Statistical properties of a skew product with a curve of neutral points

SÉBASTIEN GOUÈZEL
IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France
(e-mail: sebastien.gouezel@univ-rennes1.fr)

(Received 9 November 2004 and accepted in final form 7 March 2006)

Abstract. We study a skew product with a curve of neutral points. We show that there exists a unique absolutely continuous invariant probability measure, and that the Birkhoff averages of a sufficiently smooth observable converge to a normal law or a stable law, depending on the average of the observable along the neutral curve.

1. Introduction
Let $T : M \to M$ be a map on a compact manifold. While uniformly hyperbolic or uniformly expanding dynamics are well understood, problems arise when there are neutral fixed points (where the differential of $T$ has an eigenvalue equal to 1). The one-dimensional case has been thoroughly studied, particularly when $T$ has only one neutral fixed point (see [LSV99] and references therein). The normal form at the fixed point dictates the asymptotics of the dynamics and, in particular, the speed of mixing and the convergence of Birkhoff sums to limit laws [Gou04, Zwe03].

In this article, we study the same type of phenomenon, but in higher dimension. In contrast to [Hu01, PY01] (where the case of isolated fixed points is considered), our models admit a whole invariant neutral curve. We show that the one-dimensional results remain essentially true.

More precisely, for $\alpha > 0$, define a map $T_\alpha$ on $[0, 1]$ by

$$T_\alpha(x) = \begin{cases} 
  x (1 + 2^\alpha x^\alpha) & \text{if } 0 \leq x \leq \frac{1}{2}, \\
  2x - 1 & \text{if } \frac{1}{2} < x \leq 1.
\end{cases}$$

It has a neutral fixed point at 0, behaving like $x (1 + x^\alpha)$. This map admits an absolutely continuous invariant measure, which is of finite mass if and only if $\alpha < 1$. To mix different such behaviours, we consider a skew product, similar to the Alves–Viana map [Via97] but where the unimodal maps are replaced by $T_\alpha$. Let $\alpha : S^1 \to (0, \infty)$ be a map with minimum $\alpha_{\text{min}}$ and maximum $\alpha_{\text{max}}$. Assume that:
\( \alpha \) is \( C^2 \);
(2) \( 0 < \alpha_{\text{min}} < \alpha_{\text{max}} < 1 \);
(3) \( \alpha \) takes the value \( \alpha_{\text{min}} \) at a unique point \( \omega_0 \in S^1 \), with \( \alpha''(\omega_0) > 0 \);
(4) \( \alpha_{\text{max}} < \frac{1}{2} \alpha_{\text{min}} \) (which implies \( \alpha_{\text{max}} < \alpha_{\text{min}} + \frac{1}{4} \)).

These conditions are, for example, satisfied by \( \alpha(\omega) = \alpha_{\text{min}} + \varepsilon(1 + \sin(2\pi \omega)) \) where \( \alpha_{\text{min}} \in (0, 1) \) and \( \varepsilon \) is small enough.

We define a map \( T \) on \( S^1 \times [0, 1] \) by

\[
T(\omega, x) = (F(\omega), T_{\alpha(\omega)}(x))
\]  

where \( F(\omega) = 4\omega \).

The qualitative behaviour of \( T \) can be described as follows. In a compact set disjoint from \( S^1 \times \{0\} \), say \( S^1 \times [\frac{1}{\pi}, 1] \), \( T \) is uniformly expanding. Hence, the interesting points are the points \( x = (\omega, \varepsilon) \) with small \( \varepsilon \). Such a point takes a long time to reach \( S^1 \times [\frac{1}{\pi}, 1] \) since each map \( T_{\alpha(\omega)} \) has a neutral fixed point at 0. The iterates of \( x \) will feel the strongest expansion essentially when they are of the form \( (\omega', \varepsilon') \) with \( \omega' \) close to \( \omega_0 \) (where the neutral point is the least neutral). Therefore, the precise behaviour of the map \( T \) will depend on a strong way on the behaviour of \( \alpha(\omega') \) for \( \omega' \) close to \( \omega_0 \), and on the value of \( \alpha_{\text{min}} \). This explains the conditions \( \alpha_{\text{min}} < 1 \) and \( \alpha''(\omega_0) > 0 \), which are really important for our analysis. On the other hand, the other conditions \( \alpha_{\text{max}} < 1 \) and \( \alpha_{\text{max}} < \frac{1}{2} \alpha_{\text{min}} \) are merely technical. They could probably be removed at the expense of greater technicalities in the proofs.

In the following, we will generalize the one-dimensional results on the maps \( T_\alpha \) to this skew product \( T \). First of all, in §2, we prove that there exists a unique absolutely continuous invariant probability measure \( m \), whose density \( h \) is, in fact, Lipschitz on every compact subset of \( S^1 \times (0, 1) \) (Theorem 2.10). In §3, we prove limit theorems for abstract Markov maps (using a method essentially due to [MT04] and recalled in Appendix A, and estimates of [AD01b] and [Gou04]). Finally, in §§4 and 5, we study the limit laws of Birkhoff sums for the skew product \( T \), and we obtain the convergence to a normal law or a stable law, depending on the value of \( \alpha_{\text{min}} \). We obtain the following theorem (see Theorem 5.1 for more details).

**THEOREM 1.1.** Set

\[
A = \frac{1}{4(\alpha_{\text{min}}^{1/2}\sqrt{\pi/2\alpha''(\omega_0)})^{1/\alpha_{\text{min}}}} \int_{S^1 \times [1/2]} h \, d \text{Leb},
\]

where \( h \) is the density of the absolutely continuous invariant probability measure.

Let \( f \) be a Lipschitz function on \( S^1 \times [0, 1] \), with \( \int f \, dm = 0 \). Write \( c = \int_{S^1 \times [0, 1]} f \, d \text{Leb} \) and \( S_n f = \sum_{k=0}^{n-1} f \circ T^k \). Then:

- if \( \alpha_{\text{min}} < \frac{1}{2} \), there exists \( \sigma^2 \geq 0 \) such that \( (1/\sqrt{n}) S_n f \to \mathcal{N}(0, \sigma^2) \);
- if \( \frac{1}{2} \leq \alpha_{\text{min}} < 1 \) and \( c = 0 \), then there exists \( \sigma^2 \geq 0 \) such that \( (1/\sqrt{n}) S_n f \to \mathcal{N}(0, \sigma^2) \);
- if \( \alpha_{\text{min}} = \frac{1}{2} \) and \( c \neq 0 \), then \( S_n f / \sqrt{(c^2 A^4/4) n/(\ln n)^2} \to \mathcal{N}(0, 1) \);
- if \( \frac{1}{2} < \alpha_{\text{min}} < 1 \) and \( c \neq 0 \), then \( S_n f / n^{\alpha_{\text{min}}} \sqrt{\alpha_{\text{min}}} \ln n \to Z \), where the random variable \( Z \) has an explicit stable distribution.
An interesting feature of this example is that its study involves the sophisticated mixing properties of \( F \), particularly a multiple decorrelation property, proved in Appendix B using [Pën02].

**Remark.** Theorems of [Gou04] could be used instead of the method of [MT04] to get the limit laws. However, the proof of [Gou04] is much more complicated than the elementary method of [MT04], and less versatile. Among others, an advantage of this new method is that it can easily be extended to stable laws of index 1, in contrast to [Gou04].

In fact, the previous results remain true for a much larger family of maps. Although we will only give the proofs for the previous maps for the sake of simplicity, we indicate now the more general results that can be proved with the same arguments.

We first define the generalizations of the maps \( T_\alpha \). For \( \alpha \in (0, 1) \), consider a map \( \tilde{T}_\alpha : [0, 1] \to [0, 1] \) such that \( \tilde{T}_\alpha \) is an increasing diffeomorphism between \([0, x_\alpha)\) and \([0, 1)\) (for some \( 0 < x_\alpha < 1 \)) and between \([x_\alpha, 1] \) and \([0, 1] \). Assume that \( \alpha \mapsto x_\alpha \) is \( C^1 \), that the map \( (x, \alpha) \mapsto \tilde{T}_\alpha(x) \) is \( C^1 \) on the sets \([0 < x < x_\alpha]\) and \([x_\alpha \leq x \leq 1]\), and that \( \tilde{T}_\alpha(x) > 1 \) for all \( x \neq 0 \). We also need to prescribe the behaviour of \( \tilde{T}_\alpha \) close to 0. Let \( \epsilon_0 > 0 \). Assume that \( \tilde{T}_\alpha(x) = x + c_\alpha x^{1+\alpha}(1 + f_\alpha(x)) \) for \( x \in [0, \epsilon_0] \), where \( c_\alpha > 0 \) depends continuously on \( \alpha \), \( f_\alpha(0) = 0 \) and \( (x, \alpha) \mapsto f_\alpha(x) \) is continuous on \([0, \epsilon_0] \times (0, 1) \). Finally, assume that \( \tilde{T}_\alpha \) is \( C^3 \) on \((0, \epsilon_0)\) with non-negative Schwarzian derivative and that the partial derivatives of the function \( (x, \alpha) \mapsto \tilde{T}_\alpha(x) \) are bounded by \( C_\alpha x^{\alpha-1} \) on \((0, \epsilon_0) \times (\epsilon, 1 - \epsilon) \) for all \( \epsilon > 0 \).

Let \( \alpha : S^1 \to (0, 1) \) be a \( C^1 \) map. Let \( \tilde{F} : S^1 \to S^1 \) be a \( C^2 \) uniformly expanding map, such that \( \tilde{F}^n(\omega) = \tilde{T}_{\alpha(n)}(x) \) for all \( \omega \in S^1 \) and \( x \in [0, 1] \). This ensures that the map \( \tilde{T} \) defined on \( S^1 \times [0, 1] \) by \( \tilde{T}(\omega, x) = (\tilde{F}^n(\omega), \tilde{T}_{\alpha(n)}(x)) \) is partially hyperbolic. The arguments of §2 apply to \( \tilde{T} \), and show that \( \tilde{T} \) admits an absolutely continuous invariant probability measure \( \tilde{m} \), which is ergodic and whose density is Lipschitz on every compact subset of \( S^1 \times (0, 1] \).

To obtain limit theorems, we need additional assumptions. Let \( \alpha_{\min} \) be the minimal value taken by the function \( \alpha \), and \( \alpha_{\max} \) its maximal value. Assume that \( \alpha_{\max} < \frac{1}{2} \alpha_{\min} \) and that Leb\( \{ \omega \in S^1 \mid |\alpha(\omega) - \alpha_{\min}| < \epsilon \} \sim C\epsilon^{\gamma} \) for some \( C > 0 \) and \( \gamma \geq 0 \). This is, for example, the case when \( \alpha \) is \( C^2 \) and takes the value \( \alpha_{\min} \) at a unique point \( \alpha(\omega) > 0 \) (and, in this case, \( \gamma = \frac{1}{2} \)). This holds more generally if \( \alpha'(\omega_0), \ldots, \alpha'(\omega_n) = 0 \) and \( \alpha'(\omega) > 0 \) for some \( p \in \mathbb{N} \) (and, in this case, \( \gamma = 1/p \)). The following analogue of Theorem 1.1 then holds. Let \( f \) be a Lipschitz function on \( S^1 \times [0, 1] \). Denote by \( \mu \) the probability measure on \( S^1 \) which is absolutely continuous and \( \tilde{F} \)-invariant. Let \( \epsilon = \int_{S^1} f d\mu \). If \( \alpha_{\min} < \frac{1}{2} \), or \( \frac{1}{2} \leq \alpha_{\min} < 1 \) and \( c = 0 \), then \( S_n f / \sqrt{n} \) converges in distribution to a normal law \( \mathcal{N}(0, \sigma^2) \) for some \( \sigma^2 \geq 0 \). On the other hand, if \( \alpha_{\min} = \frac{1}{2} \) and \( c \neq 0 \), then \( S_n f / n^{1/2} \) converges in distribution to a normal law, and if \( \alpha_{\min} > \frac{1}{2} \) and \( c \neq 0 \), then \( S_n f / n^{1/2} \) converges in distribution to a stable law, which can be explicitly given in terms of \( \mu \) and of the density of \( \tilde{m} \).

**Remark.** An important assumption of our arguments is the fact that the maps \( T_\alpha \) are Markov. This is heavily used in our computations of return times. With the present techniques, it is unlikely that this assumption could be removed.
In this article, $a(n) \sim b(n)$ means that $a(n)/b(n) \to 1$ when $n \to \infty$. The integral with respect to a probability measure will sometimes be denoted by $E(\cdot)$. Finally, $\lfloor x \rfloor$ will denote the integer part of $x$. From this point on, we will only deal with the skew product $T$, and not its generalization $\bar{T}$.

2. Invariant measure
An important property of the map $T$, that will be used thoroughly in what follows, is that it is Markov: there exists a partition of the space such that every element of this partition is mapped by $T$ on a union of elements of this partition. In fact, we will consider $TY$ (the induced map on $Y = S^1 \times (\frac{1}{2}, 1]$), which is also Markov and expanding, contrary to $T$. We will apply to $TY$ classical results on expanding Markov maps (also called Gibbs–Markov maps), which we recall below.

2.1. Markov maps and invariant measures. Let $(Y, B, m_Y)$ be a standard probability space, endowed with a bounded metric $d$. A non-singular map $TY$ defined on $Y$ is said to be a Markov map if there exists a finite or countable partition $\alpha$ of $Y$ such that for all $a \in \alpha$, $m_Y(a) > 0$, $TY(a)$ is a union (mod 0) of sets of $\alpha$, and $TY : a \to TY(a)$ is invertible. In this case, $\alpha$ is a Markov partition for $TY$.

A Markov map $TY$ (with a Markov partition $\alpha$) is a Gibbs–Markov map [Aar97] if:

1. $TY$ has the big image property: $\inf_{a \in \alpha} m_Y(TY(a)) > 0$;
2. there exists $\lambda > 1$ such that for all $a \in \alpha$, for all $x, y \in a$, $d(TYx, TYy) \geq \lambda d(x, y)$;
3. let $g$ be the inverse of the Jacobian of $TY$, i.e. on a set $a \in \alpha$, $g(x) = (\frac{dm_Y}{d(m_Y \circ (TY)_\alpha)}(x))$, then there exists $C > 0$ such that for all $a \in \alpha$, for almost all $x, y \in a$,

$$\left| 1 - \frac{g(x)}{g(y)} \right| \leq Cd(TYx, TYy).$$

This definition is slightly more general than the definition of [Aar97]: the distance $d = d_\tau$ considered there is given by $d_\tau(x, y) = \tau(s(x, y))$ where $\tau < 1$ and $s(x, y)$ is the separation time of $x$ and $y$, i.e.

$$s(x, y) = \inf\{n \in \mathbb{N} \mid \exists a \in \alpha, T^n x \in a, T^n y \in a\}.$$

The proof of [Aar97, Theorem 4.7.4] still works in our context, and gives the following.

