Commun. Math. Phys. 263, 461–512 (2006)
Digital Object Identifier (DOI) 10.1007/s00220-005-1511-6

Communications in **Mathematical Physics**

Limit Theorems in the Stadium Billiard

Péter Bálint[1], **Sébastien Gouëzel**[2]

[1] Institute of Mathematics, Budapest University of Technology and Economics, 1111 Egry József u. 1, Budapest, Hungary. E-mail: bp@renyi.hu
[2] IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France.
E-mail: sebastien.gouezel@univ-rennes1.fr

Received: 22 February 2005 / Accepted: 10 September 2005
Published online: 3 February 2006 –

Abstract: We prove that the Birkhoff sums for "almost every" relevant observable in the stadium billiard obey a non-standard limit law. More precisely, the usual central limit theorem holds for an observable if and only if its integral along a one-codimensional invariant set vanishes, otherwise a $\sqrt{n \log n}$ normalization is needed. As one of the two key steps in the argument, we obtain a limit theorem that holds in Young towers with exponential return time statistics in general, an abstract result that seems to be applicable to many other situations.

Introduction

The subject of this article, the stadium billiard, belongs to the class of dynamical systems that are sometimes referred to as intermittent ones. This name is related to the *weakly chaotic* nature of the time evolution that accounts for a modified, relaxed appearance of the behavior characteristic to systems with uniform hyperbolicity. In particular, the mathematically rigorous investigation of the stadium started with [Bun79] where Bunimovich showed (with respect to the natural invariant measure) that the Lyapunov exponents are almost everywhere non-zero, and that the system is ergodic. Thus in that respect the stadium billiard resembles dispersing billiards, however, when finer statistical properties are discussed, deviations start to show up. Recent works by Markarian ([Mar04]) and Chernov–Zhang ([CZ05]) have obtained an upper bound on the rate of mixing: given two sufficiently smooth (Hölder or Lipschitz continuous) observables, their correlations decay as $O((\log n)^2/n)$. Although this upper bound is most likely not sharp, it is definitely not far from the optimal either (see Corollary 1.3). In this paper we investigate the issue of probabilistic limit laws and provide further evidence of the intermittent nature of the dynamics. Namely we show that the limit behavior of a sufficiently smooth observable with zero mean, to be denoted by f_0, is characterized by a quantity I (cf. (1)), its average along the one dimensional set of trajectories bouncing forever along the straight segments. In the typical case $I \neq 0$, the Birkhoff sums of f_0

satisfy a non-standard limit theorem – convergence in distribution to the Gaussian law can be obtained with a $\sqrt{cn \log n}$ normalization, where the constant c is a multiple of I^2, see Theorem 1.1. On the other hand the central limit theorem in its usual form applies if $I = 0$, see Theorem 1.4. These results have some almost immediate corollaries: we obtain the analogous limit theorems for the billiard flow (Corollary 1.6) and, though in a very weak form, some lower bounds on the rate of correlation decay (Corollary 1.3).

The issue of probabilistic limit laws in dynamical systems has a long history. In the chaotic setting the possibly most frequently applied method is Gordin's martingale argument (see [Gor69], or [You98] and references therein) that roughly states that under quite general conditions, whenever the correlations decay at a summable rate, the usual central limit theorem holds. This technique, however, cannot treat non-standard limit behavior or non-summable decay rates. Recently Aaronson and Denker have proposed an approach to the issue of non-standard limit theorems, see e.g. [AD01]. The dynamical systems they study, the so-called Gibbs-Markov maps, possess some important features characteristic to uniformly expanding Markov maps of the interval, in particular, they are strongly chaotic. However, the functions f for which limit theorems are proved are unbounded, and do not even belong to L^2. This setting allows for the use of Perron-Frobenius techniques: there is a one parameter family of transfer operators the spectra of which give precise information on the limit behavior of the observable. In particular, the Birkhoff sums satisfy exactly the same limit theorem that an i.i.d. sequence of random variables with the distribution of f would have. For details see [AD01] and Sect. 3.1 of the present paper.

The above ideas can be implemented to treat limit laws for bounded functions in weakly chaotic systems $T_0 : X_0 \to X_0$ in case the following scenario applies. Let us assume that the source of non-uniformity in hyperbolicity is a well-distinguishable geometric effect. Then one may consider a subset $X \subset X_0$ such that the first return map onto X is uniformly hyperbolic, however, our observable induces an unbounded function on X. Thus we arrive at a setting close to that of [AD01]. This line of approach has been successfully applied to systems for which the induced map is Gibbs-Markov (see e.g. [Gou04]), which, however, is not exactly the case of the stadium billiard.

What replaces Gibbs-Markov property in billiards is the presence of a Young tower, an object that has turned out to be very effective when estimating the rate of the decay of correlations. There are two versions of Young towers: those with exponential return time statistics ensure rapid mixing – exponential decay of correlations – via Perron-Frobenius techniques ([You98]), while those with polynomial return time statistics give polynomial bounds on the rate of correlation decay – slow mixing rates – via coupling techniques ([You99]). As to the case of the stadium billiard, the Young towers constructed in [Mar04] and [CZ05] have polynomial return time statistics with respect to the original map, and exponential return time statistics with respect to the induced map. The aim of the present paper is, in addition to present our results on the stadium billiard, to demonstrate that Young towers, originally designed to estimate mixing rates, are almost equally powerful when the issue of various limit laws is investigated. Note that this fact has already been observed and emphasized by Szász and Varjú in the papers [SV04a] and [SV04b].

The proof of Theorem 1.1 consists of two clearly distinguishable ingredients. On the one hand, via Perron-Frobenius techniques, we prove Theorem 3.4, a general result in Young towers with exponential return time statistics. This concerns the limit behavior of the Birkhoff sums of observables belonging to the non-standard domain of attraction of the Gaussian law. It is important to note that, as the Gibbs-Markov property is replaced by a Young tower, a new effect shows up that typically rescales the normalizing sequence with a constant multiplicator. We would also like to emphasize that this first ingredient of the proof is completely general and could be applied to many other situations. On the other hand, the second ingredient is directly related to the stadium billiard. We rely on [Mar04] and [CZ05] when considering a suitable induced map that allows for a Young tower with exponential return time statistics. However, in order to "pull back" the limit theorem from the Young tower to the phase space of the billiard, and in order to give a transparent interpretation in terms of quantities easy to calculate, we need to perform a finer and more detailed geometric analysis of the stadium than the one presented in the above two papers.

We strongly do believe that our line of approach could be applied to obtain non-standard limit theorems in many other hyperbolic dynamical systems, in particular, in certain billiards with slow mixing rates. One of the most interesting candidates, the infinite horizon Lorentz process, for which the significance of the limit behavior is further emphasized as it may give an effective tool to discuss recurrence properties, is investigated by Szász and Varjú ([SV05]). Among others, it is also worth mentioning skewed stadia (see [CZ05]) and dispersing billiards with cusps ([Mac83]). We plan to turn back to these systems in separate papers.

The article has five sections. In the first one we state our main results and fix some basic notation. The second section is devoted to general results on the stadium billiard. We essentially recall the existence of Young towers for an induced map, proved by Markarian in [Mar04]. In the third part, we study abstract Young towers and establish a spectral perturbation estimate. In particular, to get a limit theorem, it is sufficient to study an integral with sufficient precision. In Sect. 4, we come back to the stadium billiard map, and describe geometrically this integral. With a careful study of the singularities of the stadium map, this gives an accurate description of this integral. Finally, in Sect. 5, we use together the abstract results of Sect. 3 and the explicit estimate of Sect. 4, to prove Theorem 1.1.

1. Results

Let $\ell > 0$. We consider a region in the plane delimited by two semicircles of radius 1, joined by two horizontal segments of length ℓ, tangent to the semicircles. To a point on the boundary of this set and a vector pointing inwards, we associate an image by the usual billiard reflection law. This defines the stadium billiard map $T_0 : X_0 \to X_0$. This map admits a unique absolutely continuous invariant probability measure μ_0.

A point in the phase space X_0 is given by (r, θ), where $r \in \mathbb{R}/(2\pi + 2\ell)\mathbb{Z}$ is the position on the boundary, and $\theta \in (-\pi/2, \pi/2)$ is the angle with respect to the normal to this boundary at r. The invariant measure μ_0 is given by

$$\mathrm{d}\mu_0 = \frac{\cos\theta \,\mathrm{d}r \,\mathrm{d}\theta}{2(2\pi + 2\ell)}.$$

We will assume that $r = 0$ corresponds to the lower endpoint of the right semi-circle, and that the boundary is oriented counterclockwise. Hence, the semicircles correspond to $0 \leqslant r \leqslant \pi$ and $\pi + \ell \leqslant r \leqslant 2\pi + \ell$.

Let $f_0 : X_0 \to \mathbb{R}$ be a Hölder function. We will be interested in the asymptotic behavior of the Birkhoff sums of f_0. The map T_0 is slowly mixing, by [Mar04] and [CZ05]: its correlations decay (at least) like $O((\log n)^2/n)$. This estimate is not summable, whence the usual Gordin martingale argument to get a central limit theorem does not apply. We will indeed prove that the usual central limit theorem does not hold.

Let

$$I = \frac{1}{2\ell}\left[\int_{r\in[\pi,\pi+\ell]} f_0(r,0)\,\mathrm{d}r + \int_{r\in[2\pi+\ell,2\pi+2\ell]} f_0(r,0)\,\mathrm{d}r\right]. \tag{1}$$

This is the average of f_0 along the trajectories bouncing perpendicularly to the segments of the stadium.

In this article, we prove the following theorem:

Theorem 1.1. *Let $f_0 : X_0 \to \mathbb{R}$ be Hölder continuous, satisfying $\int f_0\,\mathrm{d}\mu_0 = 0$ and $I \neq 0$. Then*

$$\frac{\sum_{k=0}^{n-1} f_0 \circ T_0^k}{\sqrt{cn\log n}} \to \mathcal{N}(0,1),$$

where

$$c = \frac{4+3\log 3}{4-3\log 3} \cdot \frac{\ell^2 I^2}{4(\pi+\ell)}.$$

Remark 1.2. Note that here, and throughout the paper, log means logarithm with respect to the natural base e. This is related to the fact that, at the level of concrete calculations, c is obtained as a sum that approximates the Riemann integral of the function $\frac{1}{x}$, see Sect. 4.

Corollary 1.3. *Under the assumptions of Theorem 1.1, the quantity $n\int f_0 \cdot f_0 \circ T_0^n$ does not tend to zero.*

Proof. We have

$$\int\left[\sum_{k=0}^{n-1} f_0 \circ T_0^k\right]^2 = n\int f_0^2 + 2\sum_{i=1}^{n-1}(n-i)\int f_0 \cdot f_0 \circ T^i.$$

If $\int f_0 \cdot f_0 \circ T^i = o(1/i)$, we obtain $\int\left[\sum_{k=0}^{n-1} f_0 \circ T_0^k\right]^2 = o(n\log n)$. In particular, the variance of the random variable $\frac{\sum_{k=0}^{n-1} f_0\circ T_0^k}{\sqrt{n\log n}}$ tends to zero. This implies that this random variable tends to zero in probability, which is in contradiction with Theorem 1.1. □

This corollary indicates that the upper bound $O((\log n)^2/n)$ on the decay of correlations, proved by Markarian and Chernov-Zhang, is close to optimal (it may probably be replaced by $O(1/n)$, since the $(\log n)^2$ seems to be due to the technique of proof).

We also obtain the following (easier) result:

Theorem 1.4. *Let $f_0 : X_0 \to \mathbb{R}$ be Hölder continuous, satisfying $\int f_0\,\mathrm{d}\mu_0 = 0$ and $I = 0$. Then there exists $\sigma^2 \geqslant 0$ such that*

$$\frac{\sum_{k=0}^{n-1} f_0 \circ T_0^k}{\sqrt{n}} \to \mathcal{N}(0,\sigma^2). \tag{2}$$

Moreover, $\sigma^2 = 0$ if and only if there exists a measurable function $\chi_0 : X_0 \to \mathbb{R}$ such that $f_0 = \chi_0 - \chi_0 \circ T_0$ almost everywhere.

Hence, when $I = 0$, the Birkhoff sums of f_0 satisfy a usual central limit theorem.

Before going into the details of the proof we consider one particularly interesting observable: the free path. Given $x = (r, \theta)$, we denote $T_0 x = (r_1, \theta_1)$ and define $\tau_0(x)$ as the planar distance of r and r_1. In other words, $\tau_0(x)$ is the length of the trajectory segment the point particle follows until the next collision. To investigate the limit behavior of the free path $\tau_0 : X_0 \to \mathbb{R}$, we have to subtract its mean $E(\tau_0) = \int \tau_0 \, \mathrm{d}\mu_0$, thus we define $\tau_0^*(x) = \tau_0(x) - E(\tau_0)$. There is a remarkably simple formula for $E(\tau_0)$ that can be obtained by comparing the invariant measures for the billiard map and the billiard flow (see [Che97]):

$$E(\tau_0) = \frac{\pi(\pi + 2\ell)}{2\ell + 2\pi}. \tag{3}$$

On the other hand, we may easily calculate (1) as we have $\tau_0(r, 0) = 2$ whenever $r \in [\pi, \pi + \ell]$ or $r \in [2\pi + \ell, 2\pi + 2\ell]$, thus $I_{\tau_0} = 2$ and $I_{\tau_0^*} = 2 - \frac{\pi(\pi+2\ell)}{2\ell+2\pi}$. This means there is a "best" stadium with $\ell = \ell^* = \frac{4\pi-\pi^2}{2\pi-4} \approx 1.18$ for which $I_{\tau_0^*} = 0$ and consequently, by Theorem 1.4 the (centralized) free path satisfies the usual central limit theorem. However, whenever $\ell \neq \ell^*$, we have $I_{\tau_0^*} \neq 0$ and, by Theorem 1.1 a stronger normalization is needed.

It is interesting to know when the central limit theorem of Theorem 1.4 is degenerate, i.e., when $\sigma = 0$. The coboundary condition $f_0 = \chi_0 - \chi_0 \circ T_0$ is not easy to manipulate, since it is valid only almost everywhere, but it is nevertheless quite restrictive, by Livšic-like arguments. For example, we can prove the following:

Proposition 1.5. *In the stadium with $\ell = \ell^*$, the free path satisfies a non-degenerate central limit theorem, i.e., $\sigma \neq 0$.*

This proposition will be proved in Subsect. 5.3.

Our interest in τ_0 is also related to the fact that the billiard flow may be considered as a suspension above the billiard map with the roof function $\tau_0(x)$. By [MT04] suspension flows do inherit some statistical properties from the base transformation, in particular limit theorems, under quite general conditions. Let us denote the billiard flow by

$$X_\tau = \{(x, u) \mid x \in X_0, 0 \le u \le \tau_0(x)\}/\sim, \; (x, \tau_0(x)) \sim (T_0 x, 0),$$
$$S_t(x, u) = (x, u + t), \; \mu_\tau = \mu_0 \times \frac{\text{Leb}}{E(\tau_0)},$$

where the action of the flow is understood modulo identifications. Consider a Hölder observable $\Phi : X_\tau \to \mathbb{R}$ satisfying $\int \Phi \, \mathrm{d}\mu_\tau = 0$, and define

$$\Phi_T(x) = \int_0^T \Phi(S_t x) \, \mathrm{d}t;$$
$$J_\Phi = \frac{1}{2\ell} \left[\int_{r \in [\pi, \pi+\ell] \cup [2\pi+\ell, 2\pi+2\ell]} \int_{t \in [0,2]} \Phi(r, 0, t) \, \mathrm{d}t \, \mathrm{d}r \right].$$

Corollary 1.6. *1. If $J_\Phi \neq 0$, then*

$$\frac{\Phi_T}{\sqrt{\frac{c}{E(\tau_0)} T \log T}} \to \mathcal{N}(0, 1).$$

Here c is the constant from Theorem 1.1, with I replaced by J_Φ.

2. If $J_\Phi = 0$, then

$$\frac{\Phi_T}{\sqrt{T}} \to \mathcal{N}(0, \sigma_\Phi^2)$$

for some $\sigma_\Phi^2 \geqslant 0$.

Proof. Define $f_0 : X_0 \to \mathbb{R}$ as $f_0(x) = \int_0^{\tau_0(x)} \Phi(S_t(x, 0))\,\mathrm{d}t$. The function f_0 is Hölder, $\int f_0 \,\mathrm{d}\mu_0 = E(\tau_0) \int \Phi \,\mathrm{d}\mu_\tau = 0$ and $I_{f_0} = J_\Phi$. Hence, depending on the value of J_Φ, one of our two main theorems applies. To show that Φ inherits the limit behavior from f_0, we apply the flow version of [Gou03, Theorem A1] recalled as Theorem 5.1 in this paper (see also Remark 5.2). We only need to check that the three conditions of this theorem are satisfied. In case $J_\Phi \neq 0$ (and even if $J_\Phi = 0$ and $\ell = \ell^*$) Conditions 1 and 3 are satisfied with $b = 1$. Then Condition 2 is merely the Birkhoff ergodic theorem, thus the first statement is established. If $J_\Phi = 0$, the appropriate normalization for τ_0 may be $\sqrt{n \log n}$ as opposed to $\sqrt{n}$ needed for f_0. Thus Conditions 1 and 3 of Theorem 5.1 are satisfied for any $0 < b < 1$, but not for $b = 1$. This means Condition 2 has to be established for some $b < 1$, but this is merely our Remark 5.6. This completes the proof of the second statement. □

We will say that a Hölder continuous function $f_0 : X_0 \to \mathbb{R}$ with vanishing integral satisfies $(P1)$ if $I \neq 0$ and f_0 vanishes on the set of points x such that x, $T_0(x)$ and $T_0^{-1}(x)$ belong to the same semicircle, and that f_0 satisfies $(P2)$ if $I = 0$. We will in fact prove Theorem 1.1 for functions satisfying $(P1)$, and Theorem 1.4 for functions satisfying $(P2)$. This will imply Theorem 1.1 in full generality. Namely, if f_0 is Hölder continuous and satisfies $I \neq 0$, then we may write it as $f_0 = f_1 + f_2$, where f_1 satisfies $(P1)$ and f_2 satisfies $(P2)$. By Theorem 1.4, $\frac{S_n f_2}{\sqrt{n \log n}} \to 0$. Hence, it is equivalent to have Theorem 1.1 for f_0 or f_1. We will comment on the technical reason for introducing the classes (P1) and (P2) in Remark 2.4 below.

In this paper, C will denote a generic constant, that can change from one occurrence to the next. Some constants, which will be used at different places in the paper, will be denoted by $C_1, C_2, \dots$ and will have a fixed value.

2. Background Material on the Stadium Billiard

2.1. Geometric description of the initial map and of an induced map. The map T_0 has almost everywhere two nonzero Lyapunov exponents. However, the expansion in the unstable cone (and the contraction in the stable cone) are not uniform: points bouncing many times between the segments, or sliding along the circles, have an expansion arbitrarily close to 1.

To get uniform expansion, we follow [Mar04] and [CZ05]. First let us note that the phase space X_0, defined by two variables, the periodic position coordinate r, and the angular velocity coordinate θ with values in the interval $(-\pi/2, \pi/2)$, has the shape of a cylinder. Let X be the set of points x in X_0 such that x belongs to a semicircle and $T^{-1}(x)$ does not belong to this semicircle. It is not hard to check geometrically that

$$X = \bigcup_{r \in (0,\pi)} \{r\} \times (-r/2, \pi/2 - r/2) \cup \bigcup_{r \in (0,\pi)} \{r + \pi + \ell\} \times (-r/2, \pi/2 - r/2).$$

In particular,

$$\mu_0(X) = 2\int_{r=0}^{\pi}\int_{\theta=-r/2}^{\pi/2-r/2} \frac{\cos\theta \,\mathrm{d}r\,\mathrm{d}\theta}{4\pi+4\ell} = \frac{2}{\pi+\ell}. \tag{4}$$

Define a new probability measure on X by

$$\mathrm{d}\mu = \frac{\cos\theta\,\mathrm{d}r\,\mathrm{d}\theta}{8}.$$

For $x \in X$, let $\varphi_+(x) = \inf\{n \geqslant 1, T_0^n(x) \in X\}$. This is the return time of x. Let $T : X \to X$ be the first return map, induced by T_0 on X, i.e., $T(x) = T_0^{\varphi_+(x)}(x)$. This map preserves the probability measure μ on X. Moreover, it is uniformly hyperbolic in the following sense:

Proposition 2.1. *There exists a continuous family of closed cones $\mathcal{C}^u(x)$ for $x \in X$, such that $DT(x)(\mathcal{C}^u(x)) \subset \mathcal{C}^u(Tx)$. Moreover, there exist constants $\Lambda > 1$ and $C > 0$ such that, for all $x \in X$, for all $v \in \mathcal{C}^u(x)$, for all $n \in \mathbb{N}$ such that T^n is defined and differentiable at x,*

$$\left\| DT^n(x)v \right\| \geqslant C\Lambda^n \left\| v \right\|.$$

Moreover, these cones are uniformly bounded away from the horizontal and vertical directions (i.e., $\{\mathrm{d}\theta = 0\}$ and $\{\mathrm{d}r = 0\}$). In the same way, there exist stable cones $\mathcal{C}^s(x)$, which satisfy the same properties for T^{-1}, except that they are not bounded away from the horizontal direction.

This proposition can be found in [Mar04] and [CZ05]. For future reference we recall some ideas from its proof. In billiard theory there are two different ways of measuring the length of a stable or an unstable vector $v = (\delta r, \delta\theta)$. The euclidean metric is defined as $\|v\| = \sqrt{|\delta r|^2 + |\delta\theta|^2}$ while the p-metric as $\|v\|_p = |\delta r| \cos\theta$. Geometrically, the p-metric measures distances orthogonally to the flow direction along the wavefront that starts out of the one parameter family of phase points corresponding to the tangent vector v. Furthermore, despite being degenerate on the full tangent space, it is well-defined on both the stable and the unstable cone and it satisfies $\|v\|_p \leqslant \|v\|$. Actually, the uniform expansion of Proposition 2.1 is proved for the p-metric in [Mar04] and [CZ05]. However, it is easy to check that there exists $C > 0$ such that, for all $x \in X$, for all $v \in \mathcal{C}^u(x)$, $\|DT(x)v\|_p \geqslant C\|v\|$. Hence, the uniform expansion in the p-metric implies the same statement for the euclidean metric, up to a constant C.

Convention 2.2. Unless otherwise stated, in billiard related calculations we use the euclidean metric throughout the paper.

There are two different types of points for which $\varphi_+(x)$ can be large: they correspond to points bouncing many times between the segments, or sliding many times along the circles. We will need to describe rather precisely the hyperbolic behavior of T in bouncing regions. The proposition below can be checked by direct calculation, see also [CZ05] and references therein.

Proposition 2.3. *If x is a bouncing point satisfying $\varphi_+(x) = n$, then T contracts the p-metric of vectors in the stable cone at least by a factor $\frac{C}{n}$, while T^{-1} contracts the p-metric of vectors in the unstable cone at least by a factor $\frac{C}{n}$. Moreover, Tx and $T^{-1}x$ are bouncing points with $\varphi_+(Tx) \geqslant n/4$, $\varphi_+(T^{-1}x) \geqslant n/4$ if n is large enough. This implies, in turn, that the above contraction estimates are valid in the euclidean metric as well.*

Remark 2.4. Note that if x is a sliding point satisfying $\varphi_+(x) = n$, then we can only guarantee that Tx and $T^{-1}x$ are sliding points with $\varphi_+(Tx) \geqslant C\sqrt{n}$ and $\varphi_+(T^{-1}x) \geqslant C\sqrt{n}$. This has an unfortunate consequence: we can only apply the coboundary arguments of Sect. 2.3 to functions vanishing along sliding trajectories. Essentially this is the technical reason for introducing the classes (P1) and (P2). The proof of Theorem 1.1 relies heavily on Perron–Frobenius techniques, and thus requires an expanding setting, which implies that collapsing along stable manifolds – coboundary arguments – are essential. Now for the class (P2) it is enough to prove the usual central limit theorem (Theorem 1.4) which can be carried out in a roundabout way in the hyperbolic setting, see Sect. 5.2.

