LIMIT THEOREMS FOR DISPERSING BILLIARDS WITH CUSPS

P. BÁLINT, N. CHERNOV, AND D. DOLGOPYAT

Abstract. Dispersing billiards with cusps are deterministic dynamical systems with a mild degree of chaos, exhibiting "intermittent" behavior that alternates between regular and chaotic patterns. Their statistical properties are therefore weak and delicate. They are characterized by a slow (power-law) decay of correlations, and as a result the classical central limit theorem fails. We prove that a non-classical central limit theorem holds, with a scaling factor of $\sqrt{n \log n}$ replacing the standard $\sqrt{n}$. We also derive the respective Weak Invariance Principle, and we identify the class of observables for which the classical CLT still holds.

Keywords: Dispersing billiards, Sinai billiards, cusps, decay of correlations, central limit theorem, weak invariance principle.

1. Introduction

We study billiards, i.e., dynamical systems where a point particle moves in a planar domain $\mathcal{D}$ (the billiard table) and bounces off its boundary $\partial\mathcal{D}$ according to the classical rule "the angle of incidence is equal to the angle of reflection". The boundary $\partial\mathcal{D}$ is assumed to be a finite union of C^3 smooth compact curves that may have common endpoints.

Between collisions at $\partial\mathcal{D}$, the particle moves with a unit speed and its velocity vector remains constant. At every collision, the velocity vector changes by

$$\mathbf{v}^+ = \mathbf{v}^- - 2\langle \mathbf{v}^-, \mathbf{n}\rangle \mathbf{n} \tag{1.1}$$

where $\mathbf{v}^-$ and $\mathbf{v}^+$ denote the velocities before and after collision, respectively, $\mathbf{n}$ stands for the inward unit normal vector to $\partial\mathcal{D}$, and $\langle\cdot,\cdot\rangle$ designates the scalar product.

If the boundary $\partial\mathcal{D}$ is entirely smooth and concave, and the curvature of $\partial\mathcal{D}$ does not vanish, the billiard is said to be dispersing. Such billiards were studied by Sinai [22], and now they are known as Sinai billiards. A classical example is a unit torus $\mathbb{T}^2$ with finitely many fixed disjoint convex obstacles $\mathbb{B}_i$, $i = 1, \dots, k$, i.e., the table is $\mathcal{D} = \mathbb{T}^2 \setminus \cup_i(\operatorname{int}\mathbb{B}_i)$.

Sinai proved that the resulting billiard dynamics in $\mathcal{D}$ is uniformly hyperbolic, ergodic, and K-mixing. By uniform hyperbolicity we mean that the expansion rates of unstable vectors are uniform, i.e., they expand exponentially fast. Gallavotti and Ornstein [15] proved that Sinai billiards are Bernoulli. Young [24] proved that correlations decay exponentially fast. The central limit theorem and other limit laws were derived in [4, 6].

All these results have been extended to dispersing billiards with piecewise smooth boundaries, i.e., to tables with corners, provided the boundary components intersect each other transversally, i.e., the angles made by the walls at corner points are positive.

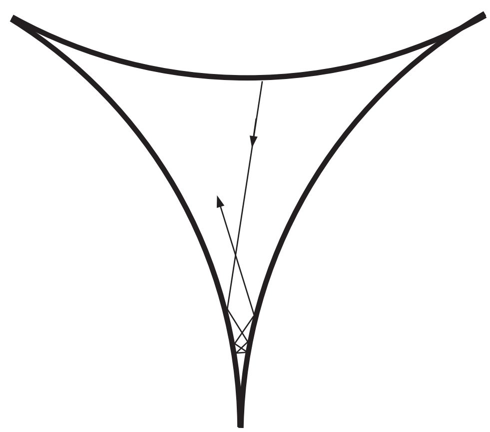

FIGURE 1. Billiard table with three cusps.

A very different picture arises if some boundary components converge tangentially at a corner, i.e., make a cusp. Dispersing billiards with cusps were first studied by Machta [19] who investigated a billiard table made by three identical circular arcs tangent to each other at their points of contact (Fig. 1). He found (based on heuristic arguments) that correlations for the collision map decay slowly (only as $1/n$, where n denotes the collision counter). The hyperbolicity is non-uniform meaning that there are no uniform bounds on the expansion of unstable vectors: when the trajectory falls into a cusp (Fig. 1), it may be trapped there for quite a while, and during long series of collisions in the cusp unstable vectors expand very slowly.

Rigorous bounds on the decay of correlations were derived recently in [11, 13]. It was shown that if A is a Hölder continuous function ("observable") on the collision space M, then for all $n \in \mathbb{Z}$

$$\zeta_n(A) := \left|\mu\big(A \cdot (A \circ F^n)\big) - \big[\mu(A)\big]^2\right| = \mathcal{O}(1/|n|) \tag{1.2}$$

where $F\colon M \to M$ denotes the collision map (billiard map) and μ its invariant measure; we use standard notation $\mu(A) = \int_M A\,d\mu$. We refer the reader to [10] for a comprehensive coverage of the modern theory of dispersing billiards and to [11] for a detailed description of dispersing billiards with cusps.

