{j_0}\varepsilon_i\gamma_{d_i}^{9/10}\Big)\Big(\prod_{i=j_0+1}^{j}\varepsilon_i\gamma_{d_i}\Big)(\max|k_i|)\theta^{40MNj}\|v^0\|_{C^1}\\ &\leq C\theta^{40MNj}(\max|k_i|)^2\Big(\prod_{i=1}^{j}\varepsilon_i\gamma_{d_i}^{9/10}\Big)\|v^0\|_{C^1}.\end{aligned}$$

Assume first that $\max|k_i|\leq 2(|k_0|+jK)$. As $\theta^{40MNj}j^2\leq C\theta^{30MNj}$, we obtain the conclusion of the lemma by bounding directly $\big(\prod_{i=1}^{j}\gamma_{d_i}\big)^{9/10}$ by $\big(\prod_{i=1}^{j}\gamma_{d_i}\big)^{1/3}$.

Assume now that $\max|k_i|>2(|k_0|+jK)$. We have $|k_0|+\sum|d_i|\geq\max|k_i|$. Denote by J the set of indexes ≥ 1 for which $|d_i|>K$. Then

$$\sum_{i\in J}|d_i|\geq\max|k_i|-|k_0|-jK\geq\max|k_i|/2. \tag{4.34}$$

By (4.22), $\gamma_d\leq 1/(1+|d|)^{60/17}$ for any $|d|>K$. We get

$$\Big(\prod\gamma_{d_i}\Big)^{17/30}\leq\Big(\frac{1}{\prod_{i\in J}(1+|d_i|)^{60/17}}\Big)^{17/30}\leq\Big(\frac{1}{\sum_{i\in J}|d_i|}\Big)^2\leq 4/(\max|k_i|)^2.$$

Finally,

$$(\max|k_i|)^2\Big(\prod_{i=1}^{j}\gamma_{d_i}\Big)^{9/10}=(\max|k_i|)^2\Big(\prod_{i=1}^{j}\gamma_{d_i}\Big)^{17/30}\cdot\Big(\prod_{i=1}^{j}\gamma_{d_i}\Big)^{1/3}\leq 4\Big(\prod_{i=1}^{j}\gamma_{d_i}\Big)^{1/3}.$$

This yields again the conclusion of the lemma. □

5. Proof of the local limit theorem

We fix a C^6 function $\psi : X\times\mathbb{S}^1\to\mathbb{R}$ with vanishing average, and a real number $t_0>0$. We study the operators $\hat{\mathscr{T}}_t := \hat{\mathscr{T}}(e^{it\psi}\cdot)$ for $|t|\leq t_0$. We first choose M, A, a sequence γ_d, and an integer K so that the results of Section 4.3 apply. All these choices depend on ψ and t_0.

5.1. Choosing the constants

Let ψ_Y be the function defined in (4.11). There exists a constant $C(\psi)$ such that $|S_n^Y\psi_Y(x,\omega)|\leq C(\psi)r^{(n)}(x)$. More generally, as $\mathscr{T}$ is an isometry in the fiber direction $\mathbb{S}^1$, we even have, for $0\leq j\leq 4$,

$$\left|\frac{\partial^j}{\partial\omega^j}S_n^Y\psi_Y(x,\omega)\right|\leq C(\psi)r^{(n)}(x). \tag{5.1}$$

In particular, for any $|t|\leq t_0$, we have

$$\left|\frac{\partial^4}{\partial\omega^4}e^{itS_n^Y\psi_Y(x,\omega)}\right|\leq C(t_0,\psi)r^{(n)}(x)^4. \tag{5.2}$$

Let us denote by $F_d^{(n,t)}$ the dth Fourier coefficient of $e^{itS_n^Y\psi_Y}$ in the circle direction. Making 4 integrations by parts in the circle direction and using the previous equation yields

$$|F_d^{(n,t)}(x)|\leq\frac{C(t_0,\psi)r^{(n)}(x)^4}{1+|d|^4}\leq\frac{C'(t_0,\psi)e^{\varepsilon r^{(n)}(x)}}{1+|d|^4}. \tag{5.3}$$

There also exists $C(n,t_0,\psi)$ such that, for any $h\in\mathscr{H}_n$,

$$\|D(F_d^{(n,t)}\circ h)(x)\|\leq C(n,t_0,\psi)\frac{e^{\varepsilon r^{(n)}(hx)}}{1+|d|^4}. \tag{5.4}$$

Once and for all, we fix an integer M such that

$$\theta^{20MN}\sum_{d\in\mathbb{Z}}\min\Big(1,\frac{C'(t_0,\psi)}{1+|d|^4}\Big)^{1/3}<\theta^{10MN} \tag{5.5}$$

and

$$\theta^{100MN} \sum_{d\in\mathbb{Z}} \min\Big(1, \frac{C'(t_0,\psi)}{1+|d|^4}\Big) < 1/4. \tag{5.6}$$

Let $\gamma_d = \min(1, C'(t_0,\psi)/(1+|d|^4))$. By (5.4), we can then choose a constant A such that

$$\|F_d^{(MN,t)}\|_{\mathscr{C}_{MN}^{A,\varepsilon}} \leq \gamma_d \tag{5.7}$$

for any $d \in \mathbb{Z}$. Finally, we choose K satisfying (4.22) – (4.24).

All the constants C that we consider through the remainder of Section 5 may depend on $M, A, \{\gamma_d\}, K$. We work on the space $X^{(MN)}$ with the map $U = U^{(MN)}$ to prove Theorem 1.12 for $t \in [-t_0, t_0]$. We freely use all the results that we proved in Section 3. Formally, we proved these results for $X^{(N)}$, but the same arguments hold verbatim in $X^{(MN)}$.

As in the proof of Theorem 1.7, we assume until the end of the proof that $d^{(MN)} = 1$ (i.e., that $U^{(MN)}$ is mixing). Only at the end of the proof do we give the modifications to be done to handle the general case.

5.2. The renewal process

As in Section 4.1, let us define a space $\mathscr{B}_K = \bigoplus_{|k|\leq K} C^1(Y)$ endowed with the norm of the supremum of the C^1 norms of the different components. We see an element v of $\mathscr{B}_K$ as a set of functions $(v_k)_{|k|\leq K}$, where v_k corresponds to frequency k, and then $\|v\|_{\mathscr{B}_K} = \sup_{|k|\leq K} \|v_k\|_{C^1}$. We also write $\|v\|_{C^0} = \sup \|v_k\|_{C^0}$.

For $z \in \mathbb{C}$, $t \in [-t_0, t_0]$, and $\underline{k} = (k_0, \dots, k_j)$ an admissible sequence, we formally define an operator $Q_{\underline{k}}^t(z)$ on $C^1(Y)$ as follows, where $d_i = k_i - k_{i-1}$:

$$Q_{\underline{k}}^t(z)v = \mathscr{L}_{k_j}^{MN}\big(z^{r^{(MN)}} F_{d_j}^{(MN,t)} \mathscr{L}_{k_{j-1}}^{MN} z^{r^{(MN)}} \cdots \mathscr{L}_{k_1}^{MN}(z^{r^{(MN)}} F_{d_1}^{(MN,t)} v)\dots\big). \tag{5.8}$$

Intuitively, this operator applies to a function of frequency k_0 and gives a function of frequency k_j. If $\mathscr{B}$ is a Banach space of functions from $Y \times \mathbb{Z}$ to $\mathbb{C}$, it is therefore more natural to consider an operator $\bar{Q}_{\underline{k}}^t(z)$ from $\mathscr{B}$ to $\mathscr{B}$, defined by $(\bar{Q}_{\underline{k}}^t(z)v)_k = 0$ (if $k \neq k_j$) and $(\bar{Q}_{\underline{k}}^t(z)v)_{k_j} = Q_{\underline{k}}^t(z)v_{k_0}$. This applies, for instance, if $\mathscr{B} = \mathscr{B}_K$ (and $|k_0| \leq K$, $|k_j| \leq K$). We occasionally use the operators $\bar{Q}_{\underline{k}}^t(z)$, but the technical estimates are formulated in terms of $Q_{\underline{k}}^t(z)$.

LEMMA 5.1

The operator $Q_{\underline{k}}^t(z)$ acts continuously on $C^1(Y)$ for any $t \in [-t_0, t_0]$ and any $|z| \leq e^{2\varepsilon}$, and its norm is bounded by $C(1+k_0^2)\theta^{20MNj} \prod_{i=1}^j \gamma_{d_i}^{1/3}$. Moreover, the map $z \mapsto Q_{\underline{k}}^t(z)$ is holomorphic from $\{|z| < e^{2\varepsilon}\}$ to $\mathrm{End}(C^1(Y))$, where $\mathrm{End}(C^1(Y))$ is the set of continuous linear operators on $C^1(Y)$.

There exist $a > 0$ and $C > 0$ such that, for all $|t - t'| \leq a$, for any admissible sequence $\underline{k}$, we have

$$\| Q^t_{\underline{k}}(z) - Q^{t'}_{\underline{k}}(z) \|_{\mathrm{End}(C^1(Y))} \leq C|t - t'|(1 + k_0^2)\theta^{20MNj} \prod_{i=1}^{j} \gamma_{d_i}^{1/3}. \tag{5.9}$$

Finally, if $|t| \leq a$,

$$\| Q^t_{\underline{k}}(z) \| \leq C(1 + k_0^2)(C|t|)^{\#\{i \mid d_i \neq 0\}} \theta^{20MNj} \prod_{i=1}^{j} \gamma_{d_i}^{1/3}. \tag{5.10}$$

Proof
To estimate the norm of $Q^t_{\underline{k}}(z)$, we use the estimate given by Lemma 4.4, taking $\varepsilon_i = 1$ and $\psi_i = z^{r^{(MN)}} F^{(MN,t)}_{d_i}$. If $|z| \leq e^{2\varepsilon}$, we have $\|\psi_i\|_{\mathscr{C}^{A,3\varepsilon}_{MN}} \leq \|F^{(MN,t)}_{d_i}\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \leq \gamma_{d_i}$. We obtain

$$\| Q^t_{\underline{k}}(z) \|_{\mathrm{End}(C^1(Y))} \leq C(1 + k_0^2)\Big(\prod_{i=1}^{j} \gamma_{d_i}^{1/3} \Big) \theta^{30MNj}. \tag{5.11}$$

If $|z| < e^{2\varepsilon}$, each function $\psi_i 1_{r^{(MN)} > n}$ tends to 0 in $\mathscr{C}^{A,3\varepsilon}_{MN}$ when n tends to infinity. As a consequence, $z \mapsto Q^t_{\underline{k}}(z)$ is a uniform limit of polynomials on any compact subset of $\{|z| < e^{2\varepsilon}\}$ and is therefore holomorphic there.

To prove the rest of the lemma, we use the following inequality (which can easily be proved by four integrations by parts): there exists $C > 0$ such that, for any $t, t' \in [-t_0, t_0]$ and for any $d \in \mathbb{Z}$,

$$\| F^{(MN,t)}_d - F^{(MN,t')}_d \|_{\mathscr{C}^{A,\varepsilon}_{MN}} \leq C|t - t'|\gamma_d. \tag{5.12}$$

To prove (5.9), let us write $Q^t_{\underline{k}}(z)v - Q^{t'}_{\underline{k}}(z)v$ as

$$\sum_{b=0}^{j} \mathscr{L}^{MN}_{k_j}\Big(z^{r^{(MN)}} F^{(MN,t)}_{d_j} \mathscr{L}^{MN}_{k_{j-1}} \cdots \mathscr{L}^{MN}_{k_b}\Big(z^{r^{(MN)}} (F^{(MN,t)}_{d_b} - F^{(MN,t')}_{d_b}) \mathscr{L}^{MN}_{k_{b-1}} \\ z^{r^{(MN)}} F^{(MN,t')}_{d_{b-1}} \mathscr{L}^{MN}_{k_{b-2}}\Big(\cdots \mathscr{L}^{MN}_{k_1}(z^{r^{(MN)}} F^{(MN,t')}_{d_1} v) \cdots \Big).$$

Fix b. To estimate the corresponding term in this equation, we again use Lemma 4.4. Let $\psi_i = z^{r^{(MN)}} F^{(MN,t)}_{d_i}$ for $i > b$, let $\psi_i = z^{r^{(MN)}} F^{(MN,t')}_{d_i}$ for $i < b$, and let $\psi_b = z^{r^{(MN)}}(F^{(MN,t)}_{d_b} - F^{(MN,t')}_{d_b})$. Let also $\varepsilon_i = 1$ for $i \neq b$. Then ψ_i, ε_i satisfy the assumptions of Lemma 4.4 for $i \neq b$. Finally, let $\varepsilon_b = C|t' - t|$ (where C is as in (5.12)). If t' is close enough to t, we have $\varepsilon_b \leq 1$, and the assumptions of Lemma 4.4 are again satisfied by (5.12).

Using this lemma, we obtain (after summation over b)

$$\|Q^{t}_{\underline{k}}(z)v - Q^{t'}_{\underline{k}}(z)v\|_{C^1} \leq C(j+1)|t'-t|(1+k_0^2)\Big(\prod_{i=1}^{j}\gamma_{d_i}^{1/3}\Big)\theta^{30MNj}\|v_{k_0}\|_{C^1}. \quad (5.13)$$

As $(j+1)\theta^{30MNj} \leq C\theta^{20MNj}$, we get (5.9).

Finally, to prove (5.10), note that $F_d^{(MN,0)} = 0$ if $d \neq 0$. As a consequence, (5.12) applied to $t' = 0$ gives $\|F_d^{(MN,t)}\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \leq C|t|\gamma_d$. We can therefore apply Lemma 4.4 to $\varepsilon_i = 1$ if $d_i = 0$ and $\varepsilon_i = C|t|$ if $d_i \neq 0$ to obtain (5.10). □

Let us then define formally an operator $R(z,t)$ on $\mathscr{B}_K$ by $R(z,t) = \sum \bar{Q}^t_{\underline{k}}(z)$, where we sum over all admissible sequences $\underline{k}$ with $|k_0| \leq K$ and $|k_j| \leq K$, that is,

$$\big(R(z,t)v\big)_k = \sum_{j=1}^{\infty} \sum_{\substack{k_0,k_1,\dots,k_{j-1} \\ |k_0|\leq K \\ \underline{k}=(k_0,k_1,\dots,k_{j-1},k) \text{ admissible}}} Q^t_{\underline{k}}(z)v_{k_0}. \quad (5.14)$$

The coefficient of z^n corresponds to considering the first returns to $Y \times [-K, K]$ after a time exactly n. By (4.12), this is exactly the operator R^t_n defined in (4.9). Using the estimates in Lemma 5.1, our next goal is to prove that the operators R^t_n satisfy the assumptions of Theorem 4.2. Indeed, this theorem thus provides us with a good estimate for T^t_n (defined in (4.10)), which is the main building block of $\hat{\mathscr{U}}^n_t$.

LEMMA 5.2

The formal series $R(z,t)$ defines a holomorphic function on the disk $|z| < e^{2\varepsilon}$, uniformly bounded in $t \in [-t_0, t_0]$. In particular, there exists $C > 0$ such that, for any $t \in [-t_0, t_0]$, for any $n \in \mathbb{N}$, and for any $v \in \mathscr{B}_K$, we have

$$\|R^t_n v\|_{\mathscr{B}_K} \leq Ce^{-n\varepsilon}\,\|v\|_{\mathscr{B}_K}.$$

Moreover,

$$\|R(z,t)v - R(z,t')v\|_{\mathscr{B}_K} \leq C|t-t'|\,\|v\|_{\mathscr{B}_K}. \quad (5.15)$$

In particular, for any $n \in \mathbb{N}$, for any $v \in \mathscr{B}_K$, we have

$$\|R^t_n v - R^{t'}_n v\|_{\mathscr{B}_K} \leq C|t-t'|e^{-n\varepsilon}\,\|v\|_{\mathscr{B}_K}.$$

Proof

As $\theta^{20MN} \sum_{d\in\mathbb{Z}} \gamma_d^{1/3} < 1$, the estimates given by Lemma 5.1 are summable. This directly implies the lemma. □

LEMMA 5.3

There exists a constant C such that, for any z with $|z| \le e^{2\varepsilon}$, for any $t \in [-t_0, t_0]$, and for any $v \in \mathscr{B}_K$, we have

$$\|R(z,t)v\|_{\mathscr{B}_K} \le \frac{1}{2}\|v\|_{\mathscr{B}_K} + C\|v\|_{C^0}. \tag{5.16}$$

Proof

Fix an integer P. We define a truncated series $R(z,t,P)$ by summing as in $R(z,t)$ along admissible sequences $\underline{k} = (k_0, k_1, \dots, k_j)$ but with the additional restrictions that $\sup |k_i| \le P$ and $j \le P$. When P tends to infinity, $R(z,t,P)$ converges (in norm) to $R(z,t)$, uniformly for $(z,t) \in \{|z| \le e^{2\varepsilon}\} \times [-t_0, t_0]$. We show that, for any $P \in \mathbb{N}$, there exists $C(P)$ such that

$$\|R(z,t,P)v\|_{\mathscr{B}_K} \le \frac{1}{3}\|v\|_{\mathscr{B}_K} + C(P)\|v\|_{C^0}. \tag{5.17}$$

This implies the desired result, by choosing a large enough P.

Let $\underline{k}$ be an admissible sequence of length $j > 0$. Iterating j times the equation (2.5) (applied to the functions $\psi_i = z^{r^{(MN)}} e^{-ik_i S^Y_{MN}\phi_Y} F^{(MN,t)}_{d_i}$), we obtain a constant $C(\underline{k})$ such that, for any $v \in C^1(Y)$, we have

$$\|Q^t_{\underline{k}}(z)v\|_{C^1} \le \theta^{100MNj}\left(\prod_{i=1}^{j} \gamma_{d_i}\right)\|v\|_{C^1} + C(\underline{k})\|v\|_{C^0}. \tag{5.18}$$

The operator $R(z,t,P)$ involves only a finite number of admissible sequences. Denoting by $C(P)$ the sum of $C(\underline{k})$ over these admissible sequences, we obtain, for any $v \in \mathscr{B}_K$,

$$\begin{aligned}
\|R(z,t,P)v\|_{\mathscr{B}_K} &\le \sum_{j=1}^{P} \theta^{100MNj}\left(\sum_{d\in\mathbb{Z}} \gamma_d\right)^j \|v\|_{\mathscr{B}_K} + C(P)\|v\|_{C^0} \\
&\le \frac{\theta^{100MN}\sum \gamma_d}{1-\theta^{100MN}\sum\gamma_d}\|v\|_{\mathscr{B}_K} + C(P)\|v\|_{C^0} \\
&\le \frac{1}{3}\|v\|_{\mathscr{B}_K} + C(P)\|v\|_{C^0}
\end{aligned}$$

by (5.6). □

COROLLARY 5.4
For any $t \in [-t_0, t_0]$ *and for any* $|z| \leq e^{2\varepsilon}$, *the operator* $R(z,t)$ *acting on* $\mathscr{B}_K$ *has an essential spectral radius bounded by* $1/2$.

Proof
This is a consequence of Theorem 2.1. □

Definition 5.5
Let $\psi : X \times \mathbb{S}^1 \to \mathbb{R}$ be a C^6 function. We say that it is continuously periodic if there exist $a > 0, \lambda > 0$, and $f : X \times \mathbb{S}^1 \to \mathbb{R}/\lambda\mathbb{Z}$ measurable such that $\psi = f - f \circ \mathscr{T} + a$ mod λ almost everywhere and f is continuous on $Y \times \mathbb{S}^1$. Otherwise, we say that ψ is continuously aperiodic.

Proposition 1.10 says that aperiodicity and continuous aperiodicity are equivalent. However, we prove this equivalence only at the end of our arguments. Until then, it is more convenient to work with the notion of continuous aperiodicity.

PROPOSITION 5.6
For any $z \in \overline{\mathbb{D}} - \{1\}$, *the operator* $I - R(z,0)$ *is invertible on* $\mathscr{B}_K$. *Moreover, if the function* ψ *is continuously aperiodic, the operator* $I - R(z,t)$ *is invertible on* $\mathscr{B}_K$ *for any* $(z,t) \in (\overline{\mathbb{D}} \times [-t_0, t_0]) - \{(1,0)\}$.

Proof
Let $|z| \leq 1$, and let $t \in [-t_0, t_0]$. If the operator $I - R(z,t)$ is not invertible, its kernel contains a nonzero function $v = (v_{-K}, \dots, v_K)$ by Corollary 5.4. Let us define a function v_k for $|k| > K$ by

$$v_k = \sum_{p=1}^{\infty} \sum_{\substack{\underline{k}=(k_0,k_1,\dots,k_{j-1},k) \text{ admissible} \\ |k_0| \leq K}} Q^t_{\underline{k}}(z) v_{k_0}.$$

Lemma 5.1 implies (after summation over the admissible sequences) that $\sum_{k \in \mathbb{Z}} \|v_k\|_{C^1} < \infty$. Moreover, for any $k \in \mathbb{Z}$,

$$v_k = \sum_{l \in \mathbb{Z}} \mathscr{L}^{MN}_k \big(z^{r^{(MN)}} F^{(MN,t)}_{k-l} v_l\big). \tag{5.19}$$

This equation is indeed a consequence of the construction of the v_k's if $|k| > K$, and of the fact that v is a fixed point of $R(z,t)$ if $|k| \leq K$.

Let us define a continuous function g on $Y \times \mathbb{S}^1$ by $g(x,\omega) = \sum_{k \in \mathbb{Z}} v_k(x) e^{ik\omega}$. As v is nonzero, g is also nonzero. The invariance equation (5.19) translates into the following for g:

$$\hat{\mathscr{U}}_Y(z^{r^{(MN)}} e^{itS^Y_{MN}\psi_Y} g) = g, \tag{5.20}$$

where $\hat{\mathscr{U}}_Y$ is the transfer operator associated to the map which is induced by $\mathscr{U} = \mathscr{U}^{(MN)}$ on Y. Lemma 2.2 yields $|z| = 1$ and $g \circ \mathscr{U}_Y = e^{itS^Y_{MN}\psi_Y} z^{r^{(MN)}} g$. Let us extend g to the whole space $X^{(MN)} \times \mathbb{S}^1$ by setting

$$g(x, j, \omega) = z^j g(x, 0, \omega) \exp\Big(it \sum_{k=0}^{j-1} \psi \circ \mathscr{U}^k(x, \omega)\Big). \tag{5.21}$$

This function is bounded (since g is bounded on Y), nonzero, and satisfies $g \circ \mathscr{U} = z e^{it\psi} g$.