2.2. Young tower of T. A compact subset $R \subset X$ is a *rectangle* if there exist $x \in R$ with a local stable manifold $W^s_{loc}(x)$ and a local unstable manifold $W^u_{loc}(x)$, and two Cantor sets $C^s \subset W^s_{loc}(x)$ and $C^u \subset W^u_{loc}(x)$, such that, for any $y_s \in C^s$ and $y_u \in C^u$, then y_s has a local unstable manifold $W^u_{loc}(y_s)$ and y_u has a local stable manifold $W^s_{loc}(y_u)$. Moreover, these two local manifolds intersect at exactly one point, and this point belongs to R.

An s-subrectangle of R is a set $\left(\bigcup_{y\in C} W^s_{loc}(y)\right) \cap R$, where C is a compact subset of C^u. A u-subrectangle is defined in the same way.

[Mar04] and [CZ05] have proved that $T : X \to X$ satisfies Chernov's axioms of [Che99]. This implies that it admits a hyperbolic Young tower in the following sense: there exist a rectangle R of positive measure, a partition $R = \bigcup R_i$ (modulo 0) by s-subrectangles, and return times $r_i \in \mathbb{N}$ such that T^{r_i} is a homeomorphism on R_i, and $T^{r_i}(R_i)$ is a u-subrectangle of R. Moreover, the tails of the tower are exponentially small: there exist $\rho < 1$ and $C > 0$ such that

$$\forall n \in \mathbb{N}, \mu\left(\bigcup_{r_i > n} R_i\right) \leqslant C\rho^n.$$

We can then define an abstract space $\bar{\Delta}$ as the disjoint union of the sets $T^k(R_i)$ for $i \in \mathbb{N}$ and $k < r_i$. It is endowed with a natural projection $\pi_X : \bar{\Delta} \to X$ and a dynamics $\bar{U} : \bar{\Delta} \to \bar{\Delta}$ such that $\pi_X \circ \bar{U} = T \circ \pi_X$.

Young has proved in [You98] that it is possible to construct on $\bar{\Delta}$ a probability measure $\mu_{\bar{\Delta}}$ which is invariant under $\bar{U}$ and such that $(\pi_X)_*(\mu_{\bar{\Delta}}) = \mu$. Note however that π_X is in general strongly not injective, so that $\mu_{\bar{\Delta}}$ can not be defined as the pullback of μ. Rather, one constructs an invariant measure for $\bar{U}$, and one proves that its projection, being absolutely continuous with respect to μ and T-invariant, is necessarily μ.

It is then useful to go from this abstract hyperbolic dynamics to an abstract expanding dynamics. To do so, one identifies the points of $\bar{\Delta}$ which are on the same stable leaf in some rectangle. This defines a space Δ, together with a projection $\pi_\Delta : \bar{\Delta} \to \Delta$. Since the map $\bar{U}$ sends stable leaves to stable leaves, it gives rise to a dynamics $U : \Delta \to \Delta$ on the quotient. The measure $\mu_\Delta := (\pi_\Delta)_*(\mu_{\bar{\Delta}})$ is invariant under U. Then (Δ, U, μ_Δ) is an *expanding Young tower*, in the sense of Sect. 3.2.

2.3. Coboundary results. Let $f_0 : X_0 \to \mathbb{R}$ be a Hölder function satisfying $(P1)$, for which we want to prove a limit theorem. Since it is easier to work in an expanding and well understood setting, we will first prove results in Δ, and then go back from Δ to X_0.

For $x \in X$, let first $f(x) = \sum_{k=0}^{\varphi_+(x)-1} f_0(T_0^k x)$. This function is not bounded any more. However, if two points x and y are on a local stable manifold which is not cut by a discontinuity of T_0 during the next n iterates of T_0, and with $\varphi_+(x) = \varphi_+(y) = n$, then

$$|f(x) - f(y)| \leqslant Cnd(x, y)^\alpha \tag{5}$$

for some $\alpha > 0$. In the same way, if x and y are on a local unstable manifold which is not cut during the next n iterates of T_0, and $\varphi_+(x) = \varphi_+(y) = n$, then $|f(x) - f(y)| \leqslant Cnd(Tx, Ty)^\alpha$. Moreover, the property $(P1)$ implies that f is bounded on the set of points sliding along the semicircles.

Let $X_n \subset X$ be the set of points bouncing exactly n times between the segments. Note that $\mu(X_n) \sim \frac{\ell^2}{4n^3}$. On X_n, the function f can be interpreted by means of a Riemann sum approximating the integral of f over the set of points bouncing perpendicularly to the segments of the stadium, with better and better precision as n increases. Thus the function f is equivalent to nI on X_n (where I is defined in (1)). Since $I \neq 0$ by assumption, we obtain

$$\mu\{x \mid |f(x)| \geqslant n\} \sim \sum_{n/|I|}^{\infty} \frac{\ell^2}{4k^3} \sim \frac{I^2 \ell^2}{8n^2}. \tag{6}$$

Hence, the distribution of f is in the nonstandard domain of attraction of the Gaussian law (see Paragraph 3.1).

Define a function $\bar{f}$ on $\bar{\Delta}$ by $\bar{f} = f \circ \pi_X$. It would be easy to go finally from $\bar{\Delta}$ to Δ if $\bar{f}$ were constant along the local stable leaves in $\bar{\Delta}$ (which would mean that $\bar{f}$ would induce a function on the quotient Δ). This is in general not the case, but we will prove that $\bar{f}$ is cohomologous to such a function, using the usual cohomology trick.

For every rectangle in $\bar{\Delta}$, choose a definite unstable leaf. Define a projection $\pi : \bar{\Delta} \to \bar{\Delta}$ by sliding along stable manifolds to this specific unstable manifold. We define a function $\bar{u}(x) = \sum_{k=0}^{\infty} \left[\bar{f}(\bar{U}^k x) - \bar{f}(\bar{U}^k \pi x)\right]$. Note that, despite the fact that T contracts stable manifolds uniformly, the function $\bar{u}(x)$ may not seem well-defined at first sight, as $\bar{f}$ and its Hölder constant are unbounded. Nevertheless, whenever $\bar{f}$ is large, T contracts stable manifolds strongly, and the Hölder constant can be regained by going down the tower. This is the essence of the following lemma.

Lemma 2.5. *The function $\bar{u}$ is well defined and bounded on $\bar{\Delta}$.*

Proof. In this proof the positive constants C do depend on the Hölder exponent α, but this has no significance. Let $K \in \mathbb{N}$ be such that $\alpha K > 1$. Consider first x which is at height $\geqslant K$ in the tower. Let $y = \pi x$. Let $x' = \bar{U}^{-K} x$ and $y' = \bar{U}^{-K} y$. We will prove that

$$\forall k \in \mathbb{N}, |\bar{f}(\bar{U}^k x) - \bar{f}(\bar{U}^k y)| \leqslant Cd(\pi_X \bar{U}^k x', \pi_X \bar{U}^k y')^\alpha. \tag{7}$$

Namely, if $\varphi_+(\pi_X \bar{U}^k x) = n$, then

$$\begin{aligned} |\bar{f}(\bar{U}^k x) - \bar{f}(\bar{U}^k y)| &\leqslant Cnd(\pi_X \bar{U}^k x, \pi_X \bar{U}^k y)^\alpha \\ &= Cnd(\pi_X \bar{U}^{k+K} x', \pi_X \bar{U}^{k+K} y')^\alpha \end{aligned} \tag{8}$$

by (5). If $n = \varphi_+(\pi_X \bar{U}^k x)$ is bounded, the conclusion is trivial. If n is large and $\pi_X \bar{U}^k x$ is a sliding point, the conclusion is also trivial by $(P1)$.

Hence, assume that n is large and that $\pi_X \bar{U}^k x$ is a bouncing point. Proposition 2.3 implies that, for $0 \leqslant i < K$, $\varphi_+(\pi_X \bar{U}^{k+i} x') \geqslant n/4^{K-i} \geqslant n/4^K$. Once again by Proposition 2.3, we get

$$d(\pi_X \bar{U}^{k+i+1} x', \pi_X \bar{U}^{k+i+1} y') \leqslant \frac{C}{n} d(\pi_X \bar{U}^{k+i} x', \pi_X \bar{U}^{k+i} y').$$

Hence,

$$d(\pi_X \bar{U}^{k+K} x', \pi_X \bar{U}^{k+K} y') \leqslant \frac{C}{n^K} d(\pi_X \bar{U}^{k} x', \pi_X \bar{U}^{k} y').$$

Together with (8) and the inequality $K\alpha > 1$, this implies (7).

Since $\pi_X x'$ and $\pi_X y'$ are on a local stable manifold, $d(\pi_X \bar{U}^k x', \pi_X \bar{U}^k y')$ goes exponentially fast to zero. Hence, the series $\sum |\bar{f}(\bar{U}^k x) - \bar{f}(\bar{U}^k y)|$ is summable, and $\bar{u}(x)$ is well defined.

Suppose now that x is at height $< K$. Let $y = \pi x$. Applying the previous argument to $x' = x$ and $y' = y$, we get that $\sum_{k=K}^{\infty} |\bar{f}(\bar{U}^k x) - \bar{f}(\bar{U}^k y)|$ is bounded. Moreover, during the first K iterates, x and y remain at a bounded height in the tower, which implies that $\bar{f}(\bar{U}^k x)$ and $\bar{f}(\bar{U}^k y)$ remain uniformly bounded. This concludes the proof. □

Let $\bar{g}(x) = \bar{f}(x) - \bar{u}(x) + \bar{u}(\bar{U}x)$. Then

$$\bar{g}(x) = \bar{f}(\pi x) + \sum_{k=0}^{\infty} \Big[\bar{f}(\bar{U}^k(\bar{U}\pi x)) - \bar{f}(\bar{U}^k(\pi \bar{U} \pi x)) \Big].$$

Hence, $\bar{g}(x)$ depends only on πx, i.e., $\bar{g}$ is constant along the stable manifolds in the rectangles. Consequently, there exists a function $g : \Delta \to \mathbb{R}$ such that $\bar{g} = g \circ \pi_\Delta$.

It will be important that g is regular enough on Δ, to use functional analytic techniques. For $x_1, x_2 \in \Delta$, let $s(x_1, x_2)$ be their separation time, i.e., the number of returns to the basis before x_1 and x_2 get into different elements of the partition. To obtain the following lemma, we will use several times the same argument as in Lemma 2.5, but sometimes along unstable manifolds instead of stable ones.

Lemma 2.6. *There exist $C > 0$ and $\tau < 1$ such that, for every x_1, x_2 in the same element of partition of Δ,*

$$|g(x_1) - g(x_2)| \leqslant C \tau^{s(x_1, x_2)}.$$

Proof. Let us first prove that, if x_1, x_2 belong to the same unstable leaf in a rectangle of $\bar{\Delta}$, then

$$\bar{g}(x_1) - \bar{g}(x_2) \text{ is uniformly bounded.} \tag{9}$$

The same argument as in the proof of Lemma 2.5 shows that $\sum_{k=0}^{\infty} \big[\bar{f}(\bar{U}^k(\bar{U}\pi x)) - \bar{f}(\bar{U}^k(\pi \bar{U} \pi x)) \big]$ is bounded. Hence, it is sufficient to prove that $\bar{f}(\pi x_1) - \bar{f}(\pi x_2)$ is bounded. Let K be as in the proof of Lemma 2.5. If x_1 (and x_2) return to the basis of $\bar{\Delta}$ before time K, then $\varphi_+(x_1) = \varphi_+(x_2)$ is bounded, which implies that $\bar{f}(\pi x_1)$ and $\bar{f}(\pi x_2)$ are bounded. If x_1 (and x_2) are sliding points, then the conclusion is also a consequence of $(P1)$. Otherwise, x_1 and x_2 are bouncing points. We show as in the proof of Lemma 2.5 (but along the *unstable* leaf containing πx_1 and πx_2) that $|\bar{f}(\pi x_1) - \bar{f}(\pi x_2)| \leqslant C d(\pi_X \bar{U}^K \pi x_1, \pi_X \bar{U}^K \pi x_2)^\alpha$. Since this quantity is uniformly bounded, this concludes the proof of (9).

Take $x_1, x_2 \in \bar{\Delta}$ on the same unstable leaf, and let $s = s(\pi_\Delta(x_1), \pi_\Delta(x_2))$. We will prove that

$$|\bar{g}(x_1) - \bar{g}(x_2)| \leqslant C\lambda^{\alpha s/2} \tag{10}$$

for some $C > 0$, where $\lambda < 1$ is larger than the contraction coefficient of T along stable manifolds, and the contraction coefficient of T^{-1} along unstable manifolds.

By (9), this is trivial if $s < 2K$. Hence, we can assume $s \geqslant 2K$. Let $N = \left\lfloor \frac{s}{2} \right\rfloor \geqslant K$, then

$$\begin{aligned}\bar{g}(x_1) - \bar{g}(x_2) = {} & \bar{f}(\pi x_1) - \bar{f}(\pi x_2) \\ & + \sum_{k=0}^{N-1} \left[\bar{f}(\bar{U}^k(\bar{U}\pi x_1)) - \bar{f}(\bar{U}^k(\bar{U}\pi x_2)) \right] \\ & + \sum_{k=0}^{N-1} \left[\bar{f}(\bar{U}^k(\pi\bar{U}\pi x_2)) - \bar{f}(\bar{U}^k(\pi\bar{U}\pi x_1)) \right] \\ & + \sum_{k=N}^{\infty} \left[\bar{f}(\bar{U}^k(\bar{U}\pi x_1)) - \bar{f}(\bar{U}^k(\pi\bar{U}\pi x_1)) \right] \\ & + \sum_{k=N}^{\infty} \left[\bar{f}(\bar{U}^k(\pi\bar{U}\pi x_2)) - \bar{f}(\bar{U}^k(\bar{U}\pi x_2)) \right].\end{aligned} \tag{11}$$

Since $N + K \leqslant s$, we have for any $k < N$,

$$\begin{aligned}\left| \bar{f}(\bar{U}^k(\bar{U}\pi x_1)) - \bar{f}(\bar{U}^k(\bar{U}\pi x_2)) \right| & \leqslant Cd(\pi_X \bar{U}^{k+K}(\bar{U}\pi x_1), \pi_X \bar{U}^{k+K}(\bar{U}\pi x_2))^\alpha \\ & \leqslant C\lambda^{\alpha(s-(k+K+1))} d(\pi_X \bar{U}^s \pi x_1, \pi_X \bar{U}^s \pi x_2)^\alpha \\ & \leqslant C\lambda^{\alpha(s-k)}.\end{aligned}$$

Summing over k, we obtain

$$\left| \sum_{k=0}^{N-1} \left[\bar{f}(\bar{U}^k(\bar{U}\pi x_1)) - \bar{f}(\bar{U}^k(\bar{U}\pi x_2)) \right] \right| \leqslant C\lambda^{\alpha(s-N)} \leqslant C\lambda^{\alpha s/2}.$$

The term on the third line of (11) can be estimated in the same way, as well as the term on the first line of (11).

Since $N \geqslant K$, we also have for any $k \geqslant N$,

$$\begin{aligned}\left| \bar{f}(\bar{U}^k(\bar{U}\pi x_1)) - \bar{f}(\bar{U}^k(\pi\bar{U}\pi x_1)) \right| & \leqslant Cd(\pi_X \bar{U}^{k-K}(\bar{U}\pi x_1), \pi_X \bar{U}^{k-K}(\pi\bar{U}\pi x_1))^\alpha \\ & \leqslant C\lambda^{\alpha(k-K)} d(\pi_X(\bar{U}\pi x_1), \pi_X(\pi\bar{U}\pi x_1))^\alpha \\ & \leqslant C\lambda^{\alpha k}.\end{aligned}$$

Summing over k, we obtain

$$\left| \sum_{k=N}^{\infty} \left[\bar{f}(\bar{U}^k(\bar{U}\pi x_1)) - \bar{f}(\bar{U}^k(\pi\bar{U}\pi x_1)) \right] \right| \leqslant C\lambda^{\alpha N} \leqslant C\lambda^{\alpha s/2}.$$

The term on the fifth line of (11) is handled in the same way. □

3. Limit Theorems in Young Towers

3.1. A result by Aaronson and Denker. A function $f : \mathbb{R}_+^* \to \mathbb{R}_+^*$ is *slowly varying* if, for all $\lambda > 0$, $f(\lambda x)/f(x)$ tends to 1 when $x \to \infty$.

By classical probabilistic results, a real random variable Z is in the nonstandard domain of attraction of the Gaussian distribution $\mathcal{N}(0, 1)$ if and only if it satisfies one of the following equivalent conditions:

- The function $L(x) := E(Z^2 1_{|Z| \leqslant x})$ is unbounded and slowly varying.
- $P(|Z| > x) \sim x^{-2} l(x)$ for some function l such that $\widetilde{L}(x) := 2\int_1^x \frac{l(u)}{u} \, du$ is unbounded and slowly varying.

Remark 3.1. In this case, $\widetilde{L}(x) \sim L(x)$ when $x \to \infty$, and $l(x) = o(L(x))$. It is possible, however, that l is not slowly varying and that these conditions hold anyway.

Such a random variable belongs to L^p for all $1 \leqslant p < 2$, but not to L^2. We will say that l and L are the *tail functions* of Z. They are defined up to asymptotic equivalence. Choose a sequence $B_n \to \infty$ such that $\frac{n}{B_n^2} L(B_n) \to 1$. Then, if $Z_0, Z_1, \dots$ is a sequence of independent random variables distributed as Z, then

$$\frac{Z_0 + \cdots + Z_{n-1} - nE(Z)}{B_n} \to \mathcal{N}(0, 1).$$

More generally, if $\frac{n}{B_n^2} L(B_n) \to C > 0$, then the previous sequence converges to $\mathcal{N}(0, C)$.

In [AD01], Aaronson and Denker have proved the same kind of limit theorem when the sequence $Z_0, Z_1, \dots$ is not independent. More precisely, consider U a mixing Gibbs-Markov map (as defined in [Aar97]) on a space Δ, preserving a probability measure μ_Δ, and let $g : \Delta \to \mathbb{R}$ be a function which is locally Hölder and whose distribution with respect to μ_Δ is in the nonstandard domain of attraction of $\mathcal{N}(0, 1)$ as above. Then they prove that

$$\frac{g + g \circ U + \cdots + g \circ U^{n-1} - n \int g}{B_n} \to \mathcal{N}(0, 1)$$

as above.

The proof goes as follows: let $\widehat{U}$ be the transfer operator associated to U, and $\widehat{U}_t$ its perturbation given by $\widehat{U}_t u = \widehat{U}(e^{itg} u)$. These operators satisfy a Lasota-Yorke inequality on the space of Hölder functions, and $\left\| \widehat{U}_t - \widehat{U} \right\| = O(t)$. Hence, the eigenvalue λ_t of $\widehat{U}_t$ close to 1 satisfies $|\lambda_t - 1| = O(t)$, and the corresponding eigenfunction w_t (normalized so that $\int w_t = 1$) is such that $\|w_t - 1\| = O(t)$.

Then they prove the abstract lemma below, which has nothing to do with dynamics and could be stated on any probability space whenever g has the appropriate distribution.

Lemma 3.2. *For any bounded measurable function w on Δ,*

$$\int (e^{itg} - 1 - itg) w = -\frac{t^2}{2} \int 1_{|g| \leqslant 1/|t|} g^2 w + \|w\|_\infty o(t^2 L(1/|t|)).$$

Here, the $o(t^2 L(1/|t|))$ is uniform in w.

Applying this lemma to w_t, one gets

$$\lambda_t - 1 - it \int g w_t = \int (e^{itg} - 1 - itg) w_t = -\frac{t^2}{2} \int 1_{|g| \leqslant 1/|t|} g^2 w_t + o(t^2 L(1/|t|))$$

(where we have used the fact that w_t is bounded). Since $\|w_t - 1\|_\infty = o(1)$,

$$\frac{t^2}{2} \int 1_{|g| \leqslant 1/|t|} g^2 w_t = \frac{t^2}{2} L(1/|t|)(1 + o(1)).$$

Hence,

$$\lambda_t = 1 + it \int g w_t - \frac{t^2}{2} L(1/|t|)(1 + o(1)). \tag{12}$$

Finally, $\int g w_t = \int g + O(t)$ since $\|w_t - 1\|_\infty = O(t)$. So we get

$$\lambda_t = 1 + it \int g - \frac{t^2}{2} L(1/|t|)(1 + o(1)).$$

This expansion is sufficient to get the required limit theorem.

3.2. The result in Young towers. Let (Δ, μ_Δ) be a probability space and $U : \Delta \to \Delta$ a probability preserving map. We say that (Δ, U) is an *expanding Young tower* ([You99]) if there exist integers $r_p \in \mathbb{N}^*$ and a partition $\{\Delta_{k,p}\}_{p \in \mathbb{N}, k \in \{0,\dots,r_p - 1\}}$ of Δ such that

1. For all p and $k < r_p - 1$, U is a measurable isomorphism between $\Delta_{k,p}$ and $\Delta_{k+1,p}$, preserving μ_Δ.
2. For all p, U is a measurable isomorphism between $\Delta_{r_p - 1, p}$ and $\Delta_0 := \bigcup_m \Delta_{0,m}$.
3. Let U_0 be the first return map induced by U on Δ_0. For $x, y \in \Delta_0$, define their separation time $s(x, y) = \inf\{n \in \mathbb{N} \mid U_0^n(x)$ and $U_0^n(y)$ are not in the same $\Delta_{0,p}\}$. We extend this separation time to the whole tower in the following way: if x, y are not in the same element of partition, set $s(x, y) = 0$. Otherwise, $x, y \in \Delta_{k,p}$. Let $x', y' \in \Delta_{0,p}$ be such that $x = U^k x'$ and $y = U^k y'$, and set $s(x, y) = s(x', y')$.
   For $x \in \Delta$, let $J(x)$ be the inverse of the jacobian of U at x. We assume that there exist $\beta < 1$ and $C > 0$ such that, for all x, y in the same element of partition,

$$\left| 1 - \frac{J(x)}{J(y)} \right| \leqslant C \beta^{s(Ux, Uy)}. \tag{13}$$

Remark 3.3. Note that the definition of separation time in [You98] is in terms of the number of all iterations of U, while we follow the convention of [You99] when we define separation in terms of returns to the basis. Hence, our setting is more general than that of [You98], but it will make the proof of the spectral gap more complicated.

Let $\Delta_n = \bigcup \Delta_{n,p}$. This is the set of points at height n in the tower. We will say that (U, Δ) is an *expanding Young tower with exponentially small tail* if there exists $\rho < 1$ such that $\mu_\Delta(\Delta_n) = O(\rho^n)$.

Let $J^{(n)}$ be the inverse of the jacobian of U^n. It is standard that (13) implies that the distortion of the iterates of U is uniformly bounded, in the following sense: there exists

$C > 0$ such that, for all points x, y such that $U^k x$ and $U^k y$ remain in the same elements of the partition for $0 \leqslant k < n$,

$$\left|1 - \frac{J^{(n)}(x)}{J^{(n)}(y)}\right| \leqslant C\beta^{s(U^n x, U^n y)}. \tag{14}$$

A function $g : \Delta \to \mathbb{R}$ is *locally Hölder* if there exist $C > 0$ and $\tau < 1$ such that $|g(x) - g(y)| \leqslant C\tau^{s(x,y)}$ for all x, y in the same element of the partition. This is exactly the type of functions that arise from the stadium billiard, cf. Lemma 2.6. Note that g can very well be unbounded. Without loss of generality, we can assume $\tau \geqslant \beta$.