Billiards with cusps are among the very few physically realistic chaotic models where correlations decay polynomially as in (1.2) leading to a non-classical Central Limit Theorem. However, the proofs of limit theorems (Theorems 1, 2 and 3 presented here) require much tighter control over the underlying dynamics than the proof of (1.2) does. As for the general strategy, our arguments follow the scheme developed in [8]. Nonetheless, concerning the specifics of billiards with cusps, we implement new ideas in the following sense.

Machta's original argument [19] consists of approximating the dynamics in the cusps by differential equations. The proofs in [11] involve direct, though technically complicated, estimates of the deviations of the actual billiard trajectories from the solutions of Machta's equations. We employ a novel approach: we integrate Machta's differential equation and find a conserved quantity, then show that the corresponding dynamical quantity is within $\mathcal{O}(1)$ of that ideal quantity. This gives us the necessary tight control over the dynamics.

Here we summarize some issues that provide the main motivation for proving Theorems 1, 2 and 3. Billiards with cusps can be obtained by a continuous transformation of Sinai billiards. Suppose we enlarge the obstacles $\mathbb{B}_i$ on the torus $\mathbb{T}^2$ until they touch each other. At that moment cusps are formed on the boundary $\partial\mathcal{D}$ and the billiard ceases to be a Sinai billiard. Thus billiards with cusps appear on a natural boundary $\partial\mathfrak{S}$ of the space $\mathfrak{S}$ of all Sinai billiards. Strong statistical properties of Sinai billiards deteriorate near that boundary and one gets slow nonuniform hyperbolicity with 'intermittent chaos'.

Billiards within the class $\mathfrak{S}$ but near its boundary $\partial\mathfrak{S}$ are also interesting, because the obstacles nearly touch one another leaving narrow tunnels (of width $\varepsilon > 0$) in between. A periodic Lorentz gas with narrow tunnels was first examined by Machta and Zwanzig [20] who analyzed (heuristically) the diffusion process as $\varepsilon \to 0$. We plan to investigate billiards with tunnels rigorously, and the current work is a first step in that direction.

Our interest in billiards with cusps also comes from the studies [9] of a Brownian motion of a heavy hard disk in a container subject to a bombardment of fast light particles. When the slowly moving disk collides with a wall of the container, the area available to the light particles turns into a billiard table with (two) cusps, see [9, p. 193], and some light particles may be caught in one of them, and they would be hitting the disk at an unusually high rate. An important task is then to estimate the overall effect produced by those rapid collisions in the cusp. At each collision the light particle transfers some momentum to the heavy disk, and the total momentum transferred to the disk can be represented by a Birkhoff sum of a certain function.

We study the limit behavior of Birkhoff sums

$$S_n A = A + A \circ F + \cdots + A \circ F^{n-1} \tag{1.3}$$

for Hölder continuous functions on M. As usual, we consider centered sums, i.e., $S_nA - n\mu(A) = S_n(A - \mu(A))$, so we will always assume that $\mu(A) = 0$; otherwise we replace A with $A - \mu(A)$.

Because correlations decay as $1/n$, the central limit theorem (CLT) fails. Indeed, due to (1.2)

$$\mu\bigl([S_nA]^2\bigr) = \sum_{k=-n+1}^{n-1} (n - |k|)\zeta_k(A) = \mathcal{O}(n \log n), \tag{1.4}$$

so the proper normalization factor for S_nA must be $\sqrt{n \log n}$, rather than the classical $\sqrt{n}$. Our main goal is to establish a non-classical central limit theorem:

Theorem 1 (CLT). *Let $\mathcal{D}$ be a planar dispersing billiard table with a cusp. Let A be a Hölder continuous*[1] *function on the collision space M. Then we have a (nonclassical) Central Limit Theorem*

$$\frac{S_nA}{\sqrt{n \log n}} \Rightarrow \mathcal{N}(0, \sigma^2) \tag{1.5}$$

for some $\sigma^2 = \sigma_A^2 \ge 0$, which is given by explicit formula (1.7)*.*

The convergence (1.5) means precisely that for every $z \in \mathbb{R}$

$$\mu\left\{\frac{S_nA}{\sqrt{n \log n}} < z\right\} \to \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{z} e^{-\frac{s^2}{2\sigma^2}}\, ds \tag{1.6}$$

as $n \to \infty$. In the degenerate case $\sigma = 0$, the left hand side of (1.5) converges to zero in probability; see also Theorem 3.

