ISRAEL JOURNAL OF MATHEMATICS **139** (2004), 29–65

SHARP POLYNOMIAL ESTIMATES FOR THE DECAY OF CORRELATIONS

BY

SÉBASTIEN GOUËZEL

Laboratoire de mathématiques, bâtiment 425
Université Paris-Sud, 91405 Orsay, France
e-mail: Sebastien.Gouezel@math.u-psud.fr

ABSTRACT

We generalize a method developed by Sarig to obtain polynomial lower bounds for correlation functions for maps with a countable Markov partition. A consequence is that LS Young's estimates on towers are always optimal. Moreover, we show that, for functions with zero average, the decay rate is better, gaining a factor $1/n$. This implies a Central Limit Theorem in contexts where it was not expected, e.g., $x + Cx^{1+\alpha}$ with $1/2 \leqslant \alpha < 1$. The method is based on a general result on renewal sequences of operators, and gives an asymptotic estimate up to any precision of such operators.

1. Statement of results

In recent years, several methods have been developed to obtain polynomial upper bounds for the correlations of some dynamical systems. However, there was no general method to get polynomial lower bounds for the decay of correlations, until Omri Sarig's recent article [Sar02]. He used an abstract result on renewal sequences of operators to obtain lower bounds on the decay of correlations for Markov maps. As an application, he proved that the upper bounds obtained by Young on tower maps in [You99] are in many cases optimal. The goal of this article is to remove some unnecessary assumptions in [Sar02], and as a consequence to prove that Young's estimates are optimal in full generality.

In this article, $\mathbb{D}$ will always denote $\{z \in \mathbb{C} \mid |z| < 1\}$. The analogue of Sarig's theorem on renewal sequences that we obtain is the following:

Received February 12, 2002 and in revised form November 29, 2002

THEOREM 1.1: *Let T_n be bounded operators on a Banach space $\mathcal{L}$ such that $T(z) = I + \sum_{n\geqslant 1} z^n T_n$ converges in $\operatorname{Hom}(\mathcal{L},\mathcal{L})$ for every $z \in \mathbb{D}$. Assume that:*

1. **Renewal equation**: *for every $z \in \mathbb{D}$, $T(z) = (I - R(z))^{-1}$ where $R(z) = \sum_{n\geqslant 1} z^n R_n$, $R_n \in \operatorname{Hom}(\mathcal{L},\mathcal{L})$ and $\sum \|R_n\| < +\infty$.*
2. **Spectral gap**: *1 is a simple isolated eigenvalue of $R(1)$.*
3. **Aperiodicity**: *for every $z \in \overline{\mathbb{D}} - \{1\}$, $I - R(z)$ is invertible.*

Let P be the eigenprojection of $R(1)$ at 1. If $\sum_{k>n} \|R_k\| = O(1/n^\beta)$ for some $\beta > 1$ and $PR'(1)P \neq 0$, then for all n

$$T_n = \frac{1}{\mu}P + \frac{1}{\mu^2}\sum_{k=n+1}^{+\infty} P_k + E_n,$$

where μ is given by $PR'(1)P = \mu P$, $P_n = \sum_{l>n} PR_lP$ and $E_n \in \operatorname{Hom}(\mathcal{L},\mathcal{L})$ satisfy

$$\|E_n\| = \begin{cases} O(1/n^\beta) & \text{if } \beta > 2, \\ O(\log n/n^2) & \text{if } \beta = 2, \\ O(1/n^{2\beta-2}) & \text{if } 2 > \beta > 1. \end{cases}$$

Note that, in all cases, $\|E_n\| = o(1/n^{\beta-1})$, which is what is needed to obtain sharp asymptotics for the decay of correlations. This theorem extends Sarig's: he assumed $\beta > 2$ while we only need $\beta > 1$. Moreover, the result we obtain is slightly stronger than Sarig's even in the case $\beta > 2$ because the error term is a $O(1/n^\beta)$ instead of a $O(1/n^{\lfloor\beta\rfloor})$.

Finally, our aperiodicity assumption is weaker than Sarig's who needed to suppose that the spectral radius of $R(z)$ was < 1 for every $z \neq 1$. Our assumption is necessary because other eigenvalues equal to 1 would generate other terms in the asymptotic expression of T_n (which could be calculated using the same methods as in the following proof, and would involve the spectral projection at these points). For example, if $R(z) = z^2$, then $T_{2n} = 1$ while $T_{2n+1} = 0$, which shows that the conclusions of the theorem are not valid any more (there is a periodicity problem). This less restrictive aperiodicity hypothesis will be useful, for example, when applied to tower maps (see Corollary 1.6).

It is in fact possible to give an asymptotic estimate of T_n up to an error term $O(1/n^\beta)$ even when $\beta \leqslant 2$. However, the result is quite technical to state, and will be deferred to Section 5. The following consequence of Theorem 5.4 will be sufficient for most dynamical applications.

THEOREM 1.2: *Under the hypotheses of Theorem 1.1, if $f \in \mathcal{L}$ is such that $Pf = 0$, then $\|T_nf\| = O(1/n^\beta)$.*

These abstract results enable us to enhance the applications in [Sar02]. We state briefly the results we obtain, without recalling all the notation. In Section 6, a precise meaning will be given to all the notions involved. The following theorem is stated more precisely as Theorem 6.3.

THEOREM 1.3: *Let $(X, \mathcal{B}, m, T, \alpha)$ be a topologically mixing probability preserving Markov map, and let γ be a subset of the partition α. Denote by T_γ the map induced by T on $Y = \bigcup_{a \in \gamma} a$ — it is a Markov map for a subpartition δ of γ. Assume that the distortion of T_γ is Hölder and that T_γ has the "big image" property, i.e., the measures of the images of the elements of the partition are bounded away from 0 (which is always true when γ is finite). Assume, moreover, that $m[\varphi_\gamma > n] = O(1/n^\beta)$ for some $\beta > 1$, where φ_γ is the first return time from Y to Y.*

Then $\exists \theta \in (0,1)$, $C > 0$ such that $\forall f, g$ integrable and supported inside Y,

$$\left| \mathrm{Cor}(f, g \circ T^n) - \left(\sum_{k=n+1}^{\infty} m[\varphi_\gamma > k] \right) \int f \int g \right| \leqslant C F_\beta(n) \|g\|_\infty \|f\|_{\mathcal{L}}$$

where $F_\beta(n) = 1/n^\beta$ if $\beta > 2$, $\log n/n^2$ if $\beta = 2$ and $1/n^{2\beta-2}$ if $2 > \beta > 1$ (and $\mathcal{L}$ denotes the space of θ-Hölder functions on Y).

Moreover, if $\int f = 0$, then $\mathrm{Cor}(f, g \circ T^n) = O(1/n^\beta)$.

When $m[\varphi_\gamma > n] \asymp 1/n^\beta$ and $\int f, \int g \neq 0$, Theorem 1.3 implies that $\mathrm{Cor}(f, g \circ T^n) \asymp 1/n^{\beta-1}$. Thus, the exact speed of decay of correlations is polynomial, with exponent $\beta - 1$. Surprisingly, the decay rate is better for functions with zero integral, with a gain of 1 in the exponent. This kind of result is (to the knowledge of the author) new, and does not seem to be obtainable by more crude estimates: the methods giving only upper bounds on the speed of decay of correlations do not distinguish between functions with zero or nonzero integral, since they do not "see" the higher order terms in the expansion of T_n.

As an application, we obtain the summability of the correlations for functions with zero integral (and supported in Y) even when $\beta \leqslant 2$, which gives a Central Limit Theorem in cases where it was not expected. Note that the condition of zero integral is important and cannot be eliminated by subtracting a constant, since the functions would not remain supported in Y. In fact, the estimate in the previous theorem shows that, when $\beta \leqslant 2$, the correlations are *not* summable for a function with nonzero integral supported in Y. In the same way, this speed of decay of correlations does not hold for general functions with zero integral but not supported in Y: take a function f of nonzero integral supported in Y, the function $g := f - \int f$ has zero integral but its correlations are the same as those of f, whence they decay at a rate $\asymp 1/n^{\beta-1}$.

The following Central Limit Theorem is stated more accurately as Theorem 6.13.

THEOREM 1.4: *Under the same hypotheses as in Theorem 1.3, if $f \in \mathcal{L}$ is supported in Y and $\int f = 0$, then the sequence $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k$ converges in distribution to a Gaussian random variable of zero mean and finite variance σ^2, with*

$$\sigma^2 = -\int f^2 \mathrm{d}m + 2\sum_{n=0}^{\infty} \int f \cdot f \circ T^n \mathrm{d}m.$$

Finally, even though Theorem 1.3 describes the speed of decay of correlations only for functions f and g supported in Y, it is possible to drop this hypothesis on g. However, the results obtained are less precise and give only an upper bound on the decay of correlations, in $O(1/n^{\beta-1})$ if $\int f \neq 0$ and in $O(1/n^{\beta})$ if $\int f = 0$ (see Theorem 6.9 and Proposition 6.11). This kind of result is useful in the proof of the Central Limit Theorem.

The following corollaries are already present in weaker form in [Sar02], where the notations are explained. Some details on their proofs will be given in the last section of this article. The first corollary (stated more precisely as Corollary 7.1) deals with an explicit one-dimensional Markov map with a neutral fixed point, while the second corollary (see section 7.2 and Corollary 7.2) is essentially Theorem 1.3 expressed in the framework of LS Young towers, which are devices built up from non-Markov maps which have proved very useful in studying their statistical properties (see [You99]).

COROLLARY 1.5: *In the case of the Liverani–Saussol–Vaienti map T: $[0,1] \to [0,1]$ defined by*

$$T(x) = \begin{cases} x(1+2^{\alpha}x^{\alpha}) & \text{if } 0 \leqslant x \leqslant 1/2 \\ 2x-1 & \text{if } 1/2 < x \leqslant 1 \end{cases}$$

(see [LSV99]), if $\alpha \in (0,1)$, f is Lipschitz, g is bounded measurable, $\int f, \int g \neq 0$ and $f, g = 0$ in a neighborhood of 0, then

$$\mathrm{Cor}(f, g \circ T^n) \sim \frac{1}{4} h\left(\frac{1}{2}\right) \alpha^{-1/\alpha} \left(\frac{1}{\alpha} - 1\right)^{-1} n^{1-1/\alpha} \int f \int g$$

with respect to the invariant probability measure.

Moreover, if $\int f = 0$ (and f, g are still Lipschitz and zero in a neighborhood of 0), then $\mathrm{Cor}(f, g \circ T^n) = O(1/n^{1/\alpha})$. Consequently, f satisfies a Central Limit Theorem.

This result is in fact not specific to this particular map and can easily be extended to a class of maps admitting a neutral fixed point in 0 with a prescribed behavior, and expanding outside of any neighborhood of 0, making use of the following corollary and the techniques of [You99]. Note that Mark Holland has recently obtained upper bounds for the decay of correlations when the fixed point is more neutral ([Hol02]) — the techniques of the present article are not sufficient to prove that these upper bounds are always optimal.

COROLLARY 1.6: *Let $(\Delta, \mathcal{B}, m, F)$ be a probability preserving LS Young tower with $\gcd\{R_i\} = 1$ and $m[R > n] = O(1/n^\beta)$ where $\beta > 1$. If $f \in C_\theta(\Delta)$, $g \in L^\infty$ are supported inside $\bigcup_0^{N-1} \Delta_i$ for some N, then*

$$\mathrm{Cor}(f, g \circ F^n) = \sum_{k>n} m[R > k] \int f \int g + O(F_\beta(n)).$$

Moreover, if $\int f = 0$, then $\mathrm{Cor}(f, g \circ F^n) = O(1/n^\beta)$. Thus, f satisfies a Central Limit Theorem.

The aperiodicity hypothesis on $\gcd\{R_i\}$ is the same as Young's, and cannot be omitted. In her paper [You99], Young proved that, if $m[R > n] = O(1/n^\beta)$, then $\mathrm{Cor}(f, g \circ F^n) = O(1/n^{\beta-1})$ for any $f \in C_\theta(\Delta)$, $g \in L^\infty$ (not necessarily supported in $\bigcup_0^{N-1} \Delta_i$). Corollary 1.6 proves that this upper estimate is in fact optimal, and gives additionally a Central Limit Theorem even if $1 < \beta \leqslant 2$.

From this point on, the paper is divided into two parts: the first one (sections 2, 3, 4 and 5) is devoted to the proof of the abstract results on renewal sequences of operators, and the second one (sections 6 and 7) deals with the applications to Markov maps.

2. Preliminary results

2.1 $C^{1+\alpha}$ FUNCTIONS IN BANACH ALGEBRAS. The results in this section are mainly straightforward computations, and most of them can be found in [Sar02].

Let $\mathcal{A}$ be a Banach algebra (in the applications, $\mathcal{A} = \mathrm{Hom}(\mathcal{L}, \mathcal{L})$). Fix K a compact subset of $\mathbb{C}$, in which two points are always joined by a C^1 curve. The distance on K will not be the usual one, but the geodesic distance, i.e., $d(x,y)$ is the infimum of the lengths of C^1-paths in K joining x to y. We assume that this distance is equivalent to the usual one, which will be true for $K = \overline{\mathbb{D}}$ or $K = S^1$.

Fix some $0 < \alpha < 1$. For any $f \colon K \to \mathcal{A}$, we will say that f is C^α if there exists a constant C such that, for any $x, y \in K$ with $d(x,y) < 1$, $\|f(x) - f(y)\| \leqslant$

$Cd(x,y)^\alpha$. Let $D_\alpha(f)$ denote the least such constant. We write $\|f\|_\alpha = \|f\|_\infty + D_\alpha(f)$, and denote by $C^\alpha(K)$ the space of all functions such that $\|f\|_\alpha < +\infty$.

PROPOSITION 2.1: *The space $(C^\alpha(K), \| \ \|_\alpha)$ is a Banach algebra. In fact, we even have, for $f, g \in C^\alpha(K)$, $D_\alpha(fg) \leqslant \|f\|_\infty D_\alpha(g) + \|g\|_\infty D_\alpha(f)$.*

We say that $f\colon K \to \mathcal{A}$ is C^1 if there exists a continuous function $g\colon K \to \mathcal{A}$ such that $f(x+h) - f(x) - hg(x) = o(h)$ for any $x \in K$. The function g is unique if it exists, and we write $g = f'$. Contrary to the usual derivative of f, it is defined on the whole set K, and not just on its interior.