Let $\omega(x)$ be the height of the point x, i.e., $\omega(x) = n$ if $x \in \Delta_n$. Let $\pi_0 : \Delta \to \Delta_0$ be the projection to the basis, and define a function G on Δ by $G(x) = \sum_{k=0}^{\omega(x)-1} g(U^k \pi_0 x)$. In this setting, we get the following extension of the theorem proved by Aaronson and Denker:

Theorem 3.4. *Let $U : \Delta \to \Delta$ be an expanding Young tower with exponentially small tail, and let $g : \Delta \to \mathbb{R}$ be locally Hölder continuous. Assume that the distribution of g is in the nonstandard domain of attraction of $\mathcal{N}(0, 1)$, with tail functions l and L. Assume moreover that l and L are slowly varying, and $l(x \ln x)/l(x) \to 1$, $L(x \ln x)/L(x) \to 1$ when $x \to \infty$. Finally, assume that there exists a real number $a \neq -1/2$ such that*

$$\int g(e^{itG} - 1) = (a + o(1))itL(1/|t|)) \text{ when } t \to 0. \tag{15}$$

Write $L_1(x) = (2a + 1)L(x)$, and choose a sequence $B_n \to \infty$ such that $\frac{n}{B_n^2} L_1(B_n) \to 1$. Then

$$\frac{S_n g - n \int g}{B_n} \to \mathcal{N}(0, 1).$$

The additional assumption on l and L is satisfied in most natural cases (for example when $l = 1$ and $L = \ln$, which will be the case for the stadium billiard).

When $a = 0$, we get the same asymptotics as in Aaronson-Denker's Theorem. However, when $a \neq 0$, then there is an additional effect due to the presence of the tower. Theorem 3.4 discusses the case when the two effects are of the same order of magnitude. The constant a reflects the proportion of the two effects: its value is intrinsic and does not change, for example, when the function g is multiplied by a constant factor. In principle one could imagine $a = -1/2$ which could result in the two effects cancelling out, however, a negative a is not very likely to be realizable in a dynamical situation – this would mean that the value of g high up in the tower and its sum for the levels below the level considered are negatively correlated.

The proof will follow the same lines as Aaronson-Denker's proof: it is possible to construct a good space on which the transfer operator $\widehat{U}$ has a spectral gap. The perturbed operator $\widehat{U}_t$ also has a spectral gap, which gives an eigenvalue λ_t and an eigenfunction w_t. The main problem is that $\left\|\widehat{U}_t - \widehat{U}\right\|$ can not be $O(t)$ in general: it is easy to construct examples where $t = o\left(\left\|(\widehat{U}_t - \widehat{U})1\right\|_{L^2}\right)$, whence $t = o\left(\left\|\widehat{U}_t - \widehat{U}\right\|\right)$ as soon as the good space has a norm stronger than the L^2 norm, and contains the function 1.

Using abstract arguments by Keller and Liverani, we can nevertheless prove that $|\lambda_t - 1| = O(|t|^{1/10})$ and $\|w_t - 1\|_{L^1} = O(|t|^{1/10})$. This is (essentially) sufficient to apply Aaronson and Denker's argument and get $\lambda_t = 1 + it \int g w_t - \frac{t^2}{2} \int 1_{|g| \leqslant 1/|t|} g^2 w_t +$

$o(t^2 L(1/|t|))$ as in (12). The main difficulty is then to make the function w_t disappear in this expression, to get something more tractable. We will namely show that $\int 1_{|g|\leqslant 1/|t|} g^2 w_t \sim L(1/|t|)$ and $\int g w_t = \int g e^{itG} + o(tL(1/|t|))$, which will conclude the proof.

To do this, we need to know that $w_t - 1 = O(t)$ in some sense. To prove such an estimate, we use a roundabout technical argument relying on the fact that the induced map on the basis of the tower is uniformly expanding, to prove that $\left\| 1_{\Delta_0}(w_t - 1) \right\|_\infty = O(t)$, and then we propagate this information up in the tower, using the information we have already proved on λ_t. This propagation requires the Birkhoff sums of g to be small enough. To ensure this on a set of large measure, we use the information on the tails of g. This is the only point where the additional information on l and L is used.

3.3. Proof of Theorem 3.4. We will first prove Theorem 3.4 assuming that $\int g = 0$. In Paragraph 3.3.5, we will show that this implies the theorem in full generality. Hence, *until the end of Paragraph 3.3.4, we will assume that* $\int g = 0$.

3.3.1. Construction of the functional spaces and the transfer operators. Since the tails of the tower are exponentially small by assumption, there exists $\rho < 1$ such that $\mu_\Delta(\Delta_n) \leqslant C\rho^n$. Denote the return time to the basis from itself by φ. Take $\varepsilon > 0$ such that $e^{6\varepsilon}\rho < 1$.

For $u : \Delta \to \mathbb{C}$, write

$$\|u\|_m = \inf\{C \mid \forall n \in \mathbb{N}, \text{for almost every } x \in \Delta_n, |u(x)| \leqslant Ce^{\varepsilon n}\}$$

and

$$\|u\|_l = \inf\{C \mid \text{for almost every } x, y \text{ in the same element of the partition at height } n, \\ |u(x) - u(y)| \leqslant Ce^{\varepsilon n}\tau^{s(x,y)}\}.$$

Denote by $\mathcal{H}$ the space of measurable functions u on Δ for which $\|u\| := \|u\|_m + \|u\|_l < +\infty$. It is a Banach space included in L^1 (and even in L^6 because of the condition $e^{6\varepsilon}\rho < 1$). This inclusion is compact.

The following proposition is similar to a result of Young:

Proposition 3.5. *There exist $C > 0$ and $\theta < 1$ such that, for any $u \in \mathcal{H}$, for any $n \in \mathbb{N}$,*

$$\left\| \widehat{U}^n u \right\| \leqslant C\theta^n \|u\| + C \|u\|_{L^1}.$$

Note that our definition of separation time is not the same as in [You98], and that Young uses the fact that the return to the basis only occur after a large time N. This gives her a strong expansion, sufficient to get rid of constants easily. This is not true in our setting. Hence, the proof of the proposition will be more involved than Young's.

Proof. Take $x \in \Delta_0$. Then $\widehat{U}^n u(x) = \sum J^{(n)}(x_p)u(x_p)$, where the set $\{x_p\}_{p\in\mathbb{N}}$ is the set of all preimages of x under U^n. Let A_p containing x_p be such that $U^n : A_p \to \Delta_0$ is an isomorphism. Then $J^{(n)}(x_p) \leqslant C\mu_\Delta(A_p)$ since the distortion is bounded, by (14). Let ω_p be the height of the set A_p and r_p the number of returns of A_p to the basis before time n.

For $y \in A_p$, $s(x_p, y) \geqslant r_p$, whence

$$|u(x_p) - u(y)| \leqslant \tau^{r_p} e^{\varepsilon\omega_p} \|u\|_l .$$

Hence,

$$|u(x_p)| \leqslant \tau^{r_p} e^{\varepsilon\omega_p} \|u\|_l + \frac{1}{\mu_\Delta(A_p)} \int_{A_p} |u|. \tag{16}$$

We get

$$|\widehat{U}^n u(x)| \leqslant C \sum \mu_\Delta(A_p) \tau^{r_p} e^{\varepsilon\omega_p} \|u\|_l + C \int |u|. \tag{17}$$

Let $\omega : \Delta \to \mathbb{N}$ be the function "height", and let $\Psi_n(x)$ be the number of returns of x to the basis between time 1 and n. Then (17) implies that

$$|\widehat{U}^n u(x)| \leqslant C \|u\|_l \int_{U^{-n}\Delta_0} \tau^{\Psi_n} e^{\varepsilon\omega} + C \|u\|_{L^1}. \tag{18}$$

We will use the following technical lemma, which will be proved in the Appendix.

Lemma 3.6. *There exist $C > 0$ and $\theta < 1$ such that, for any $n \in \mathbb{N}$,*

$$\int_{U^{-n}\Delta_0} \tau^{\Psi_n} e^{\varepsilon\omega} \leqslant C\theta^n. \tag{19}$$

Increasing θ if necessary, we can assume that $e^{-\varepsilon} \leqslant \theta$.

This lemma, together with (18), implies that, for any $x \in \Delta_0$,

$$|\widehat{U}^n u(x)| \leqslant C\theta^n \|u\|_l + C \|u\|_{L^1}. \tag{20}$$

Consider now $x \in \Delta$ such that $\omega(x) = k < n$. Let x' be its projection in the basis. Then $\widehat{U}^n u(x) = \widehat{U}^{n-k} u(x')$, whence

$$\begin{aligned} e^{-\varepsilon k} |\widehat{U}^n u(x)| &= e^{-\varepsilon k} |\widehat{U}^{n-k} u(x')| \\ &\leqslant e^{-\varepsilon k} C\theta^{n-k} \|u\|_l + Ce^{-\varepsilon k} \|u\|_{L^1} \leqslant C\theta^n \|u\|_l + C \|u\|_{L^1}. \end{aligned} \tag{21}$$

Assume finally that $\omega(x) = k \geqslant n$. Let $x' = U^{-n}(x)$; it satisfies $\omega(x') = k - n$. Then

$$e^{-\varepsilon k} |\widehat{U}^n u(x)| = e^{-\varepsilon n} e^{-\varepsilon(k-n)} |u(x')| \leqslant e^{-\varepsilon n} \|u\|_m. \tag{22}$$

These equations prove that

$$\left\| \widehat{U}^n u \right\|_m \leqslant C\theta^n \|u\| + C \|u\|_{L^1}.$$

We still have to handle the Hölder norm. Consider two points x, y in the same element of partition of the basis Δ_0. Let x_p and y_p be their preimages, in sets A_p as above. Then

$$\begin{aligned} |\widehat{U}^n u(x) - \widehat{U}^n u(y)| &\leqslant \sum |J^{(n)}(x_p) u(x_p) - J^{(n)}(y_p) u(y_p)| \\ &\leqslant \sum |J^{(n)}(x_p)| |u(x_p) - u(y_p)| \\ &\quad + \sum |J^{(n)}(x_p)| \left| 1 - \frac{J^{(n)}(y_p)}{J^{(n)}(x_p)} \right| |u(y_p)|. \end{aligned}$$

In the first sum, $|J^{(n)}(x_p)| \leqslant C\mu_\Delta(A_p)$ and $|u(x_p) - u(y_p)| \leqslant \tau^{s(x,y)+r_p} e^{\varepsilon\omega_p} \|u\|_l$. Hence,

$$\sum |J^{(n)}(x_p)||u(x_p) - u(y_p)| \leqslant C\tau^{s(x,y)} \|u\|_l \int_{U^{-n}\Delta_0} \tau^{\Psi_n} e^{\varepsilon\omega} \leqslant C\theta^n \tau^{s(x,y)} \|u\|_l$$

by Lemma 3.6.

In the second sum, $|J^{(n)}(x_p)| \leqslant C\mu_\Delta(A_p)$ and $\left|1 - \frac{J^{(n)}(y_p)}{J^{(n)}(x_p)}\right| \leqslant C\tau^{s(x,y)}$ by (14). Moreover, $|u(y_p)|$ is bounded by (16). Using these inequalities, we get

$$\begin{aligned}
&\sum |J^{(n)}(x_p)| \left|1 - \frac{J^{(n)}(y_p)}{J^{(n)}(x_p)}\right| |u(y_p)| \\
&\leqslant \sum C\mu_\Delta(A_p)\tau^{s(x,y)} \left[\tau^{r_p} e^{\varepsilon\omega_p} \|u\|_l + \frac{1}{\mu_\Delta(A_p)} \int_{A_p} |u|\right] \\
&\leqslant C\tau^{s(x,y)} \|u\|_l \int_{U^{-n}\Delta_0} \tau^{\Psi_n} e^{\varepsilon\omega} + C\tau^{s(x,y)} \int |u| \\
&\leqslant C\theta^n \tau^{s(x,y)} \|u\|_l + C\tau^{s(x,y)} \|u\|_{L^1}
\end{aligned}$$

by Lemma 3.6.

To sum up, we have proved that, when x and y belong to the same partition element of the basis,

$$\frac{|\widehat{U}^n u(x) - \widehat{U}^n u(y)|}{\tau^{s(x,y)}} \leqslant C\theta^n \|u\|_l + C \|u\|_{L^1}.$$

Let now x and y belong to the same element of the partition, with $k = \omega(x) < n$. Let x' and y' be their projections in the basis. Then

$$\begin{aligned}
e^{-\varepsilon k} \frac{|\widehat{U}^n u(x) - \widehat{U}^n u(y)|}{\tau^{s(x,y)}} &= e^{-\varepsilon k} \frac{|\widehat{U}^{n-k} u(x') - \widehat{U}^{n-k} u(y')|}{\tau^{s(x',y')}} \\
&\leqslant e^{-\varepsilon k} \left[C\theta^{n-k} \|u\|_l + C \|u\|_{L^1}\right] \\
&\leqslant C\theta^n \|u\|_l + C \|u\|_{L^1}.
\end{aligned}$$

Assume finally that $k \geqslant n$. Let $x' = U^{-n}x$ and $y' = U^{-n}y$. Then

$$e^{-\varepsilon k} \frac{|\widehat{U}^n u(x) - \widehat{U}^n u(y)|}{\tau^{s(x,y)}} = e^{-\varepsilon n} e^{-\varepsilon(k-n)} \frac{|u(x') - u(y')|}{\tau^{s(x',y')}} \leqslant \theta^n \|u\|_l.$$

Summing up these equations, we get $\left\|\widehat{U}^n u\right\|_l \leqslant C\theta^n \|u\|_l + C \|u\|_{L^1}$. This concludes the proof of the proposition. □

Let g be the locally Hölder function for which we want to prove a limit theorem. It is possible that $g \notin \mathcal{H}$, since $\|g\|_m$ is not necessarily finite.

Define a perturbed transfer operator, à la Nagaev, by $\widehat{U}_t(u) = \widehat{U}(e^{itg}u)$.

Proposition 3.7. *There exist constants $C > 0$ and $\theta < 1$ such that, for all $t \in [-1, 1]$, for all $u \in \mathcal{H}$, for all $n \in \mathbb{N}$,*

$$\left\|\widehat{U}_t^n u\right\| \leqslant C\theta^n \|u\| + C \|u\|_{L^1}.$$

This proposition contains Proposition 3.5 as a special case, for $t = 0$.

Proof. Let $x \in \Delta$. Then $\widehat{U}_t^n u(x) = \sum_{U^n y = x} e^{itS_n g(y)} J^{(n)}(y) u(y)$, whence $|\widehat{U}_t^n u(x)| \leqslant \widehat{U}^n |u|(x)$. The bound on $\left\| \widehat{U}^n |u| \right\|_m$ thus implies the required bound on $\left\| \widehat{U}_t^n u \right\|_m$.

For the Hölder norm, take x and y two points in the same element of partition. Then, with the notations of the proof of Proposition 3.5,

$$\begin{aligned} \left|\widehat{U}_t^n u(x) - \widehat{U}_t^n u(y)\right| &= \left|\sum e^{itS_n g(x_p)} J^{(n)}(x_p) u(x_p) - e^{itS_n g(y_p)} J^{(n)}(y_p) u(y_p)\right| \\ &\leqslant \sum |J^{(n)}(x_p) u(x_p) - J^{(n)}(y_p) u(y_p)| \\ &\quad + \sum |e^{itS_n g(x_p)} - e^{itS_n g(y_p)}| |J^{(n)}(x_p)| |u(x_p)|. \end{aligned}$$

The first sum has already been estimated in the proof of Proposition 3.5. For the second one, $|e^{itS_n g(x_p)} - e^{itS_n g(y_p)}| \leqslant nC\tau^{s(x,y)}$. Hence, Proposition 3.5 implies that

$$\left\| \widehat{U}_t^n u \right\| \leqslant C(n+1)\theta^n \|u\| + C(n+1) \|u\|_{L^1}.$$

Choose $N > 0$ such that $\bar{\theta} := C(N+1)\theta^N < 1$. Iterating the equation $\left\| \widehat{U}_t^N u \right\| \leqslant \bar{\theta} \|u\| + C \|u\|_{L^1}$ (and using the fact that $\left\| \widehat{U}_t^N u \right\|_{L^1} \leqslant \|u\|_{L^1}$), we get

$$\left\| \widehat{U}_t^{nN} u \right\| \leqslant \bar{\theta}^n \|u\| + \frac{C}{1-\bar{\theta}} \|u\|_{L^1}.$$

This implies the conclusion of the proposition, for the constant $\bar{\theta}^{1/N} < 1$. □

Lemma 3.8. *When $t \to 0$, $\|\widehat{U}_t - \widehat{U}\|_{\mathcal{H} \to L^3} = O(|t|^{1/6})$.*

Proof. For $u \in \mathcal{H}$, $(\widehat{U}_t - \widehat{U})u = \widehat{U}((e^{itg} - 1)u)$. The transfer operator $\widehat{U}$ is a contraction in every L^p space, and in particular in L^3. Hence,

$$\left\| (\widehat{U}_t - \widehat{U})u \right\|_{L^3} \leqslant \left\| (e^{itg} - 1)u \right\|_{L^3} \leqslant \left\| e^{itg} - 1 \right\|_{L^6} \|u\|_{L^6}.$$

Note that $\|u\|_{L^6} \leqslant C \|u\|$. Hence, $\|\widehat{U}_t - \widehat{U}\|_{\mathcal{H} \to L^3} = O(\left\| e^{itg} - 1 \right\|_{L^6})$. To estimate this quantity, choose $C > 0$ such that, for all $x \in \mathbb{R}$, $|e^{ix} - 1| \leqslant C|x|^{1/6}$. Then $\int |e^{itg} - 1|^6 \leqslant C \int |tg| = O(|t|)$. Hence, $\left\| e^{itg} - 1 \right\|_{L^6} = O(|t|^{1/6})$. □

3.3.2. Definition of λ_t, first estimates. By Proposition 3.5 and Hennion's Theorem [Hen93], the operator $\widehat{U} : \mathcal{H} \to \mathcal{H}$ is quasicompact: outside of the disk $\{|z| \leqslant \theta\}$, its spectrum is composed of discrete eigenvalues of finite multiplicity. In particular, by ergodicity, 1 is a simple isolated eigenvalue of $\widehat{U}$, with multiplicity one (and the corresponding eigenfunction is the constant function 1).

Lemma 3.8 is *not* a continuity statement in $\mathcal{H}$. However, the operators $\widehat{U}$ and $\widehat{U}_t$ satisfy a uniform Lasota-Yorke inequality between $\mathcal{H}$ and L^3, by Proposition 3.7 (and the fact that $\|u\|_{L^1} \leqslant \|u\|_{L^3}$). Hence, we can apply the abstract results of [KL99, Corollary 1] (following ideas of [BY93]). We get the following:

For small enough t, $\widehat{U}_t$ has a unique eigenvalue λ_t close to 1, and it satisfies $|\lambda_t - 1| = O(|t|^{1/10})$. Let P_t be the corresponding eigenprojection. Then $\|P_t\|_{\mathcal{H} \to \mathcal{H}}$ is bounded when $t \to 0$. Moreover, $\|P_t - P_0\|_{\mathcal{H} \to L^3} = O(|t|^{1/10})$.

Remark 3.9. Here, $1/10$ could be replaced by any exponent $< 1/6$, but any positive exponent would be sufficient for our purposes.

Let $\bar{w}_t := P_t 1$, and write $w_t = \frac{\bar{w}_t}{\int \bar{w}_t}$. Then w_t is bounded in $\mathcal{H}$ and

$$\|w_t - 1\|_{L^3} = O(|t|^{1/10}). \tag{23}$$

Lemma 3.10. *When $t \to 0$,*

$$\lambda_t = 1 - \frac{t^2}{2}\int 1_{|g|\leqslant 1/|t|} g^2 w_t + it\int g w_t + o(t^2 L(1/|t|)). \tag{24}$$

Proof. By definition, $\widehat{U}_t(w_t) = \lambda_t w_t$. Integrating, we get

$$\lambda_t = \int e^{itg} w_t. \tag{25}$$

We want to use Lemma 3.2 to estimate this integral. However, this lemma applies only to bounded functions. Hence, we will have to modify w_t.

Take $x \in \Delta$ with $\omega(x) > 0$, and let $x' = U^{-1}(x)$. The equation $\widehat{U}_t w_t = \lambda_t w_t$ implies that $e^{itg(x')} w_t(x') = \lambda_t w_t(x)$. Hence, $|w_t(x)| = |\lambda_t|^{-1} |w_t(x')|$. Since w_t is uniformly bounded on the basis of the tower (since it is bounded in $\mathcal{H}$), we get

$$|w_t(x)| \leqslant C|\lambda_t|^{-\omega(x)}. \tag{26}$$

Define a function w'_t by $w'_t(x) = w_t(x)$ if $\omega(x) \leqslant |t|^{-1/10}$ and $w'_t(x) = 0$ otherwise. It belongs to $\mathcal{H}$, with $\|w'_t\| \leqslant \|w_t\|$. Since $\lambda_t = 1 + O(|t|^{1/10})$, (26) implies that

$$\|w'_t\|_\infty \leqslant C(1 + C|t|^{1/10})^{|t|^{-1/10}} \leqslant C'.$$

Lemma 3.2 applied to w'_t gives

$$\int (e^{itg} - 1 - itg) w'_t = -\frac{t^2}{2}\int 1_{|g|\leqslant 1/|t|} g^2 w'_t + o(t^2 L(1/|t|)). \tag{27}$$

Let us show that this equation is also satisfied by $w''_t := w_t - w'_t$. First,

$$\begin{aligned}\left|\int (e^{itg} - 1) w''_t\right| &\leqslant 2\int_{\omega \geqslant |t|^{-1/10}} |w_t| \\ &\leqslant 2\int_{\omega \geqslant |t|^{-1/10}} (1 + C|t|^{1/10})^\omega \leqslant C \sum_{n=|t|^{-1/10}}^{\infty} \rho^n (1 + C|t|^{1/10})^n.\end{aligned}$$

When t is small enough, $\rho(1 + C|t|^{1/10}) < \sqrt{\rho} < 1$. Hence, $|\int (e^{itg} - 1) w''_t| \leqslant C\rho^{|t|^{-1/10}/2} = o(t^2 L(1/|t|))$. In the same way, $\left|\int g w''_t\right| \leqslant \|g\|_{L^{3/2}} \|w''_t\|_{L^3}$ and $\|w''_t\|_{L^3}$ decays stretched exponentially, whence it is $o(t^2 L(1/|t|))$. Finally,

$$\left|\int 1_{|g|\leqslant 1/|t|} g^2 w''_t\right| \leqslant \frac{1}{|t|^2}\int |w''_t| = O(\rho^{|t|^{-1/10}/2}/t^2) = o(t^2 L(1/|t|)).$$

Hence, (27) holds also for w''_t. We get

$$\int (e^{itg} - 1 - itg) w_t = -\frac{t^2}{2}\int 1_{|g|\leqslant 1/|t|} g^2 w_t + o(t^2 L(1/|t|)).$$

Since $\int e^{itg} w_t = \lambda_t$ and $\int w_t = 1$, this gives the conclusion of the lemma. □

Corollary 3.11. *We have* $\lambda_t = 1 + O(|t|^{11/10})$.

Proof. In the proof of the previous lemma, we have proved that $\int_{\omega>|t|^{-1/10}} 1_{|g|\leqslant 1/|t|} g^2 w_t = O(t)$. Moreover, on $\{\omega \leqslant |t|^{-1/10}\}$, the function w_t is uniformly bounded. Hence, $\int_{\omega\leqslant|t|^{-1/10}} 1_{|g|\leqslant 1/|t|} g^2 w_t \leqslant C \int 1_{|g|\leqslant 1/|t|} g^2 \sim CL(1/|t|)$. Hence,

$$\int 1_{|g|\leqslant 1/|t|} g^2 w_t = O(L(1/|t|)).$$

Moreover, by (23) and our assumption $\int g = 0$,

$$\left|\int g w_t\right| = \left|\int g(w_t - 1)\right| \leqslant \|g\|_{L^{3/2}} \|w_t - 1\|_{L^3} = O(|t|^{1/10}).$$

This proves that, in (24), the right side is $1 + O(|t|^{11/10})$. □

3.3.3. Estimates on the basis. To proceed, we will need to know that w_t is constant on the basis up to $O(t)$. We already know that $\|w_t - 1\|_{L^3} = O(|t|^{1/10})$, but this is not sufficient to estimate precisely the terms in (24). To get such an estimate, we will need real continuity, and not only the weak continuity given by Keller-Liverani's theorem. This will be achieved by working directly on the basis. The goal of this paragraph is to prove Lemma 3.15.