Remark. Since the map F is ergodic, the limit law (1.5) is *mixing* in the sense of probability theory, i.e., the limit in (1.6) holds true if we

[1]The function A may be piecewise Hölder continuous, provided its discontinuity lines coincide with discontinuities of $F^{\pm k}$ for some $k \ge 1$.

replace μ with any measure that is absolutely continuous with respect to μ; see [14, Section 4.2].

The limit law (1.5) can be interpreted, in physical terms, as *superdiffusion*, and $\sigma^2 = \sigma_A^2$ as *superdiffusion coefficient*; see similar studies in [1, 8, 23]. In [1], an analogue of Theorem 1 was proved for another billiard model where correlations decay as $\mathcal{O}(1/n)$ – the Bunimovich stadium. In [8, 23] a similar superdiffusion law was proved for the Lorentz gas with infinite horizon, which is also characterized (after a proper reduction [12, 13]) by $\mathcal{O}(1/n)$ correlations. Our proofs follow the lines of [8].

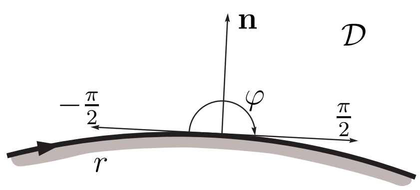

FIGURE 2. Orientation of r and φ on $\partial\mathcal{D}$

There are natural coordinates r and φ in the collision space M, where r denotes the arc length parameter on $\partial\mathcal{D}$ and φ the angle of reflection, i.e., the angle between $\mathbf{v}^+$ and $\mathbf{n}$ in the notation of (1.1). Note that $-\pi/2 \le \varphi \le \pi/2$; the orientation of r and φ is shown in Fig. 2. The billiard map F preserves measure μ on M given by

$$d\mu = c_\mu \cos\varphi \, dr, d\varphi,$$

where $c_\mu = [2\,\mathrm{length}(\partial\mathcal{D})]^{-1}$ is the normalizing factor. In these coordinates, M is a union of rectangles $[r_i', r_i''] \times [-\pi/2, \pi/2]$, where the intervals $[r_i', r_i'']$ correspond to smooth components (arcs) of $\partial\mathcal{D}$.

The Hölder continuity of a function $A\colon M \to \mathbb{R}$ means that

$$|A(r,\varphi) - A(r',\varphi')| \le \mathcal{K}_A\big(|r-r'|^{\alpha_A} + |\varphi-\varphi'|^{\alpha_A}\big)$$

for some $\alpha_A > 0$ (the Hölder exponent) and $\mathcal{K}_A > 0$ (the Hölder norm) provided r and r' belong to one interval $[r_i', r_i'']$. The function A need not change continuously from one interval to another, even if the corresponding arcs have a common endpoint.

The cusp is a common terminal point of two arcs, i_i and i_2, of $\partial\mathcal{D}$; thus the coordinate r takes two values at the cusp, $r' = r_{i_1}'$ and $r'' = r_{i_2}''$. Now the coefficient σ_A^2 is given by

$$\sigma_A^2 = \frac{c_\mu}{8\bar{a}}\left[\int_{-\pi/2}^{\pi/2} [A(r',\varphi) + A(r'',\varphi)]\,\sqrt{\cos\varphi}\,d\varphi\right]^2 \tag{1.7}$$

where $\bar{a} = (a_1 + a_2)/2$ and a_1, a_2 denote the curvatures of the two arcs making the cusp measured at the vertex of the cusp.

Remark. Our results easily extend to dispersing billiards with more than one cusp. To account for the total effect of all the cusps, σ_A^2 must be the sum of expressions (1.7), each corresponding to one cusp.

The non-classical limit theorem (1.5) leads to the following lower bound on the correlations $\zeta_n(A)$ defined by (1.2):

Corollary 1.1. *If $\sigma_A^2 \neq 0$, then the sequence $|n|\zeta_n(A)$ cannot converge to zero as $n \to \pm\infty$.*

Proof. This is an analogue of Corollary 1.3 in [1]. If we had $\zeta_n(A) = o(1/|n|)$, then the first identity in (1.4) would imply $\mu\big([S_n A]^2\big) = o(n \log n)$, hence $\frac{S_n A}{\sqrt{n \log n}}$ would converge to zero in probability. This would contradict (1.5). $\square$

Our next result reinforces the central limit theorem (1.5):

Theorem 2 (WIP)**.** *Let A satisfy the assumptions of Theorem 1 and $\sigma_A^2 \neq 0$. Then the following Weak Invariance Principle holds: the process*

$$W_N(s) = \frac{S_{sN} A}{\sqrt{\sigma_A^2 N \log N}}, \qquad 0 < s < 1, \tag{1.8}$$

converges, as $N \to \infty$, to the standard Brownian motion.

As usual, $S_{sN}A$ here is defined by (1.3) for integral values of sN and by linear interpolation in between.

The same remark as we made after Theorem 1 applies here: the limit distribution of the left hand side of (1.8) is the same with respect to any measure that is absolutely continuous with respect to μ.