PROPOSITION 2.2: *If f is C^1 on K, then $D_\alpha(f) \leqslant \|f'\|_\infty$.*

Proof: Let $x, y \in K$ with $d(x,y) < 1$. Let γ be a C^1 path in K from x to y. The Taylor–Lagrange inequality along this path gives $\|f(x) - f(y)\| \leqslant \|f'\|_\infty l(\gamma)$. ∎

We consider the geodesic distance on K instead of the usual one precisely to get the above proposition.

Let $C^{1+\alpha}(K)$ denote the space of all C^1 functions from K to $\mathcal{A}$ whose derivative is C^α, endowed with the norm $\|f\|_{1+\alpha} = \|f\|_\infty + \|f'\|_\infty + \frac{1}{2}D_\alpha(f')$.

PROPOSITION 2.3: *The space $(C^{1+\alpha}(K), \| \ \|_{1+\alpha})$ is a Banach algebra.*

The following proposition will be used systematically in Section 3, often without explicit reference.

PROPOSITION 2.4: *Let $f\colon K \to \mathcal{A}$ be a $C^{1+\alpha}$ function such that, for every $z \in K$, $f(z)$ is invertible (as an element of $\mathcal{A}$). If $g(z) = f(z)^{-1}$, then g is $C^{1+\alpha}$ and there is an inequality $\|g\|_{1+\alpha} \leqslant F(\|g\|_\infty, \|f\|_{1+\alpha})$ for some universal polynomial function F.*

Proof: Differentiating $g(z) = f(z)^{-1}$, we get $g'(z) = -g(z)f'(z)g(z)$, hence $\|g'\|_\infty \leqslant \|g\|_\infty^2 \|f'\|_\infty$.

Then we note that

$$\begin{aligned} D_\alpha(g') &\leqslant D_\alpha(g)\|f'\|_\infty\|g\|_\infty + \|g\|_\infty D_\alpha(f')\|g\|_\infty + \|g\|_\infty\|f'\|_\infty D_\alpha(g) \\ &\leqslant 2\|g'\|_\infty\|f'\|_\infty\|g\|_\infty + \|g\|_\infty^2 D_\alpha(f'). \end{aligned}$$

The control on $\|g'\|_\infty$ enables us to conclude. ∎

2.2 FOURIER SERIES IN BANACH ALGEBRAS. Let $\mathcal{A}$ be a Banach algebra. For $f\colon S^1 \to \mathcal{A}$ a continuous function, we define the nth Fourier coefficient of f to be the element of $\mathcal{A}$ defined by

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})e^{-in\theta}d\theta.$$

Let us first recall a very useful result concerning functions from S^1 to $\mathbb{C}$.

THEOREM 2.5 (Wiener Lemma): *Let $f\colon S^1 \to \mathbb{C}$ be a continuous function, everywhere nonzero, whose Fourier coefficients are summable. Then the Fourier coefficients of $1/f$ are also summable.*

The classical proof of this result, which uses commutative Banach algebra techniques, can be found, for example, in [Kat68] (see also [New75] for an elementary proof).

PROPOSITION 2.6: *If $f\colon S^1 \to \mathcal{A}$ is continuous and satisfies $\sum \|c_n(f)\| < +\infty$, then $f(e^{i\theta}) = \sum c_n(f)e^{in\theta}$, the series converging in norm.*

Proof: Replacing f by $f - \sum c_n(f)e^{in\theta}$, we can assume that $c_n(f) = 0$ for every n, and we want to prove that $f = 0$.

Suppose on the contrary the existence of z such that $f(z) \neq 0$. There exists a bounded linear functional φ on $\mathcal{A}$ with $\varphi(f(z)) \neq 0$. The linearity of φ gives $c_n(\varphi \circ f) = \varphi(c_n(f)) = 0$ for every n. As $\varphi \circ f$ is continuous and complex-valued, a classical result (proved, for example, using Parseval's equality) gives $\varphi \circ f = 0$, which is a contradiction. ∎

PROPOSITION 2.7: *If the Fourier coefficients of f and g are summable, then it is also the case of fg.*

PROPOSITION 2.8: *If $f\colon S^1 \to \mathcal{A}$ is $C^{1+\alpha}$ then*

$$\|c_n(f)\| \leqslant C\frac{\|f\|_{1+\alpha}}{n^{1+\alpha}}$$

for some universal constant C.

The classical proof for complex valued functions can be found in [Kat68] and is easily adapted to this context (see also [Sar02, Lemma 3]).

3. Proof of Sarig's first main lemma under our weaker assumptions

The following lemma, which is the analogue of Sarig's first main lemma, is crucial to the proof of Theorem 1.1.

LEMMA 3.1 (First Main Lemma): *Under the assumptions of Theorem 1.1,*

$$\sum_{n=1}^{\infty} \|T_n - T_{n+1}\| < \infty.$$

As $(I - R(z))^{-1} = \sum T_n z^n$ (where we write $T_0 = I$), we have

$$A(z) := (1-z)(I - R(z))^{-1} = \sum (T_n - T_{n-1}) z^n.$$

Our strategy is to study A on S^1, and to see that its Fourier coefficients are summable. As

$$A(z) = \left(\frac{I - R(z)}{1-z}\right)^{-1},$$

A is well controlled on S^1 outside of any neighborhood of 1. Near 1, the problem comes from the eigenvalue $\lambda(z)$ of $R(z)$ closest to 1. To use Fourier series methods to control this eigenvalue, we must be able to extend $\lambda(z)$ to the whole circle S^1; that is why we will have to modify $R(z)$ and to construct a function $\tilde{R}(z)$ on S^1, whose spectrum will be "nice."

Proof of Lemma 3.1: We will write $\beta = 1 + \alpha$. We can assume $0 < \alpha < 1$, which amounts only to weakening the hypotheses.

STEP 1: $R(z)$ is $C^{1+\alpha}$ on $\overline{\mathbb{D}}$.

Proof: As $R(z) = \sum R_n z^n$ with $\sum_{k>n} \|R_k\| = O(1/n^{1+\alpha})$, we have $R_n = O(1/n^{1+\alpha})$, and the series defining R converges in norm on all $\overline{\mathbb{D}}$. Thus, R is continuous on $\overline{\mathbb{D}}$.

The sum $F(z) = \sum n R_n z^{n-1}$ converges also in norm on $\overline{\mathbb{D}}$, as $\sum_n n\|R_n\| = \sum_{n\geqslant 1} S_n < +\infty$ (where we write $S_n = \sum_{k\geqslant n} \|R_k\|$). Hence, this is the derivative of R and R is C^1 on $\overline{\mathbb{D}}$ (in the sense of Section 2.1).

What remains to be checked is that F is C^α. Let z and $z+h$ be two points in $\overline{\mathbb{D}}$; we estimate from above $\|F(z+h) - F(z)\|$. The Taylor–Lagrange inequality gives, for every $n \in \mathbb{N}$, $|(z+h)^n - z^n| \leqslant n|h|$. Let $N \in \mathbb{N}$. We have

$$\begin{aligned}|F(z+h)-F(z)| &\leqslant \sum_{n=1}^{N} n\|R_n\||(z+h)^{n-1}-z^{n-1}| + \sum_{n=N+1}^{+\infty} 2n\|R_n\| \\ &\leqslant \sum_{n=1}^{N} n^2(S_n - S_{n+1})|h| + \sum_{n=N+1}^{+\infty} 2n(S_n - S_{n+1}) \\ &\leqslant \sum_{n=1}^{N-1} 2nS_n|h| + \sum_{n=N+2}^{+\infty} 2S_n + 2(N+1)S_{N+1}.\end{aligned}$$

As $nS_n = O(1/n^\alpha)$, $\sum_{n=1}^{N-1} nS_n = O(1/N^{\alpha-1})$, while $\sum_{n=N+1}^{+\infty} S_n = O(1/N^\alpha)$ and $NS_N = O(1/N^\alpha)$. Hence, for some constants C and D (independent of N or h),

$$|F(z+h)-F(z)| \leqslant \frac{C}{N^{\alpha-1}}|h| + \frac{D}{N^\alpha}.$$

If we choose N close to $1/|h|$, we get a bound of the order of $|h|^\alpha$. ∎

Step 2: $(R(z) - R(1))/(z-1)$ can be continuously extended to S^1, and its Fourier coefficients are summable.

Proof: For $z \neq 1$,

$$\frac{R(z)}{1-z} = \frac{\sum R_n z^n}{1-z} = \sum_{n=1}^{+\infty}\left(\sum_{k=1}^{n} R_k\right) z^n.$$

Moreover, $R(1)/(1-z) = \sum_{n=1}^{+\infty}(\sum_{k=1}^{+\infty} R_k) z^n$, hence

$$\frac{R(z)-R(1)}{z-1} = \sum_{n=0}^{+\infty}\left(\sum_{k=n+1}^{+\infty} R_k\right) z^n.$$

The last sum converges in norm, since $\sum_{k=n+1}^{+\infty} \|R_k\| = O(1/n^{1+\alpha})$ is summable. This guarantees a continuous extension to 1. Moreover, the nth Fourier coefficient is $\sum_{k=n+1}^{+\infty} R_k$, which is summable. ∎

Step 3: Construction of a function $\tilde{R}$ on S^1, equal to R in a neighborhood of 1 in S^1, $C^{1+\alpha}$, whose spectrum consists in an isolated eigenvalue $\tilde{\lambda}(z)$ close to 1 together with a compact subset of $\mathbb{C}-\{1\}$, with $\tilde{\lambda}(z) \neq 1$ for $z \neq 1$. Furthermore, for any $\varepsilon > 0$, $\tilde{R}$ can be chosen such that $\forall z \in S^1$, $\|\tilde{R}(z) - R(1)\| < \varepsilon$.

Proof: We construct two candidates for $\tilde{R}$, U and V. The second one, i.e. V, will be the good one.

Fix some $\gamma > 0$, very small. Let $\varphi + \psi$ be a C^∞ partition of unity associated to the sets $\{\theta \in [0, \gamma)\}$ and $\{\theta \in (\gamma - \eta, \pi/2]\}$ where θ is the angle on the circle (for some very small $0 < \eta < \gamma$). We define $U(z) = \varphi(z)R(z) + \psi(z)R(e^{i\gamma})$ on $\{\theta \in [0, \pi/2]\}$: U is equal to R on $\{\theta \in [0, \gamma - \eta]\}$ and to $R(e^{i\gamma})$ on $\{\theta \in [\gamma, \pi/2]\}$. In particular, the spectrum of $U(z)$ will be "almost the same" as the spectrum of $R(1)$, if γ is small enough.

We define in the same way U on $\{\theta \in [-\pi/2, 0]\}$, equal to $R(e^{-i\gamma})$ on $\{\theta \in [-\pi/2, -\gamma]\}$ and to R on $\{\theta \in [-\gamma + \eta, 0]\}$.

Finally, we construct U on the remaining half-circle by symmetrizing, i.e. $U(e^{i(\pi/2+a)}) = U(e^{i(\pi/2-a)})$, to ensure that everything fits well.

Provided γ is small enough, there is a well defined eigenvalue close to 1 for every $U(z)$, depending continuously on z, which we denote by $\rho(z)$. The problem would be solved if $\rho(z) \neq 1$ for $z \neq 1$, which is not the case since $\rho(-1) = \rho(1) = 1$. Consequently, we have to perturb ρ a little. There exists a C^∞ function ν on $\{\theta \in [\pi/2, 3\pi/2]\}$ arbitrarily close to ρ. We can assume that ν does not take the value 1. On $\{\theta \in [\pi/2 + \eta, 3\pi/2 - \eta]\}$, we define $V(z) = \frac{\nu(z)}{\rho(z)} U(z)$: its eigenvalue close to 1 is $\nu(z) \neq 1$. Finally, we glue U and V together on $\{\theta \in [\pi/2, \pi/2 + \eta]\}$ and $\{\theta \in [3\pi/2 - \eta, 3\pi/2]\}$ with a partition of unity, as above. As the spectrum of $U(e^{i\pi/2}) = R(e^{i\gamma})$ does not contain 1 by aperiodicity, the gluing will not give an eigenvalue equal to 1 if we choose η small enough and ν close enough to ρ. ∎

STEP 4: $(\tilde{R}(z) - \tilde{R}(1))/(z - 1)$ can be continuously extended to S^1 and its Fourier coefficients are summable.

Proof: As $\tilde{R}(1) = R(1)$,

$$\frac{\tilde{R}(z) - \tilde{R}(1)}{z - 1} = \frac{\tilde{R}(z) - R(z)}{z - 1} + \frac{R(z) - R(1)}{z - 1}.$$

The first term is $C^{1+\alpha}$ outside of any neighborhood of 1, and zero on a neighborhood of 1. Thus, it is $C^{1+\alpha}$, which shows that its Fourier coefficients are summable by Proposition 2.8.

The coefficients of the second term $(R(z) - R(1))/(z - 1)$ are summable by Step 2, which gives the conclusion. ∎

STEP 5: Let $\tilde{P}(z)$ denote the spectral projection of $\tilde{R}(z)$ corresponding to its eigenvalue $\tilde{\lambda}(z)$ close to 1. Then $\tilde{P}(z)$ is $C^{1+\alpha}$, and its Fourier coefficients are summable.

Proof: The projection $\tilde{P}(z)$ can be written, for δ small enough (and independent of z if, in Step 3, ε was taken small enough),

$$\tilde{P}(z) = \frac{1}{2i\pi}\int_{|u-1|=\delta} \frac{1}{uI - \tilde{R}(z)} du.$$

We already know that $\tilde{R}$ is $C^{1+\alpha}$, which is also true of $uI - \tilde{R}$ for every u, and of $(uI - \tilde{R})^{-1}$ (with a uniform bound on its $C^{1+\alpha}$ norm) by Proposition 2.4. So, we can integrate to get a $C^{1+\alpha}$ function.

The summability of the coefficients is then a corollary of Proposition 2.8. ∎

STEP 6: The function $(\tilde{P}(z) - \tilde{P}(1))/(z-1)$ can be continuously extended to S^1 and its Fourier coefficients are summable.