**Theorem 2.1.** Let $TY$ be a transitive Gibbs–Markov map (for all $a, b \in \alpha$, there exists $n \in \mathbb{N}$, $m_Y(T^n a \cap b) > 0$) such that $\text{Card}(\alpha_\ast) < \infty$, where $\alpha_\ast$ is the partition generated by the images $TY(a)$ for $a \in \alpha$. Then $TY$ is ergodic, and there exists a unique absolutely continuous (with respect to $m_Y$) invariant probability measure, denoted by $\mu_Y$.

Moreover, $\mu_Y = hm_Y$ where the density $h$ is bounded and bounded away from 0, and Lipschitz on every set of $\alpha_\ast$.

2.2. Preliminary estimates. To apply Theorem 2.1, we will construct a Markov partition, and control the distortion of the inverse branches of $TY$. 

We will write $T^0_\omega = T_{\alpha(F^0_\omega)} \circ \cdots \circ T_{\alpha(\omega)}$, whence $T^n(\omega, x) = (F^n_\omega, T^n_\omega(x))$.
Write also $d((\omega_1, x_1), (\omega_2, x_2)) = |\omega_1 - \omega_2| + |x_1 - x_2|$. A point of $S^1 \times [0, 1]$ will be denoted by $x = (\omega, x)$. Finally, set $d_{vert}((\omega_1, x_1), (\omega_2, x_2)) = |x_2 - x_1|/\alpha$.

Define $X_0(\omega) = 1$, $X_1(\omega) = \frac{1}{2}$, and for $n \geq 2$, $X_n(\omega)$ is the preimage in $[0, \frac{1}{2})$ of $X_{n-1}(F\omega)$ by $T_{\alpha}(\omega)$. These $X_n$ will be useful in the construction of a Markov partition for $T$, in §2.3.

**PROPOSITION 2.2.** There exists $C > 0$ such that for all $n \in \mathbb{N}^*$, for all $\omega \in S^1$, 

$$\frac{1}{Cn^{1/\alpha_{\min}}} \leq X_n(\omega) \leq \frac{C}{n^{1/\alpha_{\max}}}.$$ 

**Proof.** Write $Z_1 = \frac{1}{2}$ and $V(Z_{n+1}) = Z_n$ where $V(x) = x(1 + 2^{\alpha_{\max}}x^{\alpha_{\min}})$. We easily check inductively that $X_n < X_{n}(\omega)$ for every $\omega$, since $V(x) > T_{\alpha}(\omega)(x)$ for every $\omega$. It is thus sufficient to estimate $Z_n$ to get the minoration. As $V(x) > x$, the sequence $Z_n$ is decreasing, and non-negative. Hence, it tends to a fixed point of $V$, necessarily 0.

We have 

$$\frac{1}{Z_{n+1}^{\alpha_{\min}}} = \frac{1}{Z_{n+1}^{\alpha_{\min}}(1 + 2^{\alpha_{\max}}Z_{n+1}^{\alpha_{\min}})} = \frac{1}{Z_{n+1}^{\alpha_{\min}}(1 - \alpha_{\min}2^{\alpha_{\max}}Z_{n+1}^{\alpha_{\min}} + o(Z_{n+1}^{\alpha_{\min}}))} = \frac{1}{Z_{n+1}^{\alpha_{\min}}} - \alpha_{\min}2^{\alpha_{\max}} + o(1).$$

A summation gives $1/Z_{n}^{\alpha_{\min}} \sim m\alpha_{\min}2^{\alpha_{\max}}$, whence $Z_n \sim C/m^{1/\alpha_{\min}}$, which concludes the minoration.

The majoration is similar, using a sequence $Z'_n$ with $Z'_n \geq X_n(\omega)$. □

We fix once and for all a large enough constant $D$. The following definition is analogous to a definition of Viana [Via97].

**Definition 2.3.** Let $\psi : K \to [0, 1]$, where $K$ is a subinterval of $S^1$. We say that the graph of $\psi$ is an admissible curve if $\psi$ is $C^1$ with $|\psi'| \leq D$.

**PROPOSITION 2.4.** Let $\psi$ be an admissible curve, defined on $K$ with $|K| < \frac{1}{4}$, and included in $K \times [0, \frac{1}{2}]$ or $K \times (\frac{1}{2}, 1)$. Then the image of $\psi$ by $T$ is still an admissible curve.

**Proof.** Let $(u, v)$ be a tangent vector at $(\omega, x)$ with $|v| \leq D|u|$, we have to check that its image $(u', v')$ by $DT(\omega, x)$ still satisfies $|v'| \leq D|u'|$.

Assume first that $x \leq \frac{1}{2}$, whence $u' = 4u$ and $v' = (1 + (2x)^{\alpha(\omega)}(\alpha(\omega) + 1))v + x\ln(2x)\alpha'(\omega)(2x)^{\alpha(\omega)}u$. As $\alpha(\omega) \leq \alpha_{\max} \leq 1$, we get $|v'| \leq 3|v| + C|u|$ for a constant $C$ (which depends only on $\|u''\|_{\infty}$). Thus,

$$\frac{|v'|}{|u'|} \leq \frac{3|v|}{4|u|} + \frac{C}{4}.$$ 

(6)

This will give $|v'|/|u'| \leq D$ if $\frac{3}{4}D + C/4 \leq D$, which is true if $D$ is large enough.

Assume then that $x > \frac{1}{2}$. Then $u' = 4u$ and $v' = 2v$, and there is nothing to prove. □

**COROLLARY 2.5.** Let $(\omega_1, x_1)$ and $(\omega_2, x_2)$ be two points in $S^1 \times [0, \frac{1}{2})$ with $|x_1 - x_2| \leq D|\omega_1 - \omega_2|$ and $|\omega_1 - \omega_2| \leq \frac{1}{4}$. Then their images satisfy $|x_1' - x_2'| \leq D|\omega_1' - \omega_2'|.$
Proof. Use a segment between the two points: it is an admissible curve. Hence, its image is still admissible. \hfill \Box

2.3. The Markov partition. Set \( Y = S^1 \times (\frac{1}{2}, 1) \). For \( x \in Y \), set \( \varphi_Y(x) = \inf \{ n > 0 \mid T^n(x) \in Y \} \): this is the first return time to \( Y \), everywhere finite. The map \( T_Y(x) := T^{\varphi_Y(x)}(x) \) is the map induced by \( T \) on \( Y \). We will show that \( T_Y \) is a Gibbs–Markov map, by constructing an appropriate Markov partition.

If \( I \) is an interval of \( S^1 \), we will abusively write \( I \times [X_{n+1}, X_n] \) for \( \{ (\omega, x) \mid \omega \in I, x \in [X_{n+1}(\omega), X_n(\omega)] \} \). Set \( I_n(\omega) = [X_{n+1}(\omega), X_n(\omega)] \) (or \( \{ \omega \} \times [X_{n+1}(\omega), X_n(\omega)] \), depending on the context).

By definition of \( X_n \), \( T \) maps \( \{ \omega \} \times I_n(\omega) \) bijectively on \( \{ F\omega \} \times I_{n-1}(F\omega) \). Thus, the interval \( I_n(\omega) \) returns to \( \frac{1}{2}, 1 \) in exactly \( n \) steps.

Let \( Y_n(\omega) \) be the preimage in \( \frac{1}{2}, 1 \) of \( X_{n-1}(F\omega) \) under \( T_{\alpha}(\omega) \). Thus, the interval \( J_n(\omega) = [Y_{n+1}(\omega), Y_n(\omega)] \) returns to \( \frac{1}{2}, 1 \) in \( n \) steps.

We fix once and for all \( 0 < \varepsilon_0 < \frac{1}{4} \), small enough so that \( D\varepsilon_0 \) is less than the length of every interval \( I_1(\omega) \). (This condition will be useful in distortion estimates.)

Let \( q \) be large enough so that \( 1/4^q < \varepsilon_0 \), and consider \( A_{s,n} = [s/4^q+n, (s+1)/4^q+n] \times J_n \), for \( n \in \mathbb{N} \) and \( 0 \leq s \leq 4^q+n-1 \): this set is mapped by \( T^n \) on \( [s/4^q, (s+1)/4^q] \times \frac{1}{2}, 1 \). Let \( K_0 \), \( \ldots \), \( K_{4^q-1} \) be the sets \( [i/4^q, (i+1)/4^q] \times \frac{1}{2}, 1 \). Then the map \( T_Y \) is an isomorphism between each \( A_{s,n} \) and some \( K_i \). Consequently, the map \( T_Y \) is Markov for the partition \( \{ A_{s,n} \} \), and it has the big image property.

To apply Theorem 2.1, we need expansion (for (2) in the definition of Gibbs–Markov maps) and distortion control (for (3)). The expansion is given by the next proposition, and the distortion is estimated in §2.4.

On the intervals \( [X_1(\omega), X_1(\omega)] \), the derivative of \( T_{\alpha}(\omega) \) is greater than 1, whence greater than a constant \( 2 > \lambda > 1 \), independent of \( \omega \).

For \( (\omega_1, x_1) \) and \( (\omega_2, x_2) \in S^1 \times [0, 1] \), set
\[
d'(a; (\omega_1, x_1), (\omega_2, x_2)) = a|x_1 - x_2| + |a_1 - a_2| \tag{7}\]
where \( a = (1-\lambda/4)/D \).

PROPOSITION 2.6. On each \( A_{s,n} \), the map \( T^n \) is expanding by at least \( \lambda \) for the distance \( d' \).

Proof. For \( n = 1 \) (the points return directly to \( S^1 \times \frac{1}{2}, 1 \)), everything is linear and the result is clear. Assume that \( n \geq 2 \). Take \( (\omega_1, x_1) \) and \( (\omega_2, x_2) \in A_{s,n} \), with, for example, \( x_2 \geq x_1 \).

Since \( (\omega_1, x_1) \in A_{s,n} \), this point returns to \( S^1 \times \frac{1}{2}, 1 \) after exactly \( n \) iterations. Since \( x_1 \leq x_2 \) and \( (\omega_2, x_2) \) returns to \( S^1 \times \frac{1}{2}, 1 \) after exactly \( n \) iterations, the point \( (\omega_2, x_1) \) takes at least \( n \) iterations to come back to \( S^1 \times \frac{1}{2}, 1 \). Therefore, we can apply Corollary 2.5 \( n-1 \) times to \( (\omega_1, x_1) \) and \( (\omega_2, x_1) \). We get that in vertical distance,
\[
d_{\text{vert}}(T^n(\omega_1, x_1), T^n(\omega_2, x_1)) \leq D|F^n\omega_1 - F^n\omega_2| \tag{8}\]
In particular, \( T^n_{\alpha_1}(x_1) \geq T^n_{\alpha_2}(x_1) - D\varepsilon_0 \geq \frac{1}{2} - D\varepsilon_0 \). Thus, by the definition of \( \varepsilon_0 \),
\[
T^n(x_2, x_1) \in I_i(F^n\omega_2) \quad \text{for } i = 0 \text{ or } 1. \tag{9}\]
Taking the preimage under $T$, this implies that $T^{n-1}(\omega_2, x_1) \in [X_3(F^{n-1})_{\omega_2}, X_1(F^{n-1})_{\omega_2}]$. Moreover, $T^{n-1}(\omega_2, x_2) \in [X_2(F^{n-1})_{\omega_2}, X_1(F^{n-1})_{\omega_2}] \subset [X_3(F^{n-1})_{\omega_2}, X_1(F^{n-1})_{\omega_2}]$. Since each map $T_\alpha$ is expanding, we also have $d_{\text{ver}}(T^{n-1}(\omega_2, x_1), T^{n-1}(\omega_2, x_2)) \geq |x_1 - x_2|$. We apply once more $T$, which expands at least by $\lambda$ on $[X_3(F^{n-1})_{\omega_2}, X_1(F^{n-1})_{\omega_2}]$ by definition of $\lambda$, and get

$$d_{\text{ver}}(T^n(\omega_2, x_1), T^n(\omega_2, x_2)) \geq \lambda|x_1 - x_2|.$$

Finally,

$$d'(T^n(\omega_1, x_1), T^n(\omega_2, x_2))$$

$$= a d_{\text{ver}}(T^n(\omega_1, x_1), T^n(\omega_2, x_2)) + |F^n\omega_1 - F^n\omega_2|$$

$$\geq a d_{\text{ver}}(T^n(\omega_2, x_1), T^n(\omega_2, x_2)) - a d_{\text{ver}}(T^n(\omega_1, x_1), T^n(\omega_2, x_1))$$

$$+ |F^n\omega_1 - F^n\omega_2|$$

$$\geq a\lambda|x_1 - x_2| - aD|F^n\omega_1 - F^n\omega_2| + |F^n\omega_1 - F^n\omega_2|.$$

The proposition will be proved if $(1-aD)|F^n\omega_1 - F^n\omega_2| \geq \lambda|x_1 - x_2|$. Indeed, we have $(1-aD)|F^n\omega_1 - F^n\omega_2| = (1-aD)^4|x_1 - x_2| \geq (1-aD)|x_1 - x_2| = \lambda|x_1 - x_2|$. □

2.4. Distortion bounds.

**Lemma 2.7.** There exists a constant $E > 0$ such that for all $n \geq 0$, for all $\omega_1, \omega_2 \in S^1$ with $|\omega_1 - \omega_2| \leq \varepsilon_0/4^n$, for all $x_1 \in J_\alpha(\omega_1)$ with $T^{n-1}_{\omega_2} x_1 \leq \frac{1}{2}$,

$$|L(T^n_{\omega_1})'(x_1) - L(T^n_{\omega_2})'(x_1)| \leq E|F^n\omega_1 - F^n\omega_2|.$$

*(Proof. We use Corollary 2.5 $n$ times and get for $0 \leq k \leq n$ that $|T^k_{\omega_1} x_1 - T^k_{\omega_2} x_1| \leq D|F^k\omega_1 - F^k\omega_2|$.

In particular, for $k = n$, $|T^n_{\omega_1} x_1| \geq \frac{1}{2}$, whence $|T^n_{\omega_2} x_1| \geq \frac{1}{2} - D\varepsilon_0$. Consequently, $T^n(\omega_2, x_1) \in I_i(F^n\omega_2)$ for some $i \in \{0, 1\}$, by definition of $\varepsilon_0$. Applying $T^{-k}$, we get $T^{-k}(\omega_2, x_1) \in I_{i+k}(F^{-k}\omega_2)$.