Lastly we investigate the degenerate case $\sigma_A^2 = 0$ (which occurs when the integral in (1.7) vanishes).

Theorem 3 (Degenerate case)**.** *Let A satisfy the assumptions of Theorem 1 and $\sigma_A^2 = 0$. Then we have classical Central Limit Theorem*

$$\frac{S_n A}{\sqrt{n}} \Rightarrow \mathcal{N}(0, \hat{\sigma}^2) \tag{1.9}$$

for some $\hat{\sigma}^2 = \hat{\sigma}_A^2 \geq 0$ (see Theorem 6 for a precise formula).

The remark made after Theorem 1 applies here, too, i.e., the limit distribution of the left hand side of (1.9) is the same with respect to any measure that is absolutely continuous with respect to μ. Also, if a billiard table has several cusps, then $\sigma_A^2 = 0$ if and only if the expression

(1.7) vanishes for *every* cusp. Lastly, it is standard that $\hat{\sigma}_A^2 = 0$ if and only if A is a coboundary, i.e., there exists a function $g \in L^2_\mu(M)$ such that $A = g - g \circ F$ almost everywhere (this follows from general results; see, e.g., [18] and [17, Theorem 18.2.2]).

Remark. The function A is a coboundary if and only if for every periodic point $x \in M$, $F^p x = x$, we have $\sum_{i=1}^p A(F^i x) = 0$; see [7]. And in dispersing billiards periodic points are dense [5]. Thus coboundaries make a subspace of infinite codimension in the space of Hölder continuous functions, i.e., the situation $\hat{\sigma}_A^2 = 0$ is extremely rare. On the other hand, $\sigma_A^2 = 0$ occurs whenever the integral in (1.7) vanishes, hence such functions make a subspace of codimension one (for billiards with k cusps it would be a subspace of codimension k), so such functions are not so exceptional.

Remark. Correlations decay slowly in discrete time, when each collision counts as a unit of time. The picture is different in physical, continuous time: the collisions inside a cusp occur in rapid succession, thus their effect is much less pronounced. In fact, the corresponding billiard flow is *rapid mixing* in the sense that correlations for smooth observables decay faster than any polynomial rate, and a classical Central Limit Theorem holds [2].

It is worth noting that any observable $\tilde{A}$ on the phase space of the flow, $\mathcal{D} \times \mathbb{S}^1$, can be reduced to an observable $A\colon M \to \mathbb{R}$ for the billiard map by integrating $\tilde{A}$ between collisions. If $\tilde{A}$ is bounded, then clearly $A \to 0$ near cusps, hence $\sigma_A^2 = 0$, and therefore A, too (just like $\tilde{A}$), satisfies a classical CLT.

On the other hand, some physically important observables for the flow are not bounded near cusps, and for them the classical CLT may fail; for example the number of collisions during the time interval $(0, T)$, as $T \to \infty$, does not satisfy a classical CLT.

2. Induced map

It is standard in the studies of nonuniformly hyperbolic maps to reduce the dynamics onto a subset $\mathcal{M} \subset M$ so that the induced map $\mathcal{F}\colon \mathcal{M} \to \mathcal{M}$ will be strongly hyperbolic and have exponential decay of correlations.

In the present case the hyperbolicity is slow only because of the cusp. So we cut out a small vicinity of the cusp; i.e., we remove from M two rectangles, $R_1 = [r'_{i_1}, r'_{i_1} + \varepsilon_0] \times [-\pi/2, \pi/2]$ and $R_2 = [r''_{i_2} - \varepsilon_0, r''_{i_2}] \times [-\pi/2, \pi/2]$, with some small $\varepsilon_0 > 0$ and consider the induced map $\mathcal{F}$ on the remaining collision space $\mathcal{M} = M \setminus (R_1 \cup R_2)$. It preserves the conditional measure ν on $\mathcal{M}$, where $\nu(B) = \mu(B)/\mu(\mathcal{M})$ for any $B \subset \mathcal{M}$. The map $\mathcal{F}\colon \mathcal{M} \to \mathcal{M}$ is strongly hyperbolic and has exponential decay of correlations [11]. On the other hand, the induced map $\mathcal{F}$ is rather complex and has infinitely many discontinuity lines.

Now let

$$\mathcal{R}(x) = \min\{m \geq 1\colon F^m x \in \mathcal{M}\} \tag{2.1}$$

denote the return time function on $\mathcal{M}$. The domains

$$\mathcal{M}_m = \{x \in \mathcal{M}\colon \mathcal{R}(x) = m\}$$

for $m \geq 1$ are called cells; note that $\mathcal{M} = \cup_{m\geq 1}\mathcal{M}_m$. Cells are separated by the discontinuity lines of $\mathcal{F}$.