Proof: The expression of the spectral projection used in Step 5 gives

$$\frac{\tilde{P}(z) - \tilde{P}(1)}{z-1} = \frac{1}{2i\pi}\int_{|u-1|=\delta} \frac{1}{uI - \tilde{R}(z)} \frac{\tilde{R}(z) - \tilde{R}(1)}{z-1} \frac{1}{uI - \tilde{R}(1)} du.$$

Let us fix u such that $|u-1| = \delta$. We have seen in Step 5 that the coefficients of $1/(uI - \tilde{R}(z))$ were summable. Moreover, Step 4 gives the summability of the coefficients of $(\tilde{R}(z) - \tilde{R}(1))/(z-1)$. As a consequence, the coefficients of the product

$$\frac{1}{uI - \tilde{R}(z)} \frac{\tilde{R}(z) - \tilde{R}(1)}{z-1}$$

are also summable.

To obtain the summability of the coefficients of $(\tilde{P}(z) - \tilde{P}(1))/(z-1)$, we just have to integrate with respect to u, since

$$c_n\Big(\frac{\tilde{P}(z) - \tilde{P}(1)}{z-1}\Big) = \frac{1}{2i\pi}\int_{|u-1|=\delta} c_n\Big(\frac{1}{uI - \tilde{R}(z)} \frac{\tilde{R}(z) - \tilde{R}(1)}{z-1} \frac{1}{uI - \tilde{R}(1)}\Big) du.$$

To conclude, we must get a uniform summable bound on the Fourier coefficients in the integral, i.e., we have to check that all previous estimates are uniform in u, which does not present any difficulty: the norms of $(uI - \tilde{R}(z))^{-1}$, for $|u-1| = \delta$ and $z \in S^1$, are bounded by compactness, and so are the $1+\alpha$ norms of $uI - \tilde{R}(z)$. Proposition 2.4 guarantees that the $1+\alpha$ norms of $(uI - \tilde{R})^{-1}$ are bounded by a constant independent of u. Proposition 2.8 gives that $c_n((uI - \tilde{R})^{-1}) = O(1/n^{1+\alpha})$ uniformly in u, which enables us to conclude. ∎

STEP 7: $(\tilde\lambda(z)-1)/(z-1)\to\mu$ as $z\to 1$ on S^1, where $\mu\neq 0$ satisfies $P(1)R'(1)P(1)=\mu P(1)$. Hence, the function $(z-1)/(\tilde\lambda(z)-1)$ is well defined. Moreover, its Fourier coefficients are summable.

Proof: For every $z\in S^1-\{1\}$, we have

$$\begin{aligned}\frac{1-\tilde\lambda(z)}{1-z}\tilde P(z) &= \frac{I-\tilde R(z)}{1-z}\tilde P(z)\\ &= \frac{\tilde R(1)-\tilde R(z)}{1-z}\tilde P(z)+(I-R(1))\frac{\tilde P(z)-\tilde P(1)}{1-z}.\end{aligned}\tag{1}$$

If we multiply on the left by $\tilde P(z)$ and let z go to 1, the right hand term tends to $P(1)R'(1)P(1)$ (because the other term tends to $P(1)(I-R(1))P'(1)=0$, and we can drop the tildes because $R=\tilde R$ in a neighborhood of 1). But $P(1)R'(1)P(1)$ can be written $\mu P(1)$, with $\mu\neq 0$ according to the hypotheses. We get

$$\frac{1-\tilde\lambda(z)}{1-z}\tilde P(z)\xrightarrow[z\to 1]{}\mu P(1).$$

Apply a bounded linear functional φ such that $\varphi(\tilde P(z))\neq 0$ for every z (which is possible: take $\varphi(P(1))\neq 0$, and then ε small enough in the construction of $\tilde R$). We obtain the convergence of $(1-\tilde\lambda(z))/(1-z)$ to μ.

Then, we show that the Fourier coefficients of the continuous function $(1-\tilde\lambda(z))/(1-z)$ are summable. In Equation (1), all terms on the right hand side have their coefficients summable, according to the previous steps. This remains true when we apply φ, i.e., $\frac{1-\tilde\lambda(z)}{1-z}\varphi(\tilde P(z))$ has summable coefficients. In the same way, $\tilde P(z)$ has summable coefficients, and $\varphi(\tilde P(z))$ too. But this is a complex function, everywhere nonzero, so the Wiener lemma gives that its inverse $1/\varphi(\tilde P(z))$ has also summable coefficients. Multiplying, we obtain the summability of the coefficients of $(1-\tilde\lambda(z))/(1-z)$.

Using once more the Wiener lemma (since $(1-\tilde\lambda(z))/(1-z)$ is everywhere nonzero by construction of $\tilde R$), we get the conclusion. ∎

STEP 8: $(z-1)(\tilde R(z)-I)^{-1}$ can be continuously extended to 1, and its Fourier coefficients are summable.

Proof: Let $\tilde Q(z)$ denote the spectral projection $I-\tilde P(z)$. Then, for every $z\neq 1$,

$$\begin{aligned}(1-z)(I-\tilde R(z))^{-1} &= \frac{1-z}{1-\tilde\lambda(z)}\tilde P(z)+(1-z)(I-\tilde R(z))^{-1}\tilde Q(z)\\ &= \frac{1-z}{1-\tilde\lambda(z)}\tilde P(z)+(1-z)(I-\tilde R(z)\tilde Q(z))^{-1}\tilde Q(z);\end{aligned}\tag{2}$$

$I - \tilde{R}(z)\tilde{Q}(z)$ is everywhere invertible on S^1 and is $C^{1+\alpha}$ (this is true for $\tilde{Q}$ because $\tilde{P}$ is $C^{1+\alpha}$ by Step 5 and $\tilde{Q} = I - \tilde{P}$). Proposition 2.4 gives that its inverse is $C^{1+\alpha}$, hence its coefficients are summable, which remains true when it is multiplied by $\tilde{Q}(z)$ which is $C^{1+\alpha}$.

To conclude, we have to show that $\frac{1-z}{1-\tilde{\lambda}(z)}\tilde{P}(z)$ has summable Fourier coefficients. We already know this for $\tilde{P}(z)$ (Step 5) and $(1-z)/(1-\tilde{\lambda}(z))$ (Step 7). As the product of functions with summable coefficients also has summable coefficients, this enables us to conclude. ∎

STEP 9: $(z-1)(R(z)-I)^{-1}$ can be continuously extended on all $\overline{\mathbb{D}}$, and its Fourier coefficients (on S^1) are summable.

Proof: We have already proved that $(z-1)(\tilde{R}(z)-I)^{-1}$ can be continuously extended to 1 on S^1. As R and $\tilde{R}$ coincide in a neighborhood of 1, it shows that $(z-1)(R(z)-I)^{-1}$ can be continuously extended to 1 on S^1. Since we are interested in an extension to the whole disc $\overline{\mathbb{D}}$, we must check that the previous arguments work well on $\overline{\mathbb{D}}$, which does not present any difficulty: dropping the tildes, Equation (1) is valid for z in a neighborhood of 1 in $\overline{\mathbb{D}}$, whence $(1-\lambda(z))/(1-z)$ tends to μ when $z \to 1$ in $\overline{\mathbb{D}}$; using Equation (2), this gives the desired extension to 1.

On S^1,

$$(z-1)(R(z)-I)^{-1} = (z-1)(\tilde{R}(z)-I)^{-1} \cdot (\tilde{R}(z)-I)(R(z)-I)^{-1}.$$

Step 8 shows that $(z-1)(\tilde{R}(z)-I)^{-1}$ has its Fourier coefficients summable. Moreover, $(\tilde{R}(z)-I)(R(z)-I)^{-1}$ is $C^{1+\alpha}$ outside of any neighborhood of 1, and equal to I on a neighborhood of 1. Hence, it is $C^{1+\alpha}$ on S^1 and has its coefficients summable. To conclude, we apply Proposition 2.7 which tells that the product of functions with summable Fourier coefficients still has summable coefficients. ∎

STEP 10: $\sum \|T_{n+1} - T_n\| < +\infty$.

Proof: Let $A(z) = (1-z)(I-R(z))^{-1}$. For $|z| < 1$, $A(z) = \sum (T_n - T_{n-1})z^n$, so, when $r < 1$,

$$T_n - T_{n-1} = \frac{1}{2\pi r^n}\int_0^{2\pi} A(re^{i\theta})e^{-in\theta}\mathrm{d}\theta.$$

As A can be continuously extended on $\overline{\mathbb{D}}$, we can let r tend to 1 and obtain $T_n - T_{n-1} = c_n(A)$. But we have already proved in the previous step that the coefficients of A were summable. ∎

4. Proof of the main theorem

Once we have obtained the first main lemma, the rest of the proof of Theorem 1.1 is very similar to Sarig's arguments. We will reproduce here only the parts which need to be modified to fit the current context.

To obtain the asymptotic expansion of T_n, the main idea is to write

$$T(z) = \frac{1}{1-z} S(z)^{-1}, \text{ where } S(z) = \frac{I - R(z)}{1-z},$$

to decompose $S = S_B + (S - S_B)$ where $S_B(z)$ is a well controlled polynomial and $S - S_B$ a small remainder, and to make a perturbative development of S^{-1} using this decomposition. This amounts to writing

$$(3) \quad T(z) = \frac{1}{1-z} S_B^{-1} + \frac{1}{1-z} S_B^{-1}(S_B - S) S_B^{-1} + \frac{1}{1-z}[S_B^{-1}(S_B - S)]^2 S^{-1}.$$

The term $\frac{1}{1-z} S_B^{-1}(z)$ will give the contribution $\frac{1}{\mu} P$ in the expansion of T_n, while the second one will give the term $\frac{1}{\mu^2} \sum_{k=n+1}^{+\infty} P_k$ and the third one will give the error term.

Write S_B as $(I - R_B(z))/(1-z)$ where

$$R_B(z) = \sum_{n=1}^{N} z^n R_n + \sum_{n=N+1}^{\infty} R_n + (z-1) \sum_{n=N+1}^{\infty} n R_n.$$

This expression is such that $R_B(1) = R(1)$ and $R_B'(1) = R'(1)$. For Equation (3) to be valid for $z \in \overline{\mathbb{D}} - \{1\}$, we have to check that S_B is invertible, i.e., $I - R_B$ is invertible. Following [Sar02, Proof of the Second Main Lemma], this is implied by the first main lemma proved in the previous section as soon as N is large enough.

We recall without proof Sarig's second main lemma, which is a consequence of the first main lemma.

LEMMA 4.1 (Second Main Lemma): *In the setting of Theorem 1.1, if P is the eigenprojection of $R(1)$ at 1 and μ is given by $PR'(1)P = \mu P$, then there exists $R_B \colon \mathbb{C} \to \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ with the following properties:*

1. *R_B is holomorphic, $R_B(1) = R(1)$ and $R_B'(1) = R'(1)$.*
2. *$\frac{R(1) - R_B}{1-z}$ and $\frac{1}{1-z}[\frac{R(1)-R_B}{1-z} - R'(1)]$ are polynomials in z.*
3. *$I - R_B(z)$ has a bounded inverse in $\mathrm{Hom}(\mathcal{L}, \mathcal{L})$ for every $z \in \overline{\mathbb{D}} - \{1\}$.*
4. *$\forall z \in \mathbb{D}$, $(\frac{I-R_B}{1-z})^{-1} = \frac{1}{\mu} P + (1-z) \sum_{n \geqslant 0} z^n A_n$ where $\|A_n\| = O(\kappa^n)$ for some $0 < \kappa < 1$.*

Equation (3) together with the following lemma (extending Sarig's Lemma 7 to the case $1 < \beta \leqslant 2$ and sharpening it for $\beta > 2$) gives Theorem 1.1.

LEMMA 4.2: *Under the assumptions of Theorem 1.1, if P is the eigenprojection of $R(1)$ at 1 and μ is given by $PR'(1)P = \mu P$, then*

1. $\frac{1}{1-z}S_B^{-1} = \frac{1}{\mu}\sum_{n\geqslant 0} z^n(P+\varepsilon_n)$ *where* $\|\varepsilon_n\| = O(\kappa^n)$ *for some* $0<\kappa<1$.
2. $\frac{1}{1-z}S_B^{-1}(S_B - S)S_B^{-1} = \frac{1}{\mu^2}\sum_{n\geqslant 0} z^n(\sum_{k>n} P_k + \varepsilon'_n)$ *where* $\|\varepsilon'_n\| = O(1/n^\beta)$ *and* $P_n = \sum_{l>n} PR_lP$.
3. $\frac{1}{1-z}[S_B^{-1}(S_B - S)]^2 S^{-1} = \sum_{n\geqslant 0} z^n E_n$ *where*

$$\|E_n\| = \begin{cases} O(1/n^\beta) & \text{if } \beta > 2,\\ O(\log n/n^2) & \text{if } \beta = 2,\\ O(1/n^{2\beta-2}) & \text{if } 2>\beta>1.\end{cases}$$

To prove the estimates in Lemma 4.2, we will need some results on the convolution of sequences. If a_n and b_n are sequences, put $c_n = \sum_{k+l=n} a_k b_l$. We write $c = a\star b$. For the next lemma, see [Rog73, Theorem 2].

LEMMA 4.3: *If $a_n = O(1/n^\alpha)$ and $b_n = O(1/n^\beta)$ for some $\alpha \leqslant \beta \in \mathbb{R}$, then*

$$(a\star b)_n = \begin{cases} O(1/n^\alpha) & \text{if } \beta > 1,\\ O(\log n/n^\alpha) & \text{if } \beta = 1,\\ O(1/n^{\alpha+\beta-1}) & \text{if } \beta < 1.\end{cases} \tag{4}$$

In particular, for $\alpha > 1$ or $\beta > 1$ (without assuming $\alpha \leqslant \beta$), $(a\star b)_n = O(1/n^\alpha) + O(1/n^\beta)$.

Proof: We prove the result for $\beta < 1$, the other cases being treated in the same way. If $c_n = \sum_{k=0}^n a_k b_{n-k}$, we have

$$|c_n| \leqslant \Big(\max_{0\leqslant k<n/2} |b_{n-k}|\Big) \sum_{0\leqslant k<n/2} |a_k| + \Big(\max_{n/2\leqslant k\leqslant n} |a_k|\Big) \sum_{n/2\leqslant k\leqslant n} |b_{n-k}|.$$

The sums can be estimated from above by $O(1/n^{\alpha-1})$ and $O(1/n^{\beta-1})$ respectively, while the maxima are $O(1/n^\beta)$ and $O(1/n^\alpha)$. This gives the conclusion. ∎

Let us state another lemma which will be useful later in Section 5. The proof is exactly the same, cutting the sum in the middle, and using $\sum_{0\leqslant k<n/2} |a_k| = O((\log n)^{u+1}/n^{\alpha-1})$ (and the analogous estimate for $\sum_{0\leqslant k<n/2} |b_k|$).