For $x \leq \frac{1}{2}$ and $\omega \in S^1$, write $G(\omega, x) = L(T'^{\omega}(x)) = \ln(1 + (\alpha(\omega) + 1)(2x)^{\alpha(\omega)})$.

Then

$$\frac{\partial G}{\partial x}(\omega, x) = \frac{\alpha(\omega) + 1)(\alpha(\omega) + 1)(2x)^{\alpha(\omega)} - 1}{1 + (\alpha(\omega) + 1)(2x)^{\alpha(\omega)}} \leq Cx^\alpha(\omega)$$

and

$$\left|\frac{\partial G}{\partial \omega}(\omega, x)\right| = \frac{|\alpha(\omega) + 1)(\alpha(\omega) + 1)(\alpha(\omega) + 1)(2x)^{\alpha(\omega)}|}{1 + (\alpha(\omega) + 1)(2x)^{\alpha(\omega)}} \leq C.$$

Note that $T^k(\omega_1, x_1) \in I_{n-k}(F^k\omega_1)$ and $T^k(\omega_2, x_1) \in I_{n-k+i}(F^k\omega_2)$ with $i \leq 1$. Hence, Proposition 2.2 shows that the second coordinates of $T^k(\omega_1, x_1)$ and $T^k(\omega_2, x_1)$ are at least $1/C(n-k+1)^{1/\alpha_{\min}}$. On the set of points $(\omega, x)$ with $x \geq 1/C(n-k+1)^{1/\alpha_{\min}}$, the estimates on the partial derivatives of $G$ show that this function is $C(n-k+1)^{1/\alpha_{\min}-1}$ Lipschitz. Therefore,

$$|G(T^k(\omega_1, x_1)) - G(T^k(\omega_2, x_1))| \leq C(n-k+1)^{1/\alpha_{\min}-1}d(T^k(\omega_1, x_1), T^k(\omega_2, x_1))$$

$$\leq C(n-k+1)^{1/\alpha_{\min}-1}(1 + D)|F^k\omega_1 - F^k\omega_2|$$

$$\leq C(n-k+1)^{1/\alpha_{\min}-1}(1 + D)|F^k\omega_1 - F^k\omega_2|.$$
Finally, 
\[
|\ln(T_{\omega_1}^n)'(x_1) - \ln(T_{\omega_2}^n)'(x_1)| \leq \sum_{k=0}^{n-1} |G(T^k(\omega_1), x_1)) - G(T^k(\omega_2), x_1))|
\]
\[
\leq C4^n|\omega_1 - \omega_2| \sum_{k=0}^{n-1} (n - k + 1)^{1/\alpha_{\min} - 1}4^{-k - n}
\]
\[
\leq C|f^n\omega_1 - f^n\omega_2| \sum_{l=1}^{\infty} (l + 1)^{1/\alpha_{\min} - 1}4^{-l}.
\]

The last sum is finite, which concludes the proof. 

For \( n \geq 2 \), write \( J_+^n(\omega) = [Y_{n+2}(\omega), Y_n(\omega)] \). Thus, if \( n \geq 1 \), \( J_+^{n+1}(\omega) \) is the preimage of \( I_+^n(F\omega) \), defined by \( I_+^n(F\omega) = [X_{n+2}(\omega), X_n(F\omega)] \). These intervals will appear naturally in distortion controls, since we have seen in the proof of Lemma 2.7 that, if we move away horizontally from a point in \( J_n(\omega_1) \), we find a point in \( J_{n+1}(\omega_2) \) for \( i \in [0, 1] \), i.e. in \( J_+^n(\omega_2) \).

**Lemma 2.8.** There exists a constant \( C \) such that for all \( n > 0 \), for all \( \omega \in S^1 \), for all \( x, y \in J_+^n(\omega) \),
\[
|\ln(T_{\omega}^n)'(x) - \ln(T_{\omega}^n)'(y)| \leq C|T_{\omega}^n(x) - T_{\omega}^n(y)|.
\]
**Proof.** Recall that the Schwarzian derivative of an increasing diffeomorphism \( g \) of class \( C^3 \) is
\[
Sg(x) = \frac{g''''(x)}{g'''(x)} - \frac{3}{2} \left( \frac{g''(x)}{g'(x)} \right)^2.
\]
The composition of two functions with non-positive Schwarzian derivative still has a non-positive Schwarzian derivative.

For \( \tau > 0 \), the Koebe principle [dMV93, Theorem IV.1.2] states that, if \( Sg \leq 0 \) and \( J \subset J' \) are two intervals such that \( g(J') \) contains a \( \tau \)-scaled neighbourhood of \( g(J) \) (i.e. the intervals on the left and on the right of \( g(J) \) in \( g(J') \) have length at least \( \tau |g(J)| \)), then there exists a constant \( K(\tau) \) such that
\[
|\ln g'(x) - \ln g'(y)| \leq K(\tau)|x - y|/|J|, \quad \forall x, y \in J.
\]
This implies that the distortion of \( g \) is bounded on \( J \). Hence it is possible to replace the bound on the right-hand side with \( K'(\tau)(|g(x) - g(y)|/|g(J)|) \).

In our case, if \( 0 < \alpha < 1 \), the left branch of \( T_\alpha \) has non-positive Schwarzian derivative, since \( T_\alpha'' < 0 \) and \( T_\alpha'' > 0 \). In particular, let \( g \) be the composition of the (analytic extensions to \((0, +\infty)\)) of the left branches of \( T_\alpha(F\omega_1), \ldots, T_\alpha(F\omega_b) \), and of the right branch of \( T_\alpha(\omega_0) \). Then, on \( J_+^n \), we have \( T_{\omega}^n = g \), and \( g : [1/2, +\infty) \to (0, +\infty) \) has non-positive Schwarzian derivative.

We want to see that \( |\ln(T_{\omega}^n)'(x) - \ln(T_{\omega}^n)'(y)| \leq C|T_{\omega}^n(x) - T_{\omega}^n(y)| \). For this, we apply the Koebe principle to \( J = J_+^n \) and \( J' = [1/2 + \delta, 2] \) for \( \delta \) very small. Then \( g(J) = [X_2, 1] \) while \( g(J') \) contains \([\delta', 2]\), where \( \delta' > 0 \) is arbitrarily small if \( \delta \) is small enough. As the \( X_2 \) are uniformly bounded away from 0, there exists \( \tau > 0 \) (independent of \( \omega \) and \( n \)) such
that $g(J')$ contains a $\tau$-scaled neighbourhood of $g(J)$. The Koebe principle then gives the desired result.

**Proposition 2.9.** There exists a constant $C$ such that, for every $A_s,n$, for every $(\omega_1, x_1)$ and $(\omega_2, x_2) \in A_s,n$,

$$\frac{|\det DT^n(\omega_1, x_1)|}{|\det DT^n(\omega_2, x_2)|} - 1 \leq Cd(T^n(\omega_1, x_1), T^n(\omega_2, x_2)).$$

**(Proof.** The matrix $DT^n(\omega, x)$ is upper triangular, with $4^n$ in the upper left corner. Thus, we have to show that

$$|\ln(T^n_{\omega_1})'(x_1) - \ln(T^n_{\omega_2})'(x_2)| \leq Cd(T^n(\omega_1, x_1), T^n(\omega_2, x_2)).$$

Assume, for example, that $x_2 \geq x_1$, which implies that $T^n_{\omega_2}(x_1) \leq J^1 \omega_2$ for $k = 0, \ldots, n - 1$. Lemma 2.7 can be applied to $x_1, \omega_1$ and $\omega_2$. Moreover, (9) implies that $x_1 \in J^n_{\omega_2}(\omega_2)$.

Write

$$|\ln(T^n_{\omega_2})'(x_2) - \ln(T^n_{\omega_1})'(x_1)| \leq |\ln(T^n_{\omega_2})'(x_2) - \ln(T^n_{\omega_2})'(x_1)|$$

$$+ |\ln(T^n_{\omega_1})'(x_1) - \ln(T^n_{\omega_1})'(x_1)|$$

$$\leq Cd(T^n(\omega_2, x_2), T^n(\omega_2, x_1)) + E|F^n\omega_2 - F^n\omega_1|$$

by Lemmas 2.8 and 2.7. For the first term,

$$d(T^n(\omega_2, x_2), T^n(\omega_2, x_1)) \leq d(T^n(\omega_2, x_2), T^n(\omega_1, x_1)) + d(T^n(\omega_1, x_1), T^n(\omega_2, x_1))$$

$$\leq d(T^n(\omega_2, x_2), T^n(\omega_1, x_1)) + (D + 1)|F^n\omega_1 - F^n\omega_2|$$

using admissible curves.

As $|F^n\omega_1 - F^n\omega_2| \leq d(T^n(\omega_1, x_1), T^n(\omega_2, x_2))$, we get the conclusion.

2.5. **Construction of the invariant measure.** The previous estimates and Theorem 2.1 easily give that $T_Y$ admits an invariant measure, with Lipschitz density. Inducing gives an invariant measure for $T$, whose density is Lipschitz on each set $S^1 \times (X_{n+1}, X_n)$. However, this does not exclude discontinuities on $S^1 \times X_n$, which is not surprising since $T$ itself has a discontinuity on $S^1 \times (\frac{1}{2}, 1]$, and $T^n$ is discontinuous on $S^1 \times X_n$.

However, in the one-dimensional case, Liverani et al. [LSV99] have proved that the density is really continuous everywhere, since they constructed it as an element of a cone of continuous functions. This fact remains true here, as shown in the following.

**Theorem 2.10.** The map $T$ admits a unique absolutely continuous invariant probability measure $dm$. Moreover, this measure is ergodic. Finally, the density $h = dm/d\text{Leb}$ is Lipschitz on every compact subset of $S^1 \times (0, 1]$.

**(Proof.** Consider the map $T_Y$ induced by $T$ on $Y = S^1 \times (\frac{1}{2}, 1]$. It is Markov for the partition $\alpha = \{A_s,n\}$, and transitive for this partition since $T^n_Y(a) = Y$ for all $a \in \alpha$. Moreover, it is expanding for $d'$ on each set of the partition (Proposition 2.6) and its distortion is Lipschitz (Proposition 2.9, and $d$ equivalent to $d'$).

Theorem 2.1 shows that $T_Y$ admits a unique absolutely continuous invariant probability measure $dm_Y = h \text{dLeb}$, which is ergodic. Moreover, the density $h$ is Lipschitz
(for the distance $d'$, whence for the usual one) on each element of the partition $\alpha_n$ generated by the sets $T_Y(a)$, i.e. on the sets $K_i$.

To construct an invariant measure for the initial map $T$, we use the classical induction process [Aar97, §1.1.5]: let $\varphi_Y$ be the return time to $Y$ under $T$, then $\mu = \sum_{n=0}^{\infty} T^n_\ast (m_Y|\varphi_Y > n)$ is invariant. To check that the new measure has finite mass, we have to see that

$$\sum m_Y(\varphi_Y > n) < \infty.$$ 

As $d m_Y$ and $d \operatorname{Leb}$ are equivalent, we check it for $d \operatorname{Leb}$. We have

$$\operatorname{Leb}(\varphi_Y > n) = \operatorname{Leb}\left( S^1 \times \left[ \frac{1}{2^n}, Y_{n+1} \right] \right) = \frac{1}{2} \operatorname{Leb}(S^1 \times [0, X_n]) \leq \frac{1}{2} \frac{C}{q^1 \alpha_{\max}}.$$ 

using Proposition 2.2. As $\alpha_{\max} < 1$, this is summable.

We know that $h$ is Lipschitz on the sets $[s/4^n, (s + 1)/4^n] \times [\frac{1}{4}, 1]$, we have to prove the continuity on $[s/4^n] \times [\frac{1}{4}, 1]$, which is not hard: these numbers $s/4^n$ are artificial, since they depend on the arbitrary choice of a Markov partition on $S^1$. We can do the same construction using sets other than the $A_{s,n}$. For example, set $A'_{s,n} = [\frac{1}{2} + s / 4^{n+1}, \frac{1}{2} + (s + 1)/4^{n+1}] \times J_n$, and $K'_i = [\frac{1}{4} + i/4^n, \frac{1}{4} + (i + 1)/4^n]$. Since $\frac{1}{4}$ is a fixed point of $F$, the map $T_F$ is Markov for the partition $\{A'_{s,n}\}$, and each of these sets is mapped on a set $K'_i$. Thus, the same arguments as above apply, and prove that $h$ is Lipschitz on each set $K'_i$. Since the boundaries of the sets $K_i$ and $K'_i$ are different, this shows that $h$ is, in fact, Lipschitz on $S^1 \times [\frac{1}{4}, 1]$. We show now that $h$ is Lipschitz on $S^1 \times [X_2, 1]$. Note that it is slightly incorrect to say that $h$ is Lipschitz, since $h$ is defined only almost everywhere. Nevertheless, if we prove that $|h(x) - h(y)| \leq C d(x, y)$ for almost all $x$ and $y$, then there will exist a unique version of $h$ that really is Lipschitz. Thus, all of the equalities we will write until the end of this proof will be true only almost everywhere.

Let $A^+_{i,n} = [s / 4^{n+1}, (s + 1) / 4^{n+1}] \times J_{n}^{+}$. $T^n$ is a diffeomorphism between $A^+_{i,n}$ and $K^+_i = [i/4^n, (i + 1)/4^n] \times [X_2, 1]$. We fix some $K^+_i = K^+_i = I \times [X_2, 1]$, and we show that $h$ is Lipschitz on $K^+$. Let us denote by $A^+_{s_1, n_1}, A^+_{s_2, n_2}, \ldots$ the sets $A^+_{s,n}$ whose image under $T^n$ is $K^+$, and by $U_j : K^+ \rightarrow A^+_{s_j, n_j}$ the inverse of the restriction of $T^n$ to $A^+_{s_j, n_j}$. Let $T_Y$ be the map induced by $T$ on $Y = S^1 \times [\frac{1}{4}, 1]$. Then $h$ $d \operatorname{Leb}_Y$ is invariant under $T_Y$. This implies that, for each $x \in I \times [\frac{1}{4}, 1]$,

$$h(x) = \sum J U_j(x) h(U_j x)$$

where $J U_j$ is the Jacobian of $U_j$.