If $t = 0$, we obtain $g \circ \mathscr{U} = zg$. But the map $\mathscr{U}$ is mixing (this was proved in Theorem 3.6 and in (3.41) for $\mathscr{U}^{(N)}$, and the same proof holds for $\mathscr{U}^{(MN)}$). As a consequence, $z = 1$.

If $t \neq 0$, let $f : X^{(MN)} \times \mathbb{S}^1 \to \mathbb{R}/2\pi\mathbb{Z}$ be the logarithm of g, and let a be such that $z = e^{-ia}$. Then $t\psi \circ \tilde{\pi}^{(MN)} = f \circ \mathscr{U} - f + a \mod 2\pi$, and f is continuous on $Y \times \mathbb{S}^1 \subset X^{(MN)} \times \mathbb{S}^1$ (we have reintroduced the projection $\tilde{\pi}^{(MN)}$ in the notation since we soon confront lifting problems). In general, f is not constant on the fibers of $\tilde{\pi}^{(MN)}$ and can therefore not be written as $\tilde{f} \circ \tilde{\pi}^{(MN)}$ in $\mathbb{R}/2\pi\mathbb{Z}$. However, since the fibers of $\tilde{\pi}^{(MN)}$ are countable, [G2, Theorem 1.4] shows that there exist λ of the form $2\pi/n$ for some integer n, and $\tilde{f} : X \times \mathbb{S}^1 \to \mathbb{R}/\lambda\mathbb{Z}$, such that $f = \tilde{f} \circ \tilde{\pi}^{(MN)} \bmod \lambda$ almost everywhere. As a consequence, $t\psi = \tilde{f} \circ \mathscr{T} - \tilde{f} + a \mod \lambda$ and $\tilde{f}$ has a continuous version on $Y \times \mathbb{S}^1$ (since this is the case for f). Hence, ψ is continuously periodic. □

LEMMA 5.7

The operator $R(1, 0)$ has a simple eigenvalue at 1. The corresponding spectral projection is given by $(Pv)_0 = \int_Y v_0 \, \mathrm{d}\mu_Y$, and $(Pv)_k = 0$ if $k \neq 0$. Denoting by $R'(z, t)$ the derivative with respect to z of $R(z, t)$, we have $PR'(1, 0)P = \mu^{(MN)}(X^{(MN)})P$.

Proof

We have $(R(1, 0)v)_k = \mathscr{L}^{MN}_k v_k$. It is therefore sufficient to know the spectral properties of the operators $\mathscr{L}^{MN}_k$ (for $|k| \leq K$) to conclude. For $k \neq 0$, these operators have a spectral radius less than 1, while for $k = 0$ there is a simple eigenvalue at 1, the corresponding eigenprojection being given by integration (as we saw in the proofs of Lemma 3.2 and Corollary 3.5). This yields the desired formula for P.

As $PR^0_jP = \mu_Y\{r^{(MN)} = j\}P$ for $j \geq 1$, we have

$$PR'(1, 0)P = \sum j\mu_Y\{r^{(MN)} = j\}P = \mu^{(MN)}(X^{(MN)})P \tag{5.22}$$

by Kac's formula. □

5.3. Estimate of the perturbed eigenvalue

In this section, we prove this theorem (which is necessary to apply Theorem 4.2).

THEOREM 5.8

Denote by $\lambda(1,t)$ the eigenvalue close to 1 of $R(1,t)$ for small t. Then

$$\lambda(1,t) = 1 - \mu^{(MN)}(X^{(MN)})\frac{\sigma^2 t^2}{2} + O(t^3), \tag{5.23}$$

where σ^2 is given by (1.11).

The proof of Theorem 5.8 takes the rest of Section 5.3. We write $R(t)$ and $\lambda(t)$ instead of $R(1,t)$ and $\lambda(1,t)$ since we consider only $z=1$.

Let f^t be the eigenfunction (in $\mathscr{B}_K$) of $R(t)$ for the eigenvalue $\lambda(t)$, normalized so that $\int f_0^t = 1$ (this is possible since $\int f_0^0 = 1$ and f^t converges to f^0 in $\mathscr{B}_K$). Note that $f^t = f^0 + O(t)$ and that $\lambda(t) = 1 + O(t)$ (since $R(t) = R(0) + O(t)$ and the simple isolated eigenvalues, as well as the corresponding eigenfunctions, depend in a Lipschitz way on the operator). Moreover, $f_0^0 = 1$, and $f_k^0 = 0$ for $k \neq 0$.

LEMMA 5.9

We have $\lambda(t) = 1 + O(t^2)$.

Proof

We have $\bigl(R(t)f^t\bigr)_0 = \sum Q_{\underline{k}}^t(1) f_{k_0}^t$, where the summation is over the admissible sequences $\underline{k} = (k_0,\dots,k_j)$, with $|k_0| \leq K$ and $k_j = 0$. If $j \geq 2$, there are at least two nonzero differences $d_i = k_i - k_{i-1}$, and the sum of the corresponding terms is therefore bounded by Ct^2 by (5.10). If $j=1$ but $k_0 \neq 0$, the difference is nonzero, which gives a $O(t)$ factor. As $f_{k_0}^t = O(t)$, the resulting term is therefore also $O(t^2)$. It remains $\bigl(R(t)f^t\bigr)_0 = Q_{(0,0)}^t(1) f_0^t + O(t^2)$. As $R(t)f^t = \lambda(t) f^t$ and $\int f_0^t = 1$, we obtain after integration

$$\begin{aligned}\lambda(t) &= \int_Y Q_{(0,0)}^t(1) f_0^t + O(t^2) = \int_Y \mathscr{L}^{MN}(F_0^{(MN,t)} f_0^t) + O(t^2)\\ &= \int_{Y\times\mathbb{S}^1} e^{itS_{MN}^Y \psi_Y(x,\omega)} f_0^t(x) + O(t^2).\end{aligned}$$

As $\int f_0^t = 1$, we get

$$\lambda(t) = 1 + \int (e^{itS_{MN}^Y\psi_Y} - 1)(f_0^t - 1) + \int (e^{itS_{MN}^Y\psi_Y} - 1) + O(t^2). \tag{5.24}$$

Since $f_0^t = f_0^0 + O(t) = 1 + O(t)$, the first integral is $O(t^2)$. For the second one,

$$\int (e^{itS^Y_{MN}\psi_Y} - 1) = it \int S^Y_{MN}\psi_Y + O(t^2) = MNit \int_{X\times\mathbb{S}^1} \psi + O(t^2) = O(t^2) \quad (5.25)$$

since $\int \psi = 0$. This finally yields $\lambda(t) = 1 + O(t^2)$. □

Define a function g_k on Y by $g_k(x) = \int S^Y_{MN}\psi_Y(x,\omega)e^{-ik\omega}\,\mathrm{d}\omega$.

LEMMA 5.10
The function g_k belongs to $\mathscr{C}^{1,\varepsilon}_{MN}$. Moreover, there exists a constant $C > 0$ such that, for any small enough t and for any $k \in \mathbb{Z}$,

$$\|F_k^{(MN,t)} - 1_{k=0} - itg_k\|_{\mathscr{C}^{1,\varepsilon}_{MN}} \le \frac{Ct^2}{1+k^4}. \quad (5.26)$$

Proof
Write

$$\begin{aligned} F_k^{(MN,t)}(x) - 1_{k=0} - itg_k(x) &= \int_{\mathbb{S}^1} \big(e^{itS^Y_{MN}\psi_Y(x,\omega)} - 1 - itS^Y_{MN}\psi_Y(x,\omega)\big)e^{-ik\omega}\,\mathrm{d}\omega \\ &= -t^2 \int_{v=0}^1 (1-v)\Big(\int_{\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)^2 \\ &\qquad \times e^{itS^Y_{MN}\psi_Y(x,\omega)v}e^{-ik\omega}\,\mathrm{d}\omega\Big)\,\mathrm{d}v. \end{aligned}$$

This gives (5.26) after four integrations by parts with respect to ω. □

LEMMA 5.11
For any $|k| \le K$, we have in $C^1(Y)$

$$f_k^t = f_k^0 + it\sum_{n=1}^{\infty} \mathscr{L}_k^{MNn}(g_k) + O(t^2). \quad (5.27)$$

Note that g_k belongs to $\mathscr{C}^{1,\varepsilon}_{MN}$, which implies that $\mathscr{L}_k^{MN}g_k \in C^1(Y)$ by Theorem 2.3. The series $\sum_{n\in\mathbb{N}} \mathscr{L}_k^{MNn}\mathscr{L}_k^{MN}g$ is therefore convergent in $C^1(Y)$: for $k \neq 0$, the spectral radius of $\mathscr{L}_k^{MN}$ on $C^1(Y)$ is less than 1, and the convergence is trivial. For $k = 0$, there is still exponential convergence for functions with zero average, which is the case of g_0 because $\int \psi = 0$.

Proof of Lemma 5.11
As $\lambda(t) = 1 + O(t^2)$, we have

$$\begin{aligned}\frac{f^t - f^0}{t} &= \frac{\lambda(t) f^t - f^0}{t} + O(t) = \frac{R(t) f^t - R(0) f^0}{t} + O(t)\\ &= \bigl(R(t) - R(0)\bigr)\frac{f^t - f^0}{t} + R(0)\frac{f^t - f^0}{t} + \frac{R(t) - R(0)}{t} f^0 + O(t).\end{aligned}$$

Since $R(t) - R(0) = O(t)$ and since $f^t - f^0 = O(t)$, we obtain, after moving $R(0)\frac{f^t - f^0}{t}$ to the left-hand side,

$$\bigl(I - R(0)\bigr)\frac{f^t - f^0}{t} = \frac{R(t) - R(0)}{t} f^0 + O(t). \tag{5.28}$$

The operator $R(0)$ simply acts by $(R(0)v)_k = \mathscr{L}_k^{MN} v_k$. Let us study $(R(t) f^0)_k = \sum_{\underline{k}} Q_{\underline{k}}^t(1)1$, where $\underline{k}$ is an admissible sequence beginning by 0 and ending by k. If the length of this admissible sequence is at least 2, there are two nonzero differences, and we obtain a term bounded by $O(t^2)$. Hence,

$$\bigl(R(t) f^0\bigr)_k = Q_{(0,k)}^t(1)1 + O(t^2) = \mathscr{L}_k^{MN}(F_k^{(MN,t)}) + O(t^2). \tag{5.29}$$

Applying Lemma 5.10 and using the fact that $\mathscr{L}_k^{MN}$ is continuous from $\mathscr{C}_{MN}^{1,\varepsilon}$ to $C^1(Y)$, we get, in $C^1(Y)$,

$$\bigl(R(t) f^0\bigr)_k = 1_{k=0} + it\mathscr{L}_k^{MN} g_k + O(t^2) = \bigl(R(0) f^0\bigr)_k + it\mathscr{L}_k^{MN} g_k + O(t^2). \tag{5.30}$$

Let $h_k = \sum_{n>0} \mathscr{L}_k^{MNn} g_k$. Denote by h the corresponding element in $\mathscr{B}_K$ so that the kth component of $(I - R(0))h$ is equal to $\mathscr{L}_k^{MN} g_k$. Equations (5.28) and (5.30) imply that

$$\bigl(I - R(0)\bigr)\Bigl(\frac{f^t - f^0}{t} - ih\Bigr) = O(t). \tag{5.31}$$

As $I - R(0)$ is invertible on the set of elements v of $\mathscr{B}_K$, with $\int v_0 = 0$, this shows that $(f^t - f^0)/t - ih = O(t)$, which is the desired conclusion. □

Let $\mathscr{U}_Y$ be the map induced by $\mathscr{U} = \mathscr{U}^{(MN)}$ on $Y \times \mathbb{S}^1$. The associated transfer operator $\hat{\mathscr{U}}_Y$ acts on each frequency k by $\mathscr{L}_k^{MN}$. From the spectral properties of the operators $\mathscr{L}_k^{MN}$, we obtain the convergence of the series

$$\begin{aligned}\tilde{\sigma}^2 &= \int_Y (S_{MN}^Y \psi_Y)^2 + 2\sum_{n=1}^{\infty}\int_Y S_{MN}^Y \psi_Y \cdot S_{MN}^Y \psi_Y \circ \mathscr{U}_Y^n\\ &= \int_Y (S_{MN}^Y \psi_Y)^2 + 2\sum_{n=1}^{\infty}\int_Y \hat{\mathscr{U}}_Y^n S_{MN}^Y \psi_Y \cdot S_{MN}^Y \psi_Y.\end{aligned} \tag{5.32}$$

LEMMA 5.12
We have $\lambda(t) = 1 - \tilde{\sigma}^2 t^2/2 + O(t^3)$.

Proof
Let us estimate $(R(t) f^t)_0$. We have

$$\big(R(t) f^t\big)_0 = \sum_{1\le|k|\le K} \sum_{\underline{k}=(k,k_1,\dots,k_{j-1},0)\text{ admissible}} Q^t_{\underline{k}}(1) f^t_k + \sum_{\underline{k}=(0,k_1,\dots,k_{j-1},0)\text{ admissible}} Q^t_{\underline{k}}(1) f^t_0.$$

In the first sum, $f^t_k = O(t)$. If there are two nonzero differences in the admissible sequence $\underline{k}$, we therefore obtain terms bounded by $O(t^3)$ by (5.10). In the second sum, we also get $O(t^3)$ unless there are at most two nonzero differences, which is possible only for the sequences $\underline{k} = (0,0)$ and $\underline{k} = (0,\ell,\dots,\ell,0)$, where ℓ is repeated a number of times, say j, and $|\ell| > K$. Hence,

$$\begin{aligned}\big(R(t) f^t\big)_0 = &\sum_{1\le|k|\le K} \mathscr{L}^{MN}(F^{(MN,t)}_{-k} f^t_k) + \mathscr{L}^{MN}(F^{(MN,t)}_0 f^t_0)\\ &+ \sum Q^t_{(0,\ell,\dots,\ell,0)}(1) f^t_0 + O(t^3).\end{aligned}$$

We have

$$Q^t_{(0,\ell,\dots,\ell,0)}(1)v = \mathscr{L}^{MN}\big(F^{(MN,t)}_{-\ell} \mathscr{L}^{MN}_\ell F^{(MN,t)}_0 \mathscr{L}^{MN}_\ell \cdots \mathscr{L}^{MN}_\ell (F^{(MN,t)}_\ell f^t_0)\cdots\big). \tag{5.33}$$

As there are two nonzero differences in these admissible sequences, the contribution of these terms to $R(t) f^t_0$ is $O(t^2)$. Moreover, $F^{(MN,t)}_0 = 1 + O(t)$. If we replace $F^{(MN,t)}_0$ by 1, we get an additional error of $O(t)$ in each term. It can be checked as in the proof of (5.9) that these errors are summable. In the same way, f^t_0 may be replaced by 1 since the error is $O(t)$. We get

$$\begin{aligned}\big(R(t) f^t\big)_0 = &\sum_{1\le|k|\le K} \mathscr{L}^{MN}(F^{(MN,t)}_{-k} f^t_k) + \mathscr{L}^{MN}(F^{(MN,t)}_0 f^t_0)\\ &+ \sum_{j>0}\sum_{|\ell|>K} \mathscr{L}^{MN}(F^{(MN,t)}_{-\ell} \mathscr{L}^{MNj}_\ell F^{(MN,t)}_\ell) + O(t^3).\end{aligned}$$

For $|\ell| > K$ and for $j > 0$, we have $\|\mathscr{L}^{MNj}_\ell v\|_{C^1} \le C(1+\ell^2)\theta^{30MNj}\|v\|_{\mathscr{C}^{1,\varepsilon}_{MN}}$ for any function v by Lemma 4.4. Hence, (5.26) enables us to replace $F^{(MN,t)}_\ell$ and $F^{(MN,t)}_{-\ell}$, respectively, with itg_ℓ and $itg_{-\ell}$, the additional errors being summable and giving a term of order $O(t^3)$. Using also the estimates on f^t_k of Lemma 5.11, we obtain

$$\begin{aligned}\big(R(t) f^t\big)_0 = &-t^2 \sum_{1\le|k|\le K}\sum_{n>0} \mathscr{L}^{MN}(g_{-k}\mathscr{L}^{MNn}_k g_k) + \mathscr{L}^{MN}(F^{(MN,t)}_0 f^t_0)\\ &- t^2 \sum_{|\ell|>K}\sum_{j>0} \mathscr{L}^{MN}(g_{-\ell}\mathscr{L}^{MNj}_\ell g_\ell) + O(t^3).\end{aligned}$$

To estimate $\mathscr{L}^{MN}(F_0^{(MN,t)} f_0^t)$, we write, in $\mathscr{C}^{1,\varepsilon}_{MN}$,

$$F_0^{(MN,t)}(x) = 1 + itg_0(x) - \frac{t^2}{2} \int_{\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)^2 \,\mathrm{d}\omega + O(t^3). \tag{5.34}$$

Consequently, by Lemma 5.11 and since $\int f_0^t = 1$, $\int g_0 = 0$, we have

$$\begin{aligned}\int_Y \mathscr{L}^{MN}(F_0^{(MN,t)} f_0^t) &= \int_Y F_0^{(MN,t)} f_0^t \\ &= 1 + \int_Y itg_0 f_0^t - \frac{t^2}{2}\int_Y \int_{\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)^2 f_0^t(x)\,\mathrm{d}\omega + O(t^3) \\ &= 1 - t^2 \sum_{n=1}^{\infty} \int_Y g_0 \mathscr{L}^{MNn} g_0 - \frac{t^2}{2}\int_{Y\times\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)^2 + O(t^3).\end{aligned}$$

Finally, as $\lambda(t) = \int_Y \lambda(t) f_0^t = \int_Y (R(t)f^t)_0$, we obtain

$$\lambda(t) = 1 - \frac{t^2}{2}\int_{Y\times\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)^2 - t^2 \sum_{k\in\mathbb{Z}} \sum_{n>0} \int_Y g_{-k}\mathscr{L}_k^{MNn} g_k + O(t^3), \tag{5.35}$$

and the sum is absolutely converging. To conclude the proof, it is therefore sufficient to show that, for any $n > 0$,

$$\sum_{k\in\mathbb{Z}} \int_Y g_{-k}\mathscr{L}_k^{MNn} g_k = \int_{Y\times\mathbb{S}^1} S^Y_{MN}\psi_Y \cdot S^Y_{MN}\psi_Y \circ \mathscr{U}_Y^n. \tag{5.36}$$

We have

$$\begin{aligned}\int_Y g_{-k}\mathscr{L}_k^{MNn} g_k &= \int g_{-k}\mathscr{L}^{MNn}\left(e^{-ik\sum_{j=0}^{n-1} S^Y_{MN}\phi_Y\circ U_Y^j} g_k\right) \\ &= \int_Y g_{-k}\circ U_Y^n e^{-ik\sum_{j=0}^{n-1} S^Y_{MN}\phi_Y\circ U_Y^j} g_k \\ &= \int_Y \left(\int_{\mathbb{S}^1} S^Y_{MN}\psi_Y(U_Y^n x,\tilde\omega)e^{ik\tilde\omega}\,\mathrm{d}\tilde\omega\right) e^{-ik\sum_{j=0}^{n-1} S^Y_{MN}\phi_Y\circ U_Y^j(x)} \\ &\qquad \times \left(\int_{\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)e^{-ik\omega}\,\mathrm{d}\omega\right)\mathrm{d}\mu_Y(x).\end{aligned}$$

Let $\omega' = \tilde\omega - \sum_{j=0}^{n-1} S^Y_{MN}\phi_Y\circ U_Y^j(x)$, so that the previous formula becomes

$$\begin{aligned}\int_Y g_{-k}\mathscr{L}_k^{MNn} g_k = \int_Y &\left(\int_{\mathbb{S}^1} S^Y_{MN}\psi_Y\circ\mathscr{U}_Y^n(x,\omega')e^{ik\omega'}\,\mathrm{d}\omega'\right) \\ &\times \left(\int_{\mathbb{S}^1} S^Y_{MN}\psi_Y(x,\omega)e^{-ik\omega}\,\mathrm{d}\omega\right)\mathrm{d}\mu_Y(x).\end{aligned} \tag{5.37}$$

For any $u, v \in L^2(Y \times \mathbb{S}^1)$, we have

$$\int_{Y\times\mathbb{S}^1} uv = \sum_{k\in\mathbb{Z}} \int_Y \left(\int_{\mathbb{S}^1} u(x,\omega')e^{ik\omega'}\,\mathrm{d}\omega'\right)\left(\int_{\mathbb{S}^1} v(x,\omega)e^{-ik\omega}\,\mathrm{d}\omega\right)\mathrm{d}\mu_Y(x), \quad (5.38)$$

where the series on the right-hand side converges absolutely. This is simply Parseval's equality in each fiber $\mathbb{S}^1$, integrated with respect to x. Together with (5.37), this yields (5.36) and concludes the proof of the lemma. □

LEMMA 5.13
We have $\tilde{\sigma}^2 = \mu^{(MN)}(X^{(MN)})\sigma^2$.

Together with Lemma 5.12, this lemma concludes the proof of Theorem 5.8.

Proof
We show that

$$\tilde{\sigma}^2 = \int_{X^{(MN)}\times\mathbb{S}^1} \psi^2\,\mathrm{d}(\mu^{(MN)}\otimes \mathrm{Leb}) + 2\sum_{n=1}^{\infty}\int_{X^{(MN)}\times\mathbb{S}^1} \psi\cdot\psi\circ\mathscr{U}^n\,\mathrm{d}(\mu^{(MN)}\otimes\mathrm{Leb}). \quad (5.39)$$

Since $\mu^{(MN)}$ projects on $\mu^{(MN)}(X^{(MN)})\tilde{\mu}$, this implies the result of the lemma.