LEMMA 4.4: *If $a_n = O((\log n)^u/n^\alpha)$ and $b_n = O((\log n)^v/n^\beta)$ for some $\alpha \leqslant 1$, $\beta \leqslant 1$ and $u, v \geqslant 0$, then $(a\star b)_n = O((\log n)^{u+v+1}/n^{\alpha+\beta-1})$.*

In fact, the $(\log n)^{u+v+1}$ can be replaced by $(\log n)^{u+v}$ whenever $\alpha < 1$ and $\beta < 1$, but we will not need it.

We recall a notation used by Sarig: if c_n is a real sequence and $F(z) = \sum F_n z^n$ a formal series with coefficients in a Banach algebra, write $F \in \Re(c_n)$ if $\|F_n\| = O(c_n)$. Abusing slightly notation, we write $\Re(1/n^\alpha)$ instead of $\Re(1/(n+1)^\alpha)$, discarding the problem for $n = 0$.

To prove Lemma 4.2, we will first show that $S(z)^{-1} \in \Re(1/n^\beta)$. In his main theorem, Sarig obtains $\lfloor\beta\rfloor$ instead of β since he proves only that $S^{-1} \in \Re(1/n^{\lfloor\beta\rfloor})$; we can avoid this loss of information with the help of Lemma 4.5, which should replace the general result on Banach spaces Sarig uses and will give indeed $S(z)^{-1} \in \Re(1/n^\beta)$.

LEMMA 4.5: *Let $\mathcal{A}$ be a Banach algebra and suppose that $F(z) = \sum F_k z^k$ where $\|F_k\| = O(1/n^\beta)$ for some $\beta > 1$. Suppose furthermore that for every $z \in \mathbb{D}$, $I + F(z)$ is invertible, and that $(I + F(z))^{-1} = \sum z^k G_k$. If $\sum \|G_k\| < \infty$, then $\|G_k\| = O(1/n^\beta)$.*

Let us explain how to derive $S(z)^{-1} \in \Re(1/n^\beta)$ from this lemma. Following Sarig, we use the identity $S^{-1} = S_B^{-1}(I + (S - S_B)S_B^{-1})^{-1}$. In order to get the result for S^{-1}, it is enough to prove that $(I + (S - S_B)S_B^{-1})^{-1} \in \Re(1/n^\beta)$ since we already know that $S_B^{-1} \in \Re(\kappa^n)$ for some $\kappa < 1$ (Lemma 4.1). Note that $(I + (S - S_B)S_B^{-1})^{-1} = S_B S^{-1} = I + (S_B - S)S^{-1}$ has summable coefficients because this is the case for S^{-1} (Lemma 3.1) and for $S_B - S$ (because

$$S_B - S = \frac{R(1) - R_B}{1 - z} - \frac{R(1) - R}{1 - z},$$

the first term being a polynomial and the second one in $\Re(1/n^\beta)$). Moreover, Lemma 4.1 gives that $I + (S - S_B)S_B^{-1} \in \Re(1/n^\beta)$ (since $S - S_B \in \Re(1/n^\beta)$ and $S_B^{-1} \in \Re(\kappa^n)$). Consequently, Lemma 4.5 applied to $F = (S - S_B)S_B^{-1}$ yields $(I + (S - S_B)S_B^{-1})^{-1} \in \Re(1/n^\beta)$, which gives the conclusion.

Proof of Lemma 4.5: Set $c_n = \sum_{i+k=n} \|G_i\|\|G_k\|$. As $\|G_n\|$ is summable, this is also the case for c_n. We will write f_n and g_n respectively for $\|F_n\|$ and $\|G_n\|$.

Equating coefficients in $[(I + F)^{-1}]' = -(I + F)^{-1}F'(I + F)^{-1}$ gives

$$ng_n \leqslant \sum_{i+j+k=n} g_i j f_j g_k = \sum_{j=0}^{n} j f_j c_{n-j} \leqslant (\sup j f_j) \sum c_k < +\infty.$$

Consequently, $g_n = O(1/n)$. Moreover, we have $(ng_n) \leqslant c \star (jf_j)$, with $jf_j = O(1/j^{\beta-1})$.

We show that $g_n = O(1/n^{1+\delta})$ for some $\delta > 0$. It is enough to prove this when $1 < \beta < 2$. As $g_n = O(1/n)$, $c = g \star g$ is such that $c_n = O(\log n/n)$ according to Lemma 4.3. Hence, $c_n = O(1/n^\gamma)$ for every $\gamma < 1$. Lemma 4.3 again gives $c \star (jf_j) = O(1/n^{\gamma+\beta-1-1})$, and $g_n = O(1/n^{\gamma+\beta-1})$. As $\beta - 1 > 0$ and γ can be chosen arbitrarily close to 1, we can impose $\gamma + \beta - 1 > 1$, which gives the conclusion.

Assume that $g_n = O(1/n^\eta)$ for some $\eta > 1$. As $c = g \star g$, we get $c_n = O(1/n^\eta)$, whence $c \star (jf_j) = O(1/n^\eta) + O(1/n^{\beta-1})$ once again by Lemma 4.3. As $(ng_n) \leqslant c \star (jf_j)$, this implies $g_n = O(1/n^{\eta+1}) + O(1/n^\beta)$.

We already know that $g_n = O(1/n^{1+\delta})$ for some $\delta > 0$. Using the previous paragraph, we show by induction that, for any integer k such that $1+\delta+k < \beta$, we have $g_n = O(1/n^{1+\delta+k+1}) + O(1/n^\beta)$. For the largest k such that $1+\delta+k < \beta$, we obtain $g_n = O(1/n^\beta)$. ∎

From this point on, we can strictly follow Sarig's proof, replacing his estimates $O(1/n^{\lfloor\beta\rfloor})$ by $O(1/n^\beta)$. This way, we can obtain Estimates (1) and (2) in Lemma 4.2. However, the proof of Estimate (3) has to be adapted.

Proof of Estimate (3) *in Lemma 4.2:* As in [Sar02, Step 4 of the proof of Theorem 1] (in fact, due to a misprint, this step is called Step 3 in Sarig's article), write $G(z) = S_B^{-1}(z)(S_B(z) - S(z)) = \sum G_k z^k$. As $S_B - S \in \Re(1/n^\beta)$ and $S_B^{-1} \in \Re(\kappa^n)$ for some $\kappa < 1$ (Lemma 4.1), we obtain that $G \in \Re(1/n^\beta)$. Moreover, $\sum G_k = 0$ (because $S_B(1) = S(1)$), hence $\frac{1}{1-z}G(z) = -\sum z^n \sum_{k>n} G_k$ and consequently $\frac{G(z)}{1-z} \in \Re(1/n^{\beta-1})$ (see [Sar02, Step 4 of the proof of Theorem 1] for more details).

Setting $E = \frac{1}{1-z}G^2 S^{-1} = \sum z^n E_n$, we want to estimate the coefficients E_n of E. We have

$$E' = \Big(\frac{G}{1-z}\Big)^2 S^{-1} + \Big[\Big(\frac{G}{1-z}\Big)G' + G'\Big(\frac{G}{1-z}\Big)\Big]S^{-1} + G\Big(\frac{G}{1-z}\Big)(S^{-1})'.$$

We know that $\frac{G}{1-z} \in \Re(1/n^{\beta-1})$, $G' \in \Re(1/n^{\beta-1})$ and $(S^{-1})' \in \Re(1/n^{\beta-1})$ (since $S^{-1} \in \Re(1/n^\beta)$). Lemma 4.3 on convolutions gives

$$\Big(\frac{G}{1-z}\Big)^2 \in \begin{cases} \Re(1/n^{\beta-1}) & \text{if } \beta > 2 \\ \Re(\log n/n) & \text{if } \beta = 2 \\ \Re(1/n^{2\beta-3}) & \text{if } \beta < 2 \end{cases}$$

and we have analogous estimates for the other terms in E'. Integrating, we get the desired estimates for E_n. This concludes the proof of Lemma 4.2 and, with it, of Theorem 1.1. ∎

5. Higher order terms in T_n

To obtain an asymptotic expansion of T_n, we have used the perturbation series of $T(z)$ up to order 2 given in Equation (3). While this is completely satisfactory when $\beta > 2$, since it implies an estimate of T_n up to order $O(1/n^\beta)$, it is possible to obtain a more precise estimate when $1 < \beta \leqslant 2$, using more terms in the same kind of perturbation series.

Fix $N \in \mathbb{N}$. Then

$$T(z) = \frac{1}{1-z} \sum_{k=0}^{N-1} [S_B^{-1}(S_B - S)]^k S_B^{-1} + \frac{1}{1-z}[S_B^{-1}(S_B - S)]^N S^{-1}. \tag{5}$$

To estimate the right hand side terms, we will use the fact that, if $G(z) = S_B^{-1}(S_B - S)$, then $G \in \Re(1/n^\beta)$ and $G/(1-z) \in \Re(1/n^{\beta-1})$, as we have seen in the proof of Estimate (3) in Lemma 4.2.

Lemma 5.1: *Let $G(z) = \sum G_n z^n$ be a formal series with coefficients in a Banach algebra $\mathcal{A}$, such that $G(z) \in \Re(1/n^\beta)$ and $\frac{G(z)}{1-z} \in \Re(1/n^{\beta-1})$ for some $\beta > 1$. Then, for any $p \in \mathbb{N}$, there exists a constant C such that for any $H_1, \dots, H_{p-1} \in \mathcal{A}$,*

$$\left\| \left(\frac{G(z) H_1 G(z) \dots H_{p-1} G(z)}{1-z} \right)_n \right\| \leqslant C \|H_1\| \cdots \|H_{p-1}\| \cdot \begin{cases} 1/n^\beta & \text{if } \beta < p(\beta-1) \\ (\log n)/n^\beta & \text{if } \beta = p(\beta - 1) \\ 1/n^{p(\beta-1)} & \text{if } \beta > p(\beta-1) \end{cases}$$

(the notation $(\ )_n$ denotes the coefficient of z^n in the formal series between the braces).

Proof of Lemma 5.1: In this proof, if $L(z) = \sum L_n z^n$ is a power series with coefficients in $\mathcal{A}$, we shall write $\|L\|$ or $\|L(z)\|$ for the sequence $(\|L_n\|)_{n \in \mathbb{N}}$. When $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ are sequences of nonnegative numbers such that $\forall n, a_n \leqslant b_n$, we shall write $a \leqslant b$. With these notations, if $L_0(z)$ and $L_1(z)$ are power series, then $\|L_0 L_1\| \leqslant \|L_0\| \star \|L_1\|$ where $\star$ means "convolution of sequences". Write also

$$F_{\beta,p}(n) = \begin{cases} 1/n^\beta & \text{if } \beta < p(\beta-1), \\ (\log n)/n^\beta & \text{if } \beta = p(\beta-1), \\ 1/n^{p(\beta-1)} & \text{if } \beta > p(\beta-1). \end{cases}$$

We will prove the lemma by induction on p. The crucial argument to eliminate the non-commutativity problem will be the following: if $L_0(z), \dots, L_r(z)$ are power series for some $r < p$, consider

$$L_0(z) G(z) L_1(z) \cdots L_{r-1}(z) G(z) L_r(z)/(1-z).$$

The induction hypothesis gives

$$\left\|\frac{G(z)H_1G(z)\dots H_{r-1}G(z)}{1-z}\right\| \leqslant CF_{\beta,r}(n)\|H_1\|\cdots\|H_{r-1}\| \quad \text{for some constant } C.$$

Then

$$\left\|\frac{L_0(z)G(z)L_1(z)\cdots L_{r-1}(z)G(z)L_r(z)}{1-z}\right\| \leqslant CF_{\beta,r}\star\|L_0\|\star\cdots\star\|L_r\|. \tag{6}$$

Indeed, in $L_0(z)G(z)L_1(z)\cdots L_{r-1}(z)G(z)L_r(z)/(1-z)$, expanding the $L_i(z)$s in series of z gives terms of the form $H_0G(z)H_1\cdots H_{r-1}G(z)H_r/(1-z)$, to which the induction hypothesis can be applied.

We prove Lemma 5.1 by induction on p. The result is part of the hypotheses when $p=1$ and easily follows from the proof of Estimate (3) in Lemma 4.2 for $p=2$ (the same argument works when a term H_1 is inserted). Assume $p>2$.

If $\beta<(p-1)(\beta-1)$, the induction gives

$$\left\|\frac{G(z)H_1G(z)\cdots H_{p-2}G(z)}{1-z}\right\| \leqslant \frac{C}{n^\beta}\|H_1\|\cdots\|H_{p-2}\|.$$

As $G(z)\in\Re(1/n^\beta)$, a multiplication on the right by H_{p-1} and a convolution with $G(z)$ using Lemma 4.3 give the desired result. Thus, we can assume that $\beta\geqslant(p-1)(\beta-1)$. As $p\geqslant 3$, this implies in particular that $\beta\leqslant 2$.

Differentiating $p-1$ times $F(z)=G(z)H_1G(z)\cdots H_{p-1}G(z)/(1-z)$ gives, for some constants $C_{i,i_1,\dots,i_p}$,

$$F^{(p-1)}(z)=\sum_{\substack{i+i_1+\cdots+i_p=p-1\\ i_1,\dots,i_p\geqslant 0}} C_{i,i_1,\dots,i_p}\frac{G(z)^{(i_1)}H_1\cdots H_{p-1}G(z)^{(i_p)}}{(1-z)^{i+1}}$$

where $G(z)^{(k)}$ denotes the function $G(z)$ differentiated k times.

Let $T(z)=T_{i,i_1,\dots,i_p}(z)$ be a term of this sum. Since $i+\sum_{j=1}^p i_j=p-1$, there are $i+r$ factors $G(z)$ which are not really differentiated, i.e., for which $i_j=0$, for some $r\geqslant 1$. Giving a factor $1/(1-z)$ to i of these factors, we write

$$T(z)=\frac{L_0(z)G(z)L_1(z)\cdots G(z)L_r(z)}{1-z}$$

where the factors $L_i(z)$ are products of factors of the form H_j, or $G^{(i_j)}$ with $i_j\geqslant 1$, or $G(z)/(1-z)$. Equation (6) then gives

$$\begin{aligned}\|T(z)\| &\leqslant CF_{\beta,r}\star\|L_0\|\star\cdots\star\|L_r\|\\ &\leqslant C\|H_1\|\cdots\|H_p\|\cdot F_{\beta,r}\star\|G^{(i_1)}\|\star\cdots\star\|G^{(i_p)}\|\star\left\|\frac{G(z)}{1-z}\right\|^{\star i}\end{aligned} \tag{7}$$

where $\|G^{(i_j)}\|$ is written in the product on the right only when $i_j > 0$.