Let U_0 be the map induced by U on the basis Δ_0 of the tower. Denote by φ the first return time to the basis, so that $U_0(x) = U^{\varphi(x)}(x)$.

Let us consider the space $\mathcal{H}_0$ of Hölder functions $u : \Delta_0 \to \mathbb{C}$ on the basis, and define an operator $R_n : \mathcal{H}_0 \to \mathcal{H}_0$ by $R_n u(x) = \sum J^{(n)}(y) u(y)$, where the sum is restricted to those $y \in \Delta_0$ with return time $\varphi(y) = n$, and $U^n(y) = x$. Set also $R_n(t)(u) = R_n(e^{itS_n g} u)$.

Lemma 3.12. *There exist* $C > 0$ *and* $\theta < 1$ *such that, for all* $n \in \mathbb{N}$ *and all* $t \in [-1, 1]$, $\|R_n(t)\| \leqslant C\theta^n$ *and* $\|R_n(t) - R_n\| \leqslant C\theta^n |t|$.

Proof. The map U_0 is Gibbs-Markov on Δ_0. Hence, [Gou04, Lemma 3.2] proves that $\|R_n\| \leqslant C\mu_\Delta(\varphi = n)$ and [Gou04, Lemma 3.5] yields $\|R_n(t) - R_n\| \leqslant C|t|n\mu_\Delta(\varphi = n) + C\int_{\{\varphi=n\}} |e^{itS_n g} - 1|$.

Since $\mu_\Delta(\varphi = n) = O(\rho^n)$, we get in particular $\|R_n\| \leqslant C\rho^n$, which decays exponentially. Moreover, on $\{\varphi = n\}$, $|S_n g|^{3/2} \leqslant n^{1/2} \sum_{k=0}^{n-1} |g \circ U^k|^{3/2}$, whence $\int_{\{\varphi=n\}} |S_n g|^{3/2} \leqslant n^{1/2} \int_\Delta |g|^{3/2} = O(n^{1/2})$. Consequently,

$$\begin{aligned}\|R_n(t) - R_n\| &\leqslant C|t|n\rho^n + C\int 1_{\varphi=n}|t||S_n g|\\ &\leqslant C|t|n\rho^n + C|t| \left\|1_{\varphi=n} S_n g\right\|_{L^{3/2}} \left\|1_{\varphi=n}\right\|_{L^3},\end{aligned}$$

which decays also exponentially. □

For $|z| < \theta^{-1}$, it is possible to define $R(z, t) := \sum z^n R_n(t)$. The operator $R(1, 0)$ is the transfer operator associated to U_0. It has a simple isolated eigenvalue at 1, and the corresponding eigenfunction is the constant function 1. Hence, for (z, t) close enough to $(1, 0)$, $R(z, t)$ has a unique eigenvalue $\lambda(z, t)$ close to 1.

Lemma 3.13. *There exists $C > 0$ such that, for all $z \in \mathbb{C}$ with $|z| \leqslant \theta^{-1/2}$ and all $t \in [-1, 1]$, we have $\|R(z,t) - R(1,0)\| \leqslant C(|t| + |z-1|)$.*

Proof. We have $R(z,t) - R(1,t) = \sum(z^n - 1)R_n(t)$. Moreover, if $|z| \leqslant \theta^{-1/2}$, $|z^n - 1| \leqslant |z-1| \sum_{k=0}^{n-1} |z|^k \leqslant C|z-1|\theta^{-n/2}$. Hence,

$$\|R(z,t) - R(1,t)\| \leqslant \sum_{n=0}^{\infty} |z^n - 1| \, \|R_n(t)\|$$
$$\leqslant C|z-1| \sum_{n=0}^{\infty} \theta^{-n/2}\theta^n \leqslant \frac{C}{1-\theta^{1/2}}|z-1|.$$

Moreover,

$$\|R(1,t) - R(1,0)\| \leqslant \sum_{n=0}^{\infty} \|R_n(t) - R_n\| \leqslant \sum_{n=0}^{\infty} C|t|\theta^n \leqslant \frac{C}{1-\theta}|t|. \quad \square$$

Lemma 3.14. *For small enough t, $R(\lambda_t^{-1}, t)(1_{\Delta_0} w_t) = 1_{\Delta_0} w_t$.*

Proof. Let $x \in \Delta_0$, let $\{x_p\}$ be the set of its preimages under U, at respective heights ω_p, and let x'_p be the projection of x_p in the basis. Since $\widehat{U}_t w_t = \lambda_t w_t$, we have $\lambda_t w_t(x) = \sum e^{itg(x_p)} J(x_p) w_t(x_p)$. Moreover, for any $y \in \Delta$ with $\omega(y) > 0$, we have $\lambda_t w_t(y) = e^{itg(U^{-1}y)} w_t(U^{-1}y)$. Hence, $\lambda_t^{\omega_p} w_t(x_p) = e^{itS_{\omega_p} g(x'_p)} w_t(x'_p)$. We get

$$w_t(x) = \sum \lambda_t^{-\omega_p - 1} J^{(\omega_p+1)}(x'_p) e^{itS_{\omega_p+1} g(x'_p)} w_t(x'_p). \tag{28}$$

The points x'_p are exactly the preimages of x under U_0, and the corresponding return time for U is $\omega_p + 1$. Hence, (28) gives the conclusion of the lemma. $\square$

We have all the necessary tools to prove the main result of this paragraph:

Lemma 3.15. *For $t \in [-1, 1]$, there exists $c(t) \in \mathbb{C}$ such that $\left\| 1_{\Delta_0}(w_t - c(t)) \right\|_\infty = O(t)$. Moreover, $c(t) \to 1$ when $t \to 0$.*

Proof. Lemma 3.14 proves that, for small enough t, $\lambda(\lambda_t^{-1}, t) = 1$, and the corresponding eigenfunction is $1_{\Delta_0} w_t$. Let Q_t be the eigenprojection of $R(\lambda_t^{-1}, t)$ corresponding to the eigenvalue 1. It satisfies

$$\|Q_t - Q_0\| = O\left(\left\| R(\lambda_t^{-1}, t) - R(1,0) \right\|\right) = O(|\lambda_t^{-1} - 1| + |t|) = O(|t|)$$

by Lemma 3.13 and Corollary 3.11. Let $b_t = Q_t 1_{\Delta_0}$. As $b_0 = 1_{\Delta_0}$, b_t satisfies $\left\| b_t - 1_{\Delta_0} \right\| = O(t)$. In particular, $b_t \to 1_{\Delta_0}$ in L^1.

The function b_t is proportional to w_t on the basis Δ_0. Hence, there exists a scalar $c(t)$ such that $1_{\Delta_0} w_t = c(t) b_t$. Since w_t goes to 1 in L^1 when $t \to 0$, we get

$$c(t) = \frac{\int 1_{\Delta_0} w_t}{\int b_t} \to \frac{\mu_\Delta(\Delta_0)}{\int b_0} = 1.$$

Finally,

$$\left\| 1_{\Delta_0}(w_t - c(t)) \right\|_\infty = |c(t)| \, \|b_t - b_0\|_\infty = O(t). \quad \square$$

3.3.4. Proof of Theorem 3.4 when $\int g = 0$. Let the function G be given by $G(x) = \sum_{k=0}^{\omega(x)-1} g(U^k \pi_0 x)$, as in Theorem 3.4.

Lemma 3.16. *When $t \to 0$,*

$$\lambda_t = 1 - (1+o(1))\frac{t^2}{2}\int 1_{|g|\leqslant 1/|t|} g^2 e^{itG} + it(1+o(1)) \int g e^{itG} + o(t^2 L(1/|t|)). \tag{29}$$

Proof. We will start from (24) and show that we can replace w_t by e^{itG}.

We have $w_t(x) = \lambda_t^{-\omega(x)} e^{itG(x)} w_t(\pi_0 x)$. Hence, by Corollary 3.11 and Lemma 3.15,

$$\begin{aligned}
|w_t(x) - c(t)e^{itG(x)}| &= |\lambda_t^{-\omega(x)} w_t(\pi_0 x) - c(t)| \\
&\leqslant |\lambda_t^{-\omega(x)} - 1||w_t(\pi_0 x)| + |w_t(\pi_0 x) - c(t)| \\
&\leqslant \left[(1 + C|t|^{11/10})^{\omega(x)} - 1\right] C + C|t| \\
&\leqslant \omega(x) C |t|^{11/10} (1 + C|t|^{11/10})^{\omega(x)} + C|t|.
\end{aligned}$$

Fix $b > 0$ large enough. For $\omega(x) \leqslant b\log(1/|t|)$, we obtain $|w_t(x) - c(t)e^{itG(x)}| \leqslant C|t|$. For $\omega(x) \geqslant b\log(1/|t|)$ and small enough t, we also get $|w_t(x) - c(t)e^{itG(x)}| \leqslant \rho^{-\omega(x)/4}$.

Hence,

$$\begin{aligned}
\int_{\omega \geqslant b\log(1/|t|)} 1_{|g|\leqslant 1/|t|} g^2 |w_t - c(t)e^{itG}| &\leqslant \int_{\omega \geqslant b\log(1/|t|)} \frac{1}{|t|^2} \rho^{-\omega(x)/4} \\
&\leqslant \frac{1}{|t|^2} C \sum_{n = b\log(1/|t|)}^{\infty} \rho^n \rho^{-n/4} = o(1)
\end{aligned}$$

if b is large enough. Moreover,

$$\int_{\omega \leqslant b\log(1/|t|)} 1_{|g|\leqslant 1/|t|} g^2 |w_t - c(t)e^{itG}| \leqslant C|t| \int 1_{|g|\leqslant 1/|t|} g^2 = C|t| L(1/|t|).$$

Hence,

$$\begin{aligned}
\int 1_{|g|\leqslant 1/|t|} g^2 w_t &= c(t) \int 1_{|g|\leqslant 1/|t|} g^2 e^{itG} + o(1) \\
&= (1+o(1)) \int 1_{|g|\leqslant 1/|t|} g^2 e^{itG} + o(1).
\end{aligned} \tag{30}$$

In the same way,

$$\int_{\omega \geqslant b\log(1/|t|)} |g||w_t - c(t)e^{itG}| \leqslant \|g\|_{3/2} \left\| 1_{\omega \geqslant b\log(1/|t|)} |w_t - c(t)e^{itG}| \right\|_{L^3} = O(t)$$

if b is large enough. Moreover,

$$\int_{\omega \leqslant b\log(1/|t|)} |g||w_t - c(t)e^{itG}| \leqslant \int |g| C|t| = O(t).$$

We get

$$\int gw_t = c(t)\int ge^{itG} + O(t) = (1+o(1))\int ge^{itG} + O(t). \tag{31}$$

Equations (30) and (31) together with (24) imply (29). □

Remark 3.17. The proof of the lemma also shows that, in (29), it is sufficient to integrate on $\{\omega(x) \leqslant b\log(1/|t|)\}$ if b is large enough, since the remaining part is in $o(t^2 L(1/|t|))$.

The following lemma will use the additional assumptions that l is slowly varying and $l(x\ln x) \sim l(x)$, $L(x\ln x) \sim L(x)$.

Lemma 3.18. *We have*

$$\int 1_{|g|\leqslant 1/|t|} g^2 e^{itG} = L(1/|t|)(1+o(1)). \tag{32}$$

Proof. It is sufficient to prove (32) on $\{\omega \leqslant b\log(1/|t|)\}$, since the remaining part can be ignored.

Take some $\varepsilon > 0$, we will prove that

$$\int_{\omega\leqslant b\log(1/|t|)} 1_{|g|\leqslant 1/|t|} g^2 |e^{itG} - 1| \leqslant 2\varepsilon L(1/|t|) \tag{33}$$

when t is small enough. This will be sufficient to conclude the proof.

Let $A_t := \{x \mid \omega(x) \leqslant b\log(1/|t|), |G(x)| \geqslant \varepsilon/|t|\}$. If $x \in A_t$, there exists y below x in the tower such that $|g(y)| \geqslant \frac{\varepsilon}{|t|b\log(1/|t|)}$. Let $B = \{x \mid |g(x)| \geqslant \frac{\varepsilon}{|t|b\log(1/|t|)}\}$, we get $\mu_\Delta(A_t) \leqslant b\log(1/|t|)\mu_\Delta(B)$.

Let Z be a random variable on $\mathbb{R}$ with the distribution of g. Then

$$\begin{aligned} P\left(\frac{1}{|t|\log(1/|t|)^2} \leqslant |Z| \leqslant 1/|t|\right) &= |t|^2\log(1/|t|)^4 l\left(\frac{1}{|t|\log(1/|t|)^2}\right) \\ &\quad -|t|^2 l(1/|t|) \\ &= |t|^2\log(1/|t|)^4 l(1/|t|)(1+o(1)) \end{aligned}$$

and

$$\begin{aligned} b\log(1/|t|)P\left(|Z| \geqslant \frac{\varepsilon}{|t|b\log(1/|t|)}\right) &= b\log(1/|t|)\frac{|t|^2 b^2\log(1/|t|)^2}{\varepsilon^2} \\ &\quad \times l\left(\frac{\varepsilon}{|t|b\log(1/|t|)}\right) \\ &= \frac{|t|^2 b^3\log(1/|t|)^3}{\varepsilon^2} l(1/|t|)(1+o(1)). \end{aligned}$$

Hence, if t is small enough, $\mu_\Delta(A_t) \leqslant P\left(\frac{1}{|t|\log(1/|t|)^2} \leqslant |Z| \leqslant 1/|t|\right)$.

We would like to estimate $\int_{A_t} 1_{|g|\leqslant 1/|t|} g^2$. Now

$$\int_{A_t} 1_{|g|\leqslant 1/|t|} g^2 = \int_{A_t} 1_{\frac{1}{|t|\log(1/|t|)^2} < |g| \leqslant 1/|t|} g^2 + \int_{A_t} 1_{|g|\leqslant \frac{1}{|t|\log(1/|t|)^2}} g^2.$$

On the one hand

$$\int_{A_t} 1_{\frac{1}{|t|\log(1/|t|)^2} < |g| \leqslant 1/|t|} g^2 \leqslant \int_{\frac{1}{|t|\log(1/|t|)^2}}^{1/|t|} x^2 \, \mathrm{d}P(x),$$

and on the other hand, by applying the above bounds we get

$$\begin{aligned}
\int_{A_t} 1_{|g| \leqslant \frac{1}{|t|\log(1/|t|)^2}} g^2 &\leqslant \frac{1}{|t|^2 \log(1/|t|)^4} \mu_\Delta(A_t) \\
&\leqslant \frac{1}{|t|^2 \log(1/|t|)^4} P\left(\frac{1}{|t|\log(1/|t|)^2} \leqslant |Z| \leqslant 1/|t| \right) \\
&\leqslant \int_{\frac{1}{|t|\log(1/|t|)^2}}^{1/|t|} x^2 \, \mathrm{d}P(x).
\end{aligned}$$

Thus we need to deal with the integral

$$\int_{\frac{1}{|t|\log(1/|t|)^2}}^{1/|t|} x^2 \, \mathrm{d}P(x),$$

which is equal to

$$L(1/|t|) - L\left(\frac{1}{|t|\log(1/|t|)^2}\right) = L(1/|t|) - L(1/|t|)(1 + o(1)) = o(L(1/|t|)).$$

Hence, for small enough t, we get

$$\int_{A_t} 1_{|g| \leqslant 1/|t|} g^2 |e^{itG} - 1| \leqslant \varepsilon L(1/|t|).$$

On $B_t := \{x \mid \omega(x) \leqslant b \log(1/|t|), x \notin A_t\}$, we have $|e^{itG(x)} - 1| \leqslant |t||G(x)| \leqslant \varepsilon$. Hence,

$$\int_{B_t} 1_{|g| \leqslant 1/|t|} g^2 |e^{itG} - 1| \leqslant \varepsilon \int 1_{|g| \leqslant 1/|t|} g^2 = \varepsilon L(1/|t|).$$

These two equations imply (33). This concludes the proof. □

Since $\int g = 0$, Lemma 3.16 gives

$$\begin{aligned}
\lambda_t &= 1 - \frac{t^2}{2} L(1/|t|)(1 + o(1)) + (1 + o(1)) it \int g e^{itG} \\
&= 1 - \frac{t^2}{2} L(1/|t|)(1 + o(1)) + (1 + o(1)) it \int g(e^{itG} - 1) \\
&= 1 - \frac{t^2}{2} L_1(1/|t|)(1 + o(1)),
\end{aligned}$$

since $\int g(e^{itG} - 1) = i(a + o(1)) t L(1/|t|)$ by assumption.

This asymptotic expansion readily implies the conclusion of Theorem 3.4, for g such that $\int g = 0$.

3.3.5. Proof of Theorem 3.4 in the general case. Let $g : \Delta \to \mathbb{R}$ belong to $L^p(\Delta)$ for any $p < 2$ (this is in particular the case if g satisfies the assumptions of Theorem 3.4). Set $G(x) = \sum_{k=0}^{\omega(x)-1} g(U^k \pi_0 x)$.

Lemma 3.19. *For any $p < 2$, the function G belongs to $L^p(\Delta)$.*

Proof. For $x \in \Delta$, let $\varphi(x)$ be its return time to the basis. Set also $\psi(x) = \varphi(\pi_0 x)$, where π_0 is the projection on the basis of the tower.

We have

$$|G(x)|^p = \left| \sum_{k=0}^{\omega(x)-1} g(U^k \pi_0 x) \right|^p \leqslant \omega(x)^{p-1} \sum_{k=0}^{\omega(x)-1} |g(U^k \pi_0 x)|^p.$$

Changing variables, we get

$$\int |G(x)|^p \leqslant \int |g(y)|^p \sum_{k=1}^{\varphi(y)-1} \omega(U^k y)^{p-1} \leqslant \int |g(y)|^p \psi(y)^p.$$

Since the tower has an exponentially small tail, the function ψ belongs to L^q for any $q < \infty$. Using the Hölder inequality with a sufficiently large q, we obtain $\int |G(x)|^p < \infty$. □

Let g' be another function on Δ. Define also $G'(x) = \sum_{k=0}^{\omega(x)-1} g'(U^k \pi_0 x)$.

Lemma 3.20. *If $g - g'$ is bounded, then*

$$\int g(e^{itG} - 1) = \int g'(e^{itG'} - 1) + O(t)$$

when $t \to 0$.

Proof. We have

$$\int g(e^{itG} - 1) - \int g'(e^{itG'} - 1) = \int (g - g')(e^{itG'} - 1) + \int g(e^{itG} - e^{itG'}).$$

Since $g - g'$ is bounded, the first integral satisfies

$$\left| \int (g - g')(e^{itG'} - 1) \right| \leqslant C|t| \int |G'|,$$

which is $O(t)$ since G' is integrable by Lemma 3.19. For the second integral, $|G(x) - G'(x)| \leqslant C\omega(x)$. Hence,

$$\left| \int g(e^{itG} - e^{itG'}) \right| \leqslant \int |g| C|t|\omega \leqslant C|t| \, \|g\|_{L^{3/2}} \, \|\omega\|_{L^3} = O(t). \quad \square$$

Proof (Proof of Theorem 3.4). Let g satisfy the assumptions of Theorem 3.4. Write $g' = g - \int g$. Then g' is still in the nonstandard domain of attraction of the normal law, and its tail functions l' and L' satisfy $l' \sim l$ and $L' \sim L$. Moreover,

$$\int g'(e^{itG'} - 1) = (a + o(1))itL(1/|t|)),$$

since g satisfies the same estimate and Lemma 3.20 applies.

We have already proved Theorem 3.4 for functions of zero integral. This applies to g', and gives $\frac{S_n g'}{B_n} \to \mathcal{N}(0, 1)$. Since $S_n g' = S_n g - n \int g$, this concludes the proof. □

Remark 3.21. Lemmas 3.19 and 3.20 do not involve the dynamics of the returns to the basis. Hence, they also hold in hyperbolic Young towers.

4. Estimate of the Integral in the Stadium Billiard

Let us turn back to the study of the stadium. We will use the notations of the first two sections. In particular, starting from a fixed function $f_0 : X_0 \to \mathbb{R}$ satisfying $(P1)$, we have obtained a function $g : \Delta \to \mathbb{R}$. According to Theorem 3.4, if we want to obtain a limit theorem for g, the quantity to be estimated is $\int g(e^{itG} - 1)$. The main result of this section is the following proposition.

Proposition 4.1. *Let $y = \frac{1}{1 - \frac{3}{4}\log 3}$, and recall the definition of I from* (1). *We have*

$$\int_\Delta g(e^{itG} - 1)\, \mathrm{d}\mu_\Delta = i\frac{I^2(y-1)\ell^2}{4} t\log(1/|t|) + o(t\log(1/|t|)).$$

Our proof approximates the left-hand side with an integral explicitly given in the phase space of the stadium. This later integral can be estimated with sufficient precision due to the geometric properties of the billiard map.

4.1. Preliminary estimates. First we show that the relevant expression can be pulled back to the hyperbolic Young tower. Let $\omega(x)$ be the height of the point x in $\bar{\Delta}$, and let $\bar{\pi}_0 : \bar{\Delta} \to \bar{\Delta}_0$ be the projection on the basis. We define two functions $\bar{F}$ and $\bar{G}$ on $\bar{\Delta}$ by $\bar{F}(x) = \sum_{k=0}^{\omega(x)-1} \bar{f}(\bar{U}^k \bar{\pi}_0 x)$ and $\bar{G}(x) = \sum_{k=0}^{\omega(x)-1} \bar{g}(\bar{U}^k \bar{\pi}_0 x)$.

Lemma 4.2. *We have*

$$\int_\Delta g(e^{itG} - 1) = \int_{\bar{\Delta}} \bar{f}(e^{it\bar{F}} - 1) + O(t).$$

Proof. As $(\pi_\Delta)_*(\mu_{\bar{\Delta}}) = \mu_\Delta$, $\int_\Delta g(e^{itG} - 1) = \int_{\bar{\Delta}} \bar{g}(e^{it\bar{G}} - 1)$ automatically. As $\bar{g} - \bar{f}$ is bounded by Lemma 2.5, Lemma 3.20 implies the statement. □

Note that $\bar{F}$ is essentially a Birkhoff sum of f for the *inverse map* T^{-1}. Thus, if we switch from T to T^{-1}, we may investigate our integral by dynamical tools.

For all $x \in X$, let $h(x) = f(T^{-1}x)$. Introduce $\bar{h} = h \circ \pi_X$. For $x \in \bar{\Delta}$ with $\omega(x) > 0$, let $\bar{H}(x) = \sum_{k=1}^{\omega(x)-1} \bar{h}(\bar{U}^{-k}x)$, or equivalently, $\bar{H}(x) = \bar{F}(\bar{U}^{-1}x)$. We fix $\bar{H}(x) = 0$ on Δ_0.