It is easy to convince oneself of (5.39) by expanding the expression of $S^Y_{MN}\psi_Y$ in $\tilde{\sigma}^2$ and then gluing back together the different pieces to get the right member of (5.39). However, this process involves series which are *a priori* not convergent, which is a problem. We therefore do the computation in a different way, inspired by [G1, Proposition 4.8].

Let us define a function c on $X^{(MN)}\times\mathbb{S}^1$ by $c = \sum_{n=1}^{\infty}\hat{\mathscr{U}}^n(\psi)$. This series converges by Theorem 3.6, and defines a function belonging to $L^p(X^{(MN)}\times\mathbb{S}^1)$ for any p. Moreover, $c = \hat{\mathscr{U}}\psi + \hat{\mathscr{U}}c$. Let a be the restriction of c to Y. The previous equation implies that $a = \hat{\mathscr{U}}_Y S^Y_{MN}\psi_Y + \hat{\mathscr{U}}_Y a$. As a consequence, the function $\tilde{a} = a - \int a$ is equal to $\sum_{n=1}^{\infty}\hat{\mathscr{U}}^n_Y(S^Y_{MN}\psi_Y)$ (and this series is indeed converging, since $\int S^Y_{MN}\psi_Y = 0$). In particular,

$$\begin{aligned}\tilde{\sigma}^2 &= \int_{Y\times\mathbb{S}^1}(S^Y_{MN}\psi_Y)^2 + 2\int_{Y\times\mathbb{S}^1} S^Y_{MN}\psi_Y\cdot\tilde{a}\\ &= \int_{Y\times\mathbb{S}^1}(S^Y_{MN}\psi_Y)^2 + 2\int_{Y\times\mathbb{S}^1} S^Y_{MN}\psi_Y\cdot a. \quad (5.40)\end{aligned}$$

The explicit relationship between a and c then makes it possible to show (as in the proof of [G1, Proposition 4.8]) that this quantity is equal to $\int_{X^{(MN)}\times\mathbb{S}^1}(\psi^2 + 2\psi c)$, which proves (5.39) thanks to the definition of c. □

5.4. Reconstruction of $\hat{\mathscr{U}}_t^n$
Let us assume from now on that $\sigma^2 > 0$.

We proved in Sections 5.2 and 5.3 that the sequence R_n^t is a perturbed renewal sequence of operators with exponential decay in the sense of Definition 4.1 and that it is aperiodic if the function ψ itself is continuously aperiodic. We can therefore apply Theorem 4.2 and get the following estimate on T_n^t (defined in (4.10)).

PROPOSITION 5.14
Let P be the operator on $\mathscr{B}_K$ defined in Lemma 5.7. There exist $\tau_0 > 0$, $c > 0$, $C > 0$, and $\bar{\theta} < 1$ such that, for any $n \in \mathbb{N}$, for any $t \in [-\tau_0, \tau_0]$, and for any $v \in \mathscr{B}_K$, we have

$$\left\| T_n^t v - \frac{1}{\mu^{(MN)}(X^{(MN)})}\left(1 - \frac{\sigma^2 t^2}{2}\right)^n P v \right\|_{\mathscr{B}_K} \leq C\left(\bar{\theta}^n + |t|(1-ct^2)^n\right) \|v\|_{\mathscr{B}_K}. \quad (5.41)$$

Moreover, if ψ is continuously aperiodic, we also have, for any $|t| \in [\tau_0, t_0]$,

$$\|T_n^t v\|_{\mathscr{B}_K} \leq C\bar{\theta}^n \|v\|_{\mathscr{B}_K}. \quad (5.42)$$

We recall that T_n^t is also given by $T_n^t v = 1_{Y\times[-K,K]}\hat{\mathscr{K}}^{t,n}(1_{Y\times[-K,K]}v)$. As we have a good control on $\hat{\mathscr{K}}^t$ outside $Y \times [-K, K]$, the information given by Proposition 5.14 therefore makes it possible to reconstruct precisely $\hat{\mathscr{K}}^{t,n}$. As a first step, we estimate $P_n^t v := 1_{Y\times\mathbb{Z}}\hat{\mathscr{K}}^{t,n}(1_{Y\times\mathbb{Z}}v)$. As in Section 3.2, we thus define operators A_n^t, B_n^t, and C_n^t using the kernel $\mathscr{K}^t$ along trajectories of the "random walk" of length n, starting and ending in $Y \times \mathbb{Z}$, with the following additional restrictions. For the operator A_n^t, we only sum over the trajectories that enter in $Y \times [-K, K]$ after a time exactly n, for the operator B_n^t over the trajectories starting in $Y \times [-K, K]$ and staying out of it for the next n iterates , and for the operator C_n^t over the trajectories spending all their iterates outside of $Y \times [-K, K]$. Formally, for $n > 0$, we have

$$A_n^t v(x,k) = \sum_{p\geq 0} \sum_{\substack{k_0\in[-K,K], k_1,\dots,k_{p-1}, k_p = k\notin[-K,K] \\ x_0, x_1, \dots, x_{p-1}, x_p = x \\ \sum_{i=0}^{p-1} r^{(MN)}(x_i) = n}} \mathscr{K}^{t,Y}_{(x_p,k_p)\to(x_{p-1},k_{p-1})} \cdots \mathscr{K}^{t,Y}_{(x_1,k_1)\to(x_0,k_0)} v(x_0,k_0),$$

and B_n^t, C_n^t are defined in an analogous way.

By construction, the operator P_n^t satisfies

$$P_n^t = C_n^t + \sum_{a+i+b=n} A_a^t T_i^t B_b^t \quad (5.43)$$

as long as this expression makes sense. We therefore need to introduce different Banach spaces of functions from $Y \times \mathbb{Z}$ to $\mathbb{C}$ such that the operators A_n^t, B_n^t, and C_n^t are well defined between these spaces. In addition to $\mathscr{B}_K$, let us denote by $\mathscr{B}_Y^2$ the set of functions v from $Y \times \mathbb{Z}$ to $\mathbb{C}$ such that $\sum_{k\in\mathbb{Z}}(1+k^2)\,\|v_k\|_{C^1(Y)} < \infty$ with its canonical norm, and let us denote by $\mathscr{B}_Y^0$ the set of functions v from $Y \times \mathbb{Z}$ to $\mathbb{C}$ such that $\sum_{k\in\mathbb{Z}} \|v_k\|_{C^1(Y)} < \infty$. We consider A_a^t as an operator from $\mathscr{B}_K$ to $\mathscr{B}_Y^0$, B_b^t as an operator from $\mathscr{B}_Y^2$ to $\mathscr{B}_K$, and C_n^t as an operator from $\mathscr{B}_Y^2$ to $\mathscr{B}_Y^0$. Of course, it should be checked that these operators are bounded for these respective norms. This is done in the following lemma.

LEMMA 5.15
There exists $C > 0$ such that, for any $n \in \mathbb{N}^*$ and any $t \in [-t_0, t_0]$, we have

$$\|A_n^t\|_{\mathscr{B}_K\to\mathscr{B}_Y^0} \leq C|t|e^{-\varepsilon n}, \qquad \|B_n^t\|_{\mathscr{B}_Y^2\to\mathscr{B}_K} \leq C|t|e^{-\varepsilon n}, \qquad \|C_n^t\|_{\mathscr{B}_Y^2\to\mathscr{B}_Y^0} \leq Ce^{-\varepsilon n}. \tag{5.44}$$

Proof
Let us start with A_n^t. If $\underline{k} = (k_0, \ldots, k_j)$ is an admissible sequence, we have defined an operator $\bar{Q}_{\underline{k}}^t(z)$ in Section 5.2 by $(\bar{Q}_{\underline{k}}^t(z)v)_k = 0$ if $k \neq k_j$, and $(\bar{Q}_{\underline{k}}^t(z)v)_{k_j} = Q_{\underline{k}}^t(z)v_{k_0}$. We define an operator $A(z,t)$ from $\mathscr{B}_K$ to $\mathscr{B}_Y^0$ by

$$A(z,t) = \sum_{j=1}^{\infty} \sum_{\substack{\underline{k}=(k_0,k_1,\ldots,k_{j-1},k_j)\text{ admissible} \\ |k_0|\leq K, |k_j|>K}} \bar{Q}_{\underline{k}}^t(z). \tag{5.45}$$

By construction, A_n^t is the coefficient of z^n in this series. Moreover, summing the estimates of Lemma 5.1 over admissible sequences with $|k_0| \leq K$ and $|k_j| > K$, we obtain that $A(z,t)$ is holomorphic on the disk $\{|z| < e^{2\varepsilon}\}$ (as a function from $\mathscr{B}_K$ to $\mathscr{B}_Y^0$). Summing the estimates (5.10) for small t, we also get that $A(z,t)$ is bounded by $C|t|$ (since the number of differences in such an admissible sequence is at least 1). As a consequence, $A(z,t)$ is bounded by $C|t|$ for $t \in [-t_0, t_0]$ since this inequality is trivial outside of a neighborhood of zero. Thus, the coefficient of z^n in $A(z,t)$ decays at least like $C|t|e^{-\varepsilon n}$. This concludes the proof of the estimate of A_n^t.

For B_n^t, we argue in the same way, using the fact that it is the coefficient of z^n in the series

$$\sum_{j=1}^{\infty} \sum_{\substack{\underline{k}=(k_0,k_1,\ldots,k_{j-1},k_j)\text{ admissible} \\ |k_0|>K, |k_j|\leq K}} \bar{Q}_{\underline{k}}^t(z). \tag{5.46}$$

As $\|Q_{\underline{k}}^t(z)\|_{C^1(Y)\to C^1(Y)} \leq C|t|(1+k_0^2)\theta^{20MNj}\prod_{i=1}^j \gamma_{d_i}^{1/3}$ by Lemma 5.1, we also have

$$\|\bar{Q}_{\underline{k}}^t(z)\|_{\mathscr{B}_Y^2\to\mathscr{B}_K} \leq C|t|\theta^{20MNj}\prod_{i=1}^{j}\gamma_{d_i}^{1/3}. \tag{5.47}$$

Since this quantity is summable with respect to $\underline{k}$, the series (5.46) is holomorphic on the disk $\{|z| < e^{2\varepsilon}\}$ and bounded by $C|t|$. We conclude as above.

Finally, C_n^t is the coefficient of z^n in the series

$$\sum_{j=1}^{\infty} \sum_{\substack{\underline{k}=(k_0,k_1,\dots,k_{j-1},k_j) \text{ admissible} \\ |k_0|>K, |k_j|>K}} \bar{Q}_{\underline{k}}^t(z), \tag{5.48}$$

which defines a holomorphic function from $\mathscr{B}_Y^2$ to $\mathscr{B}_Y^0$ in the disk $\{|z| < e^{2\varepsilon}\}$ (by summing the estimates of Lemma 5.1). This yields the desired estimate for C_n^t. □

We have defined a projection P on $\mathscr{B}_K$, which can be extended to an operator from $\mathscr{B}_Y^2$ to $\mathscr{B}_Y^0$, as follows: $(Pv)_k = 0$ if $k = 0$, and $(Pv)_0 = \int_Y v_0 \, \mathrm{d}\mu_Y$.

COROLLARY 5.16

There exist constants $\tau_0 > 0$, $c > 0$, $C > 0$, and $\bar{\theta} < 1$ such that, for any $n \in \mathbb{N}$, $t \in [-\tau_0, \tau_0]$, and $v \in \mathscr{B}_Y^2$, we have

$$\left\| P_n^t v - \frac{1}{\mu^{(MN)}(X^{(MN)})}\left(1 - \frac{\sigma^2 t^2}{2}\right)^n P v \right\|_{\mathscr{B}_Y^0} \le C\left(\bar{\theta}^n + |t|(1-ct^2)^n\right) \|v\|_{\mathscr{B}_Y^2} . \tag{5.49}$$

Moreover, if ψ is continuously aperiodic, one also has, for any $|t| \in [\tau_0, t_0]$,

$$\|P_n^t v\|_{\mathscr{B}_Y^0} \le C\bar{\theta}^n \|v\|_{\mathscr{B}_Y^2} . \tag{5.50}$$

Proof

We write $P_n^t = A_0^t T_n^t B_0^t + C_n^t + \sum_{a+i+b=n,\ i<n} A_a^t T_i^t B_b^t$ as an operator from $\mathscr{B}_Y^2$ to $\mathscr{B}_Y^0$. The term $A_0^t T_n^t B_0^t$ gives the desired asymptotics by Proposition 5.14 (and since A_0^t and B_0^t are simply trivial extension and restriction operators). The term C_n^t is $O(\bar{\theta}^n)$ by Lemma 5.15. Hence, we should estimate the sum $\sum_{a+i+b=n,\ i<n} A_a^t T_i^t B_b^t$ whose norm is bounded by

$$C|t| \sum_{a+i+b=n} e^{-\varepsilon a}\left(\bar{\theta}^i + (1-ct^2)^i\right)e^{-\varepsilon b}, \tag{5.51}$$

again by Lemma 5.15 and Proposition 5.14. The term $\sum e^{-\varepsilon a}\bar{\theta}^i e^{-\varepsilon b}$ is exponentially small in n, while the remaining term is bounded by

$$\begin{aligned} |t| \sum_{i+j=n} (j+1)e^{-\varepsilon j}(1-ct^2)^i &\le C|t|(1-ct^2)^n \sum_{j=0}^{n} \left((1-ct^2)^{-1}e^{-\varepsilon}\right)^j \\ &\le \frac{C|t|(1-ct^2)^n}{1-(1-ct^2)^{-1}e^{-\varepsilon}}. \end{aligned}$$

This is bounded by $C|t|(1-ct^2)^n$ if t is small enough.

When ψ is continuously aperiodic, the equation (5.50) is proved in the same way by combining (5.42) and Lemma 5.15. $\square$

The next step in the reconstruction of $\hat{\mathscr{K}}^{t,n}$ is to understand $\tilde{P}_n^t v := 1_{Y\times\mathbb{Z}}\hat{\mathscr{K}}^{t,n}(v)$. We let this operator act on the space $\mathscr{B}^3$ of functions v from $X^{(MN)}\times\mathbb{Z}$ to $\mathbb{C}$ such that $\sum_{k\in\mathbb{Z}}(1+|k|^3)\,\|v_k\|_{C^1(X^{(MN)})}<\infty$, and we let it take its values in $\mathscr{B}_Y^0$. Let us also define an operator $\tilde{P}$ from $\mathscr{B}^3$ to $\mathscr{B}_Y^0$ by $(\tilde{P}v)_k=0$ for $k\neq 0$, and $(\tilde{P}v)_0=\int_{X^{(MN)}} v_0\,\mathrm{d}\tilde{\mu}^{(MN)}$ (recall that $\tilde{\mu}^{(MN)}$ is a probability measure on $X^{(MN)}$ whose restriction to Y is $\mu_Y/\mu^{(MN)}(X^{(MN)})$).

PROPOSITION 5.17
There exist constants $\tau_0>0$, $c>0$, $C>0$, and $\bar{\theta}<1$ such that, for any $n\in\mathbb{N}$, for any $t\in[-\tau_0,\tau_0]$, and for any $v\in\mathscr{B}^3$, we have

$$\left\|\tilde{P}_n^t v-\left(1-\frac{\sigma^2t^2}{2}\right)^n\tilde{P}v\right\|_{\mathscr{B}_Y^0}\le C\left(\bar{\theta}^n+|t|(1-ct^2)^n\right)\|v\|_{\mathscr{B}^3}. \tag{5.52}$$

Moreover, if ψ is continuously aperiodic, one also has, for any $|t|\in[\tau_0,t_0]$,

$$\|\tilde{P}_n^t v\|_{\mathscr{B}_Y^0}\le C\bar{\theta}^n\,\|v\|_{\mathscr{B}^3}. \tag{5.53}$$

Proof
Let us define an operator D_n^t, which corresponds to considering the trajectories of the “random walk” starting from $Y\times\mathbb{Z}$ and staying outside of $Y\times\mathbb{Z}$ during a time n, so that $\tilde{P}_n^t=\sum_{i+j=n}P_i^tD_j^t$. Formally, for $x\in Y$, we have

$$D_n^t v(x,k)=\sum_{\substack{k_0,\dots,k_n=k\\ x_0,\dots,x_n=x\\ x_i\notin Y\text{ for }0\le i<n}}\mathscr{K}^t_{(x_n,k_n)\to(x_{n-1},k_{n-1})}\cdots\mathscr{K}^t_{(x_1,k_1)\to(x_0,k_0)}v(x_0,k_0). \tag{5.54}$$

We first study D_n^t as an operator from $\mathscr{B}^3$ to $\mathscr{B}_Y^2$. As the dynamics of U between two returns to Y is trivial, D_n^t can be explicitly described as follows. Recall that a point x in $X^{(MN)}$ is a pair (y,i) where $y\in Y$ and $i<r^{(MN)}(y)$. The preimages of $(x,0)$ under U^n which do not enter Y in between are exactly the points $(hx, r^{(MN)}(hx)-n)$, where $h\in\mathscr{H}_{MN}$ is an inverse branch of T_Y^{MN} whose return time $r^{(MN)}\circ h$ is greater than n. Let $v\in\mathscr{B}^3$. For $k,l\in\mathbb{Z}$, let us define a function $v_{k,l}^n$ on Y by

$$v_{k,l}^n(y)=1_{r^{(MN)}(y)>n}v_l\big(y,r^{(MN)}(y)-n\big)e^{-ikS_n\phi(y,r^{(MN)}(y)-n)}(e^{itS_n\psi})_{k-l}\big(y,r^{(MN)}(y)-n\big).$$

Here, $(y,r^{(MN)}(y)-n)$ is a point in $X^{(MN)}$, $e^{-ikS_n\phi}$ is a function on $X^{(MN)}$, and $(e^{itS_n\psi})_{k-l}$ is the $k-l$th Fourier coefficient (in the ω direction) of the function $e^{itS_n\psi}$

on $X^{(MN)} \times \mathbb{S}^1$, so it is also a function on $X^{(MN)}$. We have defined $v_{k,l}^n$ so that $D_n^t v(x,k) = \sum_l \mathscr{L}^{MN} v_{k,l}^n(x)$.

Let us now estimate $\|D_n^t v\|_{\mathscr{B}_Y^2}$ in terms of $\|v\|_{\mathscr{B}^3}$. As ψ belongs to $C^{5,1}$, the $k-l$th Fourier coefficient of $e^{itS_n\psi}$ is bounded by $Cn^5/(1+|k-l|^5)$. As $r^{(MN)}(x) > n$, we get

$$|v_{k,l}^n(x)| \leq C \|v_l\|_{C^0} \frac{n^5}{1+|k-l|^5} \leq C \|v_l\|_{C^0} e^{-\varepsilon n} \frac{e^{2\varepsilon r^{(MN)}(x)}}{1+|k-l|^5}, \tag{5.55}$$

and, for any inverse branch h, we get

$$\begin{aligned} \|D(v_{k,l}^n \circ h)\|_{C^0} &\leq C \|v_l\|_{C^1} (1+|k|) n \frac{n^5}{1+|k-l|^5} \\ &\leq C \|v_l\|_{C^1} (1+|k|) e^{-\varepsilon n} \frac{e^{2\varepsilon r^{(MN)}(x)}}{1+|k-l|^5}. \end{aligned} \tag{5.56}$$

As a consequence,

$$\|v_{k,l}^n\|_{\mathscr{C}_{MN}^{1,2\varepsilon}} \leq \frac{C(1+|k|)}{1+|k-l|^5} \|v_l\|_{C^1} e^{-\varepsilon n}. \tag{5.57}$$

By Theorem 2.3, $\|\mathscr{L}^{MN} v_{k,l}^n\|_{C^1(Y)} \leq C \|v_{k,l}^n\|_{\mathscr{C}_{MN}^{1,2\varepsilon}}$. Finally,

$$\|D_n^t v\|_{\mathscr{B}_Y^2} = \sum_k (1+|k|^2) \|(D_n^t v)_k\|_{C^1(Y)} \leq C e^{-\varepsilon n} \sum_{k,l} \frac{1+|k|^3}{1+|k-l|^5} \|v_l\|_{C^1}. \tag{5.58}$$

If l is fixed,

$$\sum_k \frac{1+|k|^3}{1+|k-l|^5} = \sum_j \frac{1+|j+l|^3}{1+|j|^5} \leq C \sum_j \frac{1+|j|^3+|l|^3}{1+|j|^5} \leq C(1+|l|^3). \tag{5.59}$$

Consequently,

$$\|D_n^t v\|_{\mathscr{B}_Y^2} \leq C e^{-\varepsilon n} \|v\|_{\mathscr{B}^3}. \tag{5.60}$$

In the equality $\tilde{P}_n^t v = \sum_{i+j=n} P_i^t D_j^t v$, let us replace P_i^t with $(1-\sigma^2 t^2/2)^i P/\mu^{(MN)}(X^{(MN)}) + E_i^t$, where E_i^t is an error term. The control of E_i^t given by Corollary 5.16, combined with the computation made at the end of the proof of this lemma, gives

$$\sum_{i+j=n} \|E_i^t D_j^t\|_{\mathscr{B}^3 \to \mathscr{B}_Y^0} \leq C \sum_{i+j=n} \left(\bar{\theta}^i + |t|(1-ct^2)^i\right) e^{-\varepsilon j} \leq C\left(\bar{\theta}^n + |t|(1-ct^2)^n\right). \tag{5.61}$$

Hence there is only one term left to be estimated in $\tilde{P}_n^t v$, with frequency zero, given by

$$I_n^t := \frac{1}{\mu^{(MN)}(X^{(MN)})} \sum_{i+j=n} \left(1 - \frac{\sigma^2 t^2}{2}\right)^i \int_Y (D_j^t v)_0 \, \mathrm{d}\mu_Y. \tag{5.62}$$