Let us distinguish 3 cases:

1. If $(p-1)(\beta-1) \leqslant \beta < p(\beta-1)$.

   For every term $T(z) = T_{i,i_1,\dots,i_p}(z)$, in which $i+r$ factors $G(z)$ are not differentiated, we have $r < p$ (otherwise, $i + r \leqslant p$ would imply $i = 0$, and nothing would be differentiated), whence $r(\beta-1) \leqslant \beta$, and we are in the second or third case of the induction.

   We are in fact in the second case only if $r = p-1$, which means that $i = 0$ and that one i_j is equal to $p-1$, the other ones being 0. In Equation (7),
   $$F_{\beta,p-1}(n) = O\Big(\frac{\log n}{n^{(p-1)(\beta-1)}}\Big) = O(1/n^{\gamma}) \quad \text{for some } \gamma > 1$$
   (since $p(\beta-1) > \beta$ implies $(p-1)(\beta-1) > 1$). When convolving with $\|G(z)^{(p-1)}\| \in \Re(1/n^{\beta-p+1})$ (where $\beta - p + 1 \leqslant 1$ since $\beta \leqslant 2$), Lemma 4.3 gives an expression in $\Re(1/n^{\beta-p+1})$.

   Consider now another term $T_{i,i_1,\dots,i_p}$ with $r < p-1$. As $(p-1)(\beta-1) \leqslant \beta$, we obtain $(p-2)(\beta-1) \leqslant 1$, hence $r(\beta-1) \leqslant 1$. As $F_{\beta,r}(n) = O(1/n^{r(\beta-1)})$ and $\|G^{(i_j)}\| \in \Re(1/n^{\beta-i_j})$ (with $\beta - i_j \leqslant 1$ when $i_j > 0$) and $\frac{G(z)}{1-z} \in \Re(1/n^{\beta-1})$ (with $\beta - 1 \leqslant 1$), Lemma 4.4 applied $p-r$ times ensures that the right hand side of Equation (7) is in $\Re((\log n)^u/n^{\nu})$ for some $u \geqslant 0$ and
   $$\nu = \sum_{i_j>0} (\beta - i_j) + i(\beta-1) + r(\beta-1) - (p-r) = p(\beta-1) - (p-1).$$
   Since $p(\beta-1) > \beta$, we have in fact $\nu > \beta - (p-1)$, which gives an upper estimate in $C/n^{\beta-(p-1)}$.

   Summing all terms, we obtain
   $$\|F(z)^{(p-1)}\| \leqslant C\|H_1\|\cdots\|H_{p-1}\|\frac{1}{n^{\beta-p+1}}.$$
   Integrating then $p-1$ times, we get the desired result.

2. If $\beta = p(\beta-1)$.

   Here, we have $\beta < 2$, whence $\beta - i_j < 1$ and $\beta - 1 < 1$, which implies that we will be able to use Lemma 4.3 instead of Lemma 4.4.

   We use the same reasoning as in case 1. In the case where $r = p-1$, we have $F_{\beta,p-1}(n) = O(1/n)$, which implies that the right hand side of Equation (7) consists in $(1/n) \star (1/n^{\beta-p+1})$, which is in $\Re((\log n)/n^{\beta-p+1})$ according to Lemma 4.3.

   When $r < p-1$, we use $p-r-1$ times Lemma 4.3 to estimate
   $$\|G^{(i_1)}\| \star \cdots \star \|G^{(i_p)}\| \star \|G(z)/(1-z)\|^{\star i},$$

since all the exponents are < 1, to obtain that it is in $\Re(1/n^\delta)$ where $\delta = (p-r)(\beta-1)-p+2$. Convolving once more with $F_{\beta,r}(n) = 1/n^{r(\beta-1)}$, we get a term in $\Re(1/n^\nu)$ where $\nu = p(\beta-1)-p+1 = \beta-p+1$. Summing and integrating gives the result.

3. If $\beta > p(\beta-1)$.
   We do not need to distinguish the term where $r = p-1$ any more: in all terms, all exponents are <1. A convolution gives terms in $\Re(1/n^{p(\beta-1)-p+1})$ which gives the result after integration. ∎

LEMMA 5.2 (Control of the error term): *We have*

$$\frac{1}{1-z}[S_B^{-1}(S_B-S)]^N S^{-1} \in \begin{cases} \Re(1/n^\beta) & \text{if } N(\beta-1) > \beta, \\ \Re((\log n)/n^\beta) & \text{if } N(\beta-1) = \beta, \\ \Re(1/n^{N(\beta-1)}) & \text{if } N(\beta-1) < \beta. \end{cases}$$

Proof: Set $G(z) = S_B^{-1}(S_B - S)$. Then the conditions of Lemma 5.1 are satisfied (this has been checked in the proof of Estimate (3) in Lemma 4.2). Consequently, the lemma with $p = N$ and $H_1 = \cdots = H_{p-1} = 1$ gives estimates on $G(z)^N/(1-z)$. As we already know that $S(z)^{-1} \in \Re(1/n^\beta)$ with $\beta > 1$, another convolution enables us to conclude. ∎

To use this result, it remains only to study the terms in the sum in Equation (5), i.e., the perturbative terms for $k = 0, \dots, N-1$. The method used in Sarig to estimate the first term still works: estimating S_B^{-1} by $\frac{1}{\mu}P$ gives an exponentially decreasing error, which does not matter. Moreover, we can estimate $S_B - S$ by $\sum_{k=0}^\infty (1-z^k)\sum_{n=k+1}^\infty R_k$. A formal multiplication gives finally the desired terms. More precisely, the following lemma is valid.

LEMMA 5.3 (Estimates on the perturbative terms): *For any $k \in \mathbb{N}^*$, writing $P_n = \sum_{l>n} PR_lP$, we have*

$$\frac{1}{1-z}[S_B^{-1}(S_B-S)]^k S_B^{-1} = \frac{1}{\mu^{k+1}}\frac{1}{1-z}\left[\sum_{n=0}^\infty (1-z^n)P_n\right]^k + E(z)$$

where $E \in \Re(1/n^\beta)$.

Proof: We already know that $S_B - S \in \Re(1/n^\beta)$ and S_B^{-1} also.

We write

$$\begin{aligned} S_B - S &= \left[R'(1) - \frac{R(1)-R}{1-z}\right] + \left[\frac{R(1)-R_B}{1-z} - R'(1)\right] \\ &= \left[\sum_{n=0}^\infty (1-z^n)\sum_{l=n+1}^\infty R_l\right] + (1-z)B(z) \end{aligned}$$

where $B(z)$ is a polynomial, according to Lemma 4.1. Moreover, the same lemma gives that

$$S_B^{-1} = \frac{1}{\mu}P + (1-z)A(z)$$

for some $A(z) \in \Re(\kappa^n)$ with $\kappa < 1$.

We multiply these expressions to get $\frac{1}{1-z}[S_B^{-1}(S_B - S)]^k S_B^{-1}$ and we expand the product. If we choose a term $(1-z)A(z)$ or $(1-z)B(z)$, we use it to simplify the $\frac{1}{1-z}$, and all the other terms are $\Re(1/n^\beta)$, which gives after convolution still a $\Re(1/n^\beta)$. The remaining term gives the expression stated in the lemma. ∎

Gathering the results of Lemma 5.2 and Lemma 5.3, together with (5), we obtain

THEOREM 5.4: *Under the hypotheses of Theorem 1.1, we have, for any $N \in \mathbb{N}$, writing $P_m = \sum_{k>m} PR_kP$,*

$$\sum T_n z^n = \frac{1}{1-z}\frac{1}{\mu}P + \sum_{k=1}^{N-1} \frac{1}{\mu^{k+1}}\frac{1}{1-z}\left[\sum_{m=0}^{\infty}(1-z^m)P_m\right]^k + E(z) \tag{8}$$

where

$$E(z) \in \begin{cases} \Re(1/n^\beta) & \text{if } N(\beta-1) > \beta, \\ \Re((\log n)/n^\beta) & \text{if } N(\beta-1) = \beta, \\ \Re(1/n^{N(\beta-1)}) & \text{if } N(\beta-1) < \beta. \end{cases}$$

When $\beta > 2$, this theorem adds nothing to Theorem 1.1, since the terms for $2 \leqslant k \leqslant N-1$ are already in $O(1/n^\beta)$. However, when $3/2 < \beta < 2$, the result obtained by taking $N = 3$ is more precise than Theorem 1.1, since we get an error term in $O(1/n^\beta)$ instead of $O(1/n^{2\beta-2})$. In fact, for any $\beta > 1$, it is possible to choose N such that $N(\beta-1) > \beta$, which implies that the expansion of T_n with N terms gives an estimate with an error term in $O(1/n^\beta)$ (but taking N still larger is useless, since the new terms will be in $O(1/n^\beta)$). In particular, if $Pf = 0$, we obtain $T_nf = O(1/n^\beta)$, which is exactly Theorem 1.2.

To obtain a sharp asymptotic expansion for T_n, it remains only to expand the middle terms in Equation (8). We give, for example, the theorem that we obtain for $N = 3$:

THEOREM 5.5: *Under the hypotheses of Theorem 1.1, we have*

$$T_n = \frac{1}{\mu}P + \frac{1}{\mu^2}\sum_{k=n+1}^{+\infty} P_k + \frac{1}{\mu^3}\Bigg(\sum_{k,l>n} P_kP_l - \sum_{\substack{k,l\leqslant n \\ k+l>n}} P_kP_l\Bigg) + E_n$$

where $E_n \in \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ *satisfy*

$$\|E_n\| = \begin{cases} O(1/n^\beta) & \text{if } \beta > 3/2, \\ O((\log n)/n^\beta) & \text{if } \beta = 3/2, \\ O(1/n^{3(\beta-1)}) & \text{if } 3/2 > \beta > 1. \end{cases}$$

We give for completeness the next term in the expansion: after tedious calculations, we find that it is (up to the factor $1/\mu^4$)

$$\sum_{k,l,m>n} - \sum_{\substack{0<k,l\leqslant n\\ k+l>n\\ m>n}} - \sum_{\substack{0<k,m\leqslant n\\ k+m>n\\ l>n}} - \sum_{\substack{0<l,m\leqslant n\\ l+m>n\\ k>n}} - \sum_{\substack{0<k,l,m\leqslant n\\ k+l>n\\ k+m>n}}$$
$$- \sum_{\substack{0<k,l,m\leqslant n\\ l+k>n\\ l+m>n}} - \sum_{\substack{0<k,l,m\leqslant n\\ m+k>n\\ m+l>n}} + \sum_{\substack{0<k,l,m\leqslant n\\ k+l>n\\ k+m>n\\ l+m>n}} + \sum_{\substack{k+l\leqslant n\\ k+m\leqslant n\\ l+m\leqslant n\\ k+l+m>n}} P_k P_l P_m.$$

6. Application to Markov maps

6.1 Definition of Markov maps. The definitions and results of this section are for the main part contained in [Aar97].

A Markov map is a nonsingular transformation T of a Lebesgue space $(X, \mathcal{B}, m)$ together with a measurable partition α of X such that, if $a \in \alpha$, $m(a) > 0$, Ta is a union (mod m) of elements of α, and T: $a \to Ta$ is invertible. Moreover, it is assumed that the completion of $\bigvee_0^\infty T^{-i}\alpha$ with respect to m is $\mathcal{B}$, i.e., the partition separates the points.

For $a_0, \dots, a_{n-1} \in \alpha$ define a **cylinder** by $[a_0, \dots, a_{n-1}] = \bigcap_{i=0}^{n-1} T^{-i}a_i$: two points in an identical cylinder of length n remain in the same elements of the partition up to time n. These cylinders can be used to define $d_\theta(x,y) = \theta^{t(x,y)}$, where $t(x,y) = \sup\{n \mid x, y \in [a_0, \dots, a_{n-1}]$ for some $a_0, \dots, a_{n-1} \in \alpha\}$ is the time until which x and y remain in the same elements of the partition α, and $0 < \theta < 1$ is some fixed number.

A Markov map T is said to be **irreducible** if $\forall a, b \in \alpha$, $\exists n$, $m(T^{-n}a \cap b) > 0$ (i.e., $b \subset T^n a \bmod m$). This means that there is no $\gamma \subsetneq \alpha$ such that the elements of γ are stable by T. An irreducible Markov map T is **aperiodic** if $\forall a \in \alpha, \exists N \in \mathbb{N}, \forall n \geqslant N, a \subset T^n a$. Equivalently, there exists such an a, or there exists an a such that $\gcd\{n \mid a \subset T^n a\} = 1$. An irreducible aperiodic Markov map is also said to be **topologically mixing**, i.e., $\forall a, b \in \alpha, \exists N, \forall n \geqslant N, b \subset T^n a$ (this is a definition, and does not rely on a topology on the space, even though it corresponds to the usual notion of topological mixing when d_θ is a metric).

The transfer operator $\hat{T}$ associated to T is defined by $\int(\hat{T}u)v\mathrm{d}m = \int u\cdot v\circ T\mathrm{d}m$. It can be written $\hat{T}f(x) = \sum_{Ty=x} g_m(y)f(y)$, where the weight g_m is defined by

$$g_m = \frac{\mathrm{d}m}{\mathrm{d}m\circ T}$$

(the measure $m\circ T$ is given by $m\circ T(E) = \sum_{a\in\alpha} m(T(E\cap a))$). Different regularity assumptions are possible on $\log g_m$, corresponding to different controls of the distortion.