Lemma 4.3. *We have*

$$\int \bar{f}(e^{it\bar{F}} - 1) = \int \bar{h}(e^{it\bar{H}} - 1) + O(t).$$

Proof. We have $\bar{h} \circ \bar{U} = \bar{f}$, and, apart from $\bar{U}^{-1}(\bar{\Delta}_0)$, $\bar{H} \circ \bar{U} = \bar{F}$. Thus,

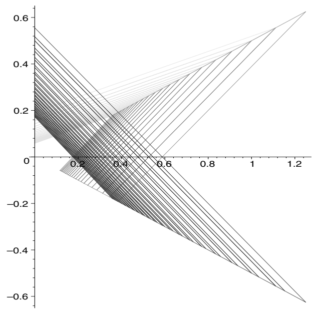
$$\begin{aligned}\int \bar{h}(e^{it\bar{H}} - 1) - \int \bar{f}(e^{it\bar{F}} - 1) &= \int \bar{h} \circ \bar{U}(e^{it\bar{H}\circ\bar{U}} - 1) - \int \bar{f}(e^{it\bar{F}} - 1)\\ &= \int_{\bar{U}^{-1}(\bar{\Delta}_0)} \left[\bar{h} \circ \bar{U}(e^{it\bar{H}\circ\bar{U}} - 1) - \bar{f}(e^{it\bar{F}} - 1)\right].\end{aligned}$$

As φ_+ is bounded on the rectangle R that defines the basis of the tower, the functions $\bar{h} \circ \bar{U}$ and $\bar{f}$ are bounded on $\bar{U}^{-1}(\bar{\Delta}_0)$. By Lemma 3.19 $\bar{F}$ and $\bar{H}$ are both integrable. This completes the proof. □

We will consider T^{-1} as the first return map of T_0^{-1} to the subspace X. The return time is $\varphi_- = \varphi_+ \circ T^{-1}$.

4.2. Geometric properties of T^{-1} in the vicinity of its singularities. The behavior of $\int \bar{h}(e^{it\bar{H}} - 1)$ is governed by the dynamical properties of T^{-1} at those parts of the phase space where it is equivalent to a long series of bounces between the parallel segments of the boundary. These sets have the following structure: the points for which T^{-1} acts as n consecutive bounces on the segments form two stripes of slope approximately -1. T^{-1} maps these two cells onto two stripes of positive slope.

The figure below describes this geometrical situation. On this figure the relevant part of the cylinder-shaped phase space X_0 is magnified: the horizontal coordinate is the position r, and the vertical is the angle θ. Recall that X is the union of two parallelograms in X_0. The origin on the figure corresponds to a corner of one of these parallelograms: a phase point with position on the junction of the segment and the semi-circle, and with velocity perpendicular to the wall ($\theta = 0$). The negatively sloped stripes terminating on the two sides of the parallelogram that meet at the origin are the regions where one application of T^{-1} consists of n applications of T_0^{-1}. Each of these is mapped by T^{-1} onto two positively sloped stripes. Both structures accumulate with increasing n at the origin, as this is the phase point in X that corresponds to infinitely long bouncing between the segments.



Actually, this figure appears four times in X (twice in both of the parallelograms that define X). The transformation T^{-1} jumps from one such region to another, however, they play the same dynamical role. Thus, to simplify matters, we pretend as if we had only one of them.

Convention 4.4. Given this geometry, we will refer to these regions as *"corners of parallelograms"*.

Besides the above described four, there are four further corners of the two parallelograms. These further corners are the accumulation points for the other type of singularities, corresponding to trajectories sliding along the semi-circles. We shall see later on that they do not play any role in the leading term behavior of $\int \bar{h}(e^{it\bar{H}} - 1)$.

Remark 4.5. We need to study the map T^{-1} and not T. These two are not isomorphic, as X is the set of points on a semi-circle for which the *previous* collision is not on that semi-circle. This definition introduces an asymmetry of past and future. The map T^{-1} is, however, isomorphic to the map induced on the set of points on a semi-circle for which the *next* collision is on another semi-circle. This later induced map has been studied by Markarian in [Mar04], where he, in particular, has shown that it satisfies Chernov's axioms from [Che99].

We fix some further conventions, to be used throughout Subsect. 4.2.

Convention 4.6. Recall that $\varphi_- = \varphi_+ \circ T^{-1}$ is the return time of T_0^{-1} to X. As a further notational simplification we ignore that there are two stripes on which $\{\varphi_- = n\}$. Let M_n stand for the stripe $\{\varphi_- = n\}$, which will be also referred to as the set of points of *return time* n. Unless otherwise stated, throughout Subsect. 4.2 *return time* is understood in this sense.

Convention 4.7. Fix $\rho < 1$ such that the tails of the original Young tower are bounded by $c\rho^n$, and $K > 0$ such that $K \log(\rho^{-1}) > 4$.

In what follows we essentially consider curves with tangent vectors in the unstable cone of T^{-1} (u-curves of T^{-1} for short). The two components of the boundary of the stripe M_n that separate it from M_{n-1} and M_{n+1} will be referred to as the long sides of M_n. The slopes of the long sides are approximately -1, with better and better precision as n increases.

Definition 4.8. *Consider the stripe M_n and its two long sides. A* **good curve** $\mathcal{C}$ *of return time n is a C^1 curve that connects these two sides. We put further requirements on the slope of $\mathcal{C}$: it should belong to the interval* $[1/4, 4]$ *for all points and it should be constant up to $1/\sqrt{n}$ precision (i.e., for all points x and y in $\mathcal{C}$, the slopes of $\mathcal{C}$ at x and y, $s(x)$ and $s(y)$, should satisfy $|s(x) - s(y)| \leqslant \frac{1}{\sqrt{n}}$).*

Remark 4.9. Note that our requirement on the slope is weaker than a bound on the curvature of the good curve. Stated in this form, it is not hard to see that good curves tend to have more and more constant slopes when iterated by T^{-1}. To see this consider a good curve with large return time n and iterate it backwards by the billiard flow. Just before collision it corresponds to a dispersing wavefront. To calculate the curvature of this wavefront we may use a well-known formula from billiard theory (see e.g. [Che99]):

$$\mathcal{B}_- = \frac{1}{\cos\theta}\left(\frac{\mathrm{d}\theta}{\mathrm{d}r} - \mathcal{K}\right),$$

where $\frac{d\theta}{dr}$ is the slope of the curve on the phase space, $\mathcal{B}_-$ is the curvature of the corresponding wavefront just before collision and $\mathcal{K}$ is the curvature of the wall at the collision point. In our case $\mathcal{K} = -1$ as we have a semicircular focusing wall of radius 1, while the slope $\frac{d\theta}{dr}$ is bounded from below, see Definition 4.8. Thus $\mathcal{B}_-$ is uniformly bounded from below and thus the dispersing wavefront, when iterated backwards by the flow, defocuses within finite time. The resulting convergent wavefront, when iterated further backwards, while experiencing many bounces with the straight walls, loses most of its curvature. As the length of this neutral flight is of order n, at the time moment just after the previous collision on the other semi-circle, this wavefront is flat up to $1/n$. Thus, any subcurve $\mathcal{C}' \subset T^{-1}\mathcal{C}$ is automatically a good curve if it connects the two long sides of a stripe.

Definition 4.10. *A* **standard** *curve is defined as a good curve of constant slope* 1. *In particular, it is a segment.*

The choice of 1 as the slope for standard curves is arbitrary. More important is the fact that the standard curves of return time n give a fixed foliation for (most of) the stripe M_n.

If $\mathcal{C}$ is a good curve of return time n, any point of $T^{-1}(\mathcal{C})$ has return time at least $n/3 - C_1$ and at most $3n + C_1$ for some constant C_1. Furthermore, there exists a constant $C_2 \in \mathbb{N}$ such that for any $i \in [n/3 + C_2, 3n - C_2]$, the set $T^{-1}(\mathcal{C}) \cap \{\varphi_- = i\}$ is a good curve of return time i (see Remark 4.9). Let us denote $\mathcal{C}_i = \{x \in \mathcal{C} \mid \varphi_-(T^{-1}x) = i\}$. We also have

$$\frac{\text{Leb}\left(\mathcal{C} \setminus \bigcup_{i\in[n/3+C_2,3n-C_2]} \mathcal{C}_i\right)}{\text{Leb}(\mathcal{C})} \leqslant \frac{C}{n} \tag{34}$$

for a universal constant C. We will say that the set $\mathcal{C} \setminus \bigcup_{i\in[n/3+C_2,3n-C_2]} \mathcal{C}_i$ is *thrown away* at the first iterate of $\mathcal{C}$. Formula (34) shows that the points which are thrown away have negligible measure.

Remark 4.11. In addition to the above observations, it is possible to estimate the transition probabilities from one stripe to the other in the following sense. There exists a sequence ε_n that tends to 0 as $n \to \infty$, such that for any good curve $\mathcal{C}$ of return time n, and for any $i \in [n/3 + C_2, 3n - C_2]$,

$$(1 - \varepsilon_n)\frac{3n}{8i^2} \leqslant \frac{\text{Leb}\{x \in \mathcal{C} \mid \varphi_-(T^{-1}(x)) = i\}}{\text{Leb}(\mathcal{C})} \leqslant \frac{3n}{8i^2}(1 + \varepsilon_n). \tag{35}$$

This can be verified by direct calculation. In other words, we go from n to i asymptotically with probability $\frac{3n}{8i^2}$ (note that $\sum_{i=n/3+C_2}^{3n-C_2} \frac{3n}{8i^2} \to 1$).

Applying the above process several times, we may iterate the good curves by T^{-1} and obtain finer and finer partitions of $\mathcal{C}$. A sequence of integers $n_0, n_1, \ldots, n_k$ is referred to as *admissible* if, for all $i < k$, $n_{i+1} \in [n_i/3 + C_2, 3n_i - C_2]$. Given a good curve of return time n_0, $\mathcal{C}$, and an admissible sequence $n_0, \ldots, n_k$, let

$$\mathcal{C}_{n_0,\ldots,n_k} = \{x \in \mathcal{C} \mid \forall i \leqslant k, \varphi_-(T^{-i}x) = n_i\}.$$

This is a subcurve of $\mathcal{C}$ mapped by T^{-k} onto a good curve of return time n_k.

Lemma 4.12. *There exists a constant* $C > 0$ *such that, for any pair of good curves of the same return time* n_0, $\mathcal{C}$ *and* $\mathcal{C}'$, *and for any fixed admissible sequence* $n_0, \dots, n_k$, *we have*

$$C^{-1} \leqslant \frac{\mathrm{Leb}(\mathcal{C}_{n_0,\dots,n_k})}{\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_k})} \leqslant C.$$

Proof. This follows from the uniform expansion and the bounded distortion properties of T^{-1} along its u-curves. □

In what follows, when we talk about iterating a good curve, we will always mean the above process of refinement, along with throwing away some part at each step. However, the number of iterations may depend on the point of $\mathcal{C}$ we are considering. This is formulated in the following definition.

Definition 4.13. *Let* $\mathcal{C}$ *be a good curve of return time* n. *Let furthermore* A *be a subset of* $\mathcal{C}$ *and* $\tau : \mathcal{C} \backslash A \to \mathbb{N}$. *Then* (A, τ) *is a* **stopping time** *on* $\mathcal{C}$ *if*

- *There exists* $p \in \mathbb{N}$ *such that* $3^{p+1} < n_0$, *with the following property: all the connected components of* $\mathcal{C}\backslash A$ *are of the form* $\mathcal{C}_{n_0,\dots,n_k}$, *where* $n_0 = n$, *the sequence* $n_0, \dots, n_k$ *is admissible, and* $n_k \in [3^p, 3^{p+1} - 1]$. *Furthermore,* τ *is uniformly equal to* k *on such a component.*
- *We have* $\mathrm{Leb}(A)/\mathrm{Leb}(\mathcal{C}) \leqslant 1/2$.

Here typically $3^p \ll n$, thus we stop at the first occasion when the return time decreases below a certain level.

Remark 4.14. If (A, τ) is a stopping time on $\mathcal{C}$, then

$$\frac{1}{2} \leqslant \frac{\mathrm{Leb}(\mathcal{C}\backslash A)}{\mathrm{Leb}(\mathcal{C})} \leqslant 1.$$

Thus in our estimates $\mathrm{Leb}(\mathcal{C}\backslash A)$ and $\mathrm{Leb}(\mathcal{C})$ may be replaced with each other. We will often use this without giving further details.

Let us define, in particular, the *standard stopping time* for a good curve $\mathcal{C}$ of return time n. Let p be the integer for which $3^p \leqslant n^{1/4} < 3^{p+1}$. We partition $\mathcal{C}$, iterate T^{-1} and throw away the negligible parts according to the process described above. We go on iterating until either the return time of the image belongs to the interval $[3^p, 3^{p+1} - 1]$, or the number of iterates exceeds $K \log n$ (here K is the constant from Convention 4.7). Thus we put into A, on the one hand, the points thrown away during this process, and, on the other hand, the intervals for which the return time does not reach $[3^p, 3^{p+1} - 1]$ before $K \log n$ iterations. On all other intervals we define τ as the first occasion when the return time belongs to $[3^p, 3^{p+1} - 1]$.

Proposition 4.15. *The standard stopping time* (A, τ) *defined this way is indeed a stopping time if* n *is large enough. Furthermore,* $\mathrm{Leb}(A)/\mathrm{Leb}(\mathcal{C}) \leqslant n^{-1/5}$.

Proof. The only non-trivial condition to be verified is $\mathrm{Leb}(A)/\mathrm{Leb}(\mathcal{C}) \leqslant n^{-1/5}$.

Let us first estimate the measure of points thrown away during the refinement process. We will denote this set by $A_0 (\subset A \subset \mathcal{C})$.

No matter which phase of the iteration we consider, the return time is $\geqslant n^{1/4}$, thus, according to (34), the points thrown away occupy at most a $Cn^{-1/4}$ proportion of the considered interval. Hence, by bounded distortion, the proportion of A_0 in $\mathcal{C}$ is at most $Cn^{-1/4}K\log n \leqslant n^{-1/5}$ for n large enough.

It remains to be shown that the overall measure of the intervals that do not reach $[3^p, 3^{p+1}-1]$ before $K\log n$ iterations is small. We have $\mathcal{C} = A_0 \cup \bigcup \mathcal{C}_i$, where each $\mathcal{C}_i$ is of the form $\mathcal{C}_{n_0,\dots,n_k}$ for some admissible sequence $n_0,\dots,n_k$, with $k \leqslant K\log n$, and $n_k < 3^{p+1}$ whenever $k < \lfloor K\log n \rfloor$. Thus it is enough to estimate the measure of $\mathcal{C}_i$-s with $\tau_{\mathcal{C}_i} = k = \lfloor K\log n \rfloor$. Let $\mathcal{C}'$ be one of our *standard* curves of return time n. We apply the same construction to $\mathcal{C}'$, and get a similar decomposition $\mathcal{C}' = A_0' \cup \bigcup \mathcal{C}_i'$. Furthermore, by Lemma 4.12, $\frac{\mathrm{Leb}\,\mathcal{C}_i}{\mathrm{Leb}\,\mathcal{C}_i'} \leqslant C$.

Recall that the standard curves of return time n foliate the major part M_n' of the stripe M_n (where $\mu(M_n')/\mu(M_n) = 1 + O(1/n)$). For a fixed $i = (n_0,\dots,n_k)$ consider B_i the subset of the stripe M_n that corresponds to the union of such $\mathcal{C}_i'$-s for all the standard curves of return time n. As the density of μ on M_n is bounded away from 0, we get $\frac{\mathrm{Leb}(\mathcal{C}_i)}{\mathrm{Leb}(\mathcal{C})} \leqslant C\frac{\mu(B_i)}{\mu(M_n')}$. Fix B as the union of all B_i-s with $\tau_i = \lfloor K\log n \rfloor$. When pulled back to the Young tower, the preimages of the points of B are all at height at least $K\log n$. As $\pi_X^*(\mu_{\bar\Delta}) = \mu$, we get $\mu(B) \leqslant C\rho^{K\log n} = O(1/n^4)$ by our choice of K (see Convention 4.7). As $\mu(M_n') \sim C/n^3$, we may put all these estimates together to conclude that

$$\frac{\sum_{\tau_i=\lfloor K\log n\rfloor}\mathrm{Leb}(\mathcal{C}_i)}{\mathrm{Leb}(\mathcal{C})} = O(1/n).$$

This completes the proof of the proposition. □

In the next proposition we consider standard curves $\mathcal{C}$ and use the notation $(A_{\mathcal{C}}, \tau_{\mathcal{C}})$ for their standard stopping times. We define a subset of the phase space, a suitable union of subcurves of standard curves, as $Y = \bigcup_{\mathcal{C}}(\mathcal{C}\setminus A_{\mathcal{C}})$. We also consider the Birkhoff sum of h with respect to T^{-1} up to standard stopping time, i.e., we fix $H(x) = \sum_{k=1}^{\tau_{\mathcal{C}}(x)-1} h(T^{-k}x)$ for $x \in Y$.

Remark 4.16. At first sight the measurability of the set Y may seem questionable. Let us show that, for n fixed, the set $Y \cap M_n$ is measurable. Consider the standard curves in M_n, all having the same return time n. Whether a point of such a curve falls into the thrown away set depends only on its past history up to the first $K\log n$ backward iterations. The singularity manifolds for the first $K\log n$ applications of T^{-1} give a finite measurable partition of M_n. By the above observation, $Y \cap M_n$ is a union of full elements of this partition, intersected with M_n'. Hence, it is measurable.

Proposition 4.17. *We have*

$$\int_{\bar\Delta} \bar h(e^{it\bar H}-1) = \int_Y h(e^{itH}-1) + O(t).$$

This proposition plays a central role as it allows us to investigate, instead of $\int_{\bar\Delta} \bar h(e^{it\bar H}-1)$ (a quantity that depends *a priori* on the choice of the Young tower), an expression which is much easier to handle, as it is completely explicitly given in terms of the phase space geometry.

Proof. Let us show first that

$$\int_{\bar\Delta} \bar h(e^{it\bar H}-1) = \int_{\pi_X^{-1}(Y)} \bar h(e^{it\bar H}-1) + O(t). \tag{36}$$

Consider $A = X \setminus Y$. The set A consists of two parts. It contains, on the one hand, the points that are not covered by standard curves and, on the other hand, those contained in $A_{\mathcal{C}}$ for some standard curve $\mathcal{C}$. These two sets will be referred to as A_1 and A_2, respectively.

We cover the set $A_1 \cap \{\varphi_- = n\}$ by two further sets. The first one contains points that slide along a semi-circle, with return time n. This set is of measure $O(1/n^4)$ as $\cos\theta$, the density of the invariant measure, is $O(1/n)$ on it. The other component in this cover is the part of M_n not covered by standard curves, which has measure $O(1/n^4)$ as well. Altogether we have $\mu(A_1 \cap \{\varphi_- = n\}) = O(1/n^4)$.

According to Proposition 4.15, we have $\mathrm{Leb}(A_{\mathcal{C}})/\mathrm{Leb}(\mathcal{C}) \leqslant n^{-1/5}$ whenever n, the return time for $\mathcal{C}$, is large enough. Integrating on the relevant standard curves we obtain $\mu(A_2 \cap \{\varphi_- = n\}) = O(1/n^{3+1/5})$.

Altogether we have

$$\mu(A \cap \{\varphi_- = n\}) = O(1/n^{3+1/5}). \tag{37}$$

For any $1/p + 1/q = 1$ we have

$$\left| \int_{\pi_X^{-1}(A)} \bar{h}(e^{it\bar{H}} - 1) \right| \leqslant \int 1_{\pi_X^{-1}(A)} |\bar{h}| t |\bar{H}| \leqslant |t| \left(\int (1_{\pi_X^{-1}(A)} |\bar{h}|)^p \right)^{1/p} \left(\int |\bar{H}|^q \right)^{1/q}.$$

Recall from Lemma 3.19 that the function $\bar{H}$ belongs to L^q for any $q < 2$, while (37) implies that $\int (1_{\pi_X^{-1}(A)} |\bar{h}|)^p$, being equal to $\int_X 1_A |h|^p$, is finite for $p < 2 + 1/5$. We can thus take $p = 2 + 1/10$ and $q = (1 - 1/p)^{-1}$, to obtain (36).

Now, to complete the proof, we need to show that

$$\int_{\pi_X^{-1}(Y)} \bar{h}(e^{it\bar{H}} - 1) = \int_{\pi_X^{-1}(Y)} \bar{h}(e^{itH\circ\pi_X} - 1) + O(t). \tag{38}$$

Consider $\mathcal{C}_i$, a connected component of $\mathcal{C} \setminus A_{\mathcal{C}}$, where $\mathcal{C}$ is a standard curve of return time n. Then the stopping time on $\mathcal{C}_i$ is an integer $\tau_i < K \log n$ such that $\mathcal{D}_i = T^{-\tau_i}(\mathcal{C}_i)$ is a good curve, with return time in the interval $[n^{1/4}/3, 3n^{1/4}]$. In the next lemma, as throughout the subsection, we use the expression "return time" in the sense of Convention 4.6.

Lemma 4.18. *There exists a constant C such that, for any large enough integer n, given any good curve $\mathcal{D}$ of return time $\in [n^{1/4}/3, 3n^{1/4}]$, the points for which the return time increases above $n^{1/2}$ within $K \log n$ iterations of T^{-1} occupy relative measure less than $Cn^{-1/4}$ in $\mathcal{D}$.*

Proof. The map T^{-1} satisfies Chernov's axioms, by [Mar04]. Consequently, we can use [Che99, Theorem 3.1], with $\delta = Z[\mathcal{D}, \mathcal{D}, 0]^{-1/\sigma}/n^{1/\sigma}$. This theorem is in fact stated for local unstable manifolds, but its proof can be straightforwardly adapted to deal with manifolds close to the unstable direction.

We obtain a decreasing sequence $W_0^1 \supset W_1^1 \supset \cdots \supset W_{\lfloor K \log n \rfloor}^1$ of subsets of $\mathcal{D}$ such that, if we denote by $Sing$ the set of singularities of T^{-1},

$$\forall c > 0, \forall 0 \leqslant p \leqslant K \log n, \ \mathrm{Leb}\{x \in W_p^1 \mid \mathrm{dist}(T^{-p}x, Sing) \leqslant cn^{-1}\} \leqslant Ccn^{-1} \tag{39}$$

(by Eq. (3.3) in [Che99]), and

$$\forall 0 \leqslant p \leqslant K \log n,\ \operatorname{Leb}(W_p^1 \backslash W_{p+1}^1) \leqslant \frac{C}{n} \operatorname{Leb}(\mathcal{D}) \tag{40}$$

(by (iv), (3.5) in [Che99] and our choice of δ).

Note that the results of [Che99] imply that (39) holds for the distance measured in the p-metric. However, we are in a region of X where $\cos\theta$ is bounded away from 0, and the stable and unstable cones are bounded away from the vertical direction by Proposition 2.1. Hence, it is equivalent to have (39) for the p-distance or for the usual distance.

If $T^{-p}(x)$ has a return time $\geqslant n^{1/2}$, then $T^{-p}x$ is at a distance at most Cn^{-1} of $Sing$. Hence, the point x belongs to one of the sets whose measure is bounded in (39) and (40). This gives a measure at most $C \log n\, n^{-1}$. Since $\operatorname{Leb}(\mathcal{D}) \geqslant Cn^{-1/2}$, this proves the lemma. □

This lemma applies to $\mathcal{D}_i$. Let us write $\mathcal{C}_i = \mathcal{C}_i^1 \cup \mathcal{C}_i^2$, where $\mathcal{C}_i^2$ corresponds to points which go to $\mathcal{D}_i$, and then reach a return time $> n^{1/2}$ in a time shorter than $K \log n$. It satisfies $\operatorname{Leb}(\mathcal{C}_i^2)/\operatorname{Leb}(\mathcal{C}_i) \leqslant Cn^{-1/4}$ by Lemma 4.18.

Let $Y_1 = \bigcup \mathcal{C}_i^1$ and $Y_2 = \bigcup \mathcal{C}_i^2$ (see Remark 4.16 on the measurability of these sets). Since $\mu(Y_2 \cap \{\varphi_- = n\}) = O(1/n^{3+1/4})$, the proof of (36) applies and gives

$$\int_{\pi_X^{-1}(Y_2)} \bar{h}(e^{it\bar{H}} - 1) = O(t); \quad \int_{\pi_X^{-1}(Y_2)} \bar{h}(e^{itH\circ\pi_X} - 1) = O(t).$$

Remark 4.19. Note that $H \circ \pi_X$ belongs to L^q for any $q < 2$ as it is smaller than a function to which Lemma 3.19 applies.