For all $u, v \in \mathbb{R}$ holds $|e^u - e^v| \leq |u - v| e^{\max(u,v)}$. As $|\int_Y (D_j^t v)_0| \leq C e^{-\varepsilon j} \|v\|_{\mathscr{B}^3}$, we obtain

$$\begin{aligned}
&\Big| \sum_{j=0}^n \left(1 - \frac{\sigma^2 t^2}{2}\right)^{n-j} \int_Y (D_j^t v)_0 \, \mathrm{d}\mu_Y - \left(1 - \frac{\sigma^2 t^2}{2}\right)^n \sum_{j=0}^n \int_Y (D_j^t v)_0 \, \mathrm{d}\mu_Y \Big| \\
&\qquad \leq C \left(1 - \frac{\sigma^2 t^2}{2}\right)^n \sum_{j=0}^n \left| \left(1 - \frac{\sigma^2 t^2}{2}\right)^{-j} - 1 \right| e^{-\varepsilon j} \|v\|_{\mathscr{B}^3} \\
&\qquad \leq C \left(1 - \frac{\sigma^2 t^2}{2}\right)^n \sum_{j=0}^n j \left| \log \left(1 - \frac{\sigma^2 t^2}{2}\right) \right| \left(1 - \frac{\sigma^2 t^2}{2}\right)^{-j} e^{-\varepsilon j} \|v\|_{\mathscr{B}^3} \\
&\qquad \leq C t^2 \left(1 - \frac{\sigma^2 t^2}{2}\right)^n \|v\|_{\mathscr{B}^3} .
\end{aligned}$$

Let us define a function f on $X^{(MN)} \times \mathbb{S}^1$ by $f(x, \omega) = \sum_k v_k(x) e^{ik\omega}$. If $Z_j \subset X^{(MN)}$ denotes the set of points in $X^{(MN)}$ which enter into Y after exactly j iterates, we have

$$\int_Y (D_j^t v)_0 \, \mathrm{d}\mu_Y = \int_{Z_j \times \mathbb{S}^1} f e^{itS_j \psi} \, \mathrm{d}(\mu^{(MN)} \otimes \mathrm{Leb}). \tag{5.63}$$

Since the measure of Z_j decays exponentially fast,

$$\Big| \int_Y (D_j^t v)_0 \, \mathrm{d}\mu_Y - \int_{Z_j \times \mathbb{S}^1} f \, \mathrm{d}(\mu^{(MN)} \otimes \mathrm{Leb}) \Big| \leq C \int_{Z_j \times \mathbb{S}^1} |t| j \, \|f\|_{C^0} \leq C |t| \bar{\theta}^j \|v\|_{\mathscr{B}^3} . \tag{5.64}$$

Finally,

$$\begin{aligned}
&\Big| \sum_{j=0}^n \int_{Z_j \times \mathbb{S}^1} f \, \mathrm{d}(\mu^{(MN)} \otimes \mathrm{Leb}) - \int_{X^{(MN)} \times \mathbb{S}^1} f \, \mathrm{d}(\mu^{(MN)} \otimes \mathrm{Leb}) \Big| \\
&\qquad \leq C \|f\|_{C^0} \sum_{j=n+1}^{\infty} \mu^{(MN)}(Z_j) \leq C \|v\|_{\mathscr{B}^3} \bar{\theta}^n.
\end{aligned}$$

Combining these different estimates, we obtain

$$
\begin{aligned}
I_n^t &= \Big(1-\frac{\sigma^2 t^2}{2}\Big)^n \frac{1}{\mu^{(MN)}(X^{(MN)})} \int_{X^{(MN)}\times\mathbb{S}^1} f \, \mathrm{d}(\mu^{(MN)} \otimes \mathrm{Leb}) + O\big(\bar\theta^n + |t|(1-ct^2)^n\big) \\
&= \Big(1-\frac{\sigma^2 t^2}{2}\Big)^n \int_{X^{(MN)}} v_0 \, \mathrm{d}\tilde\mu^{(MN)} + O\big(\bar\theta^n + |t|(1-ct^2)^n\big).
\end{aligned}
$$

This proves (5.52). Finally, (5.53) is proved in the same way, by using (5.50). □

Let $\hat{\mathscr{U}}_t$ denote the operator acting on functions on $X^{(MN)} \times \mathbb{S}^1$ by $\hat{\mathscr{U}}_t(v) = \hat{\mathscr{U}}(e^{it\psi} v)$, where $\hat{\mathscr{U}}$ is the transfer operator associated to $\mathscr{U}$.

THEOREM 5.18
Assume that $\sigma^2 > 0$. Then there exist constants $\tau_0 > 0$, $c > 0$, $C > 0$, and $\bar\theta < 1$ such that, for any $C^{5,1}$ function $v : X^{(MN)} \times \mathbb{S}^1 \to \mathbb{C}$, for any $n \in \mathbb{N}$, for any $t \in [-\tau_0, \tau_0]$, and for any $(x, \omega) \in X^{(MN)} \times \mathbb{S}^1$ such that $h(x) \le n/2$, we have

$$
\begin{aligned}
&\Big|\hat{\mathscr{U}}_t^n v(x,\omega) - \Big(1-\frac{\sigma^2 t^2}{2}\Big)^n \int v \, \mathrm{d}(\tilde\mu^{(MN)} \otimes \mathrm{Leb})\Big| \\
&\qquad \le C\big(1+h(x)\big)\big(\bar\theta^n + |t|(1-ct^2)^n\big) \|v\|_{C^{5,1}} . \qquad (5.65)
\end{aligned}
$$

Moreover, if ψ is continuously aperiodic, we also have, for any $|t| \in [\tau_0, t_0]$ and for any (x, ω) with $h(x) \le n/2$,

$$
|\hat{\mathscr{U}}_t^n v(x,\omega)| \le C\bar\theta^n \|v\|_{C^{5,1}} . \qquad (5.66)
$$

Note that this theorem implies Theorem 3.6, taking simply $t = 0$ (and a different value of $\bar\theta$).

Proof
Define w in $\mathscr{B}^3$ by $w(x,k) = \int_{\mathbb{S}^1} v(x,\omega) e^{-ik\omega} \, \mathrm{d}\omega$, so that $v(x,\omega) = \sum w(x,k) e^{ik\omega}$. As $v \in C^{5,1}$, w belongs to $\mathscr{B}^3$ and $\|w\|_{\mathscr{B}^3} \le C \|v\|_{C^{5,1}}$.

For $x \in Y$, we have $\hat{\mathscr{U}}_t^n v(x,\omega) = \sum_{k\in\mathbb{Z}} (\tilde P_n^t w)_k(x) e^{ik\omega}$ by construction of $\tilde P_n^t$. Hence Proposition 5.17 implies that, for $x \in Y$ and $t \in [-\tau_0, \tau_0]$,

$$\begin{aligned}\left|\hat{\mathcal{U}}_t^n v(x,\omega) - \left(1-\frac{\sigma^2t^2}{2}\right)^n \int v\right| &\leq \left|(\tilde{P}_n^t w)_0(x) - \left(1-\frac{\sigma^2t^2}{2}\right)^n \int w_0\right| \\ &\quad + \sum_{k\in\mathbb{Z}^*} |(\tilde{P}_n^t w)_k(x)| \\ &\leq \left\| \tilde{P}_n^t w - \left(1-\frac{\sigma^2t^2}{2}\right)^n \tilde{P} w\right\|_{\mathcal{B}_Y^0} \\ &\leq C(\bar{\theta}^n + |t|(1-ct^2)^n)\, \|w\|_{\mathcal{B}^3} \\ &\leq C(\bar{\theta}^n + |t|(1-ct^2)^n)\, \|v\|_{C^{5,1}}\,.\end{aligned}$$

This proves (5.65) for the points x with $h(x)=0$.

Assume now that $j = h(x) \in (0, n/2]$. Let x' be such that $U^j x' = x$, and let $\omega' = \omega - S_j\phi(x')$, so that $\mathcal{U}^j(x',\omega') = (x,\omega)$. Then $\hat{\mathcal{U}}_t^n v(x,\omega) = e^{itS_j\psi(x',\omega')}\hat{\mathcal{U}}_t^{n-j} v(x',\omega')$. Using the result for (x',ω'), we get

$$\left|\hat{\mathcal{U}}_t^n v(x,\omega) - e^{itS_j\psi(x',\omega')}\left(1-\frac{\sigma^2t^2}{2}\right)^{n-j} \int v\right| \leq C\left(\bar{\theta}^{n-j} + |t|(1-ct^2)^{n-j}\right) \|v\|_{C^{5,1}}\,. \tag{5.67}$$

Since $n - j \geq n/2$, this last term is bounded by $\bar{\theta}^{n/2} + |t|(1-ct^2)^{n/2}$, which is compatible with (5.65) (upon changing the values of $\bar{\theta}$ and c).

Moreover, $|e^{itS_j\psi(x',\omega')} - 1| \leq C|t|j$. Replacing $e^{itS_j\psi(x',\omega')}$ by 1 in (5.67), we add an error which is bounded by $C|t|h(x)(1-\sigma^2t^2/2)^{n/2}$. This is again compatible with (5.65). Finally,

$$\begin{aligned}\left|\left(1-\frac{\sigma^2t^2}{2}\right)^{n-j} - \left(1-\frac{\sigma^2t^2}{2}\right)^{n}\right| &\leq j\left|\log\left(1-\frac{\sigma^2t^2}{2}\right)\right|\left(1-\frac{\sigma^2t^2}{2}\right)^{n-j} \\ &\leq Cjt^2(1-ct^2)^{n/2},\end{aligned}$$

which is still compatible with (5.65). Incorporating all these substitutions, we obtain (5.65).

Finally, (5.66) is proved in the same way, by using (5.53). □

Proof of Theorem 1.12

Theorem 3.6 enabled us to prove Theorem 1.7. The same arguments make it possible to deduce Theorem 1.12 from Theorem 5.18, when $d^{(MN)} = 1$.

When $d = d^{(MN)} > 1$, let us show (1.13) ((1.14) is analogous). Applying the previous arguments to the transformation U^d, which is mixing, we almost obtain

(1.13) for times n of the form kd, with a slight difference: since σ^2 is replaced with

$$\int (S_d\psi)^2 + 2\sum_{j=1}^{\infty} \int (S_d\psi)(S_d\psi) \circ \mathscr{T}^{jd} = d\sigma^2, \tag{5.68}$$

we in fact obtain

$$\begin{aligned} \Big| \int & e^{itS_{kd}\psi} \cdot f \circ \mathscr{T}^n \cdot g \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb}) \\ & - \Big(1 - d\frac{\sigma^2 t^2}{2}\Big)^k \Big(\int f \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\Big)\Big(\int g \, \mathrm{d}(\tilde{\mu} \otimes \mathrm{Leb})\Big)\Big| \\ & \qquad \leq C(\bar{\theta}^k + |t|(1 - ct^2)^k) \, \|f\|_{L^\infty} \, \|g\|_{C^6} \, . \end{aligned}$$

To really obtain (1.13), we thus have to bound $(1 - \sigma^2 t^2/2)^{kd} - (1 - d\sigma^2 t^2/2)^k$. We have

$$\begin{aligned} \Big| \Big(1 - \frac{\sigma^2 t^2}{2}\Big)^{kd} - \Big(1 - d\frac{\sigma^2 t^2}{2}\Big)^k \Big| & \leq \Big| kd \log\Big(1 - \frac{\sigma^2 t^2}{2}\Big) - k \log\Big(1 - d\frac{\sigma^2 t^2}{2}\Big)\Big| \\ & \qquad \times \max\Big(\Big(1 - \frac{\sigma^2 t^2}{2}\Big)^{kd}, \Big(1 - d\frac{\sigma^2 t^2}{2}\Big)^k \Big) \\ & \leq Ck|t|^4(1 - ct^2)^k. \end{aligned}$$

By (4.21), this term is bounded by $Ct^2(1-ct^2/2)^k$. This concludes the proof for times $n = kd$.

If n is a general time, it can be written as $kd + r$ with $0 \leq r < d$. The theorem at time kd, applied to the functions $e^{itS_r\psi} f \circ \mathscr{T}^r$ and g (respectively, bounded and C^6) almost gives the result, the factor $(1 - \sigma^2 t^2/2)^n$ simply being replaced with $(1 - \sigma^2 t^2/2)^{kd}$. As above, one checks that the resulting additional error term is still compatible with (1.13). □

5.5. *Proof of Theorem 1.9*

Assume first that ψ is a C^6 function with $\sigma^2 > 0$. Theorem 1.12 for $f = g = 1$ shows that the characteristic function of $S_n\psi/\sqrt{n}$ converges to $e^{-\sigma^2 t^2/2}$, which is equivalent to the convergence of $S_n\psi/\sqrt{n}$ toward the Gaussian distribution $\mathscr{N}(0, \sigma^2)$. This concludes the proof in this case.

Assume now that ψ is only C^α, with zero average and with $\sigma^2 > 0$. Let ψ_ε be a C^6 function, close to ψ in $C^{\alpha/2}$, with corresponding asymptotic variance σ_ε^2. Theorem 1.7 (applied in $C^{\alpha/2}$) shows that the variance of $S_n(\psi - \psi_\varepsilon)/\sqrt{n}$ is uniformly small in n. This implies on the one hand that the distributions of $S_n\psi/\sqrt{n}$ and $S_n\psi_\varepsilon/\sqrt{n}$ are close, and on the other hand that σ_ε^2 is close to σ^2. In particular, if ε is small enough, $\sigma_\varepsilon^2 > 0$. As $S_n\psi_\varepsilon/\sqrt{n}$ converges to $\mathscr{N}(0, \sigma_\varepsilon^2)$, this implies that $S_n\psi/\sqrt{n}$ is close in distribution to $\mathcal{N}(0, \sigma^2)$ if n is large enough. Therefore, $S_n\psi/\sqrt{n}$ is indeed converging to $\mathcal{N}(0, \sigma^2)$. □

5.6. Regularity in the cohomological equation

Proof of Proposition 1.8

We proved half of the proposition in Proposition 3.9. It remains to prove that, if $\psi = f - f \circ \mathscr{T}$ for some measurable f, then $\sigma^2 = 0$. If $\sigma^2 > 0$, Theorem 1.9 implies that $S_n\psi/\sqrt{n}$ converges to a Gaussian distribution. However, $S_n\psi/\sqrt{n} = (f - f \circ \mathscr{T}^n)/\sqrt{n}$ converges in distribution to zero, which is a contradiction. Hence $\sigma^2 = 0$. □

Proof of Proposition 1.10

Let $\psi : X \times \mathbb{S}^1 \to \mathbb{R}$ be a C^6 function. We have to show that ψ is periodic if and only if ψ is continuously periodic.

If ψ is continuously periodic, it is trivially periodic. Conversely, suppose that ψ is continuously aperiodic, but it is nevertheless possible to write $\psi = u - u \circ \mathscr{T} + a \mod \lambda$, where u is measurable and $a \in \mathbb{R}$.

If σ^2 vanished, ψ would be continuously periodic by Proposition 1.8, which is a contradiction. Hence $\sigma^2 > 0$. As ψ is continuously aperiodic, it satisfies Theorem 1.12 (because (1.14) has been proved under the sole assumption of continuous aperiodicity). In particular, for $t \neq 0$ and for any functions f, g that are, respectively, bounded and C^6, $\int e^{itS_n\psi} f \circ \mathscr{T}^n g \to 0$. By density, this convergence to zero holds for any $f, g \in L^2$. However, for $t = 2\pi/\lambda$, $f = e^{itu}$, and $g = e^{-itu}$, we have

$$\int e^{itS_n\psi} f \circ \mathscr{T}^n g = \int e^{it(u - u\circ\mathscr{T}^n + na)} e^{itu\circ\mathscr{T}^n} e^{-itu} = e^{itna}, \tag{5.69}$$

which does not converge to zero. This is a contradiction. □

6. Proofs for Farey sequences

6.1. A general criterion for the weak Federer property

We would like to prove that some measures μ satisfy the weak Federer property. In the introduction, we showed that this property is quite easy to check for Lebesgue measure. However, in view of the application to Farey sequences, it is desirable to have a sufficiently simple criterion that does not apply only to absolutely continuous measures. In this section, we describe such a criterion.

Let us consider a Riemannian manifold Z endowed with a measure μ such that, for any $\rho > 0$, $\inf_{x \in Z} \mu(B(x, \rho)) > 0$. We assume that Z is partitioned in a finite number of subsets $Y_1, \ldots, Y_p$ and that each set Y_j admits a (finite or countable) subpartition modulo zero into sets $(W_{l,j})_{l \in \Lambda(j)}$. Also let $\overline{T}$ be a map that sends each set $W_{l,j}$ diffeomorphically to one of the Y_k. We can define $\mathscr{H}_n$ as the set of inverse branches of $\overline{T}^n$. Such an inverse branch h is not defined on the whole space Z but rather only on one of the sets $Y_j = Y_{j(h)}$. We assume the following.

(1) There exist $\kappa > 1$ and $C_{l,j}$ such that, for any $x \in W_{l,j}$ and v tangent at Z in x, $\kappa \|v\| \leq \|D\overline{T}(x)v\| \leq C_{l,j} \|v\|$.

(2) Let $J(x)$ be the inverse of the Jacobian of $\overline{T}$ with respect to μ. There exists $C > 0$ such that, for any $h \in \mathscr{H}_1$, $\|D((\log J) \circ h)\| \leq C$.

(3) For any $\bar{C} > 1$, there exist $\bar{D} > 1$ and $\eta_0 > 0$ such that, for any $\eta < \eta_0$ and for any $1 \leq j \leq p$, there exist disjoint balls $B(x_1, \bar{C}\eta), \dots, B(x_k, \bar{C}\eta)$ that are compactly included in Y_j; sets $A_1, \dots, A_k$ with $A_i \subset B(x_i, \bar{D}\bar{C}\eta) \cap Y_j$ such that, for any $x_i' \in B(x_i, (\bar{C}-1)\eta)$, we have $\mu(B(x_i', \eta)) \geq \mu(A_i)/\bar{D}$; and a finite number of inverse branches $h_1, \dots, h_\ell \in \mathscr{H}_1$ defined, respectively, on $Y_{j_1}, \dots, Y_{j_\ell}$ such that, for any $i \in [1, \ell]$, there exist $x \in Y_{j_i}$ and v a unit tangent vector at x with

$$\|Dh_i(x)v\| \geq \bar{C}\eta \tag{6.1}$$

such that

$$\bigcup_{i=1}^{k} B(x_i, \bar{C}\eta) \subset \bigcup_{i=1}^{k} A_i \tag{6.2}$$

and

$$Y_j = \Big(\bigcup_{i=1}^{k} A_i\Big) \sqcup \Big(\bigsqcup_{i=1}^{\ell} h_i(Y_{j_i})\Big) \quad \text{mod } 0. \tag{6.3}$$

(4) The transformation $\overline{T}$ is uniformly quasi-conformal in the following sense: there exists $K > 0$ such that, for any $h \in \bigcup_{n\in\mathbb{N}} \mathscr{H}_n$ defined on a set Y_j, for any $x, x' \in Y_j$ and any unit tangent vectors v and v', respectively, at x and x', we have

$$\|Dh(x)v\| \leq K\|Dh(x')v'\|. \tag{6.4}$$

The first two properties are uniform expansion properties, analogous to the similar requirements on T_Y in Definition 1.4. The difference is that the full shift structure has been replaced by a subshift of finite type, since such a structure naturally appears in the proofs for Farey sequences. The third property is a kind of weak Federer property, but not on the whole space, rather on the images of branches whose size is at most $\bar{C}\eta$ (by the requirement (6.1)). It turns out to be much easier to check than the true weak Federer property. Finally, the last property of uniform quasi conformality enables us to iterate the dynamics to get information at scales that are not covered by the third assumption.

PROPOSITION 6.1
Under the previous assumptions, the sets $h(Y_{j(h)})$ (for $h \in \bigcup_{n\in\mathbb{N}} \mathscr{H}_n$) uniformly have the weak Federer property (for the measure μ).

Proof
The quasi-conformality assumption shows that it is sufficient to prove that each set Y_j satisfies the weak Federer property: if sets A_i as in the definition of the weak Federer property can be constructed on Y_j, they can be transported to $h(Y_j)$ by the map h. In this process, one loses only harmless constant factors, and this implies the uniform weak Federer property. From this point on, we therefore work only on Y_j for each $1 \le j \le p$.

We want to construct sets A_i as in the definition of the weak Federer property. The third assumption of the proposition gives some of these sets, but to get the other ones we need to iterate the dynamics. Thus, our construction is inductive.

For any $1 \le j \le p$, let us fix a point $a_j \in Y_j$ and a unit tangent vector v_j at a_j. Let also $\rho > 0$ be such that the balls $B(a_j, \rho)$ are compactly included in Y_j. Fix a constant C for which one wants to prove the weak Federer property, and consider η small enough. We say that an inverse branch $h \in \mathscr{H}_n$, defined on Y_j, is (C, η)-*good*, or simply *good*, if $\|Dh(a_j)v_j\| \ge KC\eta/\rho$.

We prove the following fact: *there exists a constant M such that, if $h \in \mathscr{H}_n$ is a good branch defined on Y_j, then there exist disjoint balls $B(x_1, C\eta), \dots, B(x_k, C\eta)$ compactly included in $h(Y_j)$, sets $A_1, \dots, A_k$ with $A_i \subset h(Y_j) \cap B(x_i, MC\eta)$ such that any ball $B(x_i', \eta)$ included in $B(x_i, C\eta)$ satisfies $\mu(B(x_i', \eta)) \ge \mu(A_i)/M$, and good branches $h_1, \dots, h_\ell \in \mathscr{H}_{n+1}$ defined, respectively, on $Y_{j_1}, \dots, Y_{j_\ell}$ such that*

$$\bigcup_{i=1}^{k} B(x_i, C\eta) \subset \bigcup_{i=1}^{k} A_i \tag{6.5}$$

and

$$h(Y_j) = \Big(\bigcup_{i=1}^{k} A_i\Big) \sqcup \Big(\bigsqcup_{i=1}^{\ell} h_i(Y_{j_i})\Big). \tag{6.6}$$

This fact easily implies the proposition: we first apply it to the inverse branch Id_{Y_j} (which is obviously good if η is small enough) and then by induction to the inverse branches that are produced by the fact at the previous step. This process terminates since there is no good branch in $\mathscr{H}_n$ if n is large enough.

To prove that fact, we use the assumption (3) for the constant $\bar{C} = \max(K^2C, K^4C/\rho)$. Let $\bar{D} > 1$ and η_0 be given by (3) for this value of $\bar{C}$. Let $\eta < \eta_0$. Let $h \in \mathscr{H}_n$ be a good branch defined on a set Y_j.