For any function $\varphi : X \to \mathbb{C}$, the **variations** of φ are defined by $v_n(\varphi) = \sup\{|\varphi(x)-\varphi(y)| \mid x, y\in [a_0,\dots,a_{n-1}]$ where $a_i\in\alpha\}$. The function φ is said to have summable variations if $\sum_{n\geqslant 1} v_n(\varphi) < +\infty$, and to be Hölder continuous for the exponent θ if $\exists C>0, \forall n\geqslant 1, v_n(\varphi)\leqslant C\theta^n$ (this is a definition, which corresponds to being Lipschitz with respect to the "metric" d_θ on each element of the partition α). By a slight abuse of notation, we will say that a measurable function f has summable variations if there exists a version (mod 0) of f which has summable variations, and the notation f will denote this version.

If $\log g_m$ has summable variations, the distortion is bounded, meaning that there exists a constant C such that, for all $x, y\in[a_0,\dots,a_{n-1}]$,

$$\left|\frac{g_m^{(n)}(x)}{g_m^{(n)}(y)} - 1\right| \leqslant C,$$

where $g_m^{(n)} = \prod_{i=0}^{n-1} g_m\circ T^i$ is the weight associated to $\hat{T}^n$. In particular, this implies that

$$g_m^{(n)}(x) = C^{\pm 1}\frac{m[a_0,\dots,a_{n-1}]}{m[Ta_{n-1}]}$$
$$\left(\text{i.e. } \frac{1}{C}\frac{m[a_0,\dots,a_{n-1}]}{m[Ta_{n-1}]} \leqslant g_m^{(n)}(x) \leqslant C\frac{m[a_0,\dots,a_{n-1}]}{m[Ta_{n-1}]}\right).$$

When the **"big image" property** $\inf_{a\in\alpha} m[Ta]>0$ is satisfied, we even obtain $g_m^{(n)}(x) = D^{\pm 1}m[a_0,\dots,a_{n-1}]$.

PROPOSITION 6.1: *Let $(X,\mathcal{B},T,m,\alpha)$ be an irreducible Markov map with the big image property for which $\log g_m$ is of summable variations. Then T is conservative and ergodic.*

Proof: This is a corollary of Theorem 4.6.3 in [Aar97] (where the hypotheses are in fact weaker, since this theorem requires only the "weak distortion property"), and was originally proved in [ADU93]. ∎

If $\log g_m$ is Hölder continuous, the distortion is better controlled, which gives stronger results. In particular, the transfer operator $\hat{T}$ acting on the space of Hölder continuous bounded functions admits a spectral gap ([Aar97, Thm. 4.7.7]). More precisely, let α' denote the smallest partition such that, $\forall a \in \alpha, Ta$ is α'-measurable; the partition α' is coarser than α. For $a \in \alpha'$ and $f : X \to \mathbb{R}$, write $D_a f = \sup\{|f(x) - f(y)|/d_\theta(x,y) | \, x, y \in a\}$ the best Lipschitz constant of f on a. Finally, let $\mathcal{L}$ be the space of functions $f\colon X \to \mathbb{C}$ such that $\|f\|_{\mathcal{L}} = \|f\|_\infty + \sup_{a\in\alpha'} D_a f < +\infty$. It is the space of Lipschitz functions on X, but the norm is not the usual Lipschitz norm. When $\log g_m$ is Hölder continuous (for some exponent θ) and T has the big image property, Ruelle has proved that the essential spectral radius of $\hat{T}$ acting on $\mathcal{L}$ satisfies $r_{ess}(\hat{T}) \leqslant \theta$.

6.2 Induced Markov maps. From this point on, $(X, \mathcal{B}, m, T, \alpha)$ will be a probability preserving Markov map.

Let $\emptyset \neq \gamma \subset \alpha$. If $Y = \bigcup_{a\in\gamma} a$, the induced map $T_\gamma\colon Y \to Y$ is defined as the first return map from Y to Y, i.e., $T_\gamma = T^{\varphi_\gamma}$, where

$$\varphi_\gamma(x) = \inf\{n \geqslant 1 | \, T^n(x) \in Y\}$$

is the return time to Y. If $x \notin Y$, we set $\varphi_\gamma(x) = 0$. By the Poincaré recurrence theorem, T_γ and all its iterates are defined for m-almost every point of Y — replacing Y by this smaller set, we can assume that T_γ is in fact defined on all Y.

A measure m_γ is defined on Y by $m_\gamma = m_{|Y}$. As m is invariant by T, the measure m_γ is invariant by T_γ.

Let $\delta = \{[a, \xi_1, \ldots, \xi_{n-1}, \gamma] | \, a \in \gamma, \xi_1, \ldots, \xi_{n-1} \notin \gamma, [a, \xi_1, \ldots, \xi_{n-1}, \gamma] \neq \emptyset\}$: this is a partition of Y, for which T_γ is a Markov map. The cylinders for this partition will be denoted by $[d_0, \ldots, d_{n-1}]_\gamma$ (with $d_0, \ldots, d_{n-1} \in \delta$). If $d = [a, \xi_1, \ldots, \xi_{n-1}, \gamma] \in \delta$, its image is $T_\gamma d = T\xi_{n-1} \cap Y$ — hence, it is γ-measurable. In particular, if γ is finite, its elements have a measure $\geqslant \varepsilon > 0$, which implies that $\forall d \in \delta$, $m_\gamma(T_\gamma d) \geqslant \varepsilon$. Thus T_γ has the "big image" property.

The following straightforward lemma establishes a link between the mixing properties of T and those of the induced transformation T_γ.

Lemma 6.2: *If T is irreducible, then T_γ is irreducible.*

We will be interested in induced maps which have good distortion properties. More precisely, write

$$g_{m_\gamma} = \frac{\mathrm{d}m_\gamma}{\mathrm{d}m_\gamma \circ T_\gamma}.$$

We assume that there exist constants $C > 0$ and $\theta < 1$ such that $\forall n \geqslant 1$, $v_n(\log g_{m_\gamma}) \leqslant C\theta^n$ (where v_n is the variation with respect to the induced map T_γ): we say that $\log g_{m_\gamma}$ is locally Hölder continuous. In this case, the previous theorems on maps whose distortion has summable variations apply to T_γ.

As above, let δ' denote the smallest partition such that $\forall d \in \delta$, $T_\gamma d$ is a union of atoms of δ'. As every $T_\gamma d$ is γ-measurable, this partition is coarser than γ. For $x, y \in Y$, let $t_\gamma(x,y) = \sup\{n | \; x, y \in [d_0, \dots, d_{n-1}]_\gamma\}$ and let $\mathcal{L}$ denote the space of functions $f\colon Y \to \mathbb{C}$ such that $\|f\|_{\mathcal{L}} := \|f\|_\infty + \sup_{d \in \delta'} D_d f < +\infty$, where $D_d f$ is the least Lipschitz constant of f on d for the "distance" $d(x,y) = \theta^{t_\gamma(x,y)}$.

We now state the main theorem of this section:

THEOREM 6.3: *Let $(X, \mathcal{B}, m, T, \alpha)$ be a topologically mixing probability preserving Markov map, and $\emptyset \neq \gamma \subset \alpha$. Assume that T_γ has the big image property and that g_{m_γ} has a version such that $\log g_{m_\gamma}$ is locally θ-Hölder continuous for some $0 < \theta < 1$. Assume moreover that $m[\varphi_\gamma > n] = O(1/n^\beta)$ for some $\beta > 1$.*

Then $\exists C > 0$ such that $\forall f, g$ integrable and supported inside Y,

$$\left| \operatorname{Cor}(f, g \circ T^n) - \left(\sum_{k=n+1}^{\infty} m[\varphi_\gamma > k] \right) \int f \int g \right| \leqslant C F_\beta(n) \|g\|_\infty \|f\|_{\mathcal{L}}$$

where $F_\beta(n) = 1/n^\beta$ if $\beta > 2$, $(\log n)/n^2$ if $\beta = 2$ and $1/n^{2\beta - 2}$ if $2 > \beta > 1$ (and $\mathcal{L}$ denotes the space of θ-Hölder functions on Y).

Moreover, if $\int f = 0$, then $\operatorname{Cor}(f, g \circ T^n) = O(1/n^\beta)$.

6.3 PROOF OF THEOREM 6.3. The strategy is to apply the abstract Theorem 1.1 to "first return transfer operators". In this section, (T, α) will be a Markov map and γ a subset of α such that the hypotheses of Theorem 6.3 are satisfied. The proof of Theorem 6.3 will be quite similar to Sarig's proof of his Theorem 2 (in particular, the first three lemmas can essentially be found in [Sar02]), but there is a significant difference in the proof of the aperiodicity hypothesis (Lemma 6.7), since the hypothesis to be checked is different.

For $\underline{d} = [d_0, \dots, d_{n-1}]_\gamma \neq \emptyset$, set $M_{\underline{d}} f(x) = g_{m_\gamma}^{(n)}(d_0 \cdots d_{n-1} x) f(d_0 \cdots d_{n-1} x)$ if this point is defined, 0 otherwise.

LEMMA 6.4: *There exists a constant B such that, $\forall \underline{d} = [d_0, \dots, d_{n-1}]_\gamma$, $\forall f \in \mathcal{L}$,*

$$\|M_{\underline{d}} f\|_{\mathcal{L}} \leqslant B m[\underline{d}] \left(\theta^n \|f\|_{\mathcal{L}} + \frac{1}{m[\underline{d}]} \int_{[\underline{d}]} |f| \mathrm{d}m \right).$$

Proof: This lemma is classical and uses the distortion control to obtain explicit estimates. See, for example, [Sar02, Lemma 8] or [Aar97, Lemma 4.7.2]. ∎

Let L be the operator defined by $Lf(x) = \sum_{Ty=x} g_m(y)f(y)$: it is a version of the transfer operator $\hat{T}$, but it acts on actual functions and not on functions defined almost everywhere. In the same way, but for the induced map, set $L_\gamma f(x) = \sum_{T_\gamma y=x} g_{m_\gamma}(y)f(y)$. Write also

$$T_n f = 1_Y L^n(f 1_Y) \quad \text{and} \quad R_n f = 1_Y L^n(f 1_{\{\varphi_\gamma = n\}}).$$

The operator T_n counts all returns from Y to Y at time n, while R_n takes only the first returns at time n into account. Note that, by definition, $\varphi_\gamma = 0$ outside of Y, so R_n really counts returns to Y. For $z \in \mathbb{D}$, we set $T(z) = I + \sum T_n z^n$ and $R(z) = \sum R_n z^n$.

Lemma 6.5: *T_n and R_n are bounded operators on $\mathcal{L}$, $\|T_n\| = O(1)$, $\|R_n\| = O(m[\varphi_\gamma = n])$ and, $\forall z \in \mathbb{D}$, $T(z) = (I - R(z))^{-1}$.*

Proof: We have $R_n = \sum_{\underline{d}=[d_0]_\gamma, d_0=[a_0,\dots,a_{n-1},\gamma]} M_{\underline{d}}$. Thus, Lemma 6.4 shows that $\|R_n\| \leqslant B(1+\theta)\sum m[\underline{d}] = (1+\theta)Bm[\varphi_\gamma = n]$.

In the same way, $T_n = \sum M_{\underline{d}}$ where the sum extends to all $\underline{d} = [d_0, \dots, d_{k-1}]_\gamma$ with $d_i = [\xi_{i0}, \dots, \xi_{in_i}, \gamma]$ and $\sum(n_i + 1) = n$. Hence, $\|T_n\| \leqslant B(1+\theta)m[Y]$ (the sum is a sum of measures of disjointed sets included in Y, less than $m[Y]$).

Finally, T_n counts all returns to Y while R_n counts only the first returns. Hence, $T_n = \sum_{i_1+\dots+i_k=n} R_{i_1}\cdots R_{i_k}$, which gives the renewal equation. ∎

Lemma 6.6: *The operator $R(1)$: $\mathcal{L} \to \mathcal{L}$ has a simple isolated eigenvalue at 1, the spectral projection being given by $Pf = \frac{1}{m[Y]}\int_Y f\mathrm{d}m$.*

Proof: As $R(1)$ counts the first returns to Y, it is not hard to check that $R(1) = L_\gamma$ is the transfer operator associated to T_γ, i.e., $R(1) = \sum_{\underline{d}=[d_0]_\gamma} M_{\underline{d}}$. In fact, $R(1)^n = \sum_{\underline{d}=[d_0,\dots,d_{n-1}]_\gamma} M_{\underline{d}}$, hence Lemma 6.4 shows that

$$\|R(1)^n f\|_{\mathcal{L}} \leqslant B\theta^n\|f\|_{\mathcal{L}} + B\|f\|_1. \tag{9}$$

The injection $\mathcal{L} \to L^1(m)$ is compact by the Arzela–Ascoli theorem. Hence, the Doeblin-Fortet inequality (9) gives, with the use of Hennion's theorem ([Hen93]), that the essential spectral radius of $R(1)$ acting on $\mathcal{L}$ is $\leqslant \theta$. Thus, if 1 is an eigenvalue of $R(1)$, it is automatically isolated and of finite multiplicity.

As T_γ preserves the measure m_γ (since T preserves m), $L_\gamma 1 = 1$ and $PR(1) = R(1)P = P$. By Lemma 6.2 and Proposition 6.1, T_γ is ergodic, whence there is no other eigenfunction for the eigenvalue 1. Finally, there is no nilpotent part for this eigenvalue either, since $\|R(1)^n\|$ remains bounded. ∎

LEMMA 6.7: *$\forall z \in \overline{\mathbb{D}} - \{1\}$, $I - R(z)$ is invertible on $\mathcal{L}$.*

Proof: Summing the estimates given by Lemma 6.4 for $\underline{d}$ of length n gives that

$$\|R(z)^n f\|_{\mathcal{L}} \leqslant B|z|^n(\theta^n\|f\|_{\mathcal{L}} + \|f\|_1). \tag{10}$$

As the injection $\mathcal{L} \to L^1(m)$ is compact by the Arzela–Ascoli theorem, the theorem of Hennion ([Hen93]) ensures that, $\forall z \in \overline{\mathbb{D}}$, the essential spectral radius of $R(z)$ acting on $\mathcal{L}$ is $\leqslant \theta < 1$. To obtain the invertibility of $I - R(z)$, it is thus enough to show that 1 is not an eigenvalue of $R(z)$. The only problem is for $|z| = 1$ because otherwise, again by Equation (10), the spectral radius of $R(z)$ is $\leqslant |z| < 1$ (since $\|f\|_1 \leqslant \|f\|_{\mathcal{L}}$). So, let $z = e^{it}$ be fixed, with $0 < t < 2\pi$.