Let $Z = S^1 \times [X_2, 1]$, and $T_Z$ be the map induced by $T$ on $Z$. Since $h$ $d \operatorname{Leb}_Z$ is also invariant under $T_Z$, we have the same kind of equation as above. For $x \in I \times [X_2, \frac{1}{2}]$, all of its preimages under $T_Z$ are in $S^1 \times [\frac{1}{2}, 1]$, and the invariance gives that

$$h(x) = \sum J U_j(x) h(U_j x).$$

We have shown that, for every $x \in S^1 \times [X_2, 1]$,

$$h(x) = \sum J U_j(x) h(U_j x).$$

This means that $h$ is invariant under some kind of transfer operator, even though it is not a genuine transfer operator since the images of the maps $U_j$ are not disjoint, and since they
do not cover the space. In particular, the images of the \( U_j \) are included in \( S^1 \times [\frac{1}{4}, 1] \), and we already know that \( h \) is Lipschitz on this set.

The bounds of the previous sections still apply to the distortion of the \( U_j \) and their expansion. In particular, \( |1 - JU_j(y)/JU_j(x)| \leq C \)d(x, y) for a constant \( C \) independent of \( j \), and \( |h(U_j(x)) - h(U_j(y))| \leq C \)d(U_j,x, U_j,y) \( \leq C'd(x, y) \) (since \( h \) is Lipschitz on the image of \( U_j \)). Thus,

\[
|h(x) - h(y)| \leq \sum |JU_j(x)h(U_j(x)) - JU_j(y)h(U_j(y))| \\
\leq \sum |JU_j(x)| \left| 1 - \frac{JU_j(y)}{JU_j(x)} \right| |h(U_j(x)| + \sum |JU_j(y)||h(U_j(x)) - h(U_j(y))| \\
\leq C \)d(x, y) \sum |JU_j(x)| + C'd(x, y) \sum |JU_j(y)|.
\]

It remains to prove that \( \sum |JU_j(x)| \) is bounded. The bound on distortion gives \( JU_j(x) \approx \text{Leb}(\text{Im} \ U_j) \), whence \( \sum JU_j(x) \leq C \) \sum \text{Leb}(\text{Im} \ U_j) \), which is finite since every point of \( S^1 \times [\frac{1}{4}, 1] \) is in the image of at most two maps \( U_j \).

We have proved that \( h \) is Lipschitz on \( S^1 \times [X_2, 1] \), except maybe on \([s/4^q] \times [X_2, 1] \). As above, using another Markov partition, we exclude the possibility of discontinuities there. Thus, \( h \) is Lipschitz on \( S^1 \times [X_2, 1] \).

To prove that \( h \) is Lipschitz on \( S^1 \times [X_k, 1] \), we do exactly the same thing, except that we consider \( [Y_{n+k}, Y_n] \) instead of \( J_n^+ = [Y_{n+2}, Y_n] \). As above, writing \( U_1, U_2, \ldots \) for the inverse branches of \( T^n \) defined on a set \([s/4^q, (s+1)/4^q] \times [X_{n+k}, Y_n] \) and whose image is \( K' = [i/4^q, (i+1)/4^q] \times [X_k, 1] = I \times [X_k, 1] \), we show that \( h(x) = \sum JU_j(x)h(U_j(x)) \) for \( x \in K' \). In fact, for \( x \in I \times [X_i, X_{i-1}] \), we use the invariance of \( h \) \text{Leb} under the map induced by \( T \) on \( S^1 \times [X_i, 1] \) to prove this equality. We conclude finally as above, using the fact that \( h \) is Lipschitz on \( S^1 \times [\frac{1}{4}, 1] \), which contains the images of the \( U_j \).

This concludes the proof, since every compact subset of \( S^1 \times (0, 1] \) is contained in \( S^1 \times [X_k, 1] \) for large enough \( k \).

\[ \square \]

3. Limit theorems for Markov maps

We want to establish limit theorems for Birkhoff sums. In this direction, we give in this section an abstract result, valid for a map that induces a Gibbs–Markov map on a subset of the space (which is the case of our skew product). Related limit theorems have been proved in [Gou04], but we will show here a slightly different result, which requires more control on the return time \( \varphi \) but is more elementary, using Theorem A.1 proved in Appendix A and inspired by results of Melbourne and Török [MT04] for flows.

If \( Z_0, \ldots, Z_{n-1}, \ldots \) are independent identically distributed random variables with zero mean, the sums \( B_n = \sum_{k=0}^{n-1} Z_k \) (where \( B_n \) is a real sequence) converge to a non-trivial limit distribution in the following cases: if \( Z_k \in L^2 \), there is convergence to a normal law for \( B_n = \sqrt{n} \). There is also convergence to a normal law, but with a different normalization, if \( P(|Z_k| > x) = x^{-2}L(x) \) with \( L(x) := 2 \int_1^x (u/a)^{-2}du \) unbounded and slowly varying (i.e. \( L : (0, \infty) \to (0, \infty) \) satisfies \( \text{lim}_{x \to \infty} L(ax)/L(x) = 1 \) for all \( a > 0 \)); this is, in particular, true when \( l \) itself is slowly varying. Finally, if \( P(Z_k > x) = (c_1 + o(1))x^{-\beta}L(x) \) and \( P(Z_k < -x) = (c_2 + o(1))x^{-\beta}L(x) \), where \( L \) is
slowly varying and \( p \in (0, 2) \), we have convergence (for a good choice of \( B_n \)) to a limit law called stable law. It is a remarkable fact that, in this probabilistic setting, these sufficient conditions for convergence are also necessary (see, e.g., [Fel66, Theorem XVII.5.1a]).

In the dynamical setting, we will prove the same kind of limit theorems, still with three possible cases: \( L^2 \), normal non-standard and stable. The normalizations will, moreover, be the same as in the probabilistic setting. However, we will only give sufficient conditions for convergence, the converse seems definitely out of reach.

**Theorem 3.1.** Let \( T : X \to X \) be an ergodic transformation preserving a probability measure \( m \). Assume that there exists a subset \( Y \) of \( X \) with \( m(Y) > 0 \) and a countable partition \( \alpha \) of \( Y \), such that the first return map \( T_\alpha(x) = T^{\varphi(x)}(x) \) (where \( \varphi(x) = \inf\{n > 0 \mid T^n(x) \in Y\} \)) is Gibbs–Markov for the measure \( m|Y \) and the partition \( \alpha \). Assume, moreover, that \( \varphi \) is constant on each element of \( \alpha \).

Let \( f : X \to \mathbb{R} \) be an integrable map with \( \int f = 0 \), such that \( f_\varphi(y) := \sum_{n=0}^{\varphi(y)-1} f(T^n y) \) satisfies
\[
\sum_{a \in \alpha} m(a) Df_\varphi(a) < \infty
\] (18)
where
\[
Df_\varphi(a) = \inf\{C > 0 \mid \forall x, y \in a, |f_\varphi(x) - f_\varphi(y)| \leq C d(x, y)\}.
\] (19)

Then we have the following.

- Assume that \( f_\varphi \in L^2 \). Assume, moreover, that \( \varphi \) satisfies one of the following hypotheses:
  - \( \varphi \in L^2 \);
  - \( m(\varphi > x) = x^{-p} L(x) \) where \( L \) is slowly varying and \( p \in (1, 2] \).

Then there exists \( \sigma^2 \geq 0 \) such that \( (1/\sqrt{n}) S_n f \to \mathcal{N}(0, \sigma^2) \).

- Assume that \( m(|f_\varphi| > x) = x^{-2} L(x) \), with \( L(x) := 2 \int_1^\infty (l(u)/u) du \) unbounded and slowly varying. Assume, moreover, that \( m(\varphi > x) = (c + o(1)) x^{-\beta 2} L(x) \) with \( c > 0 \).

Let \( B_n \to \infty \) satisfy \( n L(B_n) = B_n^2 \). Then \( B_n^{-1} S_n f \to \mathcal{N}(0, 1) \).

- Assume that \( m(|f_\varphi| > x) = (c_1 + o(1)) x^{-p} L(x) \) and \( m(f_\varphi < -x) = (c_2 + o(1)) x^{-p} L(x) \) where \( L \) is a slowly varying function, \( p \in (1, 2) \), and \( c_1, c_2 \geq 0 \) with \( c_1 + c_2 > 0 \). Assume also that \( m(\varphi > x) = (c_3 + o(1)) x^{-p} L(x) \) with \( c_3 > 0 \).

Let \( B_n \to \infty \) satisfy \( n L(B_n) = B_n^p \). Then \( B_n^{-1} S_n f \to Z \) where the random variable \( Z \) has a characteristic function given by
\[
E(e^{itZ}) = e^{-c|t|^{p(1-\beta) \text{sgn}(t) \tan(p\pi/2)}}
\] (20)
with \( c = (c_1 + c_2) \Gamma(1 - p) \cos(p\pi/2) \) and \( \beta = (c_1 - c_2)/(c_1 + c_2) \).

In the second case of the theorem, when \( l \) itself is slowly varying, then \( L \) is automatically slowly varying.

**Proof.** The idea is to use Theorem A.1: we have to check all of its hypotheses. We will use the notation of this theorem and, in particular, write \( E_T(u) = \int u \, dm/m(Y) \).

We first treat the third case (stable law), using the results of [AD01b] (and the generalizations of [Gou04]). Let \( s(x, y) \) be the separation time of \( x \) and \( y \) defined in (4), \( \tau = 1/\lambda \) and \( d_\tau = \tau^\gamma \) the corresponding metric. Since every iteration of \( T_\varphi \) expands
by at least \( \lambda \), we get \( d(x,y) \leq Cd(x,y) \). In particular, we can assume without loss of generality that \( d = d_r \), which is the setting of [AD01b] and [Gou04].

Let \( P \) be the transfer operator associated to \( T_Y \) (it is defined by \( \int u \circ T_Y = \int P(u) \circ v \)), and let \( P_t(u) = P(e^{itf_Y}u) \). Let \( \mathcal{L} \) be the space of bounded Lipschitz functions (i.e., such that there exists \( C \) such that \( |g(x) - g(y)| \leq Cd(x,y) \) for all \( \alpha \) and for all \( x, y \in \alpha \)). Since \( \sum m(\alpha)Df_Y(\alpha) < \infty \), [Gou04, Theorem 3.8] (which is a strengthening of Theorem 5.1 in [AD01b]) ensures that, for small enough \( t, \) \( P_t \) acting on \( \mathcal{L} \) has an eigenvalue \( \lambda(t) = e^{-c/(m(Y))|t|^p(1 - i\beta \text{sgn}(t)\tan(\pi/2))}L(|t|^{p-1})(1 + o(1)) \) close to 1, and the remaining part of the spectrum of \( P_t \) is uniformly bounded away from 1.

We will use this information to estimate \( E_Y(\phi e^{it/B_n}S_{nm(Y)}f_Y) \). Since \( \phi \) is Lipschitz and integrable, \( P\phi \in \mathcal{L} \) by [AD01b, Proposition 1.4]. Let \( k(n) = [nm(Y)] - 1. \) Then

\[
E_Y(\phi e^{it/B_n}S_{nm(Y)}f_Y) = E_Y(P\phi \cdot e^{it/B_n}S_{nm(Y)}f_Y) = E_Y(P^{k(n)}\phi) = E_Y(\phi)(\lambda(t/B_n)^{k(n)} + o(1)).
\]

The slow variation of \( L \) implies that, for all \( t \neq 0,

\[
k(n)\frac{1}{m(Y)}\left|\frac{t}{B_n}\right|^pL(B_n/|t|) \sim |t|^p\frac{n}{B_n}L(B_n) \rightarrow |t|^p.
\]

Hence, we get

\[
\lambda\left(\frac{t}{B_n}\right)^{k(n)} \rightarrow e^{-c/\left|t\right|^p(1 - i\beta \text{sgn}(t)\tan(\pi/2))},
\]

This shows that

\[
E_Y(\phi e^{it/B_n}S_{nm(Y)}f_Y) \rightarrow E_Y(Z)E(e^{itZ}),
\]

where the random variable \( Z \) is as in the statement of the theorem. Hence,

\[
E_Y(\phi e^{it/B_n}S_{nm(Y)}f_Y) \rightarrow E_Y(Z)E(e^{itZ}).
\]

Moreover, the difference between this term and \( E_Y(\phi e^{it/B_n}S_{nm(Y)}f_Y - 1) \), which tends to 0 by dominated convergence. Thus,

\[
E_Y(\phi e^{it/B_n}S_{nm(Y)}f_Y) \rightarrow E_Y(\phi)E(e^{itZ}).
\]

This is (52). Moreover, since \( L \) is slowly varying, the equation \( nL(B_n) = B_n^p \) implies that \( \inf_{r \geq n}(B_r/B_n) > 0 \) (using for example the Potter bounds [BGT87, Theorem 1.5.6]).

Hypothesis 2 of Theorem A.1 is satisfied for \( b = 1 \), according to Birkhoff’s theorem applied to \( \phi - E_Y(\phi) \) (and because \( T_Y \) is ergodic, which is a consequence of the ergodicity of \( T \)). Finally, the hypothesis on the distribution of \( \phi \) ensures, by [AD01b, Theorem 6.1], that \( (S_{nm(Y)}^Y\phi - nm(Y)E_Y(\phi))/B_n \) converges in distribution. Thus, (53)’ is satisfied.

We can use Theorem A.1, and get that \( S_n f/B_n \rightarrow Z \). This concludes the proof of the third case of Theorem 3.1.

The proof of the second case of Theorem 3.1 is exactly the same, using Theorem 3.1 in [AD01a] instead of Theorem 5.1 in [AD01b] to show the convergence in distribution of \( S_{nm(Y)}^Yf_Y/B_n \) and \( (S_{nm(Y)}^Y\phi - nm(Y)E_Y(\phi))/B_n \).

In the first case \( f_Y \in L^2 \), the proof is again identical when \( \phi \in L^2 \), with \( B_n = \sqrt{n} \); indeed, in [GH88] it was proved that the Birkhoff sums of \( f_Y \) and \( \phi \) satisfy a classical
central limit theorem. However, when \( m(\varphi > x) = x^{-p} L(x) \), we have to check in a different way the hypotheses 2 and 3 of Theorem A.1. Theorem 6.1 of [AD01b] ensures that, if \( B'_n \) is given by

\[
n L(B'_n) = (B'_n)^p, \tag{25}
\]

then \( (S_n Y - nExY(\varphi))/B'_n \) converges in distribution. Assume for the moment that, in the natural extension of \( T_Y \), for all \( b > \frac{1}{2} \), for almost all \( x \in Y \),

\[
\frac{1}{|N|^b} \sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0 \tag{26}
\]

when \( N \to \pm \infty \). Then hypothesis 2 of Theorem A.1 is satisfied for any \( b > \frac{1}{2} \). Let \( \kappa > 0 \) be very small. As \( L \) is slowly varying, \( L(B'_n) = O((B'_n)^p) \), whence Equation (25) gives \( B'_n = O(n^{1/(p-\kappa)}) \). Thus, if \( b < p/2 \), we have \( B'_n = O(B_n^{1/b}) \), which implies (53’).