Given a function A on M we can construct the "induced" function on $\mathcal{M}$ as follows:

$$\mathcal{A}(x) = \sum_{m=0}^{\mathcal{R}(x)-1} A(F^m x). \tag{2.2}$$

We also denote by $\mathcal{S}_n\mathcal{A}$ its Birkhoff sums:

$$\mathcal{S}_n\mathcal{A} = \mathcal{A} + \mathcal{A}\circ\mathcal{F} + \cdots + \mathcal{A}\circ\mathcal{F}^{n-1}. \tag{2.3}$$

It is standard that $\nu(\mathcal{R}) = 1/\mu(\mathcal{M})$ (Kac's formula) and $\nu(\mathcal{A}) = \mu(A)/\mu(\mathcal{M})$. Since we always assume $\mu(A) = 0$, we also have $\nu(\mathcal{A}) = 0$.

If the original function A is continuous, then the induced function $\mathcal{A}$ will be continuous on each cell $\mathcal{M}_m$, but it may have countably many discontinuity lines that separate cells $\mathcal{M}_m$'s from each other. So the discontinuity lines of $\mathcal{A}$ will coincide with those of the map $\mathcal{F}$.

Theorem 4 (CLT for the induced map). *Let $A\colon M \to M$ satisfy the assumptions of Theorem 1 and $\mathcal{A}$ be the induced function on $\mathcal{M}$ constructed by* (2.2). *Then*

$$\frac{\mathcal{S}_n\mathcal{A}}{\sqrt{n\log n}} \Rightarrow \mathcal{N}(0, \sigma^2_{\mathcal{A}}), \tag{2.4}$$

where $\sigma^2_{\mathcal{A}} = \nu(\mathcal{R})\sigma^2_A$.

Remark. The function $\mathcal{R}$ itself (more precisely, its "centered" version $\mathcal{R}_0 = \mathcal{R} - \nu(\mathcal{R})$) satisfies the above limit theorem, i.e.,

$$\frac{\mathcal{S}_n\mathcal{R} - n\nu(\mathcal{R})}{\sqrt{n\log n}} \Rightarrow \mathcal{N}(0, \sigma^2_{\mathcal{R}}), \tag{2.5}$$

where

$$\sigma^2_{\mathcal{R}} = \frac{c_\mu}{2\bar{a}}\left[\int_{-\pi/2}^{\pi/2} \sqrt{\cos\varphi}\, d\varphi\right]^2. \tag{2.6}$$

Indeed, define a function A by

$$A = \begin{cases} 1 - \nu(\mathcal{R}) & \text{for } \ x \in \mathcal{M} \\ 1 & \text{for } \ x \in M \setminus \mathcal{M} \end{cases}$$

Then by (2.2) we have $\mathcal{A} = \mathcal{R} - \nu(\mathcal{R})$. Note that A is piecewise constant, with a single discontinuity line that separates $\mathcal{M}$ from $M \setminus \mathcal{M}$ (hence its discontinuity line coincides with that of $\mathcal{F}$), and $\mu(A) = 0$. Thus Theorem 4 applies and gives (2.5). The formula (2.6) follows from (1.7), because $A \equiv 1$ in the vicinity of the cusp.

The remark made after Theorem 1 applies here, too. Indeed, the ergodicity of F implies that of $\mathcal{F}$, hence the limit law (2.4) is mixing, i.e., the limit distribution of the left hand side of (2.4) is the same with respect to any measure that is absolutely continuous with respect to ν.

Proof of Theorem 1 from Theorem 4. Our argument is similar to [8, Section 3.1]; see also [10, Theorem 7.68] and [16, Theorem A.1].

First, according to the previous remark, the limit law (2.4) holds with respect to the measure $\tilde{\nu}$ defined by $d\tilde{\nu}/d\nu = \mathcal{R}/\nu(\mathcal{R})$. Our next step is to prove the limit law (1.5) with respect to $\tilde{\nu}$.

Given $n \geq 1$ we fix $\mathbf{n}_1 = [n/\nu(\mathcal{R})]$. For every $x \in \mathcal{M}$ let $\mathbf{n}_2 = \mathbf{n}_2(x)$ be the number of returns to $\mathcal{M}$ of the trajectory of x within the first n iterations, i.e., $\mathbf{n}_2$ satisfies $\mathcal{S}_{\mathbf{n}_2}\mathcal{R}(x) \leq n < \mathcal{S}_{\mathbf{n}_2+1}\mathcal{R}(x)$. Then we have

$$S_n A = \mathcal{S}_{\mathbf{n}_1}\mathcal{A} + (\mathcal{S}_{\mathbf{n}_2}\mathcal{A} - \mathcal{S}_{\mathbf{n}_1}\mathcal{A}) + (S_n A - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}) \tag{2.7}$$