Case 1. Assume that $\eta/(K\|Dh(a_j)v_j\|) \geq \eta_0$. The image of the ball $B(a_j, \rho)$ contains the ball $B(ha_j, \rho\|Dh(a_j)v_j\|/K)$, which itself contains $B(ha_j, C\eta)$ since h is good. Moreover, for $x, x' \in Y$ holds $d(hx, hx') \leq d(x, x')K\|Dh(a_j)v_j\| \leq \operatorname{diam} Y\eta/\eta_0$. In particular, if $M \geq \operatorname{diam} Y/(C\eta_0)$, we get $h(Y) \subset B(ha_j, MC\eta)$. We can thus take a ball $B(ha_j, C\eta)$ and a set $A_1 = h(Y)$. To conclude, we should check that $\mu(B(x', \eta)) \geq M^{-1}\mu(A_i)$ for any $x' \in B(ha_j, (C-1)\eta)$ if M is large enough. Since the iterates of $\overline{T}$ have a uniformly bounded distortion,

$$\frac{\mu(B(x', \eta))}{\mu(A_i)} \asymp \frac{\mu(h^{-1}B(x', \eta))}{\mu(Y)}. \tag{6.7}$$

Moreover, $h^{-1}B(x', \eta)$ contains $B(h^{-1}x', \eta/(K\|Dh(a_j)v_j\|))$, which itself contains $B(h^{-1}x', \eta_0)$. The measure of these balls is uniformly bounded from below. This concludes the proof in this case.

Case 2. Assume now that $\eta/(K\|Dh(a_j)v_j\|) < \eta_0$. Let $\eta_h = \eta/(K\|Dh(a_j)v_j\|)$; it is bounded by η_0. Hence, the assumption (3) gives sets $A_1, \dots, A_k$, balls $B(x_1, \bar{C}\eta_h), \dots, B(x_k, \bar{C}\eta_h)$, and inverse branches $h_1, \dots, h_\ell$, defined, respectively, on $Y_{j_1}, \dots, Y_{j_\ell}$. We show that the balls $B(hx_1, C\eta), \dots, B(hx_k, C\eta)$, the sets $\bar{A}_i = h(A_i)$, and the inverse branches $h \circ h_1, \dots, h \circ h_\ell$ satisfy the conclusion of the fact.

Let us first show that the inverse branch $h \circ h_i$ is good. By definition of h_i, $\|Dh_i(a_{j_i})v_{j_i}\| \geq \bar{C}\eta_h/K \geq K^2C\eta/(\rho\|Dh(a_j)v_j\|)$. We have $D(h \circ h_i)(a_{j_i})v_{j_i} = Dh(h_i a_{j_i})Dh_i(a_{j_i})v_{j_i}$. Moreover, $\|Dh(x)v\| \geq K^{-1}\|v\|\|Dh(a_j)v_j\|$. Therefore,

$$\begin{aligned}\|D(h \circ h_i)(a_{j_i})v_{j_i}\| &\geq K^{-1}\|Dh_i(a_{j_i})v_{j_i}\|\|Dh(a_j)v_j\| \\ &\geq K^{-1}\frac{K^2C\eta}{\rho\|Dh(a_j)v_j\|}\|Dh(a_j)v_j\| = KC\eta/\rho.\end{aligned}$$

This shows that $h \circ h_i$ is good.

The set $hB(x_i, \bar{C}\eta_h)$ contains the ball $B(hx_i, \bar{C}\eta_h\|Dh(a_j)v_j\|/K)$, which itself contains the ball $B(hx_i, C\eta)$ because $\bar{C} \geq K^2C$. Moreover, for any $x' \in B(hx_i, (C-1)\eta)$, the set $h^{-1}B(x', \eta)$ contains the ball $B(h^{-1}x', \eta/(K\|Dh(a_j)v_j\|)) = B(h^{-1}x', \eta_h)$. As the distortion of the iterates of $\overline{T}$ is uniformly bounded, for any $x' \in B(hx_i, (C-1)\eta)$ we obtain

$$\frac{\mu(B(x', \eta))}{\mu(\bar{A}_i)} \asymp \frac{\mu(h^{-1}B(x', \eta))}{\mu(A_i)} \geq \frac{\mu(B(hx', \eta_h))}{\mu(A_i)} \geq \bar{D}^{-1}. \tag{6.8}$$

Finally, as $A_i \subset B(x_i, \bar{D}\bar{C}\eta_h)$, $\bar{A}_i$ is contained in $B(hx_i, \bar{D}\bar{C}\eta_h K \|Dh(a_j)v_j\|) = B(hx_i, \bar{D}\bar{C}\eta)$. □

The previous criterion easily implies that Gibbs measures in dimension 1 have the uniform weak Federer property, as in the following.

PROPOSITION 6.2
Let T be a C^2 uniformly expanding map on the circle $\mathbb{S}^1$, and let μ be a Gibbs measure corresponding to a C^1 potential. Then there exists a subset Y of $\mathbb{S}^1$ such that T is nonuniformly expanding with base Y for the measure μ.

Proof
Let d be the topological degree of T, and let x_0 be a fixed point of T. Let $Y = Z = \mathbb{S}^1 - \{x_0\}$. Then $\mathbb{S}^1 - T^{-1}(x_0)$ is the union of d intervals $W_1, \dots, W_j$, each of them being sent by T onto Z. These intervals form a partition (modulo 0) of Z satisfying the first four points of Definition 1.4 (for $r_i = 1$, $1 \le i \le d$). If we can prove that T satisfies the assumptions of the previous proposition, the proof will be complete. The assumptions (1) and (2) are clear, and the fourth assumption is equivalent to the bounded distortion for Lebesgue measure since we are in one dimension. Let us check (3) for some $\bar{C} > 0$. Let η_0 be small enough so that, for any $x \in Z$ and any inverse branch $h \in \mathscr{H}$, $|h'(x)| \ge \bar{C}\eta_0$. We take no ball $B(x_i, \bar{C}\eta)$, no set A_i, and all the inverse branches $h \in \mathscr{H}$. Then (6.2) is empty, hence trivial, and (6.3) is also trivial. □

Remark 6.3
Proposition 6.2 also holds for Hölder potentials with the same proof since our argument only uses the bounded distortion properties of the measure. However, this is not sufficient to apply our main theorems, since they require the Jacobian to be really C^1.

6.2. Farey sequences
Let $r > 1$. Let T be the map on $X = [0, 1]$ given by (1.8), and let $\mathscr{T}$ be its extension to $[0, 1] \times \mathbb{R}/(\log r)\mathbb{Z}$ defined in (1.9) using a function ϕ. This function is not C^1 on $[0, 1]$, which seems to be a problem since we always worked with a function ϕ of class C^1. To avoid this problem, we can simply work with the disjoint union $X = [0, 1/2] \sqcup [1/2, 1]$, on which ϕ is C^1. All our results in the previous sections have been formulated for transformations on $X \times \mathbb{R}/2\pi\mathbb{Z}$, but the same results hold verbatim on $X \times \mathbb{R}/\gamma\mathbb{Z}$ for any $\gamma \neq 0$, and, in particular, for $\gamma = \log r$. Henceforth, we simply denote $\mathbb{R}/(\log r)\mathbb{Z}$ by $\mathbb{S}^1$ and apply without further notice the preceding results.

Let $x_0 = 1/2$, and set $x_n = h_A(x_{n-1})$ (i.e., x_n is the preimage of x_{n-1} under the left branch of T). Explicitly, $x_n = 1/(n+2)$. Let $I_j = (x_j, x_{j-1})$. Let also $\bar{I}_j = 1 - I_j$

be the symmetric of I_j with respect to $1/2$. Let $Y = (x_1, x_0) = (1/3, 1/2)$, and denote by T_Y the map induced by T on Y. Its combinatorics can be described as follows: a point of Y is sent by T in $(1/2, 1)$, spends some time $i > 0$ there, is then sent back to $(0, 1/2)$, and increases (for $j \geq 0$ iterates) before entering back in Y. The points with this combinatorics form an interval $I_{i,j} := T^{-1}(\bar{I}_i) \cap T^{-i-1}(I_{j+1})$ and $T^{i+j+1}(I_{i,j}) = Y$. Letting $r_{i,j} = i + j + 1$, we thus obtain a partition of Y that satisfies the first point of Definition 1.4.

PROPOSITION 6.4
The map T is nonuniformly expanding of base Y in the sense of Definition 1.4 for the partition $\{I_{i,j}\}_{i>0, j\geq 0}$ and for Minkowski's measure μ. Moreover, it is mixing.

Proof
The first point of Definition 1.4 is clear. For the second one, note that the Jacobian of T for Minkowski's measure is everywhere equal to 2 by definition. Hence, the Jacobian of T_Y on $I_{i,j}$ is constant (equal to 2^{i+j+1}), and $D((\log J) \circ h_{i,j}) = 0$. The third point is trivial. For the fourth one, we have, for any $\sigma > 0$,

$$\int_Y e^{\sigma r} = \sum \mu(I_{i,j}) e^{\sigma(i+j+1)} = \sum 2^{-i-j-3} e^{\sigma(i+j+1)}, \tag{6.9}$$

which is finite as soon as $\sigma < \log 2$. The mixing of T is a consequence of the equality $\gcd\{r_{i,j}\} = 1$.

Thus, we just have to prove the uniform weak Federer property. To do this, we use Proposition 6.1. Let $Y_0 = Y$, and let Y_1 be its symmetric with respect to $1/2$. Let $Z = Y_0 \cup Y_1$, and let $\overline{T}$ be the first-return map induced by T on Z. It sends each interval $T^{-1}(\bar{I}_i) \cap Y_0$ bijectively to Y_1 and each interval $T^{-1}(I_i) \cap Y_1$ bijectively to Y_0. If we prove that $\overline{T}$ satisfies the assumptions of Proposition 6.1, this concludes the proof of the uniform weak Federer property since the inverse branches of the iterates of T_Y are, in particular, inverse branches of iterates of $\overline{T}$.

Assumptions (1) and (2) of Proposition 6.1 are trivial (since J is constant on each monotonicity interval of $\overline{T}$). For assumption (4), the quickest argument is certainly to use the fact that all the inverse branches of the iterates of $\overline{T}$ are homographies (hence with vanishing Schwarzian derivative) which can be extended to the whole interval $[0, 1]$. Koebe's lemma [DV, Theorem IV.1.2] directly yields the uniform quasi conformality.

Hence we just have to check assumption (3). By symmetry, it is sufficient to check it on Y_0. If J is an interval, we denote its length by $|J|$. Then $|\bar{I}_n|$ is a decreasing sequence, with $|\bar{I}_{n+1}|/|\bar{I}_n| \to 1$ when $n \to \infty$ since $T'(1) = 1$. As a consequence,

$K_n = T^{-1}(\bar{I}_n) \cap Y_0$ satisfies $|K_{n+1}|/|K_n| \to 1$, and there exists $C > 0$ such that $|K_m| \leq C|K_n|$ for all $m \geq n$. Finally, $\mu(K_n) = 2^{-n-2}$.

We use the following fact: for any $C > 0$ there exists $D > 0$ such that, for any interval J included in an interval K_n with $|J| \geq C^{-1}|K_n|$, then $\mu(J) \geq D^{-1}\mu(K_n)$. To prove this fact, we apply the map $\overline{T}$ once, which sends K_n to Y_1, and J to an interval J' satisfying $|J'| \geq C^{-1}K^{-1}|Y_1|$ by quasi conformality. Hence, $\mu(J')$ is uniformly bounded from below. As $\mu(J')/\mu(Y_1) = \mu(J)/\mu(K_n)$, this proves the fact.

We can now prove assumption (3) of Proposition 6.1 on Y_0. Let $\bar{C} > 1$. We construct inverse branches $h_1, \dots, h_\ell$, balls $B(x_1, \bar{C}\eta), \dots, B(x_k, \bar{C}\eta)$, and sets $A_1, \dots, A_k$ as follows if η is small enough.

Let N be maximal such that $|K_n| \geq \bar{C}\eta$ for $n \leq N$. We take $\ell = N$, and we let $h_1, \dots, h_\ell$ be the inverse branches of $\overline{T}$ whose images are the intervals $K_1, \dots, K_\ell$. Then h_i is defined on Y_1, of length $1/6$, and the length of its image K_i is $\geq \bar{C}\eta$. Hence there exists a point $y_i \in Y_1$ with $h_i'(y_i) \geq 6\bar{C}\eta$. This proves (6.1).

We decompose the remaining interval as a union of intervals of length $2\bar{C}\eta$, except maybe the first one whose length belongs to $[2\bar{C}\eta, 4\bar{C}\eta)$. Let us denote this decomposition by $J_0, \dots, J_p$. Since $|K_N| = o\big(\sum_{n>N} |K_n|\big)$ when $N \to \infty$, we have $p \geq 2$ if η is small enough. Let us define sets $A_1, \dots, A_p$ by $A_i = J_i$ for $i > 1$, and $A_1 = J_0 \cup J_1$. Let $B(x_i, \bar{C}\eta) = J_{i-1}$ for $i > 1$, and let $B(x_1, \bar{C}\eta)$ be the leftmost part of J_0. For $i > 1$, the ball $B(x_i, \bar{C}\eta)$ is *not* included in the set A_i; it is strictly to its left. The balls are disjoint, and $A_i \subset B(x_i, 5\bar{C}\eta)$. Let us show that they satisfy the desired conclusion: we have to prove, for any interval J of length 2η included in $B(x_i, \bar{C}\eta)$, that $\mu(J) \geq \bar{D}^{-1}\mu(A_i)$ holds for some constant $\bar{D}$ (independent of η). Either J contains an interval K_n, or it intersects such an interval along a subinterval of length at least η. Moreover, $|K_n| \leq C|K_{N+1}| \leq C\bar{C}\eta$. In both cases, the fact we proved above implies that $\mu(J) \geq D^{-1}\mu(K_n)$.

We first deal with $i = 1$. As $|K_{n+1}| \sim |K_n|$, the set A_1 is covered by $\bigcup_{k=1}^{7} K_{N+k}$ if N is large enough (hence, if η is small enough). These seven intervals have comparable measures since $\mu(K_m) = 2^{-m-2}$, hence $\mu(A_1) \leq C\mu(K_{N+k})$ for $1 \leq k \leq 7$. As $\mu(J) \geq D^{-1}\mu(K_n)$ for at least one of these K_n's, we indeed conclude that $\mu(J) \geq C^{-1}\mu(A_1)$.

Assume now that $i > 1$. There exists an interval K_n intersecting J with $\mu(J) \geq C^{-1}\mu(K_n)$. Since A_i is located to the right of K_n, we get

$$\mu(A_i) \leq C \sum_{m=n}^{\infty} \mu(K_m) = C \sum_{m=n}^{\infty} 2^{-m-2} \leq C 2^{-n-2} \leq C\mu(K_n). \tag{6.10}$$

This also concludes the proof in this case. $\square$

LEMMA 6.5
The function ϕ is not cohomologous to a locally constant function.

Proof
Assume by contradiction that there exists a C^1 function f such that $\phi_Y - f + f \circ T_Y$ is constant on each interval $I_{i,j}$ equal to some number $a_{i,j}$. The interval $I_{1,1}$ contains the point $x = 3/2 - \sqrt{5}/2$ with $T_Y(x) = x$. Necessarily, $a_{1,1} = \phi_Y(x)$. In the same way, the interval $I_{2,1}$ contains $x' = 1 - \sqrt{3}/3$, invariant under T_Y, which gives $a_{2,1} = \phi_Y(x')$.

Let now $y = 1 - \sqrt{6}/4$. This point belongs to $I_{1,1}$, but $T_Y(y) \in I_{2,1}$, and $T_Y^2(y) = y$. Then

$$\phi_Y(y) + \phi_Y(T_Y y) = a_{1,1} + a_{2,1} = \phi_Y(x) + \phi_Y(x'). \tag{6.11}$$

However, it is possible to compute explicitly $\phi_Y(y) + \phi_Y(T_Y y) - \phi_Y(x) - \phi_Y(x')$, and check that this quantity is nonzero (approximately equal to -0.013). This is a contradiction. □

Proposition 6.4 and Lemma 6.5 show that the results of Section 1.3 apply to $\mathscr{T}$. However, this is not sufficient to prove Theorems 1.1 and 1.2 since these results are pointwise, while the results of Section 1.3 are averaged. We therefore need an additional ingredient. Let $X^{(n)}$ be the extension of X defined in Section 3.1, and let $\pi^{(n)}$, $\tilde{\pi}^{(n)}$ be the corresponding projections.

LEMMA 6.6
For any $n \in \mathbb{N}$, there exists a constant $C(n)$ such that, for any integrable function $u : X \times \mathbb{S}^1 \to \mathbb{C}$, for almost all $(x, \omega) \in X \times \mathbb{S}^1$, and for any $k \in \mathbb{N}$, we have

$$\hat{\mathscr{T}}^k u(x, \omega) = C(n) \sum_{\pi^{(n)}(x')=x} 2^{-h(x')} \hat{\mathscr{U}}^k (u \circ \tilde{\pi}^{(n)})(x', \omega). \tag{6.12}$$

Proof
Let $\mathscr{B}$ be the σ-algebra of Borel measurable subsets of $X \times \mathbb{S}^1$, and let $\mathscr{B}' = (\tilde{\pi}^{(n)})^{-1}(\mathscr{B})$. This is a sub-$\sigma$-algebra of the Borel σ-algebra on $X^{(n)} \times \mathbb{S}^1$. A function v on $X^{(n)} \times \mathbb{S}^1$ can be written as $u \circ \tilde{\pi}^{(n)}$ if and only if v is $\mathscr{B}'$-measurable.

Let us first prove that

$$(\hat{\mathscr{T}}^k u) \circ \tilde{\pi}^{(n)} = E\left(\hat{\mathscr{U}}^k (u \circ \tilde{\pi}^{(n)}) \,\middle|\, \mathscr{B}'\right). \tag{6.13}$$

To do this, let us write $E(\hat{\mathscr{U}}^k(u\circ\tilde\pi^{(n)}) \mid \mathscr{B}') = v\circ\tilde\pi^{(n)}$. As $\tilde\mu\otimes\mathrm{Leb} = \tilde\pi^{(n)}_*(\tilde\mu^{(n)}\otimes\mathrm{Leb})$, for any measurable function f on $X\times\mathbb{S}^1$ we have

$$\int_{X\times\mathbb{S}^1} vf = \int_{X^{(n)}\times\mathbb{S}^1} v\circ\tilde\pi^{(n)} f\circ\tilde\pi^{(n)} = \int_{X^{(n)}\times\mathbb{S}^1} E\big(\hat{\mathscr{U}}^k(u\circ\tilde\pi^{(n)}) \,\big|\, \mathscr{B}'\big) f\circ\tilde\pi^{(n)}. \quad (6.14)$$

As $f\circ\tilde\pi^{(n)}$ is $\mathscr{B}'$-measurable, we get

$$\begin{aligned}\int_{X\times\mathbb{S}^1} vf &= \int_{X^{(n)}\times\mathbb{S}^1} \hat{\mathscr{U}}^k(u\circ\tilde\pi^{(n)}) f\circ\tilde\pi^{(n)} = \int_{X^{(n)}\times\mathbb{S}^1} u\circ\tilde\pi^{(n)} f\circ\tilde\pi^{(n)}\circ\mathscr{U}^k \\ &= \int_{X^{(n)}\times\mathbb{S}^1} u\circ\tilde\pi^{(n)} f\circ\mathscr{T}^k\circ\tilde\pi^{(n)} = \int_{X\times\mathbb{S}^1} uf\circ\mathscr{T}^k.\end{aligned}$$

This last equality shows that $v = \hat{\mathscr{T}}^k u$ and concludes the proof of (6.13).

The set $X^{(n)}$ is endowed with a countable partition $\mathscr{A}$ such that $\pi^{(n)}$ is injective on each element of the partition. Let us define a function F on $X^{(n)}$ as follows: on each set $a\in\mathscr{A}$, let $F = \mathrm{d}\tilde\mu^{(n)}/\,\mathrm{d}(\tilde\mu\circ\pi^{(n)}_{|a})$. This is the local Radon-Nikodym derivative of $\tilde\mu^{(n)}$ with respect to $(\pi^{(n)})^*\tilde\mu$. As $\pi^{(n)}_*\tilde\mu^{(n)} = \tilde\mu$, we have $\sum_{\pi^{(n)}(x')=x} F(x') = 1$ for almost every $x\in X$. Let us show that the conditional expectation with respect to $\mathscr{B}'$ is given by

$$E(v\mid\mathscr{B}')(x,\omega) = \sum_{\pi^{(n)}(x')=\pi^{(n)}(x)} F(x')v(x',\omega). \quad (6.15)$$

Let us indeed define a function w on $X\times\mathbb{S}^1$ by

$$w(x,\omega) = \sum_{\pi^{(n)}(x')=x} F(x')v(x',\omega) = \sum_{a\in\mathscr{A}} 1_{x\in\pi^{(n)}a} F\big((\pi^{(n)}_{|a})^{-1}x\big) v\big((\pi^{(n)}_{|a})^{-1}x,\omega\big). \quad (6.16)$$

If f is a measurable function on $X\times\mathbb{S}^1$, then

$$\begin{aligned}\int_{X\times\mathbb{S}^1} fw &= \sum_{a\in\mathscr{A}} \int_{\pi^{(n)}(a)} f(x,\omega)F\big((\pi^{(n)}_{|a})^{-1}x\big)v\big((\pi^{(n)}_{|a})^{-1}x,\omega\big)\,\mathrm{d}\tilde\mu(x)\,\mathrm{d}\omega \\ &= \sum_{a\in\mathscr{A}} \int_a f(\pi^{(n)}x',\omega)v(x',\omega)\,\mathrm{d}\tilde\mu^{(n)}(x')\,\mathrm{d}\omega = \int_{X^{(n)}\times\mathbb{S}^1} f\circ\tilde\pi^{(n)} v.\end{aligned}$$

This proves (6.15). Together with (6.13), this implies the lemma if we can prove that

$$F(x') = C(n)2^{-h(x')}. \quad (6.17)$$

As T_Y is the first-return map to Y, the Jacobian of $\pi^{(1)}$ for the measure $\tilde\mu^{(1)}$ on Y is equal to 1. Since $\tilde\mu^{(n)}$ is proportional to $\tilde\mu^{(1)}$ on Y, this implies that F is constant on Y equal to a constant $C(n)$. This proves (6.17) for points with zero height.