Hence, to conclude the proof, we just have to check (26). In [Gou04, Lemma 3.4] it was proved that \( Pf_Y \in \mathcal{L} \), and has a vanishing integral. If \( T_Y \) is mixing, then \( 1 \) is the only eigenvalue of \( P \) of modulus 1, and \( P \) has a spectral gap, whence \( P^m f_Y \to 0 \) exponentially fast. In particular, \( \int f_Y \cdot T^n_Y f_Y = \int (P^m f_Y) \cdot f_Y = O((1-\delta)^n) \) for some \( 0 < \delta < 1 \). Thus, as \( f_Y \in L^2 \), [Kac96, Theorem 16] gives that, for every \( b > \frac{1}{2} \), \((1/N)^b \sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0 \) almost everywhere when \( N \to \infty \). In the natural extension, \( \int f_Y \cdot T_Y^n \cdot f_Y = \int f_Y \cdot f_Y \cdot T_Y^n \) decays also exponentially fast, whence the same argument gives that \( (1/N)^b \sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0 \) when \( N \to -\infty \). Hence, (26) is satisfied if \( T_Y \) is mixing. In the general case, there exists a decomposition \( Y = Y_1 \cup \cdots \cup Y_d \) such that \( T_Y \) maps \( Y_i \) to \( Y_{i+1} \) for \( 1 \leq i \leq d-1 \), and \( Y_d \) maps \( Y_1 \), and such that \( T^n_Y \) is mixing on each \( Y_i \) (with \( m(Y_i) = 1/d \) for \( 1 \leq i \leq d \)) (see [Aar97]).

In particular, set \( g = d \sum_{i=1}^{d} (\int f_Y f_Y) 1_{Y_i} \) and \( h = f_Y - g \); this function satisfies \( P^n h = O((1-\delta)^d n) \), using the same argument as above on each \( Y_i \), since \( f_Y h = 0 \) for all \( i \). Hence, \( P^n h = O((1-\delta)^n) \). In particular, \((1/N)^b \sum_{k=0}^{N-1} h(T^k_Y x) \to 0 \) when \( N \to \pm \infty \), as above. Moreover, \( \sum_{i=1}^{d} (1 f_Y f_Y) = \int f_Y = 0 \), whence \( (1/N)^b \sum_{k=0}^{N-1} g(T^k_Y x) \) is uniformly bounded. This implies that \((1/N)^b \sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0 \) almost everywhere when \( N \to \pm \infty \).

4. Asymptotic behaviour of \( X_n \)

We return to the study of the skew product (1). To prove limit theorems using Theorem 3.1, we will need to estimate \( m(\varphi_Y > n) \), which is directly related to the speed of convergence of \( X_n \) to 0. This section will be devoted to the proof of the following theorem.

**Theorem 4.1.** When \( n \to +\infty \),

\[
\left( \frac{n}{\sqrt{\ln n}} \right)^{1/\sigma_{\min}} X_n \to \frac{1}{(2^{\sigma_{\min}} 3^{\sigma_{\max}} \sqrt{\pi}/2 \sigma_{\min} \alpha_{\min})^{1/\sigma_{\min}}} \tag{27}
\]

almost everywhere and in \( L^1 \).

The proof of this theorem is a quite involved computation, which relies on the following lemma.
Lemma 4.2. We have
\[ E(e^{-(x_{\min}^\alpha)})w \sim \sqrt{\frac{\pi}{2e^{\alpha/2}}} \frac{1}{\sqrt{w}} \quad \text{when } w \to \infty. \] (28)

Proof. Write \( \beta = \alpha - \alpha_{\min} \), and \( f(b) = \text{Leb}\{\beta(\omega) \in [0, b]\} \). In a neighbourhood of \( \omega_0 \) (the unique point where \( \alpha \) takes its minimal value \( \alpha_{\min} \)), \( \alpha \) behaves like the parabola \( \alpha_{\min} + (\alpha'(\omega_0)/2)(\alpha - \omega_0)^2 \), whence \( f(b) \sim \sqrt{2/\alpha^2(\omega_0)} \sqrt{b} \) when \( b \to 0 \).

Writing \( P_\beta \) for the distribution of \( \beta \), an integration by parts gives
\[ E(e^{-(x_{\min}^\alpha)w}) = \int_0^\infty e^{-bw} dP_\beta(b) = w \int_0^\infty e^{-bw} f(b) dB = \int_0^\infty e^{-u f(u/w)} du = \frac{1}{\sqrt{w}} \int_0^\infty e^{-u \sqrt{w} f(u/w)} du. \]

However \( e^{-u \sqrt{w} f(u/w)} \to e^{-u \sqrt{2/\alpha^2(\omega_0)} \sqrt{u}} \) when \( u \to \infty \). There exists a constant \( E \) such that \( f(u) \leq E \sqrt{u} \) (this is clear in a neighbourhood of 0, and elsewhere since \( f \) is bounded), whence \( e^{-u \sqrt{w} f(u/w)} \leq E e^{-u \sqrt{u}} \) integrable. By dominated convergence,
\[ \int_0^\infty e^{-u \sqrt{w} f(u/w)} du \to \frac{2}{w^{\alpha/2}} \int_0^\infty e^{-u \sqrt{u}} du = \frac{2}{\alpha^2/2}. \]

Proof of Theorem 4.1. As in the proof of Proposition 2.2, we write
\[ \frac{1}{X_n(F_\omega)^{\alpha_{\min}}} = \frac{1}{X_{n+1}(\omega)^{\alpha_{\min}}} - \alpha_{\min} 2\alpha_{\min} (2X_{n+1}(\omega))^{\alpha(\omega) - \alpha_{\min}} + O(X_{n+1}(\omega)^{2\alpha(\omega) - \alpha_{\min}}). \] (29)

Proposition 2.2 gives
\[ X_{n+1}(\omega)^{2\alpha(\omega) - \alpha_{\min}} \leq X_{n+1}(\omega)^{\alpha_{\min}} \leq \frac{C}{(n + 1)^{\alpha_{\min}/\alpha_{\max}}} \leq \frac{C}{\sqrt{n + 1}} \] (30)

as \( \alpha_{\min}/\alpha_{\max} \geq \frac{1}{2} \) by hypothesis. Thus,
\[ \frac{1}{X_{n+1}(\omega)^{\alpha_{\min}}} - \frac{1}{X_n(F_\omega)^{\alpha_{\min}}} = 2\alpha_{\min} \alpha_{\min} (2X_{n+1}(\omega))^{\alpha(\omega) - \alpha_{\min}} + O(1/\sqrt{n}). \]

Summing from 1 to \( n \), we get a constant \( P \) (independent of \( \omega \)) such that
\[ \frac{1}{X_n(\omega)^{\alpha_{\min}}} \geq 2\alpha_{\min} \alpha_{\min} \left[ \sum_{k=1}^n (2X_k(F^{n-k}_\omega))^{\alpha(F^{n-k}_\omega) - \alpha_{\min}} - P\sqrt{n} \right] \] (31)
\[ \frac{1}{X_n(\omega)^{\alpha_{\min}}} \leq 2\alpha_{\min} \alpha_{\min} \left[ \sum_{k=1}^n (2X_k(F^{n-k}_\omega))^{\alpha(F^{n-k}_\omega) - \alpha_{\min}} + P\sqrt{n} \right]. \] (32)

Equation (31) and Proposition 2.2 imply that
\[ \frac{\sqrt{\ln n}}{n} \frac{1}{2^{\alpha_{\min}} \alpha_{\min} X_n(\omega)^{\alpha_{\min}}} \geq \frac{\sqrt{\ln n}}{n} \sum_{k=1}^{C^{-1}} \left( \frac{2^{1/k}}{\alpha_{\min}} \right)^{\alpha(F^{n-k}_{\omega}) - \alpha_{\min}} - P\sqrt{n} \frac{\ln n}{n} =: A_\beta(\omega). \] (33)
We first study the convergence of $A_n$. The functions $\alpha$ and $\alpha \circ F^{n-k}$ have the same distribution since $F$ preserves Lebesgue measure. Thus, by Lemma 4.2,

$$E \left( \frac{2C^{-1}}{k^{1/\alpha_{\min}}} \alpha_{F^{n-k}} - \alpha_{\min} \right) \sim \sqrt{\frac{\pi}{2\alpha^n(x_0)}} \frac{1}{\sqrt{\ln(k^{1/\alpha_{\min}}) - \ln(2C^{-1})}} \sim \sqrt{\frac{\pi}{2\alpha^n(x_0)}} \frac{1}{\sqrt{\ln k}}$$

Summing, we get that

$$E(A_n) \to C_1 := \sqrt{\frac{\pi}{2\alpha^n(x_0)}}, \quad (34)$$

since $\sum_{k=2}^n (1/\sqrt{\ln k}) \sim n/\sqrt{\ln n}$.

We will need $L^p$ estimates, for $p \geq 1$. To get them, we use a result of Pène [Pène02], recalled in Appendix B. Let us denote by $\|g\|$ the Lipschitz norm of a function $g : S^1 \to \mathbb{R}$, i.e. $\|g\| = \sup_{x \in S^1} |g(x)| + \sup_{x \neq y} |g(x) - g(y)|/|x - y|$.

We define $fk(\omega) = (2C^{-1}/k^{1/\alpha_{\min}}) \alpha'(\omega) - \alpha_{\min}$, and $g_k = f_k - E(f_k)$. Thus, $A_n = (\sqrt{n}/n) \sum_{k=1}^n f_k \circ F^{n-k} - P\sqrt{\ln n}/n$. As $g'_k = \ln(2C^{-1}/k^{1/\alpha_{\min}}) \alpha'(f_k)$, there exists a constant $L$ such that, for $k \leq n$, $\|g_k\| \leq L n$. As a consequence, Theorem B.1 applied to $g_k/(L \ln n)$ gives

$$\|A_n - E(A_n)\|_p = \sqrt{\frac{\ln n}{n}} L n \ln \left[ \sum_{k=1}^n g_k \circ F^{n-k}/(L \ln n) \right]_p \leq \sqrt{\frac{\ln n}{n}} L \ln n K P \sqrt{n},$$

i.e.

$$\|A_n - E(A_n)\|_p \leq L P \sqrt{\frac{\ln^3 n}{n}}. \quad (35)$$

This implies, in particular, that $A_n$ converges almost everywhere to $C_1$. Namely, if $\delta > 0$,

$$\text{Leb}[|A_n - E(A_n)| > \delta] \leq \int |A_n - E(A_n)|^4 \frac{L^4}{\delta^4} \leq \frac{L^4}{\delta^4} \left( \frac{\ln n}{n} \right)^{4/2}$$

which is summable, and $E(A_n) \to C_1$.

We have

$$A_n(\omega) \geq \frac{\sqrt{\ln n}}{n} \sum_{k=1}^n \left( \frac{2C^{-1}}{k^{1/\alpha_{\min}}} \alpha_{\max} - \alpha_{\min} \right) - P\sqrt{n} \geq \frac{\sqrt{\ln n}}{n} [K n^{2-\alpha_{\max}} / \alpha_{\min} - P \sqrt{n}]$$

since $\alpha_{\max} / \alpha_{\min} < \frac{3}{4}$. Thus,

$$\|A_n\|_\infty \leq K' n^{\alpha_{\max} / \alpha_{\min} - 1} \sqrt{\ln n}. \quad (37)$$

Note that $E(A_n)$ tends to $C_1 \neq 0$, whence $1/E(A_n)$ is bounded. Thus,

$$\left\| \frac{1}{A_n} - \frac{1}{E(A_n)} \right\|_p \leq \left\| \frac{1}{A_n} \right\|_\infty \frac{1}{E(A_n)} \|A_n - E(A_n)\|_p \leq K' n^{\alpha_{\max} / \alpha_{\min} - 1} L P \sqrt{\frac{\ln^3 n}{n}} \left[ \frac{\ln n}{n} \right]^{4/2}$$

$$= M \frac{\ln n}{n^x}$$
where \( \kappa = \frac{3}{2} - \frac{\alpha_{\max}}{\alpha_{\min}} > 0 \). In particular, \( 1/A_n \) tends to \( 1/C_1 \) in every \( L^p \). Equation (33) shows that

\[
\left( \frac{n}{\sqrt{\ln n}} \right)^{1/\alpha_{\min}} X_n \leq \frac{1}{(2^{a_{\min}} \alpha_{\min} A_n)^{1/\alpha_{\min}}}.
\]  

(38)

The right-hand side tends to

\[
C_2 := \frac{1}{(2^{a_{\min}} \alpha_{\min}^{3/2} \sqrt{\pi} / 2^{a''(x_0)})^{1/\alpha_{\min}}}
\]  

(39)

in every \( L^p \) and, in particular, in \( L^1 \). Thus,

\[
\lim_{n \to \infty} E \left( \left( \frac{n}{\sqrt{\ln n}} \right)^{1/\alpha_{\min}} X_n \right) \leq C_2.
\]  

(40)

Moreover, \( A_n \) converges almost everywhere to \( C_1 \), whence (38) yields that, almost everywhere,

\[
\lim_{n \to \infty} \left( \frac{n}{\sqrt{\ln n}} \right)^{1/\alpha_{\min}} X_n(\omega) \leq C_2.
\]  

(41)

Set \( Q = \sup_n (1/E(A_n)) + 1 \), we estimate \( \text{Leb}\{1/A_n \geq Q\} \). If \( p \geq 1 \),

\[
\text{Leb}\left\{ \frac{1}{A_n} \geq Q \right\} \leq \text{Leb}\left\{ \left| \frac{1}{A_n} - \frac{1}{E(A_n)} \right| \geq 1 \right\} \leq E \left( \left| \frac{1}{A_n} - \frac{1}{E(A_n)} \right|^p \right) \leq \left( M_p \frac{\ln n}{n^2} \right)^p.
\]  

In particular, choosing \( p \) large enough gives

\[
\text{Leb}\left\{ \frac{1}{A_n} \geq Q \right\} \leq \frac{M}{n^5}.
\]  

(42)

Setting \( Q' = Q / 2^{a_{\min}} \alpha_{\min} \), (38) thus yields that

\[
\text{Leb}\left\{ X_n \geq \left( \frac{Q' \sqrt{\ln n}}{n} \right)^{1/\alpha_{\min}} \right\} \leq \frac{M}{n^5}.
\]  

(43)

Consequently,

\[
U_n := \left\{ \omega \mid \exists \sqrt{n} \leq k \leq n \text{ with } X_k(F^{n-k}\omega) \geq \left( \frac{Q' \sqrt{\ln k}}{k} \right)^{1/\alpha_{\min}} \right\}
\]

has a measure at most \( \sum_n \sigma(M/k^5) \leq M'/n^2 \) (since \( \text{Leb} \) is invariant under \( F^{n-k} \)). Finally, Borel–Cantelli ensures that there is a full measure subset of \( S^1 \) on which \( \omega \not\in U_n \) for large enough \( n \).