Due to Theorem 4, we have

$$\frac{\mathcal{S}_{\mathbf{n}_1}\mathcal{A}}{\sqrt{n \log n}} \Rightarrow \mathcal{N}(0, \sigma_A^2), \tag{2.8}$$

thus it is enough to show that the other two terms in (2.7) are negligible, i.e.

$$\chi_1 = \frac{\mathcal{S}_{\mathbf{n}_2}\mathcal{A} - \mathcal{S}_{\mathbf{n}_1}\mathcal{A}}{\sqrt{n \log n}}, \qquad \chi_2 = \frac{S_n A - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}}{\sqrt{n \log n}} \tag{2.9}$$

both converge to zero in probability. It is enough to prove the convergence to zero with respect to ν, because $\tilde{\nu}$ is an absolutely continuous measure. To deal with χ_1 we use (2.5), which implies that for any $\varepsilon > 0$ there is a $C = C_\varepsilon > 0$ such that

$$\nu\big(|\mathbf{n}_2 - \mathbf{n}_1| \leq C\sqrt{n \log n}\big) \geq 1 - \varepsilon \tag{2.10}$$

Now the desired result $\chi_1 \to 0$ would follow if both expressions

$$\max_{1 \leq j \leq C\sqrt{n \log n}} \frac{1}{\sqrt{n \log n}} \left| \sum_{i=\mathbf{n}_1}^{\mathbf{n}_1+j} \mathcal{A}(\mathcal{F}^i x) \right| \tag{2.11}$$

and

$$\max_{1\le j\le C\sqrt{n\log n}} \frac{1}{\sqrt{n\log n}} \left| \sum_{i=\mathbf{n}_1-j}^{\mathbf{n}_1} \mathcal{A}(\mathcal{F}^i x) \right| \tag{2.12}$$

converged to zero in probability. Because $\mathcal{F}$ preserves the measure ν, we can replace $\mathbf{n}_1$ with 0 in the above expressions. Then their convergence to zero easily follows from the Birkhoff Ergodic Theorem.

To deal with χ_2 we note that A is bounded, hence $|S_n A - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}| \le \|A\|_\infty (n - \mathcal{S}_{\mathbf{n}_2}\mathcal{R})$. Note that $n - (\mathcal{S}_{\mathbf{n}_2}\mathcal{R})(x) = k$ if and only if $F^n(x) \in F^k(\mathcal{M}_m)$ for some $m > k$, hence

$$\nu\big(x\colon n - (\mathcal{S}_{\mathbf{n}_2}\mathcal{R})(x) = k\big) = \sum_{m=k}^{\infty} \nu(\mathcal{M}_m)$$

Thus the distribution of $n - \mathcal{S}_{\mathbf{n}_2}\mathcal{R}$ does not depend on n, which implies $\chi_2 \to 0$ in probability.

We just proved (1.5) with respect to the measure $\tilde\nu$. The latter is defined on $\mathcal{M}$, but it corresponds to a representation of the space (M,μ) as a tower over $\mathcal{M}$, whose levels are made by the images $F^i(\mathcal{M}_m)$, $0 \le i < m$. Thus the space $(\mathcal{M},\tilde\nu)$ is naturally isomorphic to (M,μ), and the limit law (1.5) on the space $(\mathcal{M},\tilde\nu)$ can be restated as follows: with respect to the measure μ on M we have

$$\frac{S_n A \circ \Pi}{\sqrt{n\log n}} \Rightarrow \mathcal{N}(0,\sigma^2), \tag{2.13}$$

where $\Pi(x) = F^{\zeta(x)}(x)$ and $\zeta(x) = \max\{m \le 0\colon F^m x \in \mathcal{M}\}$; so Π plays the role of the "projection" on the base of the tower. Lastly, the effect of Π is negligible and can be handled in the same way as the difference $S_n A - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}$ above. □

Theorem 5 (WIP for the induced map)**.** *Let A satisfy the assumptions of Theorem 1 and $\sigma_A^2 \ne 0$. Let $\mathcal{A}$ be the induced function on $\mathcal{M}$ constructed by* (2.2)*. Then the following Weak Invariance Principle holds: the process*

$$\mathcal{W}_N(s) = \frac{\mathcal{S}_{sN}\mathcal{A}}{\sqrt{\sigma_A^2 N \log N}}, \qquad 0 < s < 1, \tag{2.14}$$

converges, as $N \to \infty$, to the standard Brownian motion.

We derive Theorem 2 from Theorem 5 in Section 8.

Theorem 6 (Degenerate CLT for the induced map)**.** *Let A satisfy the assumptions of Theorem 1 and $\sigma_A^2 = 0$. Let $\mathcal{A}$ be the induced function* *on $\mathcal{M}$ constructed by* (2.2)*. Then we have a classical Central Limit Theorem*

$$\frac{\mathcal{S}_n\mathcal{A}}{\sqrt{n}} \Rightarrow \mathcal{N}(0, \hat{\sigma}^2_{\mathcal{A}}) \tag{2.15}$$

where $\hat{\sigma}^2_{\mathcal{A}} = \nu(\mathcal{R})\hat{\sigma}^2_A$. The latter satisfies standard Green-Kubo formula

$$\hat{\sigma}^2_{\mathcal{A}} = \sum_{n=-\infty}^{\infty} \nu(\mathcal{A} \cdot (\mathcal{A} \circ \mathcal{F}^n)). \tag{2.16}$$

This series converges exponentially fast.