Suppose that $R(z)f = f$ for some nonzero $f \in \mathcal{L}$. We will write, for $u, v \in L^2(m_\gamma)$, $\langle u, v\rangle = \int \overline{u} v dm_\gamma$. Define the operator W: $L^\infty(m_\gamma) \to L^\infty(m_\gamma)$ by $Wu = e^{-it\varphi_\gamma} u \circ T_\gamma$. As $R(z)v = R(1)(e^{it\varphi_\gamma}v)$, this operator W satisfies

$$\begin{aligned}\langle u, R(z)v\rangle = &\int \overline{u}\, R(z)v = \int \overline{u}\, R(1)(e^{it\varphi_\gamma}v) = \int \overline{u} \circ T_\gamma\, e^{it\varphi_\gamma} v = \int \overline{Wu} \cdot v \\ =&\langle Wu, v\rangle.\end{aligned}$$

We show that f is an eigenfunction of W for the eigenvalue 1:

$$\begin{aligned}\|Wf - f\|_2^2 =&\|Wf\|_2^2 - 2\operatorname{Re}\langle Wf, f\rangle + \|f\|_2^2 = \|Wf\|_2^2 - 2\operatorname{Re}\langle f, R(z)f\rangle + \|f\|_2^2 \\ =&\|Wf\|_2^2 - 2\operatorname{Re}\langle f, f\rangle + \|f\|_2^2 = \|Wf\|_2^2 - \|f\|_2^2.\end{aligned}$$

As T_γ preserves the measure m_γ, we have $\|Wf\|_2^2 = \int |f|^2 \circ T_\gamma = \int |f|^2 = \|f\|_2^2$, which gives $\|Wf - f\|_2^2 = 0$. Hence, the function $Wf - f$ is zero m_γ-almost everywhere. As $f \in \mathcal{L}$ and m_γ is nonzero on every cylinder, the function f is continuous, thus $Wf - f = 0$ everywhere.

We have a function f such that $e^{-it\varphi_\gamma} f \circ T_\gamma = f$. Taking the modulus, the ergodicity of T_γ gives that $|f|$ is constant almost everywhere, hence everywhere by continuity. As $f \not\equiv 0$, this constant is nonzero, and we get $e^{-it\varphi_\gamma} = f/f \circ T_\gamma$. We can apply Theorem 3.1 in [AD01] and obtain that f is δ^*-measurable, where δ^* is the smallest partition such that $\forall d \in \delta$, $T_\gamma d$ is contained in an atom of δ^*. As every $T_\gamma d$ is a union of sets of γ, this implies in particular that f is constant (almost everywhere, hence everywhere by continuity) on each set of γ.

Let $a \in \gamma$. On $[a]$, f is equal to a constant c. As T is topologically mixing, there exists N such that, $\forall n \geqslant N$, $[a] \subset T^n[a]$. Let $n \geqslant N$, and $x \in [a]$ be such that $T^n x \in [a]$. Let $T^{k_1}x, T^{k_2}x, \ldots, T^{k_p}x$ be the successive returns of x to Y, with $k_p = n$. Then $T^n x = T_\gamma^p x$ and $n = \sum_{k=0}^{p-1} \varphi_\gamma(T_\gamma^k x)$. Thus,

$$e^{-itn} = e^{-it\sum_{k=0}^{p-1}\varphi_\gamma(T_\gamma^k x)} = \frac{f(x)}{f(T_\gamma x)}\frac{f(T_\gamma x)}{f(T_\gamma^2 x)} \cdots \frac{f(T_\gamma^{p-1}x)}{f(T_\gamma^p x)} = \frac{f(x)}{f(T^n x)} = \frac{c}{c} = 1.$$

This is true for any $n \geqslant N$. Taking, for example, $n = N$ and $N+1$ and quotienting, we obtain $e^{it} = 1$, which is a contradiction. ∎

LEMMA 6.8: *We have* $PR'(1)P = \frac{1}{m[Y]}P$.

Proof: Using the explicit formula for the spectral projection P, it is not difficult to check that

$$PR_nP = \frac{m[\varphi_\gamma = n]}{m[Y]}P,$$

and consequently $PR'(1)P = \frac{1}{m[Y]}P$ by the Kac formula ([Aar97, Formula 1.5.5]). To apply this formula, we have to check that T is conservative and ergodic, knowing that this is the case for T_γ. This can be done, for example, using [Aar97, Proposition 1.5.2]. ∎

Proof of Theorem 6.3: The lemmas above show that the hypotheses of Theorem 1.1 are satisfied. Consequently, we get the existence of $E_n \in \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ with $\|E_n\| = O(F_\beta(n))$ such that $\forall f \in \mathcal{L}$,

$$1_Y \hat{T}^n f = 1_Y \Big(\int f \mathrm{d}m + \sum_{k=n+1}^{\infty} m[\varphi_\gamma > n] \int f \mathrm{d}m + E_n f \Big).$$

Multiplying by an arbitrary $g \in L^\infty(X, \mathcal{B}, m)$ supported inside Y, we have by the definition of the transfer operator

$$\int f \cdot g \circ T^n \mathrm{d}m = \int f \int g + \sum_{k=n+1}^{\infty} m[\varphi_\gamma > k] \int f \int g + \int g \cdot E_n f \mathrm{d}m.$$

The absolute value of the last term is bounded by $\|g\|_\infty \|E_n\|_{\mathcal{L}} \|f\|_{\mathcal{L}}$, which gives the result.

Finally, if $\int f = 0$, we use Theorem 1.2 and conclude in the same way, the estimates with $F_\beta(n)$ being replaced by estimates in $O(1/n^\beta)$. ∎

6.4 DECAY OF CORRELATIONS ON THE WHOLE SPACE. Theorem 6.3 gives a very sharp estimate on the decay of correlations when the functions f and g are supported in Y. It is also possible to estimate the speed of decay for a general g, not necessarily supported in Y, although the estimates will be less precise. This kind of result will be useful in the proof of the Central Limit Theorem.

THEOREM 6.9: *Under the hypotheses of Theorem 6.3, assume that f is supported in Y and that $g \in L^\infty(m)$. Then there exists a constant C (independent of f or g) such that*

$$\operatorname{Cor}(f, g\circ T^n) \leqslant \frac{C}{n^{\beta-1}} \|f\|_{\mathcal{L}} \|g\|_\infty .$$

To obtain this theorem, it is enough to prove that $\|\hat{T}^n f - \int f\|_1 \leqslant \frac{C}{n^{\beta-1}}\|f\|_{\mathcal{L}}$.

LEMMA 6.10: *There exists C such that $\forall A \in \mathcal{B}$, $|\int_A \hat{T}^n f \mathrm{d}m - m(A)\int f| \leqslant \frac{C}{n^{\beta-1}}\|f\|_{\mathcal{L}}$.*

Proof: In the course of this proof, we shall write L for the transfer operator acting on functions in $\mathcal{L}$. Write also $K_0 f = 1_Y f$ and, $\forall k \geqslant 1$, $K_k f = L^k(1_{\{\varphi_\gamma > k\}} f)$: K_k counts the first returns to Y at time k, but for points not starting in Y (contrary to R_k). It is then easy to check that $L^n f = \sum_{k=0}^n K_k T_{n-k} f$ for any f supported in Y (recall that, outside of Y, $\varphi_\gamma = 0$ by definition).

Then, writing $T_n f = \int f + \varepsilon_n$ with $\|\varepsilon_n\|_{\mathcal{L}} \leqslant \frac{C}{n^{\beta-1}}\|f\|_{\mathcal{L}}$,

$$\begin{aligned}
\int_A L^n f &= \int 1_{Y\cap A} T_n f + \sum_{k=1}^n \int 1_A L^k(1_{\{\varphi_\gamma>k\}} T_{n-k} f)\\
&= \int 1_{Y\cap A} T_n f + \sum_{k=1}^n \int 1_A \circ T^k \cdot 1_{\{\varphi_\gamma>k\}} T_{n-k} f\\
&= \int f \Big(\int 1_{Y\cap A} + \sum_{k=1}^n \int 1_A \circ T^k \cdot 1_{\{\varphi_\gamma>k\}} \Big)\\
&\quad + \Big(\int 1_{Y\cap A} \varepsilon_n + \sum_{k=1}^n \int 1_A \circ T^k \cdot 1_{\{\varphi_\gamma>k\}} \varepsilon_{n-k} \Big)\\
&= I \int f + II.
\end{aligned}$$

I can be expressed as $m(Y\cap A) + \sum_{k=1}^n m(Y \cap T^{-k}A - \bigcup_{j=1}^k T^{-j}Y)$. Thus, by Kac's Formula (see [Aar97, Lemma 1.5.4]),

$$I = m(A) - \sum_{k=n+1}^{\infty} m(Y\cap T^{-k}A - \bigcup_{j=1}^{k} T^{-j} Y).$$

As $m(Y\cap T^{-k}A - \bigcup_{j=1}^k T^{-j}Y) \leqslant m[\varphi_\gamma > k] \leqslant C/k^\beta$ by hypothesis, a summation yields $I = m(A) + O(1/n^{\beta-1})$.

In the same way, $II \leqslant \|\varepsilon_n\|_\infty + \sum_{k=1}^n m[\varphi_\gamma > k]\|\varepsilon_{n-k}\|_\infty$: this is a convolution between sequences respectively in $O(1/n^\beta)$ and $O(1/n^{\beta-1})$, whence $II = O(1/n^{\beta-1})$ by Lemma 4.3. ∎

Proof of Theorem 6.9: Lemma 6.10 yields that, $\forall A \in \mathcal{B}$,

$$\left| \int_A \left(\hat{T}^n f - \int f \right) \mathrm{d}m \right| \leqslant \frac{C}{n^{\beta-1}} \|f\|_{\mathcal{L}}.$$

Apply this estimate to $A = \{\hat{T}^n f - \int f \geqslant 0\}$, then to $A = \{\hat{T}^n f - \int f < 0\}$, and sum to obtain that

$$\int \left| \hat{T}^n f - \int f \right| \mathrm{d}m \leqslant \frac{2C}{n^{\beta-1}} \|f\|_{\mathcal{L}}. \quad ∎$$

6.5 Central Limit Theorem.

Proposition 6.11: *Under the hypotheses of Theorem 6.3, assume that f is supported in Y and that $\int f = 0$. Then there exists a constant C (independent of f) such that $\|\hat{T}^n f\|_1 \leqslant \frac{C}{n^\beta} \|f\|_{\mathcal{L}}$.*

Proof: This is an analogue of Theorem 6.9 in the case where $\int f = 0$ (which implies that there is a better bound on $\|T_n f\|_{\mathcal{L}}$, according to Theorem 6.3). The same proof works again, and is even easier because the term I in the proof of Lemma 6.10 disappears. ∎

The following lemma will be useful in the Central Limit Theorem to make precise the regularity of the cocycle in the case of zero variance.

Lemma 6.12: *Let $(X, \mathcal{B}, T, m, \alpha)$ be an irreducible probability preserving Markov map with the big image property and for which the distortion $\log g_m$ is Hölder for an exponent $\theta < 1$. Let $\mathcal{L}$ denote the space of bounded functions such that $\sup_{a \in \alpha'} D_a f < +\infty$. If $f \in \mathcal{L}$ and $g: X \to \mathbb{R}$ is measurable and satisfies $f \circ T = g \circ T - g$, then $g \in \mathcal{L}$.**

Note that g is not assumed to be integrable, whence it is not possible to apply $\hat{T}$ to g and to use spectral methods to prove this lemma.

Proof: Denote by $\alpha^n(x)$ the element of the partition $\bigvee_{i=0}^{n-1} T^{-i}\alpha$ containing x. A classical theorem on continuity points of measurable functions (true on $[0, 1]$ with the Lebesgue measure, in which X can be canonically imbedded) implies that

$$\text{for almost every } x, \forall \varepsilon > 0, \quad \frac{m\{y \in \alpha^n(x) \mid |g(y) - g(x)| > \varepsilon\}}{m[\alpha^n(x)]} \to 0 \quad \text{as } n \to \infty.$$

---

* We use the abuse of notation already mentioned in Section 6.1: this really means that there exists a version of g belonging to $\mathcal{L}$.

The points that visit infinitely many times every element of the partition α form also a set of probability 1. We fix a point x_0 satisfying these two properties.

Fix $\varepsilon > 0$. Let $n_k \to \infty$ be a sequence such that $T^{n_k} x_0$ visits infinitely often every element of α too, and

$$\sum \frac{m\{y \in \alpha^{n_k}(x_0) | \, |g(y) - g(x_0)| > \varepsilon\}}{m[\alpha^{n_k}(x_0)]} < \infty.$$

For every $k \in \mathbb{N}$, the control on the distortion implies that

$$\frac{m\{y \in X | \, \exists y' \in \alpha^{n_k}(x_0), T^{n_k} y' = y, |g(y') - g(x_0)| > \varepsilon\}}{m[T^{n_k}\alpha^{n_k}(x_0)]} \asymp \frac{m\{y' \in \alpha^{n_k}(x_0) | \, |g(y') - g(x_0)| > \varepsilon\}}{m[\alpha^{n_k}(x_0)]}.$$

Thus, $\sum_k m\{y \in X | \, \exists y' \in \alpha^{n_k}(x_0), T^{n_k} y' = y, |g(y') - g(x_0)| > \varepsilon\} < +\infty$. Consequently, $A_\varepsilon := \{y \in X | \, \exists K, \forall k \geqslant K$, if $y' \in \alpha^{n_k}(x_0)$ is such that $T^{n_k} y' = y$, then $|g(y') - g(x_0)| \leqslant \varepsilon\}$ is of full measure.

Take $y_1, y_2 \in A_\varepsilon$ such that y_1 and y_2 are in the same element of α', with $d(y_1, y_2) = \theta^n$ for some $n \geqslant 0$. Take k such that $T^{n_k} x_0$ is in the same element of α' as y_1 and y_2. If k is large enough, by definition of A_ε, the preimages y_i' of y_i in $\alpha^{n_k}(x_0)$ satisfy $|g(y_i') - g(x_0)| \leqslant \varepsilon$, hence $|g(y_1') - g(y_2')| \leqslant 2\varepsilon$. Then

$$\begin{aligned} |g(y_1) - g(y_2)| &= |g \circ T^{n_k}(y_1') - g \circ T^{n_k}(y_2')| \\ &\leqslant \sum_{i=1}^{n_k} |f \circ T^i(y_1') - f \circ T^i(y_2')| + |g(y_1') - g(y_2')| \\ &\leqslant \sum_{i=1}^{n_k} \|f\|_{\mathcal{L}} \theta^{n_k + n - i} + 2\varepsilon \leqslant \frac{\|f\|_{\mathcal{L}}}{1-\theta} d_\theta(y_1, y_2) + 2\varepsilon. \end{aligned}$$

Finally, for $y_1, y_2 \in A = \bigcap A_\varepsilon$ of full measure,

$$|g(y_1) - g(y_2)| \leqslant \frac{\|f\|_{\mathcal{L}}}{1-\theta} d(y_1, y_2).$$

Hence, there exists a unique version of the function g which is Lipschitz on every set of α', which we will still denote by g.