Set

\[
A'_n(\omega) = \frac{\sqrt{\ln n}}{n} \left[ \sum_{k=1}^{n} \left( \frac{2(Q' \sqrt{\ln k})^{1/\alpha_{\min}}}{k^{1/\alpha_{\min}}} \right)^{a(F^{n-k}\omega)-\alpha_{\min}} + (P + 1) \sqrt{n} \right].
\]  

(44)

As for \( A_n \), we show that \( A'_n \to C_1 \) in every \( L^p \) and almost everywhere.
Let $\omega$ be such that $\omega \notin U_n$ for large enough $n$, and $A'_n(\omega) \to C_1$ (these properties are true almost everywhere). Then, for large enough $n$, equation (32) and the fact that $X_k(F^{n-k}\omega) \leq (Q\sqrt[n]{\ln k}/k)^{1/\alpha_{\min}}$ for $\sqrt[n]{k} \leq n$, yield that

$$\frac{1}{2\alpha_{\min} \alpha_{\min}^n} \int \left( \sum_{k=1}^{n} \int \left( \sum_{k=\sqrt[n]{n}}^{n} \left( \frac{2(Q\sqrt[n]{\ln k})^{1/\alpha_{\min}}}{k^{1/\alpha_{\min}}} \right)^{n(F^{n-k}\omega)} \right) + P \sqrt[n]{n} \right)$$

$$\leq \frac{n}{\sqrt[n]{\ln n}} A'_n(\omega) \sim \frac{n}{\sqrt[n]{\ln n}} C_1.$$ 

Therefore,

$$\lim \left( \frac{n}{\sqrt[n]{\ln n}} \right)^{1/\alpha_{\min}} X_n(\omega) \geq C_2. \quad (45)$$

Equations (41) and (45) prove that $(n/\sqrt[n]{\ln n})^{1/\alpha_{\min}} X_n$ tends almost everywhere to $C_2$. We get the convergence in $L^1$ from the inequality (40) and the following elementary lemma. 

**Lemma 4.3.** Let $f_n$ be non-negative functions on a probability space, with $f_n \to f$ almost everywhere, and $\lim E(f_n) \leq E(f) < \infty$. Then $f_n \to f$ in $L^1$.

**Proof.** Write $g_n = f_n + f - |f - f_n| \geq 0$. Fatou’s lemma gives $E(\lim g_n) \leq \lim E(g_n)$. Therefore,

$$2E(f) \leq \lim E(f_n) + E(f) - \lim E(|f - f_n|).$$

Consequently, the hypotheses imply that $\lim E(|f - f_n|) \leq 0$. 

**5. Limit theorems**

Set

$$A = \frac{1}{4(\alpha_{\min}^{3/2} \sqrt{\pi/2\alpha_{\min}}(\chi_0))^{1/\alpha_{\min}}} \int_{S^1 \times [-1/2]} h \, d\text{Leb}, \quad (46)$$

where $h$ is the density of $m$ with respect to Leb.

In this section, we prove the following theorem.

**Theorem 5.1.** Let $f$ be a Hölder function on $S^1 \times [0,1]$, with $\int f \, dm = 0$. Write $c = \int_{S^1 \times [0]} f \, d\text{Leb}$. Then:

- if $\alpha_{\min} < \frac{1}{2} < \sigma$ such that $(1/\sqrt[n]{n}) S_n f \to \mathcal{N}(0, \sigma^2)$;

- if $\frac{1}{2} < \sigma < \alpha_{\min} < 1$ and $c = 0$, assume also that there exists $\gamma > (\alpha_{\max}/\alpha_{\min})(\alpha_{\min} - \frac{1}{2})$ such that $|f(\omega, x) - f(\omega, 0)| \leq C x^\gamma$; then there exists $\sigma^2 > 0$ such that $(1/\sqrt[n]{n}) S_n f \to \mathcal{N}(0, \sigma^2)$;

- if $\alpha_{\min} = \frac{1}{2}$ and $c \neq 0$, then $S_n f/\sqrt{(c^2 A)/4 n(\ln n)^2} \to \mathcal{N}(0, 1)$;

- if $\frac{1}{2} < \sigma < \alpha_{\min} < 1$ and $c \neq 0$, then $S_n f/\alpha_{\min} \sqrt{\alpha_{\min} \ln n} \to Z$, where the random variable $Z$ has a characteristic function given by

$$E(e^{itZ}) = e^{-A|c|^{1/\alpha_{\min}} \Gamma(1 - 1/\alpha_{\min}) \cos(\pi/2 \alpha_{\min}) |t|^{1/\alpha_{\min}} (1 - i \operatorname{sgn}(c) \tan(\pi/2 \alpha_{\min}))). \quad (47)$$

The random variable $Z$ in the last case has a so-called stable distribution of exponent $1/\alpha_{\min}$ and parameters $A|c|^{1/\alpha_{\min}} \Gamma(1 - 1/\alpha_{\min}) \cos(\pi/2 \alpha_{\min})$ and $\operatorname{sgn}(c)$ (see, e.g., [Fel66, Ch. XVII] for general background on stable laws).
To prove this theorem, we will use Theorem 3.1. For this, we need a control of \( m(\varphi_Y > n) \) which comes from the asymptotic behaviour of \( X_n \) proved in Theorem 4.1. It will also be necessary to estimate \( m(f_Y > x) \), through the study of the integrability of \( f_Y \) (Lemmas 5.3 and 5.4).

In the rest of this section, \( f \) will be a Hölder function on \( S^1 \times [0, 1] \), fixed once and for all. Recall that \( f_Y(y) = \sum_{k=0}^{\varphi_Y(y)-1} f(T^k y) \), where \( \varphi_Y \) is the first return time to \( Y = S^1 \times (\frac{1}{2}, 1] \).

5.1. Estimates on measures.

**Lemma 5.2.** We have

\[
m(\varphi_Y > n) \sim \left( \frac{\sqrt{\ln n}}{n} \right)^{1/\alpha_{\min}} A \tag{48}
\]

where \( A \) is given by (46).

**Proof.** We have

\[
m(\varphi_Y > n) = \int_{S^1} \int_{Y_n} h(\omega, u) \, du \, d\omega = \int_{S^1} \int_0^{X_n(F\omega)/2} h\left(\omega, \frac{1}{2} + u\right) \, du \, d\omega
\]

\[
= \int_{S^1} X_n(F\omega) \frac{h(\omega, \frac{1}{2})}{2} \, d\omega + \int_{S^1} \int_0^{X_n(F\omega)/2} \left[ h\left(\omega, \frac{1}{2} + u\right) - h\left(\omega, \frac{1}{2}\right) \right] \, du \, d\omega
\]

As

\[
\left( \frac{n}{\sqrt{\ln n}} \right)^{1/\alpha_{\min}} \to 1
\]

in \( L^1 \) and almost everywhere (Theorem 4.1) and \( h(\omega, \frac{1}{2}) \) is bounded, we get that \( I \sim (\sqrt{\ln n}/n)^{1/\alpha_{\min}} A \). Moreover, for large enough \( n \), \( |h(\omega, \frac{1}{2} + u) - h(\omega, \frac{1}{2})| \leq \epsilon \), whence \( II = o(\sqrt{\ln n}/n)^{1/\alpha_{\min}} \).

**Lemma 5.3.** If \( \alpha_{\min} < \frac{1}{2} \), then \( f_Y \in L^2(Y, dm) \).

**Proof.** We have

\[
\int f_Y^2 \, dm \leq C \sum m(\varphi_Y = n)n^2 = C \sum (m(\varphi_Y > n) - m(\varphi_Y > n))n^2
\]

which is summable since \( m(\varphi_Y > n) \sim A(\sqrt{\ln n}/n)^{1/\alpha_{\min}} \) with \( 1/\alpha_{\min} > 2 \).

**Lemma 5.4.** Assume that \( \int_{S^1 \times \{0\}} f = 0 \). Let \( \gamma \in (0, \alpha_{\max}) \) be such that \( |f(\omega, x) - f(\omega, 0)| \leq Cx^\gamma \). If \( 1 < p < \min(2/\alpha_{\min}, 1/\alpha_{\min}(1 - \gamma / \alpha_{\max})) \), then \( f_Y \in L^p(Y, dm) \).

**Proof.** As \( h \) is bounded on \( Y \), it is sufficient to prove that \( f_Y \in L^p(Y, d\text{Leb}) \).

Assume first that \( f \equiv 0 \) on \( S^1 \times \{0\} \). Then, if \( x = (\omega, x) \) satisfies \( \varphi_Y(x) = n \), we have \( f_Y(x) = \sum_0^{\varphi_Y(x)-1} f(T^k x) \). If \( k \geq 1 \), \( T_{\omega}^k(x) \leq X_{n-k}(F_{\omega}^k \omega) \leq C/(n-k)^{1/\alpha_{\max}} \),

\[
\sum_0^{\varphi_Y(x)-1} f(T^k x) \leq C \sum_0^{\varphi_Y(x)-1} (n-k)^{1/\alpha_{\max}} \]

which is summable since \( (n-k)^{1/\alpha_{\max}} \sim A(\sqrt{\ln n}/n)^{1/\alpha_{\min}} \) with \( 1/\alpha_{\min} > 2 \).
whence \(|f(T^k x)| \leq C/(n-k)^{\gamma/\alpha_{\max}}\), and a summation yields that \(|f_Y(x)| \leq C n^{1-\gamma/\alpha_{\max}}\). This bound tends to infinity when \(n \to \infty\), but sufficiently slowly so that \(f_Y\) still belongs to \(L^p\). More precisely,

\[
\int |f_Y|^p \leq C \sum m(\varphi_Y = n)n^{p(1-\gamma/\alpha_{\max})}
\leq C \sum m(\varphi_Y > n)n^{p(1-\gamma/\alpha_{\max})-1}.
\]

As \(m(\varphi_Y > n) \sim A(\sqrt{n/n})^{1/\alpha_{\max}}\), this last series is summable as soon as

\[
-\frac{1}{\alpha_{\min}} + p \left(1 - \frac{\gamma}{\alpha_{\max}}\right) - 1 < -1,
\]

which is the case by assumption on \(p\).

Assume now that \(f\) has a vanishing integral on \(S^1\). Let \(g(\omega, x) = f(\omega, 0)\). The function \(f-g\) vanishes on \(S^1 \times [0,1]\), whence \(f_Y - g_Y \in L^p\) according to the first part of this proof. Consequently, it is sufficient to prove that \(g_Y \in L^p\). Write \(\chi(\omega) = f(\omega, 0)\) and \(S_t \chi(\omega) = \sum_{k=0}^{t-1} \chi(F^k \omega)\); then \(g_Y(\omega, x) = S_t \chi(\omega)\).

Let \(M_n \chi(\omega) = \max_{s \leq n} |S_s \chi(\omega)|\). Let \(\delta > 0\), and \(l = (1+\delta)/\delta\), so that \(1/l + 1/l(1+\delta) = 1\). We have

\[
\int_{\{|g_Y| \geq 2\}} |g_Y|^p \leq \sum_{n=2}^{\infty} \int_{S^1} \int_{1/2+X_{n-1}(F \omega)/2}^1 |S_n \chi(\omega)|^p \, du \, d\omega
\leq \sum_{k=1}^{\infty} \int_{S^1} \int_{1/2+X_{2k-1}(F \omega)/2}^1 |M_{2k} \chi(\omega)|^p \, du \, d\omega
\leq \sum_{k=1}^{\infty} \int_{S^1} X_{2k-1}(F \omega)|M_{2k} \chi(\omega)|^p \, d\omega
\leq \sum_{k=1}^{\infty} \|X_{2k-1} \circ F\|_{l+\delta} \|M_{2k} \chi\|_{l^p},
\]

where the last inequality is H"older inequality. If \(\delta\) is small enough, \(lp > 2\), whence Corollary B.4 yields that \(\|M_{2k} \chi\|_{l^p} \leq C k^{(l-1)/lp} \sqrt{2^k}\). Moreover,

\[
\|X_{2k-1} \circ F\|_{l+\delta} = \|X_{2k-1}\|_{l+\delta} \leq \left(\int X_{2k-1} \right)^{1/(1+\delta)} \sim C \left(\frac{\sqrt{\ln(2k^{-1})}}{2k^{-1}}\right)^{(1+\delta)\alpha_{\min}}
\]

by Theorem 4.1. Thus, \(\int |g_Y|^p < \infty\) if \(l/(1+\delta)\alpha_{\min} > p/2\), and it is possible to choose \(\delta\) such that this inequality is true, since \(1/\alpha_{\min} > p/2\) by hypothesis. \(\square\)

5.2. **Proof of Theorem 5.1.** To apply Theorem 3.1, we first check the condition (18). Let \(\theta\) be the H"older exponent of \(f\). We will work with the distance \(d_{\lambda^{-\theta}}(x, y) = \lambda^{-\theta d(x, y)}\). For this distance, \(T_\lambda\) is a Gibbs–Markov map.

**FACT.** If \(f\) is \(\theta\)-H"older on \(S^1 \times [0,1]\), then

\[
\sum m[A_{s,n}]Df_Y(A_{s,n}) < \infty.
\]

Recall that \(Df_Y(A_{s,n})\) (defined in Theorem 3.1) is the best Lipschitz constant of \(f_Y\) on \(A_{s,n}\), here for the distance \(d_{\lambda^{-\theta}}\).
Proof of the fact. Take \((\omega_1, x_1)\) and \((\omega_2, x_2)\) \(\in A_{x,n}\) with, for example, \(x_2 \geq x_1\). Then (9) implies that \(x_1 \in J_n^+ (\omega_2)\), and Corollary 2.5 applied \(n-1\) times proves that, for \(0 < k \leq n\),
\[
d(T^k(\omega_1, x_1), T^k(\omega_2, x_1)) \leq D|F^k\omega_1 - F^k\omega_2|.
\]
Moreover, \(d(T^k(\omega_2, x_1), T^k(\omega_2, x_2)) \leq d(T^n(\omega_2, x_1), T^n(\omega_2, x_2))\) (since each map \(T_\alpha(\omega)\) is expanding).