We derive Theorem 3 from Theorem 6 in Section 7.

Remark Even in the non-degenerate case, $\sigma^2_A \neq 0$, all the terms in the series (2.16), except the one with $n = 0$, are finite and their sum converges (see our Lemma 3.2 below). Hence $\hat{\sigma}^2_{\mathcal{A}}$ given by the series (2.16) is finite if and only if its central term (with $n = 0$) is finite. The latter occurs if and only if $\sigma^2_A = 0$ (see our Lemma 4.3 and Section 7 below), so we have

$$\hat{\sigma}^2_{\mathcal{A}} < \infty \quad \Leftrightarrow \quad \sigma^2_A = 0.$$

In the bulk of the paper we prove Theorem 4. The degenerate case is treated in Section 7. The more specialized limit law (Theorem 5) is proved in the last section 8.

The underlying map $\mathcal{F}\colon \mathcal{M} \to \mathcal{M}$ is strongly hyperbolic and has exponential decay of correlations for bounded Hölder continuous functions [11]. But we have to deal with a function $\mathcal{A}$ that has infinitely many discontinuity lines and is unbounded; in fact its second moment is usually infinite; see below.

There are two strategies for proving limit theorems for such functions. One is based on Young's tower and spectral properties of the corresponding transfer operator on functional spaces [1]. The other is more direct – it truncates the unbounded function $\mathcal{A}$ and then uses probabilistic moment estimates [8]. We follow the latter approach.

In Sections 3–4 we describe the general steps of the proof, which can be applied to many similar models. In the Sections 5–6 we provide model-specific details.

3. Truncation of $\mathcal{A}$

The constructions and arguments in Sections 3–4 are rather general, they are based on a minimal collection of properties of the underlying dynamical system. Thus our arguments can be easily applied to other models. The necessary model-specific facts are stated as lemmas here; they will be all proved in Sections 5–6.

Lemma 3.1. *We have* $\nu(\mathcal{M}_m) \asymp m^{-3}$.

The notation $P \asymp Q$ means that $C_1 < P/Q < C_2$ for some positive constants $C_2 > C_1 > 0$. This lemma is in fact proved in [11].

The power 3 here is the minimal integral power for this estimate. Indeed, the sets $F^i(\mathcal{M}_m)$, $1 \le i \le m$, are disjoint, hence the series $\sum_{m=1}^{\infty} m\nu(\mathcal{M}_m)$ converges; its sum is $\nu(\mathcal{R}) = 1/\mu(\mathcal{M})$. Thus, if $\nu(\mathcal{M}_m) \asymp m^{-a}$ for some $a > 0$, then $a > 2$.

In most interesting systems with weak hyperbolicity (such as stadia, semi-dispersing billiards, etc.; see [12, 13]), we have either $\nu(\mathcal{M}_m) = \mathcal{O}(m^{-3})$ or $\nu(\mathcal{M}_m) = \mathcal{O}(m^{-4})$; in the latter case one expects the classical central limit theorem to hold.

Next we note that $\mathcal{A}|_{\mathcal{M}_m} = \mathcal{O}(m)$, because the original function A is bounded. This implies that $\nu(|\mathcal{A}|) < \infty$, but usually the second moment of $\mathcal{A}$ is infinite, i.e., $\nu(\mathcal{A}^2) = \infty$. To cope with this difficulty we will truncate the function $\mathcal{A}$ (in two different ways).

To fix our notation, for each $1 \le p \le q$ we denote

$$\mathcal{M}_{p,q} = \cup_{p \le m < q} \mathcal{M}_m$$

(occasionally we let $q = \infty$). Note that if $p \ll q$, then

$$\nu(\mathcal{M}_{p,q}) \asymp \sum_{m=p}^{q} m^{-3} \asymp p^{-2}. \tag{3.1}$$

We also put $\mathcal{A}_{p,q} = \mathcal{A} \cdot \mathbf{1}_{\mathcal{M}_{p,q}}$, where $\mathbf{1}_B$ denotes the indicator of the set B, and

$$\hat{\mathcal{A}}_{p,q} = \left(\mathcal{A} - \frac{1}{\nu(\mathcal{M}_{p,q})} \int_{\mathcal{M}_{p,q}} \mathcal{A}\, d\nu \right) \mathbf{1}_{\mathcal{M}_{p,q}}$$

the "centered" version of $\mathcal{A}_{p,q}$. Note that both $\mathcal{A}_{p,q}$ and $\hat{\mathcal{A}}_{p,q}$ vanish outside $\mathcal{M}_{p,q}$. Also, $\nu(\hat{\mathcal{A}}_{p,q}) = 0$.