The Jacobian of T for $\tilde{\mu}$ is equal to 2, while the Jacobian of U is equal to 1 on the set of points that do not come back to the basis. By induction over $h(x')$, this implies (6.17). □

COROLLARY 6.7
There exist constants $C > 0$ and $\bar{\theta} < 1$ such that, for any C^6 function $f : X \times \mathbb{S}^1 \to \mathbb{C}$ and for any $(x, \omega) \in X \times \mathbb{S}^1$, we have

$$\left| \hat{\mathscr{T}}^n f(x, \omega) - \int f \right| \le C \bar{\theta}^n \, \|f\|_{C^6} . \tag{6.18}$$

Proof
Since everything is symmetric with respect to $1/2$ and continuous, it is sufficient to prove the assertion for almost every $x \in (1/2, 1)$.

We work in $X^{(N)}$, where N is given by Theorem 2.3. Note that $d^{(N)}$ is equal to 1, since $r^{(N)}$ takes the values $2N$ and $2N + 1$. Applying Theorem 3.6 to the function $v = f \circ \tilde{\pi}^{(N)}$, we get, for any $n \in \mathbb{N}$ and for any $x' \in X^{(N)}$ with $h(x') \le n/2$,

$$\left| \hat{\mathscr{U}}^n (f \circ \tilde{\pi}^{(N)})(x', \omega) - \int f \right| \le C \bar{\theta}^n \, \|f\|_{C^6} . \tag{6.19}$$

Together with Lemma 6.6, this yields

$$\left| \hat{\mathscr{T}}^n f(x, \omega) - \int f \right| \le C \Big(\sum_{\pi^{(N)}(x') = x,\, h(x') \le n/2} \bar{\theta}^n 2^{-h(x')} + \sum_{\pi^{(N)}(x') = x,\, h(x') > n/2} 2^{-h(x')} \Big) \, \|f\|_{C^6} .$$

To conclude, it is thus sufficient to prove that, for $x \in (1/2, 1)$, the cardinality of

$$\left\{ x' \,\middle|\, \pi^{(N)}(x') = x,\ h(x') = k \right\} \tag{6.20}$$

grows at most polynomially with k. If we write a point of $X^{(N)}$ as a pair (x', j) with $x' \in Y$ and $j < r^{(N)}(x')$, it is easy to check that U^k induces a bijection between the set (6.20) and the set of points in $T^{-k}(x) \cap Y$ whose first k iterates under T spend a time $t < N$ in Y. If t is fixed, such a point is determined by the combinatorics $(i_1, j_1, \dots, i_t, j_t, i_{t+1})$ of times spent in $[1/2, 1]$, then in $[0, 1/2]$, then in $[1/2, 1]$, and so on, with the constraint that the sum of these lengths is k (recall that we assume that $x \in (1/2, 1)$). As a consequence,

$$\operatorname{Card} \left\{ x' \,\middle|\, \pi^{(N)}(x') = x,\ h(x') = k \right\} \le \sum_{t=0}^{N-1} k^{2t+1} \le C k^{2N}. \tag{6.21}$$

This quantity indeed grows polynomially with k. □

Proof of Theorem 1.1
If f is a continuous function on $[0, 1] \times \mathbb{S}^1$, then $\int f \, \mathrm{d}\bar{\mu}_n = \hat{\mathscr{T}}^n f(1, 0)$. Hence Corollary 6.7 shows the theorem for C^6 functions. The case of C^α functions is then deduced by interpolation, just as was done at the end of the proof of Theorem 1.7. □

Proof of Theorem 1.2
If ψ is a C^6 function that is not a coboundary, we show, as in the proof of Corollary 6.7 (but using Theorem 5.18 instead of Theorem 3.6), for $|t| \leq \tau_0$, that

$$\left| \hat{\mathscr{T}}_t^n f(x, \omega) - \left(1 - \frac{\sigma^2 t^2}{2}\right)^n \int f \right| \leq C\left(\bar{\theta}^n + |t|(1 - ct^2)^n\right) \|f\|_{C^6} . \tag{6.22}$$

Moreover, if ψ is aperiodic, for $\tau_0 \leq |t| \leq t_0$, then

$$\left| \hat{\mathscr{T}}_t^n f(x, \omega) \right| \leq C\bar{\theta}^n \|f\|_{C^6} . \tag{6.23}$$

As $\hat{\mathscr{T}}_t^n 1(1, 0) = E(e^{it \sum_{k=1}^n \psi(X_k)})$, this implies the limit assertions in Theorem 1.2.

The automatic regularity properties still have to be checked. If $\psi = f - f \circ \mathscr{T}$ with f measurable, let us show that f is continuous on $[0, 1]$. Proposition 1.8 shows that f is continuous on $Y \times \mathbb{S}^1$. As $\mathscr{T}$ is a homeomorphism between $Y \times \mathbb{S}^1$ and $[1/2, 1] \times \mathbb{S}^1$, we conclude from the equality $f \circ \mathscr{T} = f - \psi$ that f is continuous on $[1/2, 1] \times \mathbb{S}^1$. Finally, as $\mathscr{T}$ is a homeomorphism between $[1/2, 1] \times \mathbb{S}^1$ and $[0, 1] \times \mathbb{S}^1$, we obtain with the same argument the continuity of f on the whole space.

We argue in the same way for the cohomological equation in $\mathbb{R}/\lambda\mathbb{Z}$, by using Proposition 1.10. □

Appendix. Contraction properties of transfer operators

In this appendix, we prove Theorem 2.3 on the contraction properties (in C^1 norm or in Dolgopyat norm) of the transfer operator associated to a map T_Y, where T is a nonuniformly expanding map of base Y. The notation and assumptions used are those of Theorem 2.3.

A.1. Contraction in the C^1 norm

In this section, we introduce the tools to prove the first part of Theorem 2.3. However, the choices of the constants N and θ of Theorem 2.3 are only possible at the complete end of the proof in Section A.2.

Several times, we use the following distortion lemma. Its proof is completely standard and therefore omitted.

LEMMA A.1
Let $J^{(n)}(x)$ be the inverse of the Jacobian of T_Y^n at the point x. There exists $C>0$ (independent of n) such that, for any $h\in\mathscr{H}_n$ and for any $x,y\in Y$, we have $\|D(J^{(n)}\circ h)(x)\|\leq CJ^{(n)}\circ h(x)$ and $J^{(n)}\circ h(x)\leq CJ^{(n)}\circ h(y)$.

For small enough ε, we define an operator $\mathscr{L}_\varepsilon$ acting on functions from Y to $\mathbb{C}$ by $\mathscr{L}_\varepsilon u(x)=\sum J(hx)u(hx)e^{\varepsilon r(hx)}$. If $H_0\subset\mathscr{H}$, we also denote by $\mathscr{L}_{\varepsilon,H_0}$ the same operator with the exception that the sum is restricted to the inverse branches belonging to H_0. The following elementary estimates are used throughout the forthcoming arguments.

LEMMA A.2
There exists a function $\alpha(\varepsilon)$ that tends to zero when $\varepsilon\to 0$ such that $\|\mathscr{L}_\varepsilon\|_{L^2\to L^2}\leq e^{\alpha(\varepsilon)}$ and $\|\mathscr{L}_\varepsilon\|_{C^0\to C^0}\leq e^{\alpha(\varepsilon)}$.

Moreover, if $\varepsilon_0>0$ is small enough, then for any $\gamma>0$ there exists $H_0\subset\mathscr{H}$ with a finite complement such that $\|\mathscr{L}_{\varepsilon_0,H_0}\|_{L^2\to L^2}\leq\gamma$.

Proof
We have

$$\begin{aligned}|\mathscr{L}_{\varepsilon,H_0}u(x)|^2 &= \Big|\sum_{h\in H_0}J(hx)u(hx)e^{\varepsilon r(hx)}\Big|^2\\ &\leq\Big(\sum_{h\in H_0}J(hx)|u(hx)|^2\Big)\Big(\sum_{h\in H_0}J(hx)e^{2\varepsilon r(hx)}\Big).\end{aligned}$$

Consequently, $\|\mathscr{L}_{\varepsilon,H_0}u\|_{L^2}\leq\|u\|_{L^2}\cdot\sup_{x\in Y}\left(\sum_{h\in H_0}J(hx)e^{2\varepsilon r(hx)}\right)^{1/2}$. We have $J(hx)\leq CJ(hy)$ for any $h\in\mathscr{H}$ and all $x,y\in Y$, hence $\sum J(hx)e^{2\varepsilon r(hx)}\leq C\sum J(hy)e^{2\varepsilon r(hy)}$. Integrating this inequality with respect to y, we get

$$\sum_{h\in H_0}J(hx)e^{2\varepsilon r(hx)}\leq C\sum_{h\in H_0}\int_Y J(hy)e^{2\varepsilon r(hy)}\,\mathrm{d}\mu_Y(y)=C\int_{H_0(Y)}e^{2\varepsilon r(y)}\,\mathrm{d}\mu_Y(y). \tag{A.1}$$

This quantity is finite if ε is small enough, by assumption (4) of Definition 1.4. Taking the complement of H_0 large enough, it can even be made arbitrarily small. This proves the second point of the lemma.

For the first point, we have to be slightly more precise. For any x, we have $e^{2\varepsilon r(hx)}\leq 1+2\varepsilon r(hx)e^{2\varepsilon r(hx)}$. Hence, using the inequality $J(hx)\leq CJ(hy)$ for any

$h \in \mathscr{H}$ and $x, y \in Y$, we get

$$\begin{aligned}\sum_{h\in\mathscr{H}} J(hx)e^{2\varepsilon r(hx)} &\le \sum_{h\in\mathscr{H}} J(hx) + 2\varepsilon \sum_{h\in\mathscr{H}} J(hx)r(hx)e^{2\varepsilon r(hx)} \\ &\le 1 + C\varepsilon \sum_{h\in\mathscr{H}} J(hy)r(hy)e^{2\varepsilon r(hy)}.\end{aligned}$$

Integrating with respect to y,

$$\sum_{h\in\mathscr{H}} J(hx)e^{2\varepsilon r(hx)} \le 1 + C\varepsilon \int_Y r(y)e^{2\varepsilon r(y)}\, \mathrm{d}\mu_Y(y), \tag{A.2}$$

and this last integral is uniformly bounded if ε is small enough. This gives the desired estimate for the action of $\mathscr{L}_\varepsilon$ on L^2 and C^0. □

Let us prove a lemma that easily implies (2.5).

LEMMA A.3
There exist $\varepsilon_0 > 0$ and $\theta_0 < 1$ such that, for any $A > 0$, $n \in \mathbb{N}$, and $\varepsilon < \varepsilon_0$, there exists $C > 0$ such that, for any $\psi \in \mathscr{C}_n^{A,\varepsilon}$ and $v \in C^1(Y)$, we have

$$\|\mathscr{L}^n(\psi v)\|_{C^1} \le \theta_0^n \Big(\sup_{x\in Y} |\psi(x)|/e^{\varepsilon r^{(n)}(x)}\Big) \|v\|_{C^1} + C\, \|\psi\|_{\mathscr{C}_n^{A,\varepsilon}}\, \|v\|_{C^0}\,. \tag{A.3}$$

Proof
First, since $|\psi(x)| \le \|\psi\|_{\mathscr{C}_n^{A,\varepsilon}}\, e^{\varepsilon r^{(n)}(x)}$, we have

$$\begin{aligned}\|\mathscr{L}^n(\psi v)\|_{C^0} \le \|\mathscr{L}^n(|\psi|\cdot|v|)\|_{C^0} &\le \|\psi\|_{\mathscr{C}_n^{A,\varepsilon}}\, \|\mathscr{L}^n(e^{\varepsilon r^{(n)}(x)}|v|)\|_{C^0} \\ &= \|\psi\|_{\mathscr{C}_n^{A,\varepsilon}}\, \|\mathscr{L}_\varepsilon^n |v|\|_{C^0} \le \|\psi\|_{\mathscr{C}_n^{A,\varepsilon}}\, e^{n\alpha(\varepsilon)}\, \|v\|_{C^0}\end{aligned}$$

by Lemma A.2. This gives the desired control in the C^0 norm. For the C^1 norm, we differentiate $\mathscr{L}^n(\psi v) = \sum_{h\in\mathscr{H}_n} J^{(n)}(hx)\psi(hx)v(hx)$. If we differentiate $J^{(n)}(hx)$, we use the estimate $\|D(J^{(n)}\circ h)(x)\| \le C J^{(n)}(hx)$ given by Lemma A.1 and get the same bound as for the C^0 norm. If we differentiate $\psi(hx)$, its derivative is bounded by $A\, \|\psi\|_{\mathscr{C}_n^{A,\varepsilon}}\, e^{\varepsilon r^{(n)}(hx)}$; and using the same argument as for the C^0 norm, we obtain the same bound (with an additional factor A, which is not a problem since C is allowed to depend on A in the statement of the lemma).

Finally, if we differentiate $v \circ h$, we have $\|D(v\circ h)(x)\| \le \kappa^{-n}\, \|Dv(hx)\|$, and we therefore get a bound

$$\begin{aligned}\kappa^{-n}\, \|Dv\|_{C^0}\, \mathscr{L}^n|\psi| &\le \kappa^{-n}\, \|v\|_{C^1} \Big(\sup_{x\in Y} |\psi(x)|/e^{\varepsilon r^{(n)}(x)}\Big) \mathscr{L}^n(e^{\varepsilon r^{(n)}}) \\ &\le \kappa^{-n}\, \|v\|_{C^1} \Big(\sup_{x\in Y} |\psi(x)|/e^{\varepsilon r^{(n)}(x)}\Big) e^{n\alpha(\varepsilon)}.\end{aligned}$$

If ε is small enough, $\kappa^{-1}e^{\alpha(\varepsilon)} < 1$. This concludes the proof. □

We now turn to the proof of (2.6). As a preliminary estimate, let us first consider in the following lemma the case $\psi_i = e^{\varepsilon r^{(N)}}$ for all i.

LEMMA A.4
There exist $N_0 > 0$, $\theta_0 < 1$, $C > 0$, $\varepsilon_0 > 0$, and a function $\alpha : (0, \varepsilon_0) \to \mathbb{R}_+$ tending to 0 when $\varepsilon \to 0$, which satisfy the following property. For any $N \geq N_0$ and $\varepsilon < \varepsilon_0$, for any C^1 function $v : Y \to \mathbb{C}$, we have

$$\|D(\mathscr{L}_\varepsilon^N v)\|_{C^0} \leq \theta_0^N \|Dv\|_{C^0} + Ce^{N\alpha(\varepsilon)} \|v\|_{L^2}. \tag{A.4}$$

Proof
We have $\mathscr{L}_\varepsilon^N v = \sum_{h\in\mathscr{H}_N} J^{(N)}(hx)e^{\varepsilon r^{(N)}(hx)}v(hx)$. By Lemma A.1, $J^{(N)}(hx) \leq CJ^{(N)}(hy)$, and $\|D(J^{(N)}\circ h)(x)\| \leq CJ^{(N)}(hx)$. Moreover, since h contracts the distances by at least κ^N, $|v(hx)| \leq |v(hy)| + C\kappa^{-N}\|Dv\|$. Hence

$$\begin{aligned} J^{(N)}(hx)e^{\varepsilon r^{(N)}(hx)}|v(hx)| \ &\leq CJ^{(N)}(hy)e^{\varepsilon r^{(N)}(hy)}|v(hy)| \\ &\qquad + C\kappa^{-N}J^{(N)}(hy)e^{\varepsilon r^{(N)}(hy)}\|Dv\|_{C^0}. \end{aligned}$$

Integrating this equation over y and summing over the inverse branches, we conclude that

$$\mathscr{L}_\varepsilon^N|v|(x) \leq C\int e^{\varepsilon r^{(N)}}|v| + C\kappa^{-N}\|Dv\|_{C^0}\int e^{\varepsilon r^{(N)}}. \tag{A.5}$$

But $\int e^{\varepsilon r^{(N)}} = \int \mathscr{L}_\varepsilon^N 1 \leq e^{N\alpha(\varepsilon)}$ by Lemma A.2. In the same way,

$$\int e^{\varepsilon r^{(N)}}|v| \leq \|v\|_{L^2}\left(\int e^{2\varepsilon r^{(N)}}\right)^{1/2} \leq \|v\|_{L^2}\, e^{N\alpha(2\varepsilon)/2}. \tag{A.6}$$

We obtain (for some different function $\alpha(\varepsilon)$)

$$\mathscr{L}_\varepsilon^N|v|(x) \leq Ce^{N\alpha(\varepsilon)}\|v\|_{L^2} + C\kappa^{-N}e^{N\alpha(\varepsilon)}\|Dv\|_{C^0}. \tag{A.7}$$

Let us now bound $D(\mathscr{L}_\varepsilon^N v)$. We can differentiate $J^{(N)}(hx)$. As $\|D(J^{(N)}\circ h)(x)\| \leq CJ^{(N)}\circ h$, we obtain a term bounded by $C\mathscr{L}_\varepsilon^N|v|$. If we differentiate $v\circ h(x)$, the resulting term is bounded by

$$\kappa^{-N}\sum J^{(N)}(hx)e^{\varepsilon r^{(N)}(hx)}\|Dv\|_{C^0} \leq C\kappa^{-N}\|Dv\|_{C^0}\int e^{\varepsilon r^{(N)}}, \tag{A.8}$$

bounded by $C\kappa^{-N} e^{N\alpha(\varepsilon)} \|Dv\|_{C^0}$. We have proved that

$$\|D(\mathscr{L}_\varepsilon^N v)\|_{C^0} \leq C\kappa^{-N} e^{N\alpha(\varepsilon)} \|Dv\|_{C^0} + Ce^{N\alpha(\varepsilon)} \|v\|_{L^2}. \tag{A.9}$$

Taking ε_0 small enough so that $\kappa^{-1} e^{\alpha(\varepsilon_0)} < 1$ and taking N_0 large enough, this implies the lemma. □

The following lemma essentially proves (2.6).

LEMMA A.5
There exist $N_0 > 0$, $\theta_0 < 1$, $C > 0$, $\varepsilon_0 > 0$, and a function $\alpha : (0, \varepsilon_0) \to \mathbb{R}_+$ tending to zero when $\varepsilon \to 0$ such that, for any $N \geq N_0$ and for any $A \geq 1$, the following holds. Let $\varepsilon < \varepsilon_0$, let $\psi_1, \dots, \psi_n \in \mathscr{C}_N^{A,\varepsilon}$, and let $v : Y \to \mathbb{C}$ be a C^1 function. Let $v^0 = v$, and let $v^i = \mathscr{L}^N(\psi_i v^{i-1})$. Then

$$\|v^n\|_{C^1} \leq CA\Big(\prod_{i=1}^n \|\psi_i\|_{\mathscr{C}_N^{A,\varepsilon}}\Big)\big(\theta_0^{Nn} \|v\|_{C^1} + e^{Nn\alpha(\varepsilon)} \|v\|_{L^2}\big). \tag{A.10}$$

Proof
Note first that two points x and y of Y can be joined by a path of uniformly bounded length, since $\operatorname{diam}(Y) < \infty$. If v is a C^1 function, this implies that $|v(x)| \leq C \|Dv\|_{C^0} + |v(y)|$. Integrating with respect to y, we get

$$\|v\|_{C^0} \leq C \|Dv\|_{C^0} + \int |v|. \tag{A.11}$$

Let us first prove a preliminary inequality. For any C^1 function w and any integer i,

$$\|D(\mathscr{L}_\varepsilon^{Ni} w)\|_{C^0} \leq \theta_0^{Ni} \|Dw\|_{C^0} + Ce^{Ni\alpha(\varepsilon)} \|w\|_{L^2} \tag{A.12}$$

by Lemma A.4 (applied to the time Ni). Applying (A.11) to $\mathscr{L}_\varepsilon^{Ni} w$, we obtain

$$\|\mathscr{L}_\varepsilon^{Ni} w\|_{C^0} \leq C\theta_0^{Ni} \|Dw\|_{C^0} + Ce^{Ni\alpha(\varepsilon)} \|w\|_{L^2}. \tag{A.13}$$

Let now w be a Lipschitz function. It is a uniform limit of C^1 functions w_n, with $\|Dw_n\|_{C^0} \leq C \operatorname{Lip}(w)$. Taking limits in the previous equation for w_n, we get

$$\|\mathscr{L}_\varepsilon^{Ni} w\|_{C^0} \leq C\theta_0^{Ni} \operatorname{Lip}(w) + Ce^{Ni\alpha(\varepsilon)} \|w\|_{L^2}. \tag{A.14}$$

Finally, let v be a C^1 function. The function $|v|$ is Lipschitz, and its Lipschitz coefficient is bounded by $\|Dv\|_{C^0}$. We conclude that

$$\|\mathscr{L}_\varepsilon^{Ni} |v|\|_{C^0} \leq C\theta_0^{Ni} \|Dv\|_{C^0} + Ce^{Ni\alpha(\varepsilon)} \|v\|_{L^2}. \tag{A.15}$$

We can now prove the lemma itself. We write $\gamma_i = \|\psi_i\|_{\mathscr{C}_N^{A,\varepsilon}}$. In particular, $|\psi_i(x)| \leq \gamma_i e^{\varepsilon r^{(N)}(x)}$. Hence $|v^i| \leq \gamma_i \cdots \gamma_1 \mathscr{L}_\varepsilon^{Ni}|v^0|$. As $v^i(x) = \sum_{h \in \mathscr{H}_N} J^{(N)}(hx)\psi_i(hx)v^{i-1}(hx)$, we have

$$\begin{aligned}\|Dv^i(x)\| \leq \gamma_i\Big(&\sum \|D(J^{(N)} \circ h)(x)\| e^{\varepsilon r^{(N)}(hx)}|v^{i-1}(hx)| \\ &+ \sum J^{(N)}(hx) A e^{\varepsilon r^{(N)}(hx)}|v^{i-1}(hx)| \\ &+ \sum J^{(N)}(hx) e^{\varepsilon r^{(N)}(hx)} \|Dh(x)\| \, \|Dv^{i-1}(hx)\|\Big).\end{aligned}$$

We bound these three terms. For the first one, $\|D(J^{(N)} \circ h)(x)\| \leq C J^{(N)}(hx)$. This term is therefore bounded by $C\gamma_i \cdots \gamma_1 \|\mathscr{L}_\varepsilon^{Ni}|v^0|\|_{C^0}$, which can be estimated with (A.15). For the second term, we have a similar bound, with an additional factor A.