Thus, for \(0 \leq k \leq n\),
\[
d(T^k(\omega_1, x_1), T^k(\omega_2, x_2)) \leq d(T^k(\omega_1, x_1), T^k(\omega_2, x_1)) + d(T^k(\omega_2, x_1), T^k(\omega_2, x_2))
\]
\[
\leq D|F^k\omega_1 - F^k\omega_2| + d(T^n(\omega_2, x_1), T^n(\omega_2, x_2))
\]
\[
\leq D|F^n\omega_1 - F^n\omega_2| + d(T^n(\omega_1, x_1), T^n(\omega_2, x_1))
\]
\[
+ d(T^n(\omega_1, x_1), T^n(\omega_2, x_2))
\]
\[
\leq D|F^n\omega_1 - F^n\omega_2| + d(T^n(\omega_1, x_1), T^n(\omega_2, x_2))
\]
\[
\leq (1 + 2D)d(T^n(\omega_1, x_1), T^n(\omega_2, x_2)).
\]

We deduce that
\[
|f_T(\omega_1, x_1) - f_T(\omega_2, x_2)| \leq \sum_{k=0}^{n-1} |f(T^k(\omega_1, x_1)) - f(T^k(\omega_2, x_2))|
\]
\[
\leq \sum_{k=0}^{n-1} Cd(T^k(\omega_1, x_1), T^k(\omega_2, x_2))^0
\]
\[
\leq C'n d(T^n(\omega_1, x_1), T^n(\omega_2, x_2))^0.
\]

As \(T_\alpha\) is expanding for the distance \(d'\) (defined in (7), and equivalent to \(d\)), we have
\[
d(T^n(\omega_1, x_1), T^n(\omega_2, x_2)) \leq C\lambda d^{-1}(\omega_1, x_1, \omega_2, x_2),
\]
whence
\[
d(T^n(\omega_1, x_1), T^n(\omega_2, x_2))^0 \leq C\lambda^{-\alpha}(\omega_1, x_1, \omega_2, x_2).
\]

Thus, \(Df_T(A_{x,n}) \leq Cn\), and
\[
\sum m(A_{x,n}) Df_T(A_{x,n}) \leq C \sum m(\varphi_T = n) \lambda = C < +\infty,
\]
by Kac's formula. \(\square\)

Proof of Theorem 5.1. In the case \(\alpha_{\min} < \frac{1}{2}\), Lemma 5.3 gives that \(f_T \in L^2\). Moreover, \(\varphi \in L^2\) (since \(\varphi = g_T\) for \(g \equiv 1\), whence Lemma 5.3 also applies). We have already checked the condition (18), so we can apply (the first case of) Theorem 3.1. This yields the central limit theorem for \(f_T\).

Assume now that \(\frac{1}{2} < \alpha_{\min} < 1\) and that \(c = 0\). Under the assumptions of the theorem, we can apply Lemma 5.4 with \(p = 2\), and get that \(f_T \in L^2\). Moreover, Lemma 5.2 shows that \(m[\varphi_T > x] \sim (\sqrt{\ln x}/x)^{1/\alpha_{\min}}\). We have checked all of the hypotheses of the first case of Theorem 3.1. Applying this theorem, we conclude the proof of the second case.

The third and fourth cases are analogous. Let us prove, for example, the fourth case, i.e. \(\frac{1}{2} < \alpha_{\min} < 1\) and \(c \neq 0\). Assume, for example, that \(c > 0\). We estimate \(m(f_T > x)\).

FACT. We have
\[
m(f_T > x) \sim \left(\frac{c\sqrt{\ln x}}{x}\right)^{1/\alpha_{\min}} A \quad \text{and} \quad m(f_T < -x) = o\left(\frac{\sqrt{\ln x}}{x}\right)^{1/\alpha_{\min}}.
\]
Proof. We prove the estimate on $m(f_Y > x)$, the other being similar.

Let $g = c$ on $S^1 \times [0, 1]$. Then $g_Y = cg$ on $[\psi_Y = n]$, which implies that $m(g_Y > nc) = m(\psi_Y > n) \sim (\sqrt{n \log n})^{1/\alpha}$ by Lemma 5.2.

In the general case, consider $j = f - g$, and let us prove that $m(\lvert j_Y \rvert > x) = o(\sqrt{n \log n})^{1/\alpha}$. As $f_Y = g_Y + j_Y$, it will give $m(g_Y > x(1 + \varepsilon)) - m(\lvert j_Y \rvert > x\varepsilon) \leq m(f_Y > x) \leq m(g_Y > x(1 - \varepsilon)) + m(\lvert j_Y \rvert > x\varepsilon)$, which gives the conclusion.

Let $\gamma > 0$ with $\gamma < \min(\theta, \alpha_{\max})$ (where $\theta$ is the Hölder coefficient of $f$). Lemma 5.4 gives that $j_Y \in L^p$ if $p < \min(2/\alpha_{\min}, 1/\alpha_{\min}(1 - \gamma/\alpha_{\max}))$. We can, in particular, choose $p > 1/\alpha_{\min}$. Then $m(\lvert j_Y \rvert > x) \leq \int (\lvert j_Y \rvert / x)^p \leq O(x^{-\gamma})$, which concludes the proof of the fact.

The same fact holds for $\psi_Y$, with the same proof. Therefore, the assumptions of the third case of Theorem 3.1 are satisfied. This implies the desired result.

A. Appendix. Induced maps and limit theorems

The aim of this section is to prove very general results stating that, if a function satisfies a limit theorem for an induced map, it also satisfies one for the initial map. Similar theorems have been proved in [Gou04], by spectral methods, under strong technical assumptions. We will describe here a more elementary method, essentially due to Melbourne and Török [MT04]. Zweimüller has also used the same kind of arguments to study limit laws in dimension 1, see [Zwe03]. This new method does not imply all the results of [Gou04], but it can be used in settings where [Gou04] can not be applied.

If $Y$ is a subset of a probability space $(X, m)$, $T : X \rightarrow X$, and $T_Y$ is the induced map on $Y$, we will write $S^*_{\mu} g = \sum_{k=0}^{\mu} g \circ T^k_Y$ : this is the Birkhoff sum of $g$, for the transformation $T_Y$. We will also write $E_Y(g) = \int_T g/d\mu$. Finally, for $t \in \mathbb{R}$, $\lfloor t \rfloor$ denotes the integer part of $t$.

**Theorem A.1.** Let $T : X \rightarrow X$ be an ergodic endomorphism of a probability space $(X, m)$, and $f : X \rightarrow \mathbb{R}$ an integrable function with vanishing integral. Let $Y \subset X$ have positive measure. For $y \in Y$, write $\phi(y) = \inf\{n > 0 \mid T^n(y) \in Y\}$ and $f_Y(y) = \sum_{k=0}^{\phi(y)-1} f(T^k y)$.

We assume the following properties.

1. There exists a sequence $B_n \rightarrow +\infty$, with $\inf_{n \geq 1}(B_n / B_n) > 0$, such that $(f_Y, \phi)$ satisfies a mixing limit theorem for the normalization $B_n$: there exists a random variable $Z$ such that, for every $t \in \mathbb{R}$,

   $$E_Y(\phi^{i(S^*_{\inf(\phi)} f_Y / B_n)}) \rightarrow E_Y(\phi)E(e^{itZ}).$$

2. There exists $b > 0$ such that, in the natural extension of $T_Y$, $(1/N^b) \sum_{0}^{N-1} f_Y(T^k_Y)$ tends almost everywhere to 0 when $N \rightarrow \pm\infty$.

3. The sequence $(S^*_{\phi} \phi - n E_Y(\phi))/B_{n}^{1/b}$ is tight, in the following sense: for every $\varepsilon > 0$, there exists $A > 0$ and $N_0$ such that, for every $n \geq N_0$,

   $$m\left\{ y \in Y \mid \left| \frac{S^*_{\phi} \phi - n E_Y(\phi)}{B_{n}^{1/b}} \right| \geq A \right\} \leq \varepsilon.$$
Statistical properties of a skew product

Then the function $f$ also satisfies a limit theorem:

$$E(e^{itS_n f/B_n}) \to E(e^{itZ}),$$

(i.e. $S_n f/B_n$ tends in distribution to $Z$).

The hypotheses of the theorem are tailor-made so that the following proof works, but they are, in fact, often satisfied in natural cases. Let us comment on these three hypotheses.

1. The convergence (52) is very often satisfied when $f_Y$ satisfies a limit theorem. Namely, the martingale proofs or spectral proofs of limit theorems automatically give this kind of convergence†.

2. The natural extension is useful so that we can let $N$ tend to $-\infty$, and consider $T_{-1}^{-1} Y$ in the proof. Generally, Birkhoff’s theorem yields that this assumption is satisfied for $b = 1$. This is often sufficient. However, sometimes, it is important to have better estimates. It is then possible to use [Kac96, Theorem 16], for example: this theorem ensures that, if the correlations of $f_Y \in L^2$ decay at least like $O(1/n)$, then the hypothesis is satisfied for any $b > 1/2$ (for $N \to -\infty$, use the fact that $\int f_Y \cdot f_Y \circ T^n Y = \int f_Y \circ T_{-1}^{-n} \cdot f_Y$, and apply the result to $T_{-1}^{-1} Y$).

3. The third assumption is weaker than $\exists B_n' = O(B_n^{1/b})$ such that $S_n Y \varphi - n E_Y(\varphi)/B_n$ converges in distribution. (53')

Moreover, $\varphi$ is often simpler than $f_Y$. Since $f_Y$ satisfies a limit theorem (this is more or less the first hypothesis), this is also often the case of $\varphi$, which implies (53'). Thus, (53'), and hence (53), are satisfied quite generally.

**Proof of Theorem A.1.** We can assume that $m(Y) < 1$.

Without loss of generality, we can replace $T$ by its natural extension and assume that $T$ is invertible. We will identify $X$ with $\{ (y, i) \mid y \in Y, i \in \{0, \ldots, \varphi(y) - 1\} \}$. In this notation, for $i < \varphi(y) - 1$, $T(y, i) = (y, i + 1)$, while $T(y, \varphi(y) - 1) = (T_Y(y), 0)$. Note that $E_Y(\varphi) = 1/m(Y)$ by Kac’s formula. Let $\pi$ be the projection from $X$ to $Y$, given by $\pi(y, i) = y$.

In this proof, we will write $S_t f(x)$, even when $t$ is not an integer, for $S_{\lfloor t \rfloor} f(x)$. In the same way, $T^t$ should be understood as $T^{\lfloor t \rfloor}$. We also extend $B_n$ to $\mathbb{R}_+$, setting $B_t := B_{\lfloor t \rfloor}$.

As $T$ is ergodic, $T_Y$ is also ergodic [Aar97, Proposition 1.5.2]. Birkhoff’s theorem gives that

$$S_n^Y \varphi = \frac{n}{m(Y)} + o(n)$$

almost everywhere on $Y$. For $y \in Y$ and $N \in \mathbb{N}$, let $n(y, N)$ be the greatest integer $n$ such that $S_n^Y \varphi(y) \leq N$. If $y$ is such that $S_n^Y \varphi(y) = n/m(Y) + o(n)$ (which is true almost everywhere), then $n(y, N)$ is finite for every $N$, and $n(y, N)/m(Y) \sim N$, i.e.

$$\frac{n(y, N)}{Nm(Y)} \to 1.$$  

† In fact, Zweimüller pointed out that the distributional convergence of $S_{[nm(Y)]} f_Y / B_n$ to $Z$ always implies (52): this is a consequence of the proof of [Aar97, Proposition 3.6.1].
Then

Proof. By Lemma A.2, it is sufficient to prove it on $\mathbb{S}$. Gouezel

Set $\mathbb{B}$ iterations of $T$. This will give that $S_N f/BN \to 0$ in probability.

Take to prove that they tend to $0$ on $\mathbb{X}$. The last term, equal to $(SN f_{Y,N} f_Y) \circ \pi$, satisfies a limit theorem by (57). To conclude the proof, we will see that the three other terms, divided by $BN$, tend to $0$ in probability.

The second and third terms depend only on $y$. Thus, the following lemma will be useful to prove that they tend to $0$ on $\mathbb{X}$.

**Lemma A.2.** Let $f_n$ be a sequence of functions on $Y$, tending to $0$ in probability on $Y$. Then $f_n \circ \pi$ tends to $0$ in probability on $X$.

Proof. Take $\varepsilon > 0$. As $f_n \to 0$ in probability, the measure of $E_n := \{y \in Y \mid \|f_n(y)\| \geq \varepsilon\}$ tends to $0$. As $\varphi \in L^1$, dominated convergence yields that $\int_{E_n} \varphi \to 0$, i.e. the measure of $\pi^{-1}(E_n)$ tends to $0$. However, $\pi^{-1}(E_n)$ is exactly the set where $|f_n \circ \pi| \geq \varepsilon$.

**Fact.** $BN^{-1}(SN f(y,i) - SN f(y,0))$ tends to $0$ in probability on $X$.

Proof. Set $V_N(y) = \sum_{i=0}^N |f \circ T^N(y,i)|$ on $Y$. Then $\|V_N\|_{L^1(Y)} = \|f \circ T^N\|_{L^1(X)} = \|f\|_{L^1(X)}$ since $T$ preserves the measure. Thus, $V_N/BN$ tends to $0$ in $L^1(Y)$, and in probability. Lemma A.2 yields that $BN^{-1}V_N \circ \pi$ tends to $0$ in probability on $X$.

As $SN f(y,i) - SN f(y,0) = \sum_{i=0}^{N-i-1} f(T^k(y,0)) - \sum_{i=0}^{i-1} f(T^k(y,0))$, we get $|SN f(y,i) - SN f(y,0)| \leq V_N(y) + V_0(y)$. Thus, $BN^{-1}(SN f(y,i) - SN f(y,0))$ is bounded by a function going to $0$ in probability.

**Fact.** $BN^{-1}(SN f(y,0) - SN f_{Y,N}(f_Y(y)))$ tends to $0$ in probability on $X$.

Proof. By Lemma A.2, it is sufficient to prove it on $Y$. Set $H(y,i) = \sum_{j=0}^{i-1} f(y,j)$. Then

$$|SN f(y,0) - SN f_{Y,N}(f_Y(y))| = H \circ T^N(y,0).$$

(58)

Since $T$ preserves the measure $m$, for any $a > 0$,

$$m\left\{y \in Y \mid \frac{1}{BN} H \circ T^N(y,0) \geq a\right\} \leq m\left\{x \in X \mid \frac{1}{BN} H \circ T^N(x) \geq a\right\} = m\left\{x \in X \mid \frac{1}{BN} H(x) \geq a\right\}.
$$

Since $H$ is measurable and $BN \to \infty$, this measure tends to $0$ when $N \to \infty$.