Hence, it is sufficient to prove (38) on $\pi_X^{-1}(Y_1)$. Let us write $\pi_X^{-1}(Y_1) = Z_1 \cup Z_2$, where

$$Z_1 = \{x \in \pi_X^{-1}(Y_1) \mid \omega(x) < K \log(\varphi_-(\pi_X x))\}$$

and $Z_2 = \pi_X^{-1}(Y_1) \backslash Z_1$. For $n > 0$,

$$\mu_{\bar{\Delta}}\{x \in Z_2 \mid \varphi_-(\pi_X x) = n\} \leqslant \mu_{\bar{\Delta}}\{x \in \bar{\Delta}, \omega(x) \geqslant K \log n\} = O(1/n^4).$$

Hence, we get once again $\int_{Z_2} \bar{h}(e^{it\bar{H}} - 1) = O(t)$ and $\int_{Z_2} \bar{h}(e^{itH\circ\pi_X} - 1) = O(t)$.

On $Z_1 \cap \{\varphi_- \circ \pi_X = n\}$, the functions $\bar{H}$ and $H \circ \pi_X$ differ by at most $\|f_0\|_\infty K \log n\, n^{1/2}$ (corresponding to at most $K \log n$ iterations with a return time $< n^{1/2}$). Hence,

$$\begin{aligned} &\left| \int_{Z_1} \bar{h}(e^{it\bar{H}} - 1) - \bar{h}(e^{itH\circ\pi_X} - 1) \right| \\ &\quad \leqslant |t| \int_{Z_1} |\bar{h}||\bar{H} - H \circ \pi_X| \leqslant C|t| \sum_n \mu\{\varphi_- = n\} n \log n\, n^{1/2} \leqslant C|t| \end{aligned}$$

since $\mu\{\varphi_- = n\} = O(1/n^3)$. This proves (38), and concludes the proof of Proposition 4.17. □

4.3. An upper bound on H. The aim of this subsection is to estimate the average of the function H on a good curve $\mathcal{C}$ of return time n. We obtain the following upper bound:

Proposition 4.20. *Let $s \in [1, 2)$. Consider a good curve $\mathcal{C}$ of return time n_0, and a stopping time (A, τ) on $\mathcal{C}$. Then*

$$\frac{\int_{\mathcal{C}\setminus A} \sum_{k=0}^{\tau(x)-1} |h(T^{-k}x)|^s}{\mathrm{Leb}(\mathcal{C}\setminus A)} \leqslant C(s)n_0^s,$$

where the constant $C(s)$ depends only on s.

Let us fix some notation first. There is an integer p_0 such that the return time n_0 for our good curve $\mathcal{C}$ belongs to $[3^{p_0}, 3^{p_0+1}-1]$. By the definition of stopping times, there exists another integer $p_1 < p_0$ such that, for any $x \in \mathcal{C}\setminus A$, $\varphi_-(T^{-\tau(x)}(x)) \in [3^{p_1}, 3^{p_1+1}-1]$. Now consider an intermediate p, $p_1 < p \leq p_0$. In the course of the proof first we investigate, in a series of lemmas, what happens while the return time descends from $[3^p, 3^{p+1}-1]$ to $[3^{p-1}, 3^p-1]$. Then we sum up for $p_1 < p \leq p_0$. In the first part of the proof the value of p is fixed and $n \approx 3^p$, while in the second part p varies from p_1 to p_0. The value of $s \in [1, 2)$ is fixed throughout the subsection.

According to this plan, let us fix $p \in \mathbb{N}$ large enough. Given $x \in X$, we define $\tau_p(x)$ as the first time $k \geqslant 1$ for which $\varphi_-(T^{-k}x) < 3^p$, and $\Phi_p(x) = \sum_{k=0}^{\tau_p(x)-1} |\varphi_-(T^{-k}x)|^s$. Since $|h| \leqslant C\varphi_-$, it is sufficient to prove Proposition 4.20 for $h = \varphi_-$ to conclude.

Define $R \subset X$ as the union of all standard curves with return time from the interval $[3^p/2, 3^p-1]$.

Lemma 4.21. *There exists a constant C such that*

$$\int_R \Phi_p \leqslant C\mu(R)3^{ps}.$$

Proof. Let $R_1 = \{x \in R \mid \varphi_-(T^{-1}x) < 3^p\}$ and $R_2 = \{x \in R \mid \varphi_-(T^{-1}x) \geqslant 3^p\}$. On R_1 we have $\Phi_p(x) = |\varphi_-(x)|^s$, thus

$$\int_{R_1} \Phi_p \leqslant C\mu(R_1)3^{ps}.$$

Let us define $\varphi'(x) = \varphi_-(x)$ for x with $\varphi_-(x) \geqslant 3^{p-1}$ and $\varphi'(x) = 0$ otherwise. Note that $\Phi_p(x) = \sum_{k=0}^{\tau_p(x)-1} |\varphi'(T^{-k}x)|^s$ for $x \in R_2$.

Consider $Z \subset X$, $Z := \{3^{p-1} - C_1 \leqslant \varphi_- < 3^p\}$, and define $\tau_Z : Z \to \mathbb{N}$ as the first return time to Z. By Kac's formula,

$$\int_Z \sum_{k=0}^{\tau_Z(x)-1} |\varphi'(T^{-k}x)|^s = \int_X |\varphi'|^s \leqslant C \sum_{k\geqslant 3^{p-1}} \mu(\varphi_- = k)|k|^s$$
$$\leqslant C \sum_{k\geqslant 3^{p-1}} \frac{1}{k^3}k^s \leqslant C\frac{3^{ps}}{3^{2p}}.$$

Now $R_2 \subset Z$ and for $x \in R_2$ we have $\tau_Z(x) = \tau_p(x)$. Thus

$$\int_{R_2} \Phi_p = \int_{R_2} \sum_{k=0}^{\tau_Z(x)-1} |\varphi'(T^{-k}x)|^s \leqslant \int_Z \sum_{k=0}^{\tau_Z(x)-1} |\varphi'(T^{-k}x)|^s.$$

By Remark 4.11, $\frac{1}{3^{2p}} = O(\mu(R_2))$. This completes the proof. □

If $\mathcal{C}$ is a good curve of return time $n \in [3^p, 3^{p+2} - 1]$, τ_p defines a stopping time on $\mathcal{C}$, with the corresponding thrown-away set that we denote by A_p. To see that it is indeed a stopping time we only need to show that $\mathrm{Leb}(A_p) \leqslant \mathrm{Leb}(\mathcal{C})/2$. Now consider the standard stopping time $\tau_{\mathcal{C}}$ with its thrown away set $A_{\mathcal{C}}$. Then $A_p \subset A_{\mathcal{C}}$ while $\mathrm{Leb}(A_{\mathcal{C}}) \leqslant n^{-1/5}\,\mathrm{Leb}(\mathcal{C})$ by Proposition 4.15, which gives the claim.

The first step in the proof of Proposition 4.20 is the estimate

$$\frac{\int_{\mathcal{C}\backslash A_p} \Phi_p}{\mathrm{Leb}(\mathcal{C}\backslash A_p)} \leqslant C3^{ps} \tag{41}$$

for a good curve $\mathcal{C}$ with return time $n \in [3^p, 3^{p+1} - 1]$. To show this, we will relate the average of Φ_p on $\mathcal{C}$ to its average on R.

Consider $B = \bigcup(\mathcal{C}\backslash A_p)$, where the union is taken over all standard curves of return time from the interval $[3^p, 3^{p+2} - 1]$.

Lemma 4.22. *There is a constant C such that, for any good curve $\mathcal{C}$ of return time $n \in [3^p, 3^{p+1} - 1]$,*

$$\frac{\int_{\mathcal{C}\backslash A_p} \Phi_p}{\mathrm{Leb}(\mathcal{C}\backslash A_p)} \leqslant C\frac{\int_B \Phi_p}{\mu(B)} + C3^{ps}. \tag{42}$$

Proof. Let $U = \{x \in \mathcal{C} \mid \varphi_-(T^{-1}x) \geqslant 3^p\}$. On $\mathcal{C}$, we have $\Phi_p(x) = |\varphi_-(x)|^s + 1_U(x)\Phi_p(T^{-1}x)$. To prove (42), it is enough to show

$$\int_{U\cap(\mathcal{C}\backslash A_p)} \Phi_p \circ T^{-1} \leqslant C\frac{\int_B \Phi_p}{\mu(B)}\,\mathrm{Leb}(\mathcal{C}\backslash A_p).$$

By bounded distortion, this can be further reduced to

$$\frac{\int_{T^{-1}(\mathcal{C}\backslash A_p)\cap\{\varphi_-\geqslant 3^p\}} \Phi_p}{\mathrm{Leb}(T^{-1}(\mathcal{C}\backslash A_p)\cap\{\varphi_- \geqslant 3^p\})} \leqslant C\frac{\int_B \Phi_p}{\mu(B)}. \tag{43}$$

Let q be the maximal possible return time the points of $T^{-1}(\mathcal{C}\backslash A_p)$ have. It satisfies $3^{p+2} > q \geqslant 3^{p+1} - C_2$. By Lemma 4.12

$$\frac{\int_{T^{-1}(\mathcal{C}\backslash A_p)\cap\{\varphi_-\geqslant 3^p\}} \Phi_p}{\mathrm{Leb}(T^{-1}(\mathcal{C}\backslash A_p)\cap\{\varphi_- \geqslant 3^p\})} \leqslant C\frac{\int_{B\cap\{3^p\leqslant\varphi_-\leqslant q\}} \Phi_p}{\mu(B\cap\{3^p \leqslant \varphi_- \leqslant q\})}.$$

As $q \geqslant 3^{p+1} - C_2$, by Remark 4.11 $\mu(B) \leqslant C\mu(B\cap\{3^p \leqslant \varphi_- \leqslant q\})$. This implies (43) and completes the proof. □

Now B is not exactly R, we need to "widen up" the estimate of Lemma 4.21 from R to B to obtain (41).

Let $B_1 = B\cap\{3^p \leqslant \varphi_- < 3^{p+1}/2\}$, $B_2 = B\cap\{3^{p+1}/2 \leqslant \varphi_- < 3^{p+1}\}$, $B_3 = B\cap\{3^{p+1} \leqslant \varphi_- < 3^{p+2}/2\}$ and $B_4 = B\cap\{3^{p+2}/2 \leqslant \varphi_- < 3^{p+2}\}$.

Lemma 4.23. *There exists a constant C such that, for any good curve $\mathcal{C}$ of return time $n \in [3^p, 3^{p+1}/2)$,*

$$\frac{\int_{B_1} \Phi_p}{\mu(B_1)} \leqslant C\frac{\int_{\mathcal{C}\backslash A_p} \Phi_p}{\mathrm{Leb}(\mathcal{C}\backslash A_p)}.$$

Proof. The curve $T^{-1}(\mathcal{C})$ crosses all stripes of return time between 3^p and $3^{p+1}/2$. This allows us to apply the argument of Lemma 4.22 with reversed inequalities. □

Lemma 4.24. *There exists a constant C such that*

$$\frac{\int_{B_1} \Phi_p}{\mu(B_1)} \leqslant C 3^{ps}.$$

Proof. Let $\mathcal{C}$ be a standard curve of return time $n \in [3^p/2, 3^p - 1]$. For $i \in [3^p, 3^{p+1}/2)$, put $\mathcal{C}_i = \{x \in \mathcal{C}, \varphi_-(T^{-1}x) = i\}$ and let $\mathcal{D}_i$ be its image by T^{-1}. This is a good curve of return time i and, by bounded distortion,

$$\frac{\int_{\mathcal{D}_i} \Phi_p}{\mathrm{Leb}(\mathcal{D}_i)} \leqslant C \frac{\int_{\mathcal{C}_i} \Phi_p}{\mathrm{Leb}(\mathcal{C}_i)}.$$

Furthermore, applying Lemma 4.23 to $\mathcal{D}_i$,we get

$$\mathrm{Leb}(\mathcal{C}_i) \frac{\int_{B_1} \Phi_p}{\mu(B_1)} \leqslant C \int_{\mathcal{C}_i} \Phi_p.$$

As by Remark 4.11 the good curves $\mathcal{C}_i$ occupy a fixed proportion of $\mathcal{C}$, we may sum up

$$\mathrm{Leb}(\mathcal{C}) \frac{\int_{B_1} \Phi_p}{\mu(B_1)} \leqslant C \int_{\mathcal{C}} \Phi_p.$$

Integrating over all standard curves of return time $\in [3^p/2, 3^p - 1]$, we obtain

$$\mu(R) \frac{\int_{B_1} \Phi_p}{\mu(B_1)} \leqslant C \int_{R} \Phi_p.$$

We may conclude by Lemma 4.21. □

Lemma 4.25. *There is a constant C such that for any $l = 2, 3, 4$,*

$$\frac{\int_{B_l} \Phi_p}{\mu(B_l)} \leqslant C 3^{ps}.$$

Proof. As the three cases are essentially identical we give the argument only for one of them, for $l = 3$, say. The proof is analogous to that of the previous lemma, we only need to apply a bit more iterations. Let $\mathcal{C}$ be a standard curve with return time from $[3^p/2, 3^p - 1]$. Given $i \in [3^p, 3^{p+1}/2)$, let $\mathcal{C}_i$ be the set of points in $\mathcal{C}$ the images of which have return time i. For $j \in [3^{p+1}/2, 3^{p+1})$, let $\mathcal{C}_{ij}$ be the set of points in $\mathcal{C}_i$ the T^{-2}-images of which have return time j. Finally, for $k \in [3^{p+1}, 3^{p+2}/2)$, we define $\mathcal{C}_{ijk}$ analogously.

By Remark 4.11, at each step we keep a fixed proportion of the previous set. Thus, there exists a constant C such that

$$\mathrm{Leb}(\mathcal{C}) \leqslant C \sum_{i,j,k} \mathrm{Leb}(\mathcal{C}_{ijk}).$$

Following the lines of the proof of Lemma 4.23 we may show that given any good curve $\mathcal{D}$ of return time from the interval $[3^{p+1}, 3^{p+2}/2)$, we have $\frac{\int_{B_3} \Phi_p}{\mu(B_3)} \leqslant C \frac{\int_{\mathcal{D}} \Phi_p}{\mathrm{Leb}(\mathcal{D})}$. This applies, in particular, to $\mathcal{D} = T^{-3}(\mathcal{C}_{ijk})$ and gives

$$\frac{\int_{B_3} \Phi_p}{\mu(B_3)} \leqslant C \frac{\int_{T^{-3}\mathcal{C}_{ijk}} \Phi_p}{\mathrm{Leb}(T^{-3}\mathcal{C}_{ijk})} \leqslant C \frac{\int_{\mathcal{C}_{ijk}} \Phi_p}{\mathrm{Leb}(\mathcal{C}_{ijk})},$$

by bounded distortion. We may apply Lemma 4.21, just as we did in the proof of Lemma 4.24, to get the desired conclusion. □

Lemmas 4.24, 4.25 and 4.22 altogether imply the bound (41) for any good curve of return time $n \in [3^p, 3^{p+1} - 1]$. We apply this bound in the second (much easier) step of the proof of Proposition 4.20.

Proof of Proposition 4.20. Recall the notations from the beginning of the subsection: $\mathcal{C}$ is a good curve of return time $n_0 \in [3^{p_0}, 3^{p_0+1} - 1]$, for some large p_0, and the stopping time τ is related to another integer p_1 ($p_0 > p_1$): $\varphi_-(T^{-\tau(x)}(x)) \in [3^{p_1}, 3^{p_1+1} - 1]$ for all $x \in \mathcal{C} \backslash A$.

To simplify notation in this proof we define $\tau_{p_0+1}(x) = 0$ for $x \in \mathcal{C} \backslash A$. For $x \in \mathcal{C} \backslash A$ we have

$$\sum_{k=0}^{\tau(x)-1} |\varphi_-(T^{-k}x)|^s = \sum_{p=p_1+1}^{p_0} \Phi_p(T^{-\tau_{p+1}(x)}x).$$

Let $p_1 + 1 \leqslant p \leqslant p_0$ and $x \in \mathcal{C} \backslash A$. Then there is a subcurve $\mathcal{C}_i \subset \mathcal{C}$ that contains x and for which $T^{-\tau_{p+1}(x)}(\mathcal{C}_i)$, to be denoted by $\mathcal{D}_i$, is a good curve of return time from $[3^p, 3^{p+1} - 1]$. By bounded distortion

$$\frac{\int_{\mathcal{C}_i \backslash A} \Phi_p(T^{-\tau_{p+1}} y)}{\mathrm{Leb}(\mathcal{C}_i)} \leqslant C \frac{\int_{T^{-\tau_{p+1}}(\mathcal{C}_i \backslash A)} \Phi_p}{\mathrm{Leb}(\mathcal{D}_i)} \leqslant C \frac{\int_{\mathcal{D}_i \backslash A_p} \Phi_p}{\mathrm{Leb}(\mathcal{D}_i)}.$$

Now according to (41) this final quantity is bounded from above by $C 3^{ps}$. Summing up for all intervals $\mathcal{C}_i$ we obtain

$$\int_{\mathcal{C} \backslash A} \Phi_p(T^{-\tau_{p+1}(y)} y) \leqslant C 3^{ps} \, \mathrm{Leb}(\mathcal{C}).$$

Summation on p from $p_1 + 1$ to p_0 implies the statement. □

Corollary 4.26. *We have*

$$\int_Y h(e^{itH} - 1) = it \int_Y h 1_{\varphi_- \leqslant 1/|t|} H + o(t \log(1/|t|)).$$

Proof. We have

$$\left| \int_Y h 1_{\varphi_- \leqslant 1/|t|} (e^{itH} - 1 - itH) \right| \leqslant C \int_Y |h| 1_{\varphi_- \leqslant 1/|t|} |t|^{3/2} |H|^{3/2}.$$

We may estimate $|H(x)|^{3/2}$ as

$$|H(x)|^{3/2} \leqslant \left(\sum_1^{\tau(x)-1} |h(T^{-k}x)| \right)^{3/2} \leqslant \tau(x)^{1/2} \sum_0^{\tau(x)-1} |h(T^{-k}x)|^{3/2}.$$

Now put $\Phi(x) = \sum_0^{\tau(x)-1} |h(T^{-k}x)|^{3/2}$. Then for $x \in Y$ of return time n we get $|H(x)|^{3/2} \leqslant (K \log n)^{1/2} \Phi(x)$, as the standard stopping time satisfies $\tau(x) \leqslant K \log n$.

By Proposition 4.20 the average of the function Φ on $Y \cap \{\varphi_- = n\}$ is less than $Cn^{3/2}$. Putting these estimates together,

$$\begin{aligned}\int_Y |h| 1_{\varphi_- \leqslant 1/|t|} |t|^{3/2} |H|^{3/2} &\leqslant C|t|^{3/2} \sum_{n=1}^{1/|t|} \mu(\varphi_- = n) n \sqrt{\log n}\, n^{3/2} \\ &\leqslant C|t|^{3/2} \sqrt{\log(1/|t|)} |t|^{-1/2} = o(|t| \log(1/|t|)),\end{aligned}$$

while

$$\left| \int_Y h 1_{\varphi_- > 1/|t|} (e^{itH} - 1) \right| \leqslant C \int \varphi_- 1_{\varphi_- > 1/|t|} \leqslant C \sum_{n > 1/|t|} \mu(\varphi_- = n) n = O(t)$$

as $\mu(\varphi_- = n) = O(1/n^3)$. □

4.4. Exact asymptotics for H. Recall the value of I from (1), and the fact that on $Y \cap \{\varphi_- = n\}$ the function h is equivalent to nI.

Lemma 4.27. *Let* $y = \frac{1}{1 - \frac{3}{4} \log 3}$. *For any* $\varepsilon > 0$ *there exists* $N_0 \in \mathbb{N}$ *such that, for all* $n \geqslant N_0$, *for all good curves* $\mathcal{C}$ *with return time n,*

$$\left| \frac{\int_{\mathcal{C} \setminus A_{\mathcal{C}}} H}{\mathrm{Leb}(\mathcal{C} \setminus A_{\mathcal{C}})} - n(y-1)I \right| \leqslant \varepsilon n.$$

Proof. Recall the asymptotic expressions for the transition probabilities from Remark 4.11. These allow us to regard the map T^{-1} as a Markov chain. Then the statement of the lemma can be guessed by the expectation value with respect to the invariant distribution of this chain.

The rigorous proof is inductive. Note that first we fix $\varepsilon > 0$, that will correspond to the required precision in the asymptotics, and then we may choose n arbitrarily large. Let $L \in \mathbb{N}$ be an integer for which $(9/10)^L \leqslant \varepsilon$. This integer L is the number of inductive steps needed to obtain ε-precision. More precisely, if $n_0, \dots, n_L$ is an admissible sequence (here $n_0 = n$), then n_L is typically much smaller than n_0. The Birkhoff sum of h for the times between n_L and the stopping time can be estimated by the upper bound coming from Lemma 4.20, which roughly means that we only need to take care of the sum for the first L steps. This estimate will be the starting point of our induction. Then we place our standard curve "high enough" (i.e., choose n large enough) to ensure that the transition probabilities of Remark 4.11 are accurate with very good precision. These transition probabilities are responsible for the appearance of y as we decrease the length of the admissible sequence $n_0, \dots, n_i$ from $i = L$ to $i = 0$ in the induction.

Let $\mathcal{C}$ be a standard curve of return time n with the standard stopping time $(A_{\mathcal{C}}, \tau_{\mathcal{C}})$ on it. If $n_0, \dots, n_i$ is admissible with $n_0 = n$ and $i \leqslant L$, the set $\mathcal{C}_{n_0,\dots,n_i}$ is not empty, and we may consider $\mathcal{C}'_{n_0,\dots,n_i} = \mathcal{C}_{n_0,\dots,n_i} \cap (\mathcal{C} \backslash A_{\mathcal{C}})$. For $\mathcal{D} = T^{-i}(\mathcal{C}_{n_0,\dots,n_i})$, define $A = \mathcal{D} \backslash T^{-i}(\mathcal{C}'_{n_0,\dots,n_i})$ and $\tau(T^{-i}x) = \tau_{\mathcal{C}}(x) - i$. Then, for large enough n, (A, τ) is a stopping time on $\mathcal{D}$. To see this we note that $\mathrm{Leb}(A) \leqslant \mathrm{Leb}(\mathcal{D})/2$ as the number of iterations is bounded from above by L while $\mathrm{Leb}(A_{\mathcal{C}})/\mathrm{Leb}(\mathcal{C}) \to 0$ as $n \to +\infty$.

By increasing n if necessary, we may assume that for any $p > n/3^L$, and for any x with $\varphi_-(x) = p$ we have $|h(x) - pI| \leqslant p/(L3^L)$. Thus for $x \in \mathcal{C}'_{n_0,\dots,n_L}$ we have

$$|H(x) - (n_1 + \cdots + n_{L-1})I| \leqslant \sum_{k=1}^{L-1} |h(T^{-k}x) - n_k I| + \sum_{k=L}^{\tau_{\mathcal{C}}(x)-1} |h(T^{-k}(x)|, \quad (44)$$

where the first term satisfies

$$\sum_{k=1}^{L-1} |h(T^{-k}x) - n_k I| \leqslant \sum_{k=1}^{L-1} n_k/(L3^L) \leqslant n_L,$$

as $n_k \leqslant 3^L n_L$. On the other hand if we integrate the second term in (44), we may use the upper bound of Lemma 4.20. We get, for some constant C_3:

$$\left| \frac{\int_{\mathcal{C}'_{n_0,\dots,n_L}} H - (n_1 + \cdots + n_{L-1})I}{\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_L})} \right| \leqslant C_3 n_L. \quad (45)$$

Choose n large enough to ensure that (i) all the ε_p from Remark 4.11 are less than ε whenever $p > n/3^L$, and that (ii) the distortion of any $T^{-i}\ |_{\mathcal{C}_{n_0,\dots,n_i}}$, $i \leq L$, is bounded from above by ε.