To see that $g \in \mathcal{L}$, it remains to prove that g is bounded. Let $\eta > 0$ be such that $\forall a \in \alpha, m[Ta] > \eta$. There exists $a_1, \dots, a_N \in \alpha$ a finite number of partition sets such that $\sum_{i=1}^N m[a_i] > 1 - \eta$. Thus, $\forall a \in \alpha$, Ta contains one of the sets a_i. On each of these sets, g is Lipschitz, hence bounded by a constant C_i. If $x \in [a]$ has its image in a_i, then $|g(x)| = |g \circ T(x) - f \circ T(x)| \leqslant C_i + \|f\|_\infty \leqslant \max_i C_i + \|f\|_\infty =: C$.

Finally, for $y \in [a]$,

$$|g(y)| \leqslant |g(x) - g(y)| + |g(y)| \leqslant \frac{\|f\|_{\mathcal{L}}}{1-\theta} + C. \qquad \blacksquare$$

THEOREM 6.13: *Under the hypotheses of Theorem 6.3, if $f \in \mathcal{L}$ is supported in Y and $\int f = 0$, then the sequence $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k$ converges in distribution to a Gaussian random variable of zero mean and finite variance σ^2, with*

$$\sigma^2 = -\int f^2 \mathrm{d}m + 2 \sum_{n=0}^{\infty} \int f \cdot f \circ T^n \mathrm{d}m.$$

Moreover, $\sigma = 0$ if and only if there exists a measurable function g such that $f \circ T = g \circ T - g$. Such a function g automatically satisfies $g_{|Y} \in \mathcal{L}$ and $\forall x \in Y, \forall n < \varphi_Y(x), g(T^n x) = g(x)$.

We will use an abstract result due to Liverani [Liv96, Theorem 1.1] inspired by Kipnis–Varadhan to obtain this Central Limit Theorem. We recall for the convenience of the reader the version of this theorem that will be useful in our setting.

THEOREM 6.14: *Let $(X, \mathcal{B}, T, m)$ be a non-singular probability preserving dynamical system. Let also $f \in L^\infty(X)$, $\int f = 0$ be such that*

1. $\sum_{n=0}^{\infty} |\int f \cdot f \circ T^n| < \infty$.
2. *The series $\sum_{n=0}^{\infty} \hat{T}^n f$ converges absolutely in L^1.*

Then the sequence $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k$ converges in distribution to a Gaussian random variable of zero mean and finite variance σ^2, with

$$\sigma^2 = -\int f^2 \mathrm{d}m + 2 \sum_{n=0}^{\infty} \int f \cdot f \circ T^n \mathrm{d}m.$$

Moreover, $\sigma = 0$ if and only if there exists a measurable function g such that $f \circ T = g \circ T - g$.

Proof of Theorem 6.13: It is enough to show that the hypotheses of Theorem 6.14 are satisfied. As we have formulated this theorem, the first hypothesis is in fact a consequence of the second one, since

$$\left| \int f \cdot f \circ T^n \mathrm{d}m \right| = \left| \int \hat{T}^n f \cdot f \mathrm{d}m \right| \leqslant \|\hat{T}^n f\|_1 \|f\|_\infty.$$

Consequently, it remains only to check that $\sum \|\hat{T}^n f\|_1 < +\infty$. By Proposition 6.11, $\|\hat{T}^n f\|_1 = O(1/n^\beta)$ with $\beta > 1$, thus the series is summable.

To obtain the regularity results on g when $\sigma = 0$, we use the fact that $f = 0$ outside of Y. As $f \circ T = g \circ T - g$, this implies that $g(x) = g \circ T(x)$ when $T(x) \notin Y$. In particular, $\forall x \in Y, \forall n < \varphi_\gamma(x), g(x) = g(T^n x)$. Using once more the cocycle relation gives that $f \circ T_\gamma(x) = g \circ T_\gamma(x) - g(x)$. Thus, Lemma 6.12 applied to (Y, T_γ) shows that $g_{|Y} \in \mathcal{L}$. ∎

7. Applications to specific maps

7.1 THE LIVERANI–SAUSSOL–VAIENTI MAP. The Liverani–Saussol–Vaienti map is the map $T\colon [0,1] \to [0,1]$ defined by

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

It is an analogue of the Pomeau–Manneville map $x \mapsto \{x + x^{1+\alpha}\}$, but the second branch of the map is affine, which simplifies to some extent the computations (our results also apply directly to the Pomeau–Manneville map). It is shown in [LSV99] that, when $0 < \alpha < 1$, T admits an integrable invariant density h which is Lipschitz outside of any neighborhood of 0.

COROLLARY 7.1: *If $\alpha \in (0,1)$, f is Lipschitz, g is bounded measurable, $\int f, \int g \neq 0$ and $f, g = 0$ in a neighborhood of 0, then*

$$\operatorname{Cor}(f, g \circ T^n) \sim \frac{1}{4} h\Big(\frac{1}{2}\Big) \alpha^{-1/\alpha} \Big(\frac{1}{\alpha} - 1\Big)^{-1} n^{1-1/\alpha} \int f \int g$$

with respect to the invariant probability measure.

Moreover, if $\int f = 0$ (and f, g are still zero in a neighborhood of 0, f Lipschitz), then $\operatorname{Cor}(f, g \circ T^n) = O(1/n^{1/\alpha})$. Consequently, f satisfies a Central Limit Theorem.

Proof: If $x_0 = 1/2$ and $x_{i+1} = T^{-1}(x_i) \cap [0, 1/2]$, the partition

$$\alpha = \{(x_{i+1}, x_i]\} \cup (1/2, 1]$$

is a Markov partition for T, which makes it possible to apply the results of the previous section to this map. The distortion of the induced map on $(1/2, 1]$ is locally Hölder continuous for the density h, whence Theorem 6.3 applies and gives a precise asymptotic on the speed of decay of correlations for functions supported in $(1/2, 1]$, which can be calculated precisely (see [Sar02]).

As the distortion from $(x_{i+1}, x_i]$ to $(x_i, x_{i-1}]$ is bounded, it is not hard to check that the induced map on $\gamma = \{(x_{i+1}, x_i] | \ i < N\} \cup \{(1/2, 1]\}$ has still a Hölder continuous distortion for any N. Thus, Theorem 6.3 gives also estimates on the decay of correlations of functions supported in $(x_N, 1]$. More precisely, for functions $f \in \mathcal{L}$ and $g \in L^\infty$ supported in $(x_N, 1]$,

$$\mathrm{Cor}(f, g \circ T^n) \sim \Big(\sum_{k=n+1}^{\infty} m[\varphi_\gamma > k] \Big) \int f \int g. \tag{11}$$

For functions supported in $(1/2, 1]$, Sarig has shown, estimating $m[\varphi_{(1/2,1]}] > n$, that

$$\mathrm{Cor}(f, g \circ T^n) \sim \frac{1}{4} h\Big(\frac{1}{2}\Big) \alpha^{-1/\alpha} \Big(\frac{1}{\alpha} - 1\Big)^{-1} n^{1-1/\alpha} \int f \int g. \tag{12}$$

The estimate (11) can be applied in particular to functions supported in $(1/2, 1]$, which gives, using (12), that $\sum_{k=n+1}^\infty m[\varphi_\gamma > k] \sim \frac{1}{4} h(\frac{1}{2}) \alpha^{-1/\alpha} (\frac{1}{\alpha} - 1)^{-1} n^{1-1/\alpha}$. This proves the corollary. ∎

7.2 LS Young towers. A **LS Young tower** is a non-singular conservative transformation $(\Delta, \mathcal{B}, m, F)$ with a generating partition

$$\{\Delta_{l,i} | \ i \in \mathbb{N}, l = 0, \dots, R_i - 1\}$$

with the following properties:

1. $\forall l, i$ the measure of $\Delta_{l,i}$ is positive and finite. Moreover, if $\Delta_l = \bigcup \Delta_{l,i}$, $m(\Delta_0) < \infty$.
2. If $l + 1 < R_i$, F: $\Delta_{l,i} \to \Delta_{l+1,i}$ is a measurable bijection and $F_* m_{|\Delta_{l,i}} = m_{|\Delta_{l+1,i}}$.
3. If $l + 1 = R_i$, F: $\Delta_{l,i} \to \Delta_0$ is a measurable bijection.
4. Let R: $\Delta_0 \to \mathbb{N}$ be the function $R_{|\Delta_{0,i}} = R_i$, and set

$$g = \frac{\mathrm{d}m_{|\Delta_0}}{\mathrm{d}m_{|\Delta_0} \circ F^R}.$$

g has a version for which $\exists C > 0, \theta \in (0,1)$ such that $\forall i$ and $\forall x, y \in \Delta_{0,i}$,

$$\left| \frac{g(x)}{g(y)} - 1 \right| \leqslant C \theta^{s(F^R x, F^R y)}$$

where $s(x, y) = \min\{n | \ (F^R)^n x, (F^R)^n y \text{ lie in different } \Delta_{0,j}\}$.

The fourth condition corresponds exactly to saying that the induced map on the base Δ_0 of the tower has a distortion which is locally Hölder continuous.

Henceforth, we assume for simplicity that $\int R dm < +\infty$ and that m is an F-invariant probability, which is possible because m has an integrable invariant density h such that $c_0^{-1} \leqslant h \leqslant c_0$ (see [You99, Theorem 1]).

Set $C_\theta(\Delta) = \{f : \Delta \to \mathbb{C} |\ \exists C \forall x, y \in \Delta, |f(x) - f(y)| \leqslant C\theta^{s(x,y)}\}$: this is the space of locally Hölder continuous functions (s has been extended to all pairs $x, y \in \Delta$ by setting $s(x, y) = 0$ if x, y are not in the same $\Delta_{l,i}$ and, for $x, y \in \Delta_{l,i}$, $s(x, y) = s(x', y')$ where x', y' are the corresponding points in $\Delta_{0,i}$).

Corollary 7.2: *Let $(\Delta, \mathcal{B}, m, F)$ be a probability preserving LS Young tower with $\gcd\{R_i\} = 1$ and $m[R > n] = O(1/n^\beta)$ where $\beta > 1$. If $f \in C_\theta(\Delta)$, $g \in L^\infty$ are supported inside $\bigcup_0^{N-1} \Delta_l$ for some N, then $\mathrm{Cor}(f, g \circ F^n) = \sum_{k>n} m[R > k] \int f \int g + O(F_\beta(n))$.*

Moreover, if $\int f = 0$, then $\mathrm{Cor}(f, g \circ F^n) = O(1/n^\beta)$. Thus, f satisfies a Central Limit Theorem.

Proof: For the partition $\{\Delta_{l,i}\}$, F does not have the big image property. However, it is still a Markov map for the partition $\{\Delta_l\}$ composed of the points at different heights. If $\gamma = \{\Delta_l |\ l < N\}$ for some N, then γ is finite, whence the induced map T_γ has the big image property.

For the induced map, the partition δ is constructed as follows: at each height $0 < l < N - 1$, cut Δ_l in two pieces $\Delta_l \cap F^{-1}\Delta_0$ and $\Delta_l - F^{-1}\Delta_0$. Δ_0 remains intact, and Δ_{N-1} is cut into all the small pieces $\Delta_{N-1,i}$. With this explicit partition, it is not hard to check that the induced map has $\theta^{1/N}$-locally Hölder continuous distortion.

Thus, Theorem 6.3 applies and gives an estimate

$$\mathrm{Cor}(f, g \circ F^n) = \sum_{k>n} m[\varphi_\gamma > k] \int f \int g + O(F_\beta(n)). \tag{13}$$

To finish the proof of the theorem, we have to show that $\sum_{k>n} m[\varphi_\gamma > k] = \sum_{k>n} m[R > k] + O(F_\beta(n))$. If f and g are supported in Δ_0 and of nonzero integral, Estimate (13) applies. Moreover, the estimate for $N = 1$ applies also. Equating these two estimates of $\mathrm{Cor}(f, g \circ F^n)$, we get the result. ∎

References

[Aar97] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Mathematical Surveys and Monographs, Volume 50, American Mathematical Society, Providence, RI, 1997.

[AD01] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps*, Stochastic and Dynamics **1** (2001), 193–237.

[ADU93] J. Aaronson, M. Denker and M. Urbański, *Ergodic theory for Markov fibred systems and parabolic rational maps*, Transactions of the American Mathematical Society **337** (1993), 495–548.

[Hen93] H. Hennion, *Sur un théorème spectral et son application aux noyaux lipschitziens*, Proceedings of the American Mathematical Society **118** (1993), 627–634.

[Hol02] M. Holland, *Slowly mixing systems and intermittency maps*, Preprint, 2002.

[Kat68] Y. Katznelson, *An Introduction to Harmonic Analysis*, Wiley, New York, 1968.

[Liv96] C. Liverani, *Central limit theorems for deterministic systems*, in *International Conference on Dynamical Systems, Montevideo 1995*, Pitman Research Notes in Mathematics, Volume 362, Longman Sci. Tech., Harlow, 1996.

[LSV99] C. Liverani, B. Saussol and S. Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory and Dynamical Systems **19** (1999), 671–683.

[New75] D. J. Newman, *A simple proof of Wiener's $1/f$ theorem*, Proceedings of the American Mathematical Society **48** (1975), 264–265.

[Rog73] B. A. Rogozin, *Asymptotic behavior of the coefficients in Levi–Wiener theorems on absolutely converging trigonometric series*, Sibirskii Matematicheskii Zhurnal **14** (1973), 1304–1312.

[Sar02] O. Sarig, *Subexponential decay of correlations*, Inventiones Mathematicae **150** (2002), 629–653.

[You99] L.-S. Young, *Recurrence times and rates of mixing*, Israel Journal of Mathematics **110** (1999), 153–188.