For the third term, we bound $\|Dh(x)\|$ by κ^{-N}, and $\sum J^{(N)}(hx)e^{\varepsilon r^{(N)}(hx)} = \mathscr{L}_\varepsilon^N 1(x) \leq e^{N\alpha(\varepsilon)}$ by Lemma A.2. Taking ε small enough, we can ensure that $\kappa^{-1}e^{\alpha(\varepsilon)} \leq \theta_0$ (increasing θ_0, if necessary).

We have proved that

$$\|Dv^i\|_{C^0} \leq A\gamma_i \cdots \gamma_1 (C\theta_0^{Ni} \|Dv\|_{C^0} + Ce^{Ni\alpha(\varepsilon)} \|v\|_{L^2}) + \gamma_i \theta_0^N \|Dv^{i-1}\|_{C^0}. \quad \text{(A.16)}$$

Iterating this equation inductively over i yields

$$\begin{aligned}&\|Dv^n\|_{C^0} \\ &\quad\leq \Big(\prod_{i=1}^n \gamma_i\Big)\Big(A\sum_{i=1}^n \theta_0^{N(n-i)}(C\theta_0^{Ni} \|Dv\|_{C^0} + Ce^{Ni\alpha(\varepsilon)} \|v\|_{L^2}) + \theta_0^{Nn} \|Dv\|_{C^0}\Big) \\ &\quad\leq \Big(\prod_{i=1}^n \gamma_i\Big)\Big(CAn\theta_0^{Nn} \|Dv\|_{C^0} + CAe^{Nn\alpha(\varepsilon)} \|v\|_{L^2} + \theta_0^{Nn} \|Dv\|_{C^0}\Big) \\ &\quad\leq CA\Big(\prod_{i=1}^n \gamma_i\Big)\Big(\theta_0^{Nn/2} \|Dv\|_{C^0} + e^{Nn\alpha(\varepsilon)} \|v\|_{L^2}\Big).\end{aligned}$$

This gives the estimate of the lemma for $\|Dv^n\|_{C^0}$. Thanks to (A.11), this also implies the desired bound for $\|v^n\|_{C^0}$. □

The following technical lemma is needed later.

LEMMA A.6
There exists a constant $C_1 > 0$ such that, for any $n \in \mathbb{N}$ and for any $x \in Y$, we have

$$\sum_{h \in \mathscr{H}_n} J^{(n)}(hx)\|D(S_n^Y \phi_Y \circ h)(x)\|^4 \leq C_1^4.$$

Proof
If $h = h_n \circ \cdots \circ h_1$, then $S_n^Y \phi_Y(hx) = \sum_{i=1}^n (\phi_Y \circ h_i)(h_{i-1} \cdots h_1 x)$. Thus,

$$\left\| D(S_n^Y \phi_Y \circ h)(x) \right\|^4 \leq C \Big(\sum_{i=1}^n r(h_i \cdots h_1 x) \kappa^{-i+1} \Big)^4. \tag{A.17}$$

We use the convexity inequality $(\sum a_i x_i)^4 \leq (\sum a_i)^3 \sum a_i x_k^4$, which comes from the convexity of $x \mapsto x^4$ when $\sum a_i = 1$ (the general case can be reduced to that specific case). We take $a_i = \kappa^{-i+1}$ and $x_i = r(h_i \cdots h_1 x)$, and we obtain

$$\| D(S_n^Y \phi_Y \circ h)(x) \|^4 \leq C \sum \kappa^{-i} r(h_i \cdots h_1 x)^4. \tag{A.18}$$

Let $F_n(x) = \sum_{h_1, \ldots, h_n \in \mathscr{H}} \Big(\sum_{i=1}^n \kappa^{-i} r(h_i \cdots h_1 x)^4 \Big) J^{(n)}(h_n \cdots h_1 x)$; the sum that we want to estimate is bounded by $C F_n(x)$. As $J^{(n)}(h_n \cdots h_1 x) \leq C J^{(n)}(h_n \cdots h_1 y)$ by Lemma A.1, we have $F_n(x) \leq C F_n(y)$. Hence, $F_n(x) \leq C \int F_n$. Finally, a change of variables yields

$$\int F_n = \sum_{i=1}^n \kappa^{-i} \int_Y r(T_Y^{n-i} x)^4 \, \mathrm{d}\mu_Y(x) = \sum_{i=1}^n \kappa^{-i} \int_Y r^4 \leq \frac{\int_Y r^4}{\kappa - 1}. \tag{A.19}$$

□

A.2. Contraction for Dolgopyat's norms
To prove the contraction for Dolgopyat's norms, we essentially follow Dolgopyat's arguments as presented in [AGY, Section 7], with additional technical complications due to the facts that the involved functions are unbounded and that we want estimates that are uniform in M in Theorem 2.3.

We need the following lemma, proved in [AGY, Lemma 7.5].

LEMMA A.7
There exist constants $C_2 > 1$ and $C_3 > 0$ such that, for any ball $B(x, C_2 r)$ that is compactly included in Y, there exists a C^1 function $\rho : Y \to [0, 1]$, vanishing outside $B(x, C_2 r)$, equal to 1 on $B(x, r)$, and with $\|\rho\|_{C^1} \leq C_3 / r$.

Later, we use oscillatory integral arguments. To do that, it is important that the phases of $e^{ik S_N^Y \phi_Y \circ h}$ vary at various speeds when one uses different inverse branches h. This is ensured by the following lemma.

LEMMA A.8
There exist $C_4 > 0$ and an integer $N_0 > 0$ such that, for any $N \geq N_0$, there exist inverse branches $h_1, h_2 \in \mathscr{H}_N$ and a continuous unitary vector field $y(x)$ on Y such

that, for any $x \in Y$, we have

$$|D(S_N^Y \phi_Y \circ h_1)(x) \cdot y(x) - D(S_N^Y \phi_Y \circ h_2)(x) \cdot y(x)| \geq C_4. \tag{A.20}$$

Proof

Step 1. Let us show that there exist C' and N' such that, for any $N \geq N'$, there exist inverse branches $h_1, h_2 \in \mathscr{H}_N$, a point $x \in Y$, and a unit tangent vector y at x such that

$$|D(S_N^Y \phi_Y \circ h_1)(x) \cdot y - D(S_N^Y \phi_Y \circ h_2)(x) \cdot y| > C'. \tag{A.21}$$

We argue by contradiction, so assume that this is not the case.

Let us fix an inverse branch $h \in \mathscr{H}$, and consider the sequence of inverse branches h^n. Then $D(S_n^Y \phi_Y \circ h^n)(x) \cdot y = \sum_{k=1}^n D(\phi_Y \circ h)(h^{k-1}x) D h^{k-1}(x) \cdot y$. As $\|D(\phi_Y \circ h)\|$ is bounded and $\|Dh^{k-1}(x)\| \leq \kappa^{-k+1}$, this series converges normally to a continuous 1-form $\omega(x) \cdot y$. Let x_0 be any point in Y; the series $\sum_{k=1}^{\infty} (\phi_Y \circ h^k - \phi_Y \circ h^k(x_0))$ even converges in C^1, and its sum ψ is a C^1 function with $D\psi = \omega$.

Now let $h' \in \mathscr{H}$ be another inverse branch. Let us consider $h_n = h^{n-1} \circ h' \in \mathscr{H}_n$. Since we assume that (A.21) does not hold, $D(S_n^Y \phi_Y \circ h_n) - D(S_n^Y \phi_Y \circ h^n)$ converges pointwise to zero along a subsequence of the integers. But $D(S_n^Y \phi_Y \circ h_n) = D(\phi_Y \circ h') + \sum_{k=1}^{n-1} D(\phi_Y \circ h) D h^{k-1} D h'$. Letting n tend to infinity, we get

$$D\psi(x) \cdot y = D(\phi_Y \circ h')(x) \cdot y + D\psi(h'x) D h'(x) \cdot y. \tag{A.22}$$

Hence $D((\phi_Y + \psi - \psi \circ T_Y) \circ h') = 0$. Therefore, the function $\phi_Y + \psi - \psi \circ T_Y$ is constant on each set $h'(Y)$, $h' \in \mathscr{H}$. This contradicts the fact that ϕ_Y is not cohomologous to a locally constant function and concludes the proof of the first step.

Step 2. Let us fix an arbitrary branch $h \in \mathscr{H}$. Then $D(S_p^Y \phi_Y \circ h^p) = \sum_{k=0}^{p-1} D(\phi_Y \circ h) D h^k$ is uniformly bounded independently of p, by a constant c_0. Fix $N \geq N'$ (given by the first step) such that $c_0 \kappa^{-N} \leq C'/4$. Let h_1 and h_2 be the inverse branches given by the first step, at time N, and let x_0 and y_0 be a point in Y and a tangent vector at this point which satisfy the conclusions of the first step. We extend y_0 to a continuous vector field on a neighborhood U of x_0, still satisfying (A.21).

Since μ_Y has full support in Y, $\mu_Y(U) > 0$. Hence U intersects $\bigcap_{k>0} \bigcup_{h \in \mathscr{H}_k} h(Y)$ since μ_Y is supported on this last set. Let x_1 be a point in the intersection, and let $\ell_k \in \mathscr{H}_k$ be the inverse branch of T_Y^k such that $x_1 \in \ell_k(Y)$. Since the diameter of $\ell_k(Y)$ tends to zero when $k \to \infty$, $\ell_k(Y)$ is included in U for large enough k. In particular, there exist $k > 0$ and an inverse branch $\ell \in \mathscr{H}_k$ such that $\ell(Y) \subset U$.

Let $y_1(x) = D\ell(x)^{-1} \cdot y_0(\ell x)$. For any $p \in \mathbb{N}$ and for any $j \in \{1, 2\}$, we have

$$\begin{aligned}&|D(S^Y_{p+N+k}\phi_Y \circ h^p \circ h_j \circ \ell)(x) \cdot y_1(x) - D(S^Y_{N+k}\phi_Y \circ h_j \circ \ell)(x) \cdot y_1(x)| \\ &\quad = |D(S^Y_p \phi_Y \circ h^p)(h_j \ell x) D h_j(\ell x) \cdot y_0(\ell x)| \leq c_0 \|Dh_j(\ell x)\| \leq c_0 \kappa^{-N} \leq C'/4.\end{aligned}$$

Moreover,

$$\begin{aligned}&|D(S^Y_{N+k}\phi_Y \circ h_1 \circ \ell)(x) \cdot y_1(x) - D(S^Y_{N+k}\phi_Y \circ h_2 \circ \ell)(x) \cdot y_1(x)| \\ &\quad = |D(S^Y_N \phi_Y \circ h_1)(x) \cdot y_0(x) - D(S^Y_N \phi_Y \circ h_2)(x) \cdot y_0(x)| \geq C'.\end{aligned}$$

Adding these estimates, we obtain

$$|D(S^Y_{p+N+k}\phi_Y \circ h^p \circ h_1 \circ \ell)(x) \cdot y_1(x) - D(S^Y_{p+N+k}\phi_Y \circ h^p \circ h_2 \circ \ell)(x) \cdot y_1(x)| \geq C'/2.$$

We conclude the proof by taking $y(x) = y_1(x)/\, \|y_1(x)\|$. □

We recall that we defined a constant C_1 in Lemma A.6 and a constant C_4 in Lemma A.8.

We fix once and for all a constant $C_0 \geq \max(4C_1, 10)$. We also fix an integer N that is larger than the integers N_0 given by Lemmas A.5 and A.8 and such that $\kappa^{-N} \leq 1/1000$ and $C_4 \geq 20\kappa^{-N} C_0$.

For the remainder of this article, the D_k norms and the cones $\mathscr{E}_k$ are always defined with respect to the constant C_0. The following lemma essentially proves (2.8).

LEMMA A.9
There exists a function $\alpha : (0, \varepsilon_0) \to \mathbb{R}_+$ which tends to zero when ε tends to zero such that, for any $\varepsilon < \varepsilon_0$, $M > 0$, and $A > 0$, there exists $K > 0$ such that, for any $|\ell| \geq |k| \geq K$, for any C^1 function $v : Y \to \mathbb{C}$, and for any function $\psi \in \mathscr{C}^{A,\varepsilon}_{MN}$, we have

$$\|\mathscr{L}_k^{MN}(\psi v)\|_{D_\ell} \leq \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} e^{MN\alpha(\varepsilon)} \|v\|_{D_{2^M \ell}}. \tag{A.23}$$

Proof
Let u be such that $(u, v) \in \mathscr{E}_{2^M \ell}(C_0)$. Let

$$\tilde{u} = \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \Big(\sum_{h \in \mathscr{H}_{MN}} J^{(MN)}(hx) u(hx)^2 \Big)^{1/2}. \tag{A.24}$$

We show that there exists $\alpha(\varepsilon)$ (independent of M) such that $(e^{MN\alpha(\varepsilon)}\tilde{u}, \mathscr{L}_k^{MN}(\psi v)) \in \mathscr{E}_\ell(C_0)$.

We have

$$|\mathscr{L}_k^{MN}(\psi v)| \leq \sum_{h \in \mathscr{H}_{MN}} J^{(MN)}(hx)\psi(hx)u(hx). \tag{A.25}$$

We bound $\psi(hx)$ by $\|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}} e^{\varepsilon r^{(MN)}(hx)}$, and we use Cauchy-Schwarz inequality. We conclude that

$$\begin{aligned}|\mathscr{L}_k^{MN}(\psi v)| &\leq \|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}} \left(\sum J^{(MN)}(hx)e^{2\varepsilon r^{(MN)}(hx)}\right)^{1/2} \cdot \left(\sum J^{(MN)}(hx)u(hx)^2\right)^{1/2} \\ &= \mathscr{L}_{2\varepsilon}^{MN} 1(x)^{1/2} \cdot \tilde{u}(x).\end{aligned}$$

The coefficient $\mathscr{L}_{2\varepsilon}^{MN} 1(x)^{1/2}$ is bounded by a coefficient of the form $e^{MN\alpha(\varepsilon)}$ by Lemma A.2.

Let us now estimate the derivative of

$$\mathscr{L}_k^{MN}(\psi v)(x) = \sum_{h \in \mathscr{H}_{MN}} J^{(MN)}(hx)e^{-ikS_{MN}^Y \phi_Y(hx)}\psi(hx)v(hx). \tag{A.26}$$

If we differentiate $J^{(MN)}(hx)$, its derivative is bounded by $CJ^{(MN)}(hx)$ by Lemma A.1, and the resulting term is therefore bounded by par $Ce^{MN\alpha(\varepsilon)}\tilde{u}(x)$, as above. If we differentiate $e^{-ikS_{MN}^Y \phi_Y(hx)}$, we use Cauchy-Schwarz inequality and Lemma A.6 to obtain a bound

$$\begin{aligned}&|k|\,\|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}} \left(\sum J^{(MN)}(hx)\|D(S_{MN}^Y \phi_Y \circ h)(x)\|^4\right)^{1/4} \\ &\qquad \times \left(\sum J^{(MN)}(hx)e^{4\varepsilon r^{(MN)}(hx)}\right)^{1/4} \cdot \left(\sum J^{(MN)}(hx)u(hx)^2\right)^{1/2} \\ &\quad \leq C_1|k|e^{MN\alpha(\varepsilon)}\tilde{u}(x).\end{aligned}$$

The derivative of $\psi \circ h$ is bounded by $Ae^{\varepsilon r^{(MN)}(hx)}\,\|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}}$, and the resulting term is therefore bounded by $Ae^{MN\alpha(\varepsilon)}\tilde{u}(x)$. Finally, if we differentiate $v(hx)$, we use the inequality $\|Dv(hx)\| \leq C_0\kappa^{-MN}2^M|\ell|u(hx)$, so that the resulting term is bounded by $C_0\kappa^{-MN}2^M|\ell|e^{MN\alpha(\varepsilon)}\tilde{u}(x)$. Finally,

$$\left\|D\big(\mathscr{L}_k^{MN}(\psi v)\big)(x)\right\| \leq (C + A + C_1|k| + C_0\kappa^{-MN}2^M|\ell|)e^{MN\alpha(\varepsilon)}\tilde{u}(x). \tag{A.27}$$

The choice of N and C_0 implies that this term is bounded by $C_0|\ell|e^{MN\alpha(\varepsilon)}\tilde{u}(x)$ if K is large enough.

Let us finally bound the derivative of $\tilde{u}$, or rather of $\tilde{u}^2(x) = \|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}}^2 \sum J^{(MN)}(hx)u(hx)^2$. If we differentiate the Jacobian, the resulting term is bounded by $C\tilde{u}^2$. If we differentiate u^2, this is bounded by

$$2\,\|\psi\|^2_{\mathscr{C}^{A,\varepsilon}_{MN}} \sum J^{(MN)}(hx)\kappa^{-MN}u(hx)\,\|Du(hx)\|$$
$$\le 2\,\|\psi\|^2_{\mathscr{C}^{A,\varepsilon}_{MN}}\,\kappa^{-MN}\cdot 2^M|\ell|C_0 \sum J^{(MN)}(hx)u(hx)^2 = 2|\ell|2^M\kappa^{-MN}C_0\tilde{u}^2.$$

Hence

$$2\tilde{u}(x)\,\|D\tilde{u}(x)\| = \|D\tilde{u}^2(x)\| \le 2(C/2 + 2^M\kappa^{-MN}C_0|\ell|)\tilde{u}(x)^2. \tag{A.28}$$

Dividing by $2\tilde{u}(x)$ and using $\kappa^{-N} \le 1/1000$, we obtain the desired bound $\|D\tilde{u}(x)\| \le C_0|\ell|\tilde{u}(x)$ if $|\ell|$ is large enough.

We have proved that $(e^{MN\alpha(\varepsilon)}\tilde{u}, \mathscr{L}_k^{MN}(\psi v)) \in \mathscr{E}_\ell(C_0)$. Hence

$$\|\mathscr{L}_k^{MN}(\psi v)\|_{D_\ell} \le e^{MN\alpha(\varepsilon)}\,\|\tilde{u}\|_{L^4} \le e^{MN\alpha(\varepsilon)}\,\|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}}\,\|u\|_{L^4}. \tag{A.29}$$

Taking the infimum over the quantities $\|u\|_{L^4}$ for $(u, v) \in \mathscr{E}_{2^M\ell}(C_0)$, we obtain the lemma. □

From this point on, we concentrate on the proof of (2.7). For $v \in C^1(Y)$ and $\psi \in \mathscr{C}^{A,4\varepsilon}_{MN}$, we estimate $\mathscr{L}_k^{MN}(\psi v)$ by starting from ψv and applying M times the operator $\mathscr{L}_k^N$, which has good contraction properties thanks to the phase compensation phenomenon given by Lemma A.8. A technical issue in this argument is the fact that the functions $\psi v, \mathscr{L}_k^N(\psi v), \dots, \mathscr{L}_k^{(M-1)N}(\psi v)$ are not C^1 on Y, since the function ψ is quite wild at the beginning (it is only bounded by $e^{4\varepsilon r^{(MN)}(x)}$, so smoothness is only regained after application of $\mathscr{L}_k^{MN}$). To deal with this issue, we introduce intermediate degrees of smoothness, keeping track of the smoothness that has not yet been regained, as follows.

If Z is a subset of Y, $n \in \mathbb{N}$, and $\varepsilon \ge 0$, we say that $(u, v) \in \mathscr{E}_k(C_0, Z, n, \varepsilon)$ if the functions u and v are C^1 on Z and $|v| \le e^{\varepsilon r^{(n)}}u$, $\|Du\| \le C_0|k|u$ and $\|Dv\| \le C_0|k|e^{\varepsilon r^{(n)}}u$ on Z. In particular, $\mathscr{E}_k = \mathscr{E}_k(C_0, Y, 0, \varepsilon)$ for any $\varepsilon \ge 0$. We also write $\|v\|_{D_k(Z,n,\varepsilon)}$ for the infimum of $\|u\|_{L^4}$ over the functions u such that $(u, v) \in \mathscr{E}_k(C_0, Z, n, \varepsilon)$.

LEMMA A.10
There exists a function $\alpha : (0, \varepsilon_0) \to \mathbb{R}_+$ which tends to zero when $\varepsilon \to 0$ such that, for any $A > 0$, $n > 0$, $\varepsilon < \varepsilon_0$, and $Z \subset Y$, there exists $K > 0$ such that, for any $|\ell| \ge |k| \ge K$, for any pair of functions $(u, v) \in \mathscr{E}_{9\ell}(C_0, T_Y^{-N}Z, nN, \varepsilon)$, for any C^1 function $\chi : T_Y^{-N}Z \to [3/4, 1]$ with $\|D\chi\|_{C^0} \le |k|$ such that $|\mathscr{L}_k^N v(x)| \le \mathscr{L}^N(e^{\varepsilon r^{(Nn)}}\chi u)(x)$, we have

$$\big(e^{N\alpha(\varepsilon)}\mathscr{L}^N(\chi^2u^2)^{1/2}, \mathscr{L}_k^N v\big) \in \mathscr{E}_\ell\big(C_0, Z, (n-1)N, \varepsilon\big). \tag{A.30}$$

Note that the lemma also applies for $(u, v) \in \mathscr{E}_\ell(C_0, T_Y^{-N} Z, nN, \varepsilon)$ or $\mathscr{E}_{3\ell}(C_0, T_Y^{-N} Z, nN, \varepsilon)$, since these cones are contained in $\mathscr{E}_{9\ell}(C_0, T_Y^{-N} Z, nN, \varepsilon)$.