As $y\frac{3}{4}\log 3 - y + 1 = 0$ we have, for n large enough,

$$\left| y \sum_{p/3+C_2}^{3p-C_2} \frac{3}{8k} - y + 1 \right| < \varepsilon \quad (46)$$

whenever $p > n/3^L$.

To simplify notation we introduce $\alpha = \frac{9}{10}$ and another positive number $\beta > 2\log 3$ which is, however, not too big so that $\frac{3}{8}\beta < \alpha$. By further increasing n, if necessary, we may also assume that $\sum_{p/3}^{3p} \frac{1}{k} \leqslant \beta$ whenever $p > n/3^L$.

Now, by induction on decreasing i we show the following bound:

$$\left| \frac{\int_{\mathcal{C}'_{n_0,\dots,n_i}} H - (n_1 + \cdots + n_{i-1})I}{n_i\, \mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_i})} - yI \right| \leqslant \alpha^{L-i}(C_3 + y|I|) + C_4 \sum_{k=i}^{L-1} \varepsilon \alpha^{k-i}, \quad (47)$$

where C_4 is some constant. Note that for $i = 0$, when the sum $n_1 + \cdots + n_{i-1}$ is to be interpreted as $-n_0$, this bound implies the statement of Lemma 4.27. On the other hand, the case $i = L$ is already established in (45). So let us assume (47) holds for i, and show it for $i-1$. We have

$$\begin{aligned}
&\frac{\int_{\mathcal{C}'_{n_0,\dots,n_{i-1}}} H - (n_1+\dots+n_{i-2})I}{n_{i-1}\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})} - yI \\
&\quad = \frac{\int_{\mathcal{C}'_{n_0,\dots,n_{i-1}}} H - (n_1+\dots+n_{i-1})I}{n_{i-1}\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})} - yI + I \\
&\quad = \sum_{n_i=n_{i-1}/3+C_2}^{3n_{i-1}-C_2} \frac{\int_{\mathcal{C}'_{n_0,\dots,n_i}} H - (n_1+\dots+n_{i-1})I}{n_{i-1}\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})} - yI + I \\
&\quad = \sum_{n_i=n_{i-1}/3+C_2}^{3n_{i-1}-C_2} \left(\frac{\int_{\mathcal{C}'_{n_0,\dots,n_i}} H - (n_1+\dots+n_{i-1})I}{n_i\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_i})} \right. \\
&\qquad \left. \times \frac{n_i\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}{n_{i-1}\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})} - \frac{3}{8n_i} yI \right) \\
&\qquad + \left(y \sum_{n_i=n_{i-1}/3+C_2}^{3n_{i-1}-C_2} \frac{3}{8n_i} - y + 1 \right) I.
\end{aligned}$$

Note that as $n_0, \dots, n_L$ is admissible, $\mathcal{C}'_{n_0,\dots,n_i}$ has positive length for any $i \le L$, thus the denominators are never zero in these expressions. The choice of a large enough n ensures that even n_i is large enough so that (46) applies:

$$\left| y \sum_{n_i=n_{i-1}/3+C_2}^{3n_{i-1}-C_2} \frac{3}{8n_i} - y + 1 \right| \leqslant \varepsilon.$$

Now we will use the transition probabilities (35) on the curve $T^{-(i-1)}(\mathcal{C}_{n_0,\dots,n_{i-1}})$. We will also use that the distortions of $T^{-(i-1)}$, when restricted to this curve, are bounded from above by ε. Note furthermore that $\mathcal{C}'_{n_0,\dots,n_{i-1}}$ occupies at least $(1-\varepsilon)$-proportion of $\mathcal{C}_{n_0,\dots,n_{i-1}}$ if n is large enough (we may apply Proposition 4.15). The same holds for $\mathcal{C}'_{n_0,\dots,n_i}$ in $\mathcal{C}_{n_0,\dots,n_i}$. These observations allow us to obtain (note $n_{i-1}/n_i \le 3$):

$$\left| \frac{\operatorname{Leb}(\mathcal{C}_{n_0,\dots,n_i})}{\operatorname{Leb}(\mathcal{C}_{n_0,\dots,n_{i-1}})} - \frac{\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}{\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})} \right| \leqslant 2\varepsilon \frac{\operatorname{Leb}(\mathcal{C}_{n_0,\dots,n_i})}{\operatorname{Leb}(\mathcal{C}_{n_0,\dots,n_{i-1}})} \leqslant \frac{C\varepsilon}{n_i}.$$

One more reference to Remark 4.11 and to the fact that the distortions can be made smaller than ε if n is large enough implies

$$\left| \frac{n_i\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}{n_{i-1}\operatorname{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})} - \frac{3}{8n_i} \right| \leqslant \frac{C_5\varepsilon}{n_i}.$$

By the triangular inequality,

$$\begin{aligned}
&\left|\frac{\int_{\mathcal{C}'_{n_0,\dots,n_i}} H-(n_1+\dots+n_{i-1})I}{n_i\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}\frac{n_i\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}{n_{i-1}\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})}-\frac{3}{8n_i}yI\right|\\
&\quad\leqslant\left|\frac{\int_{\mathcal{C}'_{n_0,\dots,n_i}} H-(n_1+\dots+n_{i-1})I}{n_i\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}-yI\right|\frac{3}{8n_i}\\
&\qquad+\left|\frac{\int_{\mathcal{C}'_{n_0,\dots,n_i}} H-(n_1+\dots+n_{i-1})I}{n_i\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}\right|\left|\frac{n_i\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_i})}{n_{i-1}\,\mathrm{Leb}(\mathcal{C}'_{n_0,\dots,n_{i-1}})}-\frac{3}{8n_i}\right|.
\end{aligned}$$

Let B_i be the bound at step i of the induction. Then the first term is bounded from above by $\frac{3B_i}{8n_i}$, and the second term is bounded from above by $\frac{(B_i+y|I|)C_5\varepsilon}{n_i}$.

Recall the definitions of α and β, we have $\sum_{p/3}^{3p}\frac{1}{k}\leqslant\beta$ and, if ε is small enough, $(\frac{3}{8}+C_5\varepsilon)\beta<\alpha$.

Putting our estimates together we get

$$\begin{aligned}
B_{i-1}&=\varepsilon|I|+\sum_{n_{i-1}/3+C_2}^{3n_i-C_2}\left[\frac{3B_i}{8n_i}+\frac{(B_i+y|I|)C_5\varepsilon}{n_i}\right]\\
&\leqslant\varepsilon|I|+\left(\frac{3B_i}{8}+(B_i+y|I|)C_5\varepsilon\right)\beta\\
&\leqslant(|I|+C_5y|I|\beta)\varepsilon+\alpha B_i.
\end{aligned}$$

Now if (47) holds for i with $C_4=|I|+C_5y|I|\beta$, it holds for $i-1$ with the same constants.

Taking $i=0$ we get

$$\frac{1}{n}\left|\frac{\int_{\mathcal{C}\backslash A_{\mathcal{C}}} H}{\mathrm{Leb}(\mathcal{C}\backslash A_{\mathcal{C}})}-n(y-1)I\right|\leqslant C\alpha^L+C\varepsilon\leqslant C\varepsilon$$

by the choice of L. Note that the constant C depends only on I, thus it can be “swallowed” by ε. This completes the proof of the lemma. □

Proposition 4.28. *We have*

$$\int_Y h1_{\varphi_-\leqslant 1/|t|}H=\left(\frac{I^2(y-1)\ell^2}{4}+o(1)\right)\log(1/|t|).$$

Proof. First let us show that

$$\int_Y(h-\varphi_-I)1_{\varphi_-\leqslant 1/|t|}H=o(\log(1/|t|)). \tag{48}$$

Fix $\varepsilon>0$. If N is large enough we have $|h-\varphi_-I|\leqslant\varepsilon\varphi_-$ for $\varphi_-\geqslant N$. Thus we get (note that H is integrable, cf. Remark 4.19)

$$\left|\int_Y(h-\varphi_-I)1_{\varphi_-\leqslant 1/|t|}H\right|\leqslant O(1)+\sum_{N\leqslant n\leqslant 1/|t|}\varepsilon n\int_{Y\cap\{\varphi_-=n\}}|H|.$$

We may apply Proposition 4.20 with $s = 1$ to show $\int_{Y\cap\{\varphi_-=n\}} |H| \leqslant Cn\mu(\varphi_- = n) = O(1/n^2)$. Thus we get

$$\left|\int_Y (h - \varphi_- I)1_{\varphi_-\leqslant 1/|t|} H\right| \leqslant O(1) + C\varepsilon \log(1/|t|) \leqslant C'\varepsilon \log(1/|t|).$$

As the above inequality is true for any fixed $\varepsilon > 0$, we get (48).

Now we estimate

$$\int_Y \varphi_- I 1_{\varphi_-\leqslant 1/|t|} H = \sum_{n=1}^{1/|t|} nI \int_{Y\cap\{\varphi_-=n\}} H.$$

Since the estimate of Lemma 4.27 is uniform in the curve $\mathcal{C}$, it can be integrated and we have $\int_{Y\cap\{\varphi_-=n\}} H \sim (y-1)In\mu(\varphi_- = n) \sim (y-1)In\frac{\ell^2}{4n^3}$. Actually, the measure of the set $\{\varphi_- = n\}$ can be estimated by direct geometric arguments. Up to negligible terms, it is equivalent to $\frac{\ell^2}{16n^3}$ in all relevant zones of X which are "corners of parallelograms" (see Convention 4.4). As there are 4 such relevant zones we obtain the above formula.

Finally we get

$$\int_Y \varphi_- I 1_{\varphi_-\leqslant 1/|t|} H \sim \sum_{n=1}^{1/|t|} nI^2(y-1)n\frac{\ell^2}{4n^3} \sim \frac{I^2(y-1)\ell^2}{4}\log(1/|t|),$$

which completes the proof. □

Proposition 4.1 follows from the combination of Proposition 4.28, Corollary 4.26, Proposition 4.17 and Lemmas 4.3, 4.2.

5. Proof of the Main Theorems

In this section, we prove Theorems 1.1 and 1.4. The main tool will be an abstract theorem showing that, if an induced map satisfies a limit theorem, then the original map satisfies the same limit theorem. Such a result has been proved in the case of flows by [MT04], and extended to the discrete time case (and to non-polynomial normalizations) in [Gou03]. For the convenience of the reader, we state here the result we will use.

If Y is a subset of a probability space (X, m), $T : X \to X$, and T_Y is the induced map on Y, we will write $S_n^Y g = \sum_{k=0}^{n-1} g \circ T_Y^k$: this is the Birkhoff sum of g, for the transformation T_Y. We will also write $E_Y(g) = \frac{\int_Y g}{m[Y]}$. Finally, for $t \in \mathbb{R}$, $\lfloor t \rfloor$ denotes the integer part of t.

Theorem 5.1. *Let $T : X \to X$ be an ergodic endomorphism of a probability space (X, m), and $f : X \to \mathbb{R}$ an integrable function with vanishing integral. Let $Y \subset X$ have positive measure. For $y \in Y$, write $\varphi(y) = \inf\{n > 0 \mid T^n(y) \in Y\}$ and $f_Y(y) = \sum_{k=0}^{\varphi(y)-1} f(T^k y)$.*

We assume the following properties:

1. *There exists a sequence* $B_n \to +\infty$, *with* $\inf_{r \geqslant n} \frac{B_r}{B_n} > 0$, *such that* f_Y *satisfies a limit theorem for the normalization* B_n*: there exists a random variable* Z *such that, for every* $t \in \mathbb{R}$,

$$E_Y\left(e^{it\frac{S^Y_{\lfloor nm(Y)\rfloor} f_Y}{B_n}}\right) \to E\left(e^{itZ}\right). \tag{49}$$

2. *There exists* $b > 0$ *such that, in the natural extension of* T_Y, $\frac{1}{N^b}\sum_0^{N-1} f_Y(T_Y^k y)$ *tends almost everywhere to* 0 *when* $N \to \pm\infty$.
3. *There exists* $B'_n = O(B_n^{1/b})$ *such that* $\frac{S^Y_n\varphi - nE_Y(\varphi)}{B'_n}$ *converges in distribution.*

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

$$E\left(e^{it\frac{S_n f}{B_n}}\right) \to E(e^{itZ}),$$

i.e., $\frac{S_n f}{B_n}$ *tends in distribution to* Z.

The first assumption is apparently different from the first assumption in [Gou03, Theorem A.1]. However, they are equivalent by [Eag76] (see also [MT04]).

Remark 5.2. An analogous theorem holds in the case of flows, when Y is a Poincaré section of the flow and φ is the return time to this Poincaré section, with the same proof. Since a Poincaré section has usually zero measure, it has to be formulated slightly differently: E_Y will be the expectation with respect to the probability measure induced by m on Y, and in (49) $m(Y)$ should be replaced with $1/E_Y(\varphi)$. Finally, the sums (in the definition of f_Y, and in the definition of the Birkhoff sums of f) should be replaced with integrals, and correspondingly, the normalizing sequences B_n (B'_n) with appropriate functions $B(T)$ ($B'(T)$), $B : \mathbb{R}_+ \to \mathbb{R}_+$.

5.1. Proof of Theorem 1.1 for functions satisfying $(P1)$. Let $f_0 : X_0 \to \mathbb{R}$ be Hölder continuous and satisfy $(P1)$. In particular, $I \neq 0$. Define as in Sect. 2 functions f, $\bar{f}$, $\bar{g}$ and g. Since f satisfies (6) and $\bar{g} - \bar{f}$ is bounded, we obtain $\mu_\Delta(|g| > x) \sim x^{-2}l(x)$, where

$$l(x) = \frac{I^2\ell^2}{8}.$$

By Paragraph 3.1, the function g is in the nonstandard domain of attraction of the normal law. More precisely, set

$$L(x) = \frac{I^2\ell^2}{4}\log(x) \sim 2\int_1^x \frac{l(u)}{u}.$$

The functions l and L are the tail functions of g, as defined in Paragraph 3.1.

Proposition 4.1 gives

$$\int g(e^{itG} - 1) = (y - 1)itL(1/|t|) + o(tL(1/|t|)),$$

where $y = \frac{1}{1-\frac{3}{4}\log 3}$. Moreover, the function g is locally Hölder on Δ, by (10). Hence, all the assumptions of Theorem 3.4 are satisfied, for $a = y - 1 > 0$. Let

$$B_n = \sqrt{n \log n \frac{(2y-1)I^2\ell^2}{8}};$$

it satisfies $\frac{n}{B_n^2}(2a+1)L(B_n) \to 1$. Hence, by Theorem 3.4, we obtain that $\frac{\sum_{k=0}^{n-1} g\circ U^k}{B_n} \to \mathcal{N}(0, 1)$ in distribution with respect to μ_Δ. This is equivalent to the same convergence for $\bar{g}$, with respect to $\mu_{\bar{\Delta}}$, since $\bar{g} = g \circ \pi_\Delta$ and $\mu_\Delta = (\pi_\Delta)_*(\mu_{\bar{\Delta}})$. Since $\bar{f}$ is cohomologous to $\bar{g}$, we get the same convergence for $\bar{f}$. Finally, since $\bar{f} = f \circ \pi_X$ and $\mu = (\pi_X)_*(\mu_{\bar{\Delta}})$, we get that

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

on X, with respect to μ.

The same argument applies to $\varphi_+ - \int \varphi_+$, and we get that $\frac{\sum_{k=0}^{n-1} \varphi_+\circ T^k - n\int \varphi_+}{B_n}$ converges in distribution. Hence, Theorem 5.1 applies, with $b = 1$.

Set $B'_n = B_{\lfloor n\mu_0(X)\rfloor}$. Since $\mu_0(X) = \frac{2}{\pi+\ell}$ by (4), we get

$$B'_n \sim \sqrt{n \log n \frac{(2y-1)I^2\ell^2}{4(\pi+\ell)}}.$$

Theorem 5.1 yields

$$\frac{\sum_{k=0}^{n-1} f_0 \circ T_0^k}{B'_n} \to \mathcal{N}(0, 1).$$

This concludes the proof of Theorem 1.1.

5.2. Proof of Theorem 1.4. Let $f_0 : X_0 \to \mathbb{R}$ be Hölder continuous with $\int f_0 = 0$ and $I = 0$. In this case, we can not use the cohomology trick any more, since the proofs of Lemmas 2.5 and 2.6 relied heavily on the property $(P1)$. The argument will be to induce on the basis of the tower $\bar{\Delta}$, prove a central limit theorem here (using Gordin's martingale argument), and then get back to the original space by using Theorem 5.1 twice. The main difference in the inducing process with the previous paragraph is that we can no more apply Theorem 5.1 with $b = 1$. Hence, we will need to prove that $\frac{1}{|n|^b}\sum_{k=0}^{n-1} f \circ T^k$ converges almost everywhere to 0, for some $b < 1$. Many arguments of this paragraph are strongly inspired by [You98], with additional technical complications due to the fact that our functions are not bounded.

Let $\bar{\Delta}_0$ be the basis of the tower $\bar{\Delta}$, and let $\bar{U}_0$ be the induced map on $\bar{\Delta}_0$ (with a return time φ). Define a new function $\bar{f}_0$ on $\bar{\Delta}_0$, by $\bar{f}_0(x) = \sum_{k=0}^{\varphi(x)-1} \bar{f}(\bar{U}^k x)$.

Lemma 5.3. *There exists $\sigma_0^2 \geqslant 0$ such that*

$$\frac{\sum_{k=0}^{n-1} \bar{f}_0 \circ \bar{U}_0^k}{\sqrt{n}} \to \mathcal{N}(0, \sigma_0^2).$$

Proof. Since $I = 0$, it is not hard to check that there exists $\alpha_1 < 1$ such that $|f| \leqslant n^{\alpha_1}$ on the set of points bouncing n times between the segments of the stadium. This implies that there exists $\varepsilon_1 > 0$ such that $f \in L^{2+\varepsilon_1}(X)$. Hence, $\bar{f} \in L^{2+\varepsilon_1}(\bar{\Delta})$. Since the return time φ belongs to L^p for all $p < \infty$, we get $\bar{f}_0 \in L^{2+\varepsilon_2}(\bar{\Delta}_0)$ for some $\varepsilon_2 > 0$.

Let Δ_0 be obtained by identifying the points on the same stable leaf. It is the basis of the expanding Young tower Δ. Let $\pi_0 : \bar{\Delta}_0 \to \Delta_0$ be the canonical projection, and U_0 the dynamics induced by $\bar{U}_0$ on Δ_0. Let $\mathcal{B}_0$ be the σ-algebra on $\bar{\Delta}_0$ obtained by pulling by π_0 the σ-algebra on Δ_0. A measurable subset B of $\bar{\Delta}_0$ is $\mathcal{B}_0$-measurable if, for almost all $x \in B$, the stable leaf through x is contained in B.

We will prove

$$\sum_{n \geqslant 0} \left\| E(\bar{f}_0 \mid \bar{U}_0^n \mathcal{B}_0) - \bar{f}_0 \right\|_{L^2} < \infty \tag{50}$$

and

$$\sum_{n \geqslant 0} \left\| E(\bar{f}_0 \mid \bar{U}_0^{-n} \mathcal{B}_0) \right\|_{L^2} < \infty. \tag{51}$$

By Gordin's Theorem [Gor69], this will imply the conclusion of the lemma.

The basis $\bar{\Delta}_0$ corresponds to a rectangle R for the dynamics T, which is naturally partitioned as $R = \bigcup R_i$, where R_i is an s-subrectangle of R. Let $\bar{\Delta}_{0,i}$ be the corresponding subset of $\bar{\Delta}_0$, so that $\{\bar{\Delta}_{0,i}\}$ gives a partition of $\bar{\Delta}_0$. Define a function $A : \bar{\Delta}_0 \to \mathbb{R}$ by $A(x) = \sum_{k=0}^{\varphi(x)-1} \varphi_+(\pi_X \bar{U}^k x)$. It is constant on each set $\bar{\Delta}_{0,i}$, and corresponds to the number of times the original map T_0 is to be applied to R_i so that this s-subrectangle makes a full (Markov) return to the base R. Since φ belongs to every $L^p(\bar{\Delta}_0)$ for $p \geqslant 1$ and $\varphi_+ \in L^p(X)$ for $1 \leqslant p < 2$, the function A belongs to $L^p(\bar{\Delta}_0)$ for $1 \leqslant p < 2$. If x, y are on the same unstable leaf in a rectangle $\bar{\Delta}_{0,i}$, we have

$$|\bar{f}_0(x) - \bar{f}_0(y)| \leqslant CA(x)\tau^{s(x,y)} \tag{52}$$

for some constant $C > 0$ and some constant $\tau < 1$. Here, $s(x, y)$ is the separation time of x and y. Moreover, if x, y are on the same stable leaf in a rectangle $\bar{\Delta}_{0,i}$,

$$|\bar{f}_0(x) - \bar{f}_0(y)| \leqslant CA(x)d(\pi_X x, \pi_X y)^\alpha \tag{53}$$

for some $\alpha > 0$.

Since the stable leaves are contracted at each iteration by at least $\lambda < 1$, the atoms of the σ-algebra $\bar{U}_0^n \mathcal{B}_0$ have a diameter at most $C\lambda^n$. By (53), we get

$$\left| \bar{f}_0(x) - E(\bar{f}_0 \mid \bar{U}_0^n \mathcal{B}_0)(x) \right| \leqslant CA(x)\lambda^{\alpha n}. \tag{54}$$

Unfortunately, A does not belong to L^2, so a further argument is required to get (50). Let $p > 0$ be such that $\frac{1}{p} + \frac{1}{2+\varepsilon_2} = \frac{1}{2}$. By (54),

$$1_{A \leqslant n^{2p}} \left| \bar{f}_0(x) - E(\bar{f}_0 \mid \bar{U}_0^n \mathcal{B}_0)(x) \right| \leqslant Cn^{2p}\lambda^{\alpha n}.$$

Hence, this series is summable in L^2. Moreover,

$$\left\| 1_{A > n^{2p}} \bar{f}_0 \right\|_{L^2} \leqslant \left\| 1_{A > n^{2p}} \right\|_{L^p} \left\| \bar{f}_0 \right\|_{L^{2+\varepsilon_2}} \leqslant \frac{\left(\int A\right)^{1/p}}{n^2} \left\| \bar{f}_0 \right\|_{L^{2+\varepsilon_2}}.$$

The function $E(\bar f_0 \mid \bar U_0^n \mathcal{B}_0)$ is bounded in $L^{2+\varepsilon_2}$ by $\|\bar f_0\|_{L^{2+\varepsilon_2}}$. Hence, we obtain

$$\left\| 1_{A>n^{2p}} \left| \bar f_0 - E(\bar f_0 \mid \bar U_0^n \mathcal{B}_0)\right| \right\|_{L^2} = O(1/n^2),$$

which is summable. This proves (50).

Let $\bar h = E(\bar f_0 \mid \mathcal{B}_0)$. This function is constant along the stable leaves, and has zero integral (since $\bar f_0$ also has zero integral). Hence, it induces a function h on the quotient Δ_0. Since $\bar f_0 \in L^2$, it satisfies $h \in L^2(\Delta_0)$. The following lemma is an easy consequence of the Hölder properties of the invariant measure and (52), see [You98, Sublemma, p. 612] for details.