**Fact.** $BN^{-1}(SN f_{Y,N}(f_Y) - SN_{Y,m}(f_Y))$ tends to $0$ in probability on $X$ when $N \to \infty$. 

Since $\int_X e^{it(S_N f_Y)_{x\pi}} = \int_Y \varphi e^{it S_Y} \mathbb{F}$, (52) yields that

$$\left(\frac{SN f_Y(y \pi) \circ \pi}{BN}\right) \to Z.$$
Proof. By Lemma A.2, it is sufficient to prove it on \( Y \).

For \( n < 0 \), write \( S_n^Y f_Y = \sum_{r=0}^{\lfloor n \rfloor} f_Y \circ T^{-r}_Y \). Then, setting \( v(y, N) = n(y, N) - N m(Y) \),

\[
S^Y_{n(y, N)} f_Y(y) = S^Y_{\nu(y, N)} f_Y(T^{-N m(Y)}_Y(y)). \tag{59}
\]

Let \( A > 0 \) and \( N \in \mathbb{N} \), we will estimate the measure of \( \{ y \mid v(y, N) \geq A B_N^{1/b} \} \). Take \( \alpha > 0 \) such that \( m(Y) + \alpha < 1 \). Assume first that \( A B_N^{1/b} > \alpha N \). Then,

\[
\{ y \mid v(y, N) \geq A B_N^{1/b} \} \subset \{ y \mid n(y, N) \geq (m(Y) + \alpha) N \}. \tag{60}
\]

By (55), the measure of this set tends to 0 when \( N \to \infty \). Assume next that \( A B_N^{1/b} \leq \alpha N \).

Since \( E_Y(\psi) = 1/m(Y) \), we get

\[
\{ y \mid v(y, N) \geq A B_N^{1/b} \} = \{ n(y, N) \geq A B_N^{1/b} + N m(Y) \} = \{ S^Y_{A B_N^{1/b} + N m(Y)} \psi \leq N \}
\]

\[
\leq \left\{ \frac{S^Y_{A B_N^{1/b} + N m(Y)} \psi - (A B_N^{1/b} + N m(Y)) E_Y(\psi)}{(B_{A B_N^{1/b} + N m(Y)})^{1/b}} \leq \frac{A}{m(Y)} \left( \frac{B_N}{B_{A B_N^{1/b} + N m(Y)}} \right)^{1/b} \right\}.
\]

Moreover, \( A B_N^{1/b} + N m(Y) \leq (m(Y) + \alpha) N \leq N \). By assumption, there exists \( c > 0 \) such that, for all \( n \leq r \), \( B_r/B_n \geq c \). In particular, \( B_N/B_{A B_N^{1/b} + N m(Y)} \geq c \). Hence,

\[
\{ y \mid v(y, N) \geq A B_N^{1/b} \} \subset \left\{ \frac{S^Y_{A B_N^{1/b} + N m(Y)} \psi - (A B_N^{1/b} + N m(Y)) E_Y(\psi)}{(B_{A B_N^{1/b} + N m(Y)})^{1/b}} \leq \frac{Ac^{1/b}}{m(Y)} \right\}.
\]

Consequently, if \( A \) is large enough, assumption 3 yields that \( m\{ y \mid v(y, N) \geq A B_N^{1/b} \} \leq \varepsilon \) for large enough \( N \). We handle in the same way the set of points where \( v(y, N) \leq -A B_N^{1/b} \).

We have thus proved

\[
\forall \varepsilon > 0, \exists A > 0, \exists N_0 > 0, \forall N \geq N_0, \ m\{ y \mid |v(y, N)| \geq A B_N^{1/b} \} \leq \varepsilon. \tag{61}
\]

Set \( W_N(y) = B_N^{-1} S_{v(y, N)} f_Y(T^{-N m(Y)}_Y(y)) \), we will show that it tends to 0 in distribution, which will conclude the proof, by (59). Take \( a > 0 \), we will show that \( m(\{|W_N| > a\}) \to 0 \) when \( N \to \infty \).

Let \( \varepsilon > 0 \). Assumption 2 ensures that there exists \( \tilde{Y} \) with \( m(\tilde{Y}) \geq m(Y) - \varepsilon \) and \( N_1 \) such that \( (1/|N_1|) S_{N_1} f_Y \leq \varepsilon \) on \( \tilde{Y} \), for every \( |N| \geq N_1 \). Define \( Y'_N = \{ y \in Y \mid |v(y, N)| < N_1 \} \) and \( Y''_N = \{ y \in Y \mid |v(y, N)| \geq N_1 \} \). We estimate first the contribution of \( Y''_N \).

Set \( \psi(y) = \sum_{-N_1}^{-N_1} |f_Y \circ T^{-r}_Y(y)| \). If \( y \in Y'_N \), then \( |W_N(y)| \leq \psi(T^{-N m(Y)}_Y(y))/B_N \).

Therefore,

\[
m\{ y \in Y'_N \mid |W_N(y)| \geq a \} \leq m\{ y \in Y'_N \mid |\psi(T^{-N m(Y)}_Y(y))| \geq a B_N \}
\]

\[
= m\{ y \in Y'_N \mid |\psi(y)| \geq a B_N \}.
\]

Since \( \psi \) is measurable, this quantity tends to 0 when \( N \to \infty \). In particular, if \( N \) is large enough, it is at most \( \varepsilon \).

We then estimate the contribution of \( Y''_N \). Set \( \tilde{Y}''_N = Y''_N \cap T^{-N m(Y)}_Y(\tilde{Y}) \), it satisfies \( m(\tilde{Y}''_N) \geq m(Y''_N) - \varepsilon \). Thus,

\[
m(\{|W_N| > a\}) \leq m\{ y \in \tilde{Y}''_N \mid |W_N(y)| \geq a \} + 2\varepsilon. \tag{62}
\]
On $\tilde{Y}_N^\nu$, $|v(y, N)| \geq N_1$, whence $(1/|v(y, N)|^b)|S^Y_{v(y,N)}f_Y(T_N^{Nm(y)}y)| \leq \varepsilon$. Thus,

$$|W_N(y)| \leq \varepsilon \frac{|v(y, N)|^b}{B_N} \leq \varepsilon \left( \frac{|v(y, N)|}{B_N^{1/b}} \right)^b.$$ 

Consequently,

$$m(|W_N| \geq a) \leq m \left( \frac{|v(y, N)|}{B_N^{1/b}} \geq \left( \frac{a}{\varepsilon} \right)^{1/b} \right) + 2\varepsilon.$$ 

(63)

This equation together with (61) implies that $m(|W_N| \geq a) \rightarrow 0$ when $N \rightarrow \infty$. \hfill \Box

The three facts we have just proved imply that $S_N f(y, i)/B_N - S^Y_{Nm(y)}f_Y(y)/B_N \rightarrow 0$ in distribution on $X$. As $S^Y_{Nm(y)}f_Y(y)/B_N \rightarrow Z$ in distribution on $X$, by (57), this concludes the proof. \hfill \Box

B. Appendix. Multiple decorrelations and $L^p$-boundedness

The following theorem has been useful in this paper.

**Theorem B.1.** Let $F : \omega \rightarrow 4\omega$ on the circle $S^1$. Then, for every $p \in [1, \infty)$, there exists a constant $K_p$ such that, for every $n \in \mathbb{N}$, for every $f_0, \ldots, f_{n-1} : S^1 \rightarrow \mathbb{R}$ bounded by 1, of zero average and 1-Lipschitz,

$$\left\| \sum_{k=0}^{n-1} f_k \circ F^k \right\|_p \leq K_p \sqrt{n}.$$ 

(64)

This result has essentially been proved by Pène in [Pén02], in a much broader context. Her proof depends on a property of multiple decorrelations, which is implied by the spectral gap of the transfer operator.

**Lemma B.2.** Let $\| \cdot \|$ be the Lipschitz norm of the function $f$ on the circle $S^1$. Then, for every $m, m' \in \mathbb{N}$, there exist $C > 0$ and $\delta < 1$ such that, for every $N \in \mathbb{N}$, for every increasing sequences $(k_1, \ldots, k_m)$ and $(l_1, \ldots, l_{m'})$, for every Lipschitz functions $G_1, \ldots, G_m$, $H_1, \ldots, H_{m'}$,

$$\left| \text{Cov}(m \prod_{i=1}^m G_i \circ F^{k_i} \cdot \prod_{j=1}^{m'} H_j \circ F^{N+l_j}) \right| \leq C \left( \prod_{i=1}^m \| G_i \| \right) \left( \prod_{j=1}^{m'} \| H_j \| \right) \delta^{N-k_m}. \quad (65)$$

Here $\text{Cov}(u, v) = \int u v - \int u \int v$.

**Proof.** Let $\hat{F}$ be the transfer operator associated to $F$, and acting on Lipschitz functions. It is known that it admits a spectral gap and that its iterates are bounded, i.e. there exist constants $M > 0$ and $\delta < 1$ such that $\| \hat{F}^n f \| \leq M \| f \|$, and $\| \hat{F}^n f \| \leq M \delta^n \| f \|$ if $\int f = 0$.

We can assume that $N \geq k_m$ (otherwise $\delta^{N-k_m} \geq 1$, and the inequality (65) becomes trivial). Then, writing $\varphi = \prod_{i=1}^m G_i \circ F^{k_i}$ and $\psi = \prod_{j=1}^{m'} H_j \circ F^{l_j}$, we get

$$\left| \text{Cov}(\varphi, \psi \circ F^N) \right| = \left| \int \left( \varphi - \int \varphi \right) \psi \circ F^N \right| = \left| \int \hat{F}^N \left( \varphi - \int \varphi \right) \psi \right| \leq \| \hat{F}^N \left( \varphi - \int \varphi \right) \| \| \psi \|_\infty.$$
However,
\[ \hat{F}^N(\varphi) = \prod_{i=1}^{m} G_i \circ F^{k_i} = \hat{F}^{N-k_m}(G_{m-1} \hat{F}^{k_m-1-k_m-1}(\ldots \hat{F}^{k_2-k_1}(G_1)\ldots) =: \hat{F}^{N-k_m}(\chi). \]
As the iterates of \( \hat{F} \) are bounded on Lipschitz functions, we get a bound on the Lipschitz norm of \( \chi \):
\[ \| \chi \| \leq M^{m-1} \prod \| G_i \|. \]
Moreover, \( \int \chi = \int \varphi \), whence
\[ \| \hat{F}^N(\varphi) - \int \varphi \| \leq M^{m-1} \prod \| G_i \|. \]

When \( p \) is an even integer, Theorem B.1 is then a consequence of [Pén02, Lemma 2.3.4]. The Hölder inequality gives the general case.

**Remark.** The same result holds for Hölder functions instead of Lipschitz functions, with the same proof.

We will also need the following result.

**Theorem B.3.** Let \( T \) be a measure preserving transformation on a space \( X \). Let \( f : X \to \mathbb{R} \) and \( p > 2 \) be such that
\[ \exists C > 0, \forall n \in \mathbb{N}^*, \| S_n f \|_p \leq C \sqrt{n}. \] (66)
Write \( M_n f(x) = \sup_{1 \leq k \leq n} |S_k f(x)| \). Then there exists a constant \( K \) such that
\[ \forall n \geq 2, \| M_n f \|_p \leq K (\ln n)^{(p-1)/p} \sqrt{n}. \] (67)

**Proof.** Let \( n \in \mathbb{N}^* \). Let \( k < 2^n \), and write its binary decomposition \( k = \sum_{j=0}^{n-1} \varepsilon_j 2^j \), with \( \varepsilon_j \in \{0, 1\} \). Set \( q_j = \sum_{j=0}^{n-1} \varepsilon_j 2^j \) (in particular, \( q_0 = k \) and \( q_{n-1} = 0 \)). Then
\[ S_k f = \sum_{j=0}^{n-1} (S_{q_j} f - S_{q_{j+1}} f). \]
Consequently, the convexity inequality \( (a_0 + \cdots + a_{n-1})^p \leq n^{p-1} (a_0^p + \cdots + a_{n-1}^p) \) gives
\[ |S_k f|^p \leq n^{p-1} \sum_{j=0}^{n-1} |S_{q_j} f - S_{q_{j+1}} f|^p. \] (68)
Note that \( q_{j+1} \) is of the form \( \lambda 2^{j+1} \) with \( 0 \leq \lambda \leq 2^n - j - 1 \), and \( q_j \) is equal to \( q_{j+1} \) or \( q_{j+1} + 2^j \). Thus,
\[ |S_k f|^p \leq n^{p-1} \sum_{j=0}^{n-1} \left( \sum_{\lambda=0}^{2^n-j-1} |S_{\lambda 2^{j+1}+2^j} f - S_{\lambda 2^{j+1} f}|^p \right). \] (69)
The right-hand term is independent of \( k \), and gives a bound on \( |M_{2^n-1} f|^p \). Moreover,
\[ \int |S_{2^n-j} f - S_{2^{j+1}} f|^p = \int |S_{2^j} f|^p \leq C^p \sqrt{2^j}. \] (70)
Therefore, we get
\[
\int |M_{2^{n-1}} f|^p \leq n^{p-1} \sum_{j=0}^{n-1} 2^{n-j-1} C_p 2^{p/2} \leq Kn^{p-1} 2^n \frac{2(p/2-1)n}{n} = Kn^{p-1} \sqrt{2^p n}.
\]

For times of the form $2^n - 1$, this is a bound of the form $\|M_t\|_p \leq K (\ln t)^{(p-1)/p} \sqrt{t}$. To get the same estimate for an arbitrary time $t$, it is sufficient to choose $n$ with $2^n - 1 \leq t < 2^{n+1}$, and to note that $M_t \leq M_{2^n - 1}$.

**COROLLARY B.4.** Let $F : \omega \rightarrow 4\omega$ on the circle $S^1$, let $\chi : S^1 \rightarrow \mathbb{R}$ be a Hölder function with 0 average, and let $p > 2$. Write $M_n \chi(x) = \sup_{1 \leq k \leq n} |S_k \chi(x)|$. Then there exists a constant $K$ such that
\[
\|M_n \chi\|_p \leq K (\ln n)^{(p-1)/p} \sqrt{n}, \quad \forall n \geq 2.
\] (71)

**Proof.** Theorem B.1 (or rather the remark following it, for the Hölder case) shows that $\|S_n \chi\|_p \leq C \sqrt{n}$. Consequently, Theorem B.3 gives the conclusion. 

**References**