Proof of Lemma A.10
The proof is similar to the proof of Lemma A.9. One should check only that the additional terms coming from the function χ are harmless in the estimates. This is ensured by the choice of N and C_0. □

By Lemma A.8, we can fix two inverse branches h_1 and h_2 of T_Y^N as well as a vector field $y_0(x)$ satisfying the conclusion of the lemma. Smoothing it, we obtain a C^1 vector field y such that $1 \leq \|y\| \leq 2$ and such that, for any $x \in Y$,

$$|D(S_N^Y \phi_Y \circ h_1)(x) \cdot y(x) - D(S_N^Y \phi_Y \circ h_2)(x) \cdot y(x)| \geq C_4/2.$$

Since $\|Dh_j(x)\| \leq \kappa^{-N}$ and $C_4 \geq 20\kappa^{-N} C_0$, this implies that

$$\begin{aligned} &|D(S_N^Y \phi_Y \circ h_1)(x) \cdot y(x) - D(S_N^Y \phi_Y \circ h_2)(x) \cdot y(x)| \\ &\geq 5C_0 \max(\|Dh_1(x) \cdot y(x)\|, \|Dh_2(x) \cdot y(x)\|). \end{aligned}$$

Informally, this equation ensures that the difference between the arguments of $e^{-ikS_N^Y \phi_Y(h_1 x)}$ and $e^{-ikS_N^Y \phi_Y(h_2 x)}$ varies quickly when x moves slightly in the direction of $y(x)$. Using this, it is possible to prove the following lemma (see [AGY, Lemma 7.13] for a detailed proof).

LEMMA A.11
There exist $\delta > 0$ and $\zeta > 0$ satisfying the following property. Let $|k| \geq 10$ and $x_0 \in Y$ be such that the ball $B = B(x_0, (\zeta + \delta)/|k|)$ is compactly contained in Y. Consider $(u, v) \in \mathscr{E}_{3k}(C_0, h_1 B \cup h_2 B, 0, 0)$. Then there exist x_1 with $d(x_0, x_1) \leq \zeta/|k|$ and $j \in \{1, 2\}$ such that, for any $x \in B(x_1, \delta/|k|)$, we have

$$\begin{aligned} &|e^{-ikS_N^Y \phi_Y(h_j x)} J^{(N)}(h_j x) v(h_j x) + e^{-ikS_N^Y \phi_Y(h_{2-j} x)} J^{(N)}(h_{2-j} x) v(h_{2-j} x)| \\ &\leq \frac{3}{4} J^{(N)}(h_j x) u(h_j x) + J^{(N)}(h_{2-j} x) u(h_{2-j} x). \end{aligned}$$

If H is a set of inverse branches of T_Y^n, we write $H(Y) = \bigcup_{h \in H} h(Y)$.

LEMMA A.12
There exist $\theta_1 < 1$ and a function $\alpha : (0, \varepsilon_0) \to \mathbb{R}_+$ tending to zero when $\varepsilon \to 0$ satisfying the following property. Let $n > 0$, and let H be a finite subset of $\mathscr{H}_{nN}$. Denote by $H^{(n-1)N} \subset \mathscr{H}_{(n-1)N}$ the set of inverse branches $T_Y^N \circ h$ for $h \in H$. Then, for any H, there exists $K(H)$ such that, for any $|k| \geq K(H)$, for any function v, and

for any $\varepsilon < \varepsilon_0$, *we have*

$$\|\mathscr{L}_k^N v\|_{D_k(H^{(n-1)N}(Y),\varepsilon,(n-1)N)} \leq \theta_1^N e^{N\alpha(\varepsilon)} \|v\|_{D_{3k}(H(Y),\varepsilon,nN)}. \tag{A.31}$$

Proof
Increasing H if necessary, we can assume that, for any $h \in H^{(n-1)N}$, the branches $h_1 \circ h$ and $h_2 \circ h$ belong to H. Let $(u, v) \in \mathscr{E}_{3k}(C_0, H(Y), \varepsilon, nN)$.

Let $h \in H^{(n-1)N}$; we work on $h(Y)$ and use the weak Federer property for the constant $C = C_2(\zeta/\delta + 1)$ (where C_2 is given by Lemma A.7). Definition 1.3 provides us with constants $D > 0$ and $\eta_0(h(Y), C)$. Since the weak Federer property is uniform over the inverse branches of T_Y, we can even choose D depending only on C and not on h.

We apply the definition of the weak Federer property to $\eta = \delta/(C_2|k|)$. If $|k|$ is large enough, we indeed have $\eta < \eta_0(h(Y), C)$ for any $h \in H^{(n-1)N}$ (here, the finiteness of H is crucial). We obtain disjoint balls $B(x_1, C_2(\zeta/\delta + 1)\eta), \dots, B(x_k, C_2(\zeta/\delta + 1)\eta)$ compactly contained in $h(Y)$ and sets $A_1, \dots, A_k$ contained in $B(x_i, CD\eta)$, whose union covers $h(Y)$ and which are such that, for any $x_i' \in B(x_i, (C_2(\zeta/\delta + 1) - 1)\eta)$, we have $\mu_Y(B(x_i', \eta)) \geq \mu_Y(A_i)/D$.

On each ball $B = B(x_i, C_2(\zeta/\delta + 1)\eta) = B(x_i, (\zeta + \delta)/|k|)$, we apply Lemma A.11 to the pair of functions $(u(x)e^{\varepsilon r^{(nN)}(x)}, v(x))$ (which belongs to $\mathscr{E}_{3k}(C_0, T_Y^{-N}B, 0, 0)$). The conclusion of this lemma gives a ball $B_i' = B(x_i', \delta/|k|)$ as well as an index $j \in \{1, 2\}$. We write $\operatorname{type}(B_i') = j$. Let $B_i'' = B(x_i', \delta/(C_2|k|)) = B(x_i', \eta)$. By Lemma A.7, there exists a function ρ_i equal to 1 on B_i'' and vanishing outside of B_i', whose C^1 norm is bounded by $C|k|$.

Let us then define a function ρ on $T_Y^{-N}(hY)$ by $\rho = (\sum_{\operatorname{type}(B_i')=j} \rho_i) \circ T_Y^N$ on $h_j(hY)$ (for $j = 1, 2$) and $\rho = 0$ elsewhere. Finally, let $\chi = 1 - c\rho$ where c is small enough. Then $\|\chi\|_{C^1} \leq |k|$ if c is small enough and $|\mathscr{L}_k^N v| \leq \mathscr{L}^N(\chi u e^{\varepsilon r^{(nN)}})$ by construction (using Lemma A.11). Hence, Lemma A.10 implies that $(e^{N\alpha(\varepsilon)}\mathscr{L}^N(\chi^2 u^2)^{1/2}, \mathscr{L}_k^N v) \in \mathscr{E}_k(C_0, h(Y), (n-1)N, \varepsilon)$.

We glue together the different functions χ obtained by varying h to obtain a function (that we still denote by χ) on $H(Y)$. We still have $(e^{N\alpha(\varepsilon)}\mathscr{L}^N(\chi^2u^2)^{1/2}, \mathscr{L}_k^N v) \in \mathscr{E}_k(C_0, H^{(n-1)N}(Y), (n-1)N, \varepsilon)$. If we can prove that $\|\mathscr{L}^N(\chi^2u^2)^{1/2}\|_{L^4} \leq \beta \|u\|_{L^4}$, where $\beta < 1$ is a constant that is independent of everything else, then the proof is finished.

Let $\tilde{u} = \mathscr{L}^N(\chi^2u^2)^{1/2}$. We have

$$\begin{aligned}
\tilde{u}(x)^4 &= \Big(\sum_{h\in\mathscr{H}_n} J^{(N)}(hx)\chi(hx)^2 u(hx)^2\Big)^2 \\
&\leq \Big(\sum_{h\in\mathscr{H}_N} J^{(N)}(hx)\chi(hx)^4\Big)\cdot\Big(\sum_{h\in\mathscr{H}_N} J^{(N)}(hx)u(hx)^4\Big).
\end{aligned}$$

Let $Y_1 = \bigcup B_i''$, and let Y_2 be its complement. On Y_1, the factor $\sum_{h\in\mathscr{H}_N} J^{(N)}(hx)\chi(hx)^4$ is bounded by a uniform constant $\beta_0 < 1$, hence $\tilde{u}(x)^4 \leq \beta_0\mathscr{L}^N(u^4)(x)$. On Y_2, we have only $\tilde{u}(x)^4 \leq \mathscr{L}^N(u^4)(x)$.

Let $w = \mathscr{L}^N(u^4)$. Since $\|Du\| \leq 3C_0|k|u$, there exists a constant C such that $\|Dw\| \leq C|k|w$. Integrating this inequality along a path between two points yields $w(x) \leq e^{C|k|d(x,y)}w(y)$ for any x, y. In particular, since $A_i \subset B(x_i, CD\delta/(C_2|k|))$, there exists C such that, for any $x \in A_i$ and $y \in B_i''$, we have $w(x) \leq Cw(y)$. Integrating this inequality,

$$\frac{\int_{A_i} w}{\mu_Y(A_i)} \leq C\frac{\int_{B_i''} w}{\mu_Y(B_i'')}.$$

But $\mu_Y(A_i) \leq D\mu_Y(B_i'')$ by definition of the sets A_i, hence $\int_{A_i} w \leq C\int_{B_i''} w$. The balls B_i'' are pairwise disjoint, so we conclude that $\int_{Y_2} w \leq C'\int_{Y_1} w$ for some constant C'.

Let E be large enough so that $(E+1)\beta_0 + C' \leq E$. Then

$$\begin{aligned}(E+1)\int \tilde{u}^4 &\leq (E+1)\int_{Y_1}\beta_0 w + (E+1)\int_{Y_2} w\\ &\leq (E+1)\beta_0\int_{Y_1} w + E\int_{Y_2} w + C'\int_{Y_1} w \leq E\int w.\end{aligned}$$

Hence, $\|\tilde{u}\|_{L^4}^4 \leq (E/(E+1))\int w = (E/(E+1))\int u^4$. This is the desired inequality. □

LEMMA A.13

There exist $\theta_2 < 1$ and a function $\alpha : (0,\varepsilon_0) \to \mathbb{R}_+$ (which tends to zero when $\varepsilon \to 0$) satisfying the following property. For any $M > 0$, $\varepsilon < \varepsilon_0$, and $A > 0$, there exists $K > 0$ such that, for any C^1 function $v : Y \to \mathbb{C}$, for any $\psi \in \mathscr{C}_{MN}^{A,\varepsilon}$, and for any $|k| \geq K$, we have

$$\|\mathscr{L}_k^{MN}(\psi v)\|_{D_k} \leq e^{MN\alpha(\varepsilon)}\theta_2^{MN}\,\|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}}\,\|v\|_{D_{2^M k}}. \tag{A.32}$$

Proof

We give the proof for odd M (the proof for even M is analogous and even simpler).

We decompose $\mathscr{H}_{MN}$ as the union of a finite set H_1 (to which we apply Lemma A.12) and a set H_2 which yields a small enough contribution. Let $H \subset \mathscr{H}$ have finite complement. We take for H_1 the set of inverse branches in $\mathscr{H}_{MN}$ which are the composition of branches not belonging to H, and we take for H_2 its complement.

Let $w = 1_{H_1(Y)}\psi v$, and let $w' = 1_{H_2(Y)}\psi v$. We first estimate $\|\mathscr{L}_k^{MN}w'\|_{D_k}$. Let u be such that $(u,v) \in \mathscr{E}_{2^M k}$. Let $\tilde{u} = \|\psi\|_{\mathscr{C}_{MN}^{A,\varepsilon}}\left(\sum_{h\in H_2} J^{(MN)}(hx)u(hx)^2\right)^{1/2}$. The computation made in the proof of Lemma A.9 shows that $(e^{MN\alpha(\varepsilon)}\tilde{u}, \mathscr{L}_k^{MN} w') \in \mathscr{E}_k(C_0)$.

We have

$$\begin{aligned}\tilde{u}^2 &= \|\psi\|^2_{\mathscr{C}^{A,\varepsilon}_{MN}} \sum_{\substack{h_1,\dots,h_{MN}\in\mathscr{H}\\ \exists i, h_i\in H}} J^{(MN)}(h_{MN}\dots h_1 x) u(h_{MN}\dots h_1 x)^2 \\ &\le \|\psi\|^2_{\mathscr{C}^{A,\varepsilon}_{MN}} \sum_{i=1}^{MN} \sum_{\substack{h_1,\dots,h_{MN}\in\mathscr{H}\\ h_i\in H}} J^{(MN)}(h_{MN}\dots h_1 x) u(h_{MN}\dots h_1 x)^2 \\ &= \|\psi\|^2_{\mathscr{C}^{A,\varepsilon}_{MN}} \sum_{i=1}^{MN} \mathscr{L}^{i-1}\mathscr{L}_{0,H}\mathscr{L}^{MN-i} u^2(x),\end{aligned}$$

where $\mathscr{L}_{0,H}$ is similar to the operator $\mathscr{L}$, but the sum is only done over branches belonging to H (this operator has already been defined before Lemma A.2). This lemma shows that, if H is chosen small enough, then $\|\mathscr{L}_{0,H}\|_{L^2\to L^2}$ can be made arbitrarily small. Hence, if H is small enough (in terms of M and ε), we have

$$\|\mathscr{L}_k^{MN} w'\|_{D_k} \le (\theta_1^{MN/3} - \theta_1^{MN/2}) \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \|v\|_{D_{2M_k}}. \tag{A.33}$$

Let us fix such an H. Since M is odd, it can be written as $M = 2m+1$. The set H_1 is finite and fixed. In particular, there exists a constant B such that, for any $x \in H_1(Y)$, we have $\|D\psi(x)\| \le B \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}}$. If $|k|$ is large enough (in terms of B), this yields

$$\|w\|_{D_{3M_k}(H_1(Y),MN,\varepsilon)} \le \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \|v\|_{D_{2M_k}}. \tag{A.34}$$

Iterating m times Lemma A.10 (with $\chi = 1$), we obtain

$$\|\mathscr{L}_k^{mN} w\|_{D_{3k}(H_1^{(m+1)N}(Y),(m+1)N,\varepsilon)} \le e^{mN\alpha(\varepsilon)} \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \|v\|_{D_{2M_k}}. \tag{A.35}$$

We then apply inductively Lemma A.12. If $|k|$ is large enough, we obtain, for $i > m$,

$$\|\mathscr{L}_k^{iN} w\|_{D_k(H_1^{(M-i)N}(Y),(M-i)N,\varepsilon)} \le e^{iN\alpha(\varepsilon)} \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \theta_1^{(i-m)N} \|v\|_{D_{2M_k}}. \tag{A.36}$$

For $i = M = 2m+1$, we conclude

$$\|\mathscr{L}_k^{MN} w\|_{D_k} \le e^{MN\alpha(\varepsilon)} \|\psi\|_{\mathscr{C}^{A,\varepsilon}_{MN}} \theta_1^{MN/2} \|v\|_{D_{2M_k}}. \tag{A.37}$$

Adding up the inequalities (A.33) and (A.37), we get the conclusion of the lemma. $\square$

Proof of Theorem 2.3
We choose $\theta \in (2^{-1/(1010N)}, 1)$ such that θ^{100} is larger than the constants θ_0 given by Lemmas A.3 and A.5 and larger than θ_2 given by Lemma A.13. If $\varepsilon > 0$ is small enough, Lemma A.5 (applied to MN) shows (2.6). Moreover, (2.5) is implied by Lemma A.3. Finally, (2.8) is a consequence of Lemma A.9, and (2.7) follows from Lemma A.13. □

References

[A] J. AARONSON, *An Introduction to Infinite Ergodic Theory*, Math. Surveys Monogr. **50**, Amer. Math. Soc., Providence, 1997. MR 1450400 219

[Al] J. F. ALVES, *Strong statistical stability of non-uniformly expanding maps*, Nonlinearity **17** (2004), 1193 – 1215. MR 2069701 205

[ABV] J. F. ALVES, C. BONATTI, and M. VIANA, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Invent. Math. **140** (2000), 351 – 398. MR 1757000 204

[ALP] J. F. ALVES, S. LUZZATTO, and V. PINHEIRO, *Markov structures and decay of correlations for non-uniformly expanding dynamical systems*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), 817 – 839. MR 2172861 204

[An] N. ANANTHARAMAN, *Precise counting results for closed orbits of Anosov flows*, Ann. Sci. École Norm. Sup. (4) **33** (2000), 33 – 56. MR 1743718 194

[AP] V. ARAUJO and M. J. PACIFICO, *Physical measures for infinite-modal maps*, preprint, 2004. 205

[AGY] A. AVILA, S. GOUËZEL, and J.-C. YOCCOZ, *Exponential mixing for the Teichmüller flow*, Publ. Math. Inst. Hautes Études Sci. **104** (2006), 143 – 211. MR 2264836 194, 273, 278

[BV1] V. BALADI and B. VALLÉE, *Euclidean algorithms are Gaussian*, J. Number Theory **110** (2005), 331 – 386. MR 2122613 194

[BV2] ———, *Exponential decay of correlations for surface semi-flows without finite Markov partitions*, Proc. Amer. Math. Soc. **133** (2005), 865 – 874 (electronic). MR 2113938 194

[BGK] J.-B. BARDET, S. GOUËZEL, and G. KELLER, *Limit theorems for coupled interval maps*, Stoch. Dyn. **7** (2007), 17 – 36. MR 2293068 209

[BP] M. I. BRIN and J. B. PESIN, Partially hyperbolic dynamical systems (in Russian), Izv. Akad. Nauk. SSSR Ser. Mat. **38**, (1974) 170–212. MR 0343316 194

[BLV] H. BRUIN, S. LUZZATTO, and S. VAN STRIEN, *Decay of correlations in one-dimensional dynamics*, Ann. Sci. École Norm. Sup. (4) **36** (2003), 621 – 646. MR 2013929 206

[CG] J.-P. CONZE and Y. GUIVARC'H, "Densité d'orbites d'actions de groupes linéaires et propriétés d'équidistribution de marches aléatoires" in *Rigidity in Dynamics and Geometry (Cambridge, 2000)*, Springer, Berlin, 2002, 39 – 76. MR 1919392 194, 195, 196

[DV] W. DE MELO and S. VAN STRIEN, *One-dimensional Dynamics*, Ergeb. Math. Grenzgeb. (3) **25**, Springer, Berlin, 1993. MR 1239171 262

[DS] T.-C. DINH and N. SIBONY, *Decay of correlations for meromorphic maps and Hölder observables*, preprint, 2005. 223

[D1] D. DOLGOPYAT, *On decay of correlations in Anosov flows*, Ann. of Math. (2) **147** (1998), 357 – 390. MR 1626749 194, 207

[D2] ———, *On mixing properties of compact group extensions of hyperbolic systems*, Israel J. Math. **130** (2002), 157 – 205. MR 1919377 194, 207

[D3] ———, *Limit theorems for partially hyperbolic systems*, Trans. Amer. Math. Soc. **356** (2004), 1637 – 1689 (electronic). MR 2034323 194

[G1] S. GOUËZEL, *Central limit theorem and stable laws for intermittent maps*, Probab. Theory Related Fields **128** (2004), 82 – 122. MR 2027296 247

[G2] ———, *Berry-Esseen theorem and local limit theorem for non uniformly expanding maps*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), 997 – 1024. MR 2172207 203, 208, 241

[G3] ———, *Decay of correlations for nonuniformly expanding systems*, Bull. Soc. Math. France **134** (2006), 1 – 31. MR 2233699 204, 205, 206

[G4] ———, *Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes*, Ph.D. dissertation, Université Paris-Sud, 2004. 199

[GH] Y. GUIVARC'H and J. HARDY, *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov*, Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), 73 – 98. MR 0937957 193

[GR] Y. GUIVARC'H and A. RAUGI, *Actions of large semigroups and random walks on isometric extensions of boundaries*, Ann. Sci. École Norm. Sup. (4) **40** (2007), 209 – 249. MR 2339285 193, 195, 196

[H] H. HENNION, *Sur un théorème spectral et son application aux noyaux lipchitziens*, Proc. Amer. Math. Soc. **118** (1993), 627 – 634. MR 1129880 209

[HH] H. HENNION and L. HERVÉ, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Math. **1766**, Springer, Berlin, 2001. MR 1862393 193

[K] A. G. KACHUROVSKIĬ, *Rates of convergence in ergodic theorems*, Russian Math. Surveys **51** (1996), 653 – 703. MR 1422228 224

[N] F. NAUD, *Expanding maps on Cantor sets and analytic continuation of zeta functions*, Ann. Sci. École Norm. Sup. (4) **38** (2005), 116 – 153. MR 2136484 194, 203

[OW] D. ORNSTEIN and B. WEISS, *On the Bernoulli nature of systems with some hyperbolic structure*, Ergodic Theory Dynam. Systems **18** (1998), 441 – 456. MR 1619567 208

[PS] M. POLLICOTT and R. SHARP, *Error terms for closed orbits of hyperbolic flows*, Ergodic Theory Dynam. Systems **21** (2001), 545 – 562. MR 1827118 194

[R] D. J. RUDOLPH, *Classifying the isometric extensions of a Bernoulli shift*, J. Analyse Math. **34** (1978), 36 – 60. MR 0531270 208

[Ru] D. RUELLE, *Flots qui ne mélangent pas exponentiellement*, C. R. Acad. Sci. Paris Sér. I Math. **296** (1983), 191 – 193. MR 0692974 200

[S] O. SARIG, *Subexponential decay of correlations*, Invent. Math. **150** (2002), 629 – 653. MR 1946554 199, 208

[St] L. STOYANOV, *Spectrum of the Ruelle operator and exponential decay of correlations for open billiard flows*, Amer. J. Math. **123** (2001), 715 – 759. MR 1844576 194

[T] H. TRIEBEL, *Interpolation theory, function spaces, differential operators*, North-Holland Math. Library **18**, North-Holland, Amsterdam, 1978. MR 0503903 223

[Ts] M. TSUJII, *Decay of correlations in suspension semi-flows of angle-multiplying maps*, Ergodic Theory Dynam. Systems **28** (2008), 291 – 317. MR 2380311 194, 207

[W] S. WADDINGTON, *Large deviation asymptotics for Anosov flows*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), 445 – 484. MR 1404318 194

[Y1] L.-S. YOUNG, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), 585 – 650. MR 1637655 199

[Y2] ———, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153 – 188. MR 1750438 199

Institut de Recherche Mathématique de Rennes, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes CEDEX, France; sebastien.gouezel@univ-rennes1.fr