Lemma 5.4. *There exist constants $C > 0$ and $\tau < 1$ such that, for all x, y in the same unstable leaf of a set $\bar\Delta_{0,i}$,*

$$|\bar h(x) - \bar h(y)| \leqslant C A(x) \tau^{s(x.y)}.$$

The function A is integrable. Hence, by [Gou04, Lemma 3.4], this implies that the function $\widehat{U}_0 h$ is Hölder continuous on Δ_0. By [Gou04, Corollary 3.3], we get:

$$\begin{aligned}&\widehat{U}_0^n h \text{ tends exponentially fast to } 0\\ &\quad \text{in the space of Hölder continuous functions on } \Delta_0.\end{aligned} \tag{55}$$

A computation gives

$$\begin{aligned}\left\| E(\bar f_0 \mid \bar U_0^{-n} \mathcal{B}_0)\right\|_{L^2}^2 &= \int h \cdot (\widehat{U}_0^n h) \circ U_0^n \leqslant \|h\|_{L^2} \left\| (\widehat{U}_0^n h)\circ U_0^n \right\|_{L^2}\\ &= \|h\|_{L^2} \left\| \widehat{U}_0^n h\right\|_{L^2}.\end{aligned}$$

Hence, this term is exponentially small. This proves (51) and concludes the proof of Lemma 5.3. □

The return time φ also satisfies a central limit theorem, by the same argument. Hence, by Theorem 5.1 (applied with $b = 1$), there exists $\sigma_1^2 \geqslant 0$ such that

$$\frac{\sum_{k=0}^{n-1} \bar f \circ \bar U^k}{\sqrt{n}} \to \mathcal{N}(0, \sigma_1^2).$$

Going from $\bar\Delta$ to X, it implies that

$$\frac{\sum_{k=0}^{n-1} f \circ T^k}{\sqrt{n}} \to \mathcal{N}(0, \sigma_1^2). \tag{56}$$

Moreover, the return time $\varphi_+ : X \to \mathbb{N}$ satisfies a limit theorem with normalization $\sqrt{n \log n}$. Since $\sqrt{n} = o(\sqrt{n \log n})$, we can unfortunately not apply Theorem 5.1 with $b = 1$. However, if we can prove the following lemma, then this theorem applies with $b < 1$.

Lemma 5.5. *For all $b > 1/2$,*

$$\frac{1}{|n|^b} \sum_{k=0}^{n-1} f \circ T^k \to 0 \tag{57}$$

almost everywhere in X when $n \to \pm\infty$.

Proof. We first estimate the decay of correlations of $\bar{f}_0$ for $\bar{U}_0$. We will use the notations of the proof of Lemma 5.3. We have

$$\int \bar{f}_0 \cdot \bar{f}_0 \circ \bar{U}_0^{2n} = \int \bar{f}_0 \cdot E(\bar{f}_0 \circ \bar{U}_0^n \mid \mathcal{B}_0) \circ \bar{U}_0^n + \int \bar{f}_0 \cdot \left(\bar{f}_0 \circ \bar{U}_0^{2n} - E(\bar{f}_0 \circ \bar{U}_0^n \mid \mathcal{B}_0) \circ \bar{U}_0^n\right). \quad (58)$$

The contraction properties of $\bar{U}_0$ along stable manifolds and (53) give $|\bar{f}_0 \circ \bar{U}_0^n(x) - E(\bar{f}_0 \circ \bar{U}_0^n \mid \mathcal{B}_0)(x)| \leqslant CA(\bar{U}_0^n x)\lambda^{\alpha n}$. Hence, the second integral in (58) is at most

$$\int |\bar{f}_0| \cdot A \circ \bar{U}_0^{2n} \lambda^{\alpha n} \leqslant \|\bar{f}_0\|_{L^{2+\varepsilon_2}} \|A\|_{L^p} \lambda^{\alpha n},$$

where $p<2$ is chosen so that $\frac{1}{2+\varepsilon_2}+\frac{1}{p}=1$. Hence, this term decays exponentially fast.

In the first integral of (58), the function $E(\bar{f}_0 \circ \bar{U}_0^n \mid \mathcal{B}_0) \circ \bar{U}_0^n$ is $\mathcal{B}_0$-measurable (i.e., constant along stable leaves). Hence, this integral is equal to

$$\int \bar{h} \cdot E(\bar{f}_0 \circ \bar{U}_0^n \mid \mathcal{B}_0) \circ \bar{U}_0^n. \quad (59)$$

Let $\bar{h}_n = E(\bar{f}_0 \circ \bar{U}_0^n \mid \mathcal{B}_0)$, it is $\mathcal{B}_0$-measurable and defines a function h_n on the quotient Δ_0. The integral (59) is then equal to

$$\int_{\Delta_0} h \cdot h_n \circ U_0^n = \int \widehat{U}_0^n h \cdot h_n. \quad (60)$$

The L^2-norm of h_n is bounded independently of n. By (55), (60) is exponentially small. This proves that $\int \bar{f}_0 \cdot \bar{f}_0 \circ \bar{U}_0^{2n}$ decays exponentially. In the same way, $\int \bar{f}_0 \cdot \bar{f}_0 \circ \bar{U}_0^{2n+1}$ decays exponentially.

Since the correlations of $\bar{f}_0$ decay exponentially fast and $\bar{f}_0 \in L^2$, [Kac96, Theorem 16] implies that $\frac{1}{n^b} \sum_{k=0}^{n-1} \bar{f}_0 \circ \bar{U}_0^k$ tends to zero almost everywhere when $n \to +\infty$, for all $b > 1/2$.

Now to see that $\frac{1}{n^b} \sum_{k=0}^{n-1} \bar{f} \circ \bar{U}^k$ tends to zero almost everywhere in $\bar{\Delta}$ when $n \to +\infty$, for all $b > 1/2$, we use [MT04, Lemma 2.1 (a)] which gives this convergence on $\bar{\Delta}_0$. However, by the ergodicity of $\bar{U}$, the set on which this convergence holds must have either full or zero measure. As $\bar{\Delta}_0$ has positive measure, we get this convergence almost everywhere on $\bar{\Delta}$. Finally, this implies the same for f in X. We have proved (57) for any $b > 1/2$ when $n \to +\infty$.

To deal with $n \to -\infty$, we go to the natural extension. It is sufficient to prove the result for $\bar{f}_0$ in $\bar{\Delta}_0$, since the previous reasoning still applies (using the fact that the natural extension is functorial, i.e., the natural extension commutes with induction and projections). In the natural extension $\bar{\Delta}_0'$ of $\bar{\Delta}_0$, we have $\int \bar{f}_0' \cdot \bar{f}_0' \circ \bar{U}_0'^{-n} = \int \bar{f}_0 \circ \bar{U}_0^n \cdot \bar{f}_0$, which is exponentially small. Hence, [Kac96, Theorem 16] still applies and gives the desired result. □

Remark 5.6. As $\mu_0(X) > 0$, we may apply [MT04, Lemma 2.1 (a)] just as we did in the proof above to see that Lemma 5.5 implies

$$\frac{1}{|n|^b} \sum_{k=0}^{n-1} f_0 \circ T_0^k \to 0$$

almost everywhere when $n \to \pm\infty$, for any $b > 1/2$.

Proof of Theorem 1.4. The convergence (56), together with Lemma 5.5 and Theorem 5.1, implies (2).

We still have to prove the zero variance statement. If $f_0 = \chi - \chi \circ T_0$ for some measurable function χ, then $S_n f_0/\sqrt{n} = (\chi - \chi \circ T_0^n)/\sqrt{n}$ tends in probability to 0, which implies $\sigma = 0$. Conversely, assume that $\sigma = 0$. The function $\bar{f}_0$ on the basis $\bar{\Delta}_0$ of the Young tower satisfies a central limit theorem with zero variance. The proof of Gordin's theorem then ensures the existence of a measurable function $\bar{\chi}_0$ such that $\bar{f}_0 = \bar{\chi}_0 - \bar{\chi}_0 \circ \bar{U}_0$, i.e., $\bar{f}_0$ is a coboundary on $\bar{\Delta}_0$. This implies that $\bar{f}$ is a coboundary on $\bar{\Delta}$, as follows: let $\bar{\pi}_0 : \bar{\Delta} \to \bar{\Delta}_0$ be the projection on the basis of the tower. Defining $\bar{\chi} : \bar{\Delta} \to \mathbb{R}$ by

$$\bar{\chi}(x) = \bar{\chi}_0(\bar{\pi}_0 x) - \sum_{k=0}^{\omega(x)-1} \bar{f}(\bar{U}^k \bar{\pi}_0 x), \tag{61}$$

we have $\bar{f} = \bar{\chi} - \bar{\chi} \circ \bar{U}$.

Since the function $\bar{f} = f \circ \pi_X$ is a coboundary, general results on coboundaries (see e.g. [Gou05a, Theorem 1.4]) ensure that f also is a coboundary on X for T. Finally, this implies that f_0 is a coboundary on X_0 for T_0, using a formula similar to (61). □

5.3. Proof of Proposition 1.5. We work in the stadium billiard with $\ell = \ell^*$, for which the free flight τ_0 satisfies a usual central limit theorem. Define a function $\tau : X \to \mathbb{R}$ by $\tau(x) = \sum_{k=0}^{\varphi_+(x)-1} \tau_0^*(T_0^k x)$. Since the function τ_0^* does not satisfy $(P1)$, Lemma 2.5 does *a priori* not apply. Nevertheless, due to the geometric properties of the free flight, the function τ satisfies the following inequality: if $x, y \in X$ are two points sliding n times along the semicircles, then $|\tau(x) - \tau(y)| \leqslant C(d(x,y) + d(Tx, Ty))$. This estimate is sufficient to carry out the proofs of Lemmas 2.5 and 2.6. Hence, there exist two functions $\bar{u} : \bar{\Delta} \to \mathbb{R}$ and $g : \Delta \to \mathbb{R}$ such that $\tau \circ \pi_X = g \circ \pi_\Delta + \bar{u} - \bar{u} \circ \bar{U}$ on $\bar{\Delta}$, $\bar{u}$ is bounded on $\bar{\Delta}$ and g is Hölder continuous on Δ. Let Δ_0 be the basis of the tower Δ, and let g_0 be the function induced by g on Δ_0 (given by $g_0(x) = \sum_{k=0}^{\varphi_0(x)-1} g(U^k x)$, where φ_0 is the return time from Δ_0 to itself).

Assume that $\sigma = 0$. By Theorem 1.4, this implies that τ_0^* is a coboundary. In turn, arguments similar to the end of the proof of Theorem 1.4 show that g_0 itself is a coboundary. Since g is Hölder continuous, g_0 satisfies the assumptions of [Gou05b, Theorem 1.1]. This theorem implies that the function g_0 is essentially bounded.

Let us show that this is not the case. The function τ is bounded from above: since $E(\tau_0) = 2$, the function τ is $O(1)$ on the set of points bouncing n times between the segments. Moreover, on the set of points sliding n times along the circles, the function τ is equal to $-2n + O(1)$. Since $\bar{u}$ is bounded, this implies that there exists a constant C_1 such that $g \leqslant C_1$, and that $g = -2n + O(1)$ on a set of measure at least C/n^4.

Let $A_n \subset \Delta$ be the set of points in the tower where $\varphi_0(\pi_0 x) \leqslant n/(2C_1)$ (where $\pi_0 : \Delta \to \Delta_0$ is the projection on the basis) and $g \leqslant -n$. Since $\mu_\Delta\{\varphi_0(\pi_0 x) > n/(2C_1)\}$ is exponentially small in n, while $\mu_\Delta\{g \leqslant -n\} \geqslant C/n^4$, the set A_n has nonzero measure for n large enough.

If $y \in \pi_0(A_n)$, then $g_0(y) \leqslant -n + \sum_{k=0}^{\varphi_0(y)-1} C_1 \leqslant -n/2$. Since $\pi_0(A_n)$ has positive measure, this shows that the function g_0 is not bounded from below. This contradiction concludes the proof. □

Acknowledgement. We are very grateful to D. Szász and T. Varjú for useful discussions and for their valuable remarks on earlier versions of the manuscript. This paper has grown out of discussions we had at the CIRM conference on multi-dimensional non-uniformly hyperbolic systems in Marseille in May 2004, and while S. G. visited the Institute of Mathematics of the BUTE in October 2004. The hospitality of both institutions, along with the financial support of Hungarian National Foundation for Scientific Research (OTKA), grants TS040719, T046187 and TS049835 is thankfully acknowledged.

A. Proof of Lemma 3.6

Let U_0 be the map induced by U on the basis Δ_0 of the tower. Denote by φ the first return time on the basis, so that $U_0(x) = U^{\varphi(x)}(x)$. Note that $\varphi(x)$ can also be defined for $x \in \Delta \setminus \Delta_0$ as the first hitting time of the basis.

Let F be a finite subset of $\mathbb{N}$. Let $(n_i)_{i \in F}$ be positive integers. Let

$$K(F, n_i) = \{x \in \Delta_0 \mid \forall i \in F, \varphi(U_0^i x) = n_i\}.$$

Lemma A.1. *There exists a constant C such that, for all F and n_i as above,*

$$\mu_\Delta(K(F, n_i)) \leqslant \prod_{i \in F} (C\rho^{n_i}).$$

Proof. The proof is by induction on $\max F$, and the result is trivial when $F = \emptyset$. Write $F' = \{i - 1 \mid i \in F, i \geqslant 1\}$ and, for $i \in F'$, set $n'_i = n_{i+1}$. If $0 \notin F$, $K(F, n_i) = U_0^{-1}(K(F', n'_i))$. Since U_0 preserves μ_Δ and $\max F' < \max F$, we get the result. Otherwise, $0 \in F$. Then $K(F, n_i) = U_0^{-1}(K(F', n'_i)) \cap \{x \in \Delta_0, \varphi(x) = n_0\}$. By bounded distortion, we get

$$\begin{aligned} \mu_\Delta(K(F, n_i)) &\leqslant C\mu_\Delta(K(F', n'_i))\mu_\Delta\{x \in \Delta_0, \varphi(x) = n_0\} \\ &\leqslant C\mu_\Delta(K(F', n'_i))\rho^{n_0}. \quad \square \end{aligned}$$

Lemma A.2. *There exist $C > 0$ and $\theta < 1$ such that, for all $n \in \mathbb{N}$,*

$$\int_{U^{-n}\Delta_0} \tau^{\Psi_n} \leqslant C\theta^n.$$

Proof. Let $\kappa > 0$ be very small (how small will be specified later in the proof). Then

$$\begin{aligned} U^{-n}\Delta_0 \subset \{x \in \Delta \mid \Psi_n(x) \geqslant \kappa n\} \cup \{x \in \Delta \mid \varphi(x) \geqslant n/2\} \\ \cup \{x \in \Delta \mid \varphi(x) < n/2, \Psi_n(x) < \kappa n\}. \end{aligned}$$

On the first of these sets, $\tau^{\Psi_n} \leqslant \tau^{\kappa n}$, whence the integral of τ^{Ψ_n} is exponentially small. The second of these sets has exponentially small measure. Finally, the last of these sets is contained in $\bigcup_{i=0}^{n/2} U^{-i}\Gamma_n$, where

$$\Gamma_n = \{x \in \Delta_0 \mid \sum_{0 \leqslant i \leqslant \kappa n} \varphi(U_0^i x) \geqslant n/2\}.$$

To conclude the proof of the lemma, it is sufficient to prove that the measure of Γ_n is exponentially small.

Take $L \in \mathbb{N}$ such that $\forall n \geqslant L, (C\rho)^n \leqslant \rho^{n/2}$, where C is the constant given by Lemma A.1. For $x \in \Gamma_n$, let $F(x) := \{0 \leqslant i \leqslant \kappa n \mid \varphi(U_0^i x) \geqslant L\}$. Then

$$\sum_{i \in F(x)} \varphi(U_0^i x) \geqslant \frac{n}{2} - \sum_{i \notin F(x)} L \geqslant (1/2 - L\kappa)n.$$

This implies that

$$\Gamma_n \subset \bigcup_{F \subset [0, \lfloor \kappa n \rfloor]} \bigcup_{\substack{n_i \geqslant L \\ \sum_{i \in F} n_i \geqslant (1/2 - L\kappa)n}} K(F, n_i).$$

By Lemma A.1, we get

$$\begin{aligned}
\mu_\Delta(\Gamma_n) &\leqslant \sum_{F \subset [0, \lfloor \kappa n \rfloor]} \sum_{\substack{n_i \geqslant L \\ \sum_{i \in F} n_i \geqslant (1/2 - L\kappa)n}} \prod_{i \in F} (C\rho^{n_i}) \\
&\leqslant \sum_{k=0}^{\lfloor \kappa n \rfloor} \binom{\lfloor \kappa n \rfloor}{k} \sum_{\substack{n_0, \dots, n_{k-1} \geqslant L \\ \sum n_i \geqslant (1/2 - L\kappa)n}} (C\rho^{n_0}) \dots (C\rho^{n_{k-1}}) \\
&\leqslant 2^{\kappa n} \sum_{0 \leqslant k \leqslant \kappa n} \sum_{\substack{n_0, \dots, n_{k-1} \geqslant L \\ \sum n_i \geqslant (1/2 - L\kappa)n}} \rho^{\sum n_i / 2} \leqslant 2^{\kappa n} \sum_{0 \leqslant k \leqslant \kappa n} \sum_{\substack{n_0, \dots, n_{k-1} \in \mathbb{N} \\ \sum n_i \geqslant (1/2 - L\kappa)n}} \rho^{\sum n_i / 2}.
\end{aligned}$$

For $r \in \mathbb{N}$,

$$\begin{aligned}
\sum_{n_0 + \dots + n_{k-1} = r} \rho^{\sum n_i / 2} &= \rho^{r/2} \operatorname{Card}\{n_0, \dots, n_{k-1} \mid \\
&\sum n_i = r\} = \rho^{r/2} \binom{r+k}{k} \leqslant \rho^{r/2} \frac{(r+k)^k}{k!}.
\end{aligned}$$

Hence,

$$\mu_\Delta(\Gamma_n) \leqslant 2^{\kappa n} \sum_{0 \leqslant k \leqslant \kappa n} \sum_{r \geqslant (1/2 - L\kappa)n} \rho^{r/2} \frac{(r+k)^k}{k!}.$$

The sequence $u_r = \rho^{r/2} \frac{(r+k)^k}{k!}$ satisfies $\frac{u_{r+1}}{u_r} \leqslant \rho' := \rho^{1/2} e^{\frac{\kappa}{1/2 - L\kappa}}$ for all $r \geqslant (1/2 - L\kappa)n$ and $k \leqslant \kappa n$. If κ is small enough, $\rho' < 1$, and we get

$$\begin{aligned}
\mu_\Delta(\Gamma_n) &\leqslant 2^{\kappa n} \sum_{0 \leqslant k \leqslant \kappa n} \rho^{(1/2 - L\kappa)n/2} \frac{\big((1/2 - L\kappa)n + \kappa n\big)^k}{k!} \frac{1}{1 - \rho'} \\
&\leqslant \frac{2^{\kappa n}}{1 - \rho'} \rho^{(1/2 - L\kappa)n/2} \sum_{0 \leqslant k \leqslant \kappa n} \frac{n^k}{k!}.
\end{aligned}$$

The sequence $\frac{n^k}{k!}$ is increasing for $k \leqslant n$. Hence, we finally get

$$\mu_\Delta(\Gamma_n) \leqslant \frac{2^{\kappa n}}{1 - \rho'} \rho^{(1/2 - L\kappa)n/2} (\kappa n + 1) \frac{n^{\lfloor \kappa n \rfloor}}{\lfloor \kappa n \rfloor!}.$$

Using Stirling's Formula, it is easy to check that this expression is exponentially small if κ is small enough. This concludes the proof. □

Proof of Lemma 3.6. Let θ be given by Lemma A.2. Choose $\alpha > 0$ so that $e^{\varepsilon\alpha}\theta < 1$. Then

$$U^{-n}\Delta_0 \subset \{x \in \Delta \mid \omega(x) \geqslant \alpha n\} \cup \Big[\{x \in \Delta \mid \omega(x) < \alpha n\} \cap U^{-n}\Delta_0\Big].$$

Hence,

$$\int_{U^{-n}\Delta_0} e^{\varepsilon\omega}\tau^{\Psi_n} \leqslant \int_{\omega \geqslant \alpha n} e^{\varepsilon\omega} + e^{\varepsilon\alpha n}\int_{U^{-n}\Delta_0} \tau^{\Psi_n}.$$

The first term is exponentially small since $e^{\varepsilon}\rho < 1$. Lemma A.2 and the definition of α also imply that the second term is exponentially small. □

References

[Aar97] Aaronson, J.: An introduction to infinite ergodic theory. Volume **50** of *Mathematical Surveys and Monographs*. Providence RI: American Mathematical Society, 1997

[AD01] Aaronson, J., Denker, M.: A local limit theorem for stationary processes in the domain of attraction of a normal distribution. In N. Balakrishnan, I.A. Ibragimov, V.B. Nevzorov, eds., *Asymptotic methods in probability and statistics with applications. Papers from the international conference, St. Petersburg, Russia, 1998*, Basel: Birkhäuser, 2001, pp. 215–224

[Bun79] Bunimovich, L.: On the ergodic properties of nowhere dispersing billiards. Commun. Math. Phys. **65**, 295–312 (1979)

[BY93] Baladi, V., Young, L.-S.: On the spectra of randomly perturbed expanding maps. Commun. Math. Phys. **156**, 355–385 (1993)

[Che97] Chernov, N.: Entropy, Lyapunov exponents and mean free path for billiards. J. Stat. Phys. **88**, 1–29 (1997)

[Che99] Chernov, N.: Decay of correlations and dispersing billiards. J. Stat. Phys. **94**, 513–556 (1999)

[CZ05] Chernov, N., Zhang, H.: Billiards with polynomial mixing rates. Nonlinearity **18**, 1527–1554 (2005)

[Eag76] Eagleson, G.K.: Some simple conditions for limit theorems to be mixing. Teor. Verojatnost. i Primenen. **21**(3), 53–660 (1976)

[Gor69] Gordin, M.: The central limit theorem for stationary processes. Dokl. Akad. Nauk SSSR **188**, 739–741 (1969)

[Gou03] Gouëzel, S.: Statistical properties of a skew-product with a curve of neutral points. http://name.math.Univ-rennes1.fr/Sebastien.gouezel/articles/skewproduct. pdf, 2004

[Gou04] Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. Probab. Theory and Rel. Fields **128**, 82–122 (2004)

[Gou05a] Gouëzel, S.: Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. http://name.math.Univ-rennes1.fr/Sebastien.gouezel/articles/vitesse-TCL.pdf, Annales de l'IHP Probabilités et Statistiques, **41**, 997–1024

[Gou05b] Gouëzel, S.: Regularity of coboundaries for non uniformly expanding Markov maps. Proc. Am. Math. Soc. **134**(2), 391–401 (2005)

[Hen93] Hennion, H.: Sur un théorème spectral et son application aux noyaux lipschitziens. Proc. Amer. Math. Soc. **118**, 627–634 (1993)

[Kac96] Kachurovskiĭ, A.G.: Rates of convergence in ergodic theorems. Russian Math. Surveys **51**, 653–703 (1996)

[KL99] Keller, G., Liverani, C.: Stability of the spectrum for transfer operators. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28**(1), 141–152 (1999)

[Mac83] Machta, J.: Power law decay of correlations in a billiard problem. J. Statist. Phys. **32**, 555–564 (1983)

[Mar04] Markarian, R.: Billiards with polynomial decay of correlations. Ergodic Theory Dynam. Systems **24**, 177–197 (2004)

[MT04] Melbourne, I., Török, A.: Statistical limit theorems for suspension flows. Israel J. Math. **144**, 191–210, 2004

[SV04a] Szász, D., Varjú, T.: Local limit theorem for the Lorentz process and its recurrence on the plane. Ergodic Theory Dynam. Systems **24**, 257–278 (2004)
[SV04b] Szász, D., Varjú, T.: Markov towers and stochastic properties of billiards. In: *Modern dynamical systems and applications*, Cambridge: Cambridge University Press, 2004, pp. 433–445
[SV05] Szász, D., Varjú, T.: Limit laws and recurrence for the planar Lorentz process with infinite horizon. Preprint, 2005
[You98] Young, L.-S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. (2) **147**, 585–650 (1998)
[You99] Young, L.-S.: Recurrence times and rates of mixing. Israel J. Math. **110**, 153–188 (1999)

Communicated by G.Gallavotti