Now we choose a large constant ω (say, $\omega > 10$) and fix two levels at which we will truncate our function:

$$p = \frac{\sqrt{n}}{(\log n)^{\omega}} \quad \text{and} \quad q = \sqrt{n} \log\log n, \tag{3.2}$$

so that $\mathcal{A} = \mathcal{A}_{1,p} + \mathcal{A}_{p,q} + \mathcal{A}_{q,\infty}$.

Due to (3.1) we have

$$\nu\big(\exists i \le n \colon \mathcal{F}^i(x) \in \mathcal{M}_{q,\infty}\big) = \mathcal{O}\big((\log\log n)^{-2}\big) \to 0, \tag{3.3}$$

so the values of $\mathcal{A}_{q,\infty} \circ \mathcal{F}^i$ can be disregarded because their probabilities are negligibly small. Thus we can replace $\mathcal{A}$ with $\mathcal{A}_{1,q}$. Again, due to

Lemma 3.1

$$\nu(\mathcal{A}_{1,q}) = -\int_{\mathcal{M}_{q,\infty}} \mathcal{A}\, d\nu = \mathcal{O}\left(\frac{1}{\sqrt{n}\log\log n}\right).$$

Next, we show that $\mathcal{A}_{1,q}$ can be further replaced with

$$\hat{\mathcal{A}} \colon = \mathcal{A}_{1,q} - \hat{\mathcal{A}}_{p,q}. \tag{3.4}$$

To this end we need to prove that the overall contribution of the values $\hat{\mathcal{A}}_{p,q}\circ\mathcal{F}^i$ is negligible because they tend to cancel each other. Our proof involves a bound on correlations:

Lemma 3.2. *For each $k \geq 1$ and any $1 \leq p \leq q \leq \infty$ and $1 \leq p' \leq q' \leq \infty$ we have*

$$\left|\nu\big(\hat{\mathcal{A}}_{p,q} \cdot (\hat{\mathcal{A}}_{p',q'} \circ \mathcal{F}^k)\big)\right| \leq C\theta^k \tag{3.5}$$

for some $C > 0$ and $\theta \in (0,1)$ that are determined by the function $\hat{\mathcal{A}}$ but do not depend on p, q, p', q' or k.

The condition $k \geq 1$ here is essential, because for $k = 0$ the resulting integral $\nu\big(\hat{\mathcal{A}}_{p,q}\hat{\mathcal{A}}_{p',q'}\big)$ is not uniformly bounded (and it actually turns infinite for $q = q' = \infty$).

We now return to (3.4). The second moment of $\mathcal{S}_n\hat{\mathcal{A}}_{p,q} = \sum_{i=0}^{n-1} \hat{\mathcal{A}}_{p,q}\circ\mathcal{F}^i$ can be estimated by

$$\nu\big([\mathcal{S}_n\hat{\mathcal{A}}_{p,q}]^2\big) = \mathcal{O}(n\log\log n), \tag{3.6}$$

where the main contribution comes from the "diagonal" terms

$$\nu\big[\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^i\big]^2 = \mathcal{O}\big(\log(q/p)\big) = \mathcal{O}(\log\log n),$$

as all the other terms sum up to $\mathcal{O}(n)$ due to (3.5). Now by Chebyshev's inequality for any $\varepsilon > 0$

$$\nu\big(|\mathcal{S}_n\hat{\mathcal{A}}_{p,q}| \geq \varepsilon\sqrt{n\log n}\big) \leq \frac{\mathrm{const}\cdot n\log\log n}{\varepsilon^2 n\log n} \to 0. \tag{3.7}$$

Hence we can replace $\mathcal{A}_{1,q}$ with $\hat{\mathcal{A}}$ given by (3.4), i.e., Theorem 4 would follow if we prove that

$$\frac{\mathcal{S}_n\hat{\mathcal{A}}}{\sqrt{n\log n}} \Rightarrow \mathcal{N}(0, \sigma^2_{\mathcal{A}}) \tag{3.8}$$

with respect to the measure ν. We prove (3.8) in Section 4.

We record several useful facts. The function $\hat{\mathcal{A}}$ is constant on the set $\mathcal{M}_{p,q}$, and its value on this set is

$$\hat{\mathcal{A}}|_{\mathcal{M}_{p,q}} = \frac{1}{\nu(\mathcal{M}_{p,q})}\int_{\mathcal{M}_{p,q}} \mathcal{A}\, d\nu = \mathcal{O}(p).$$

The first few moments of $\hat{\mathcal{A}}$ can be roughly estimated as

$$\nu(\hat{\mathcal{A}}) = \nu(\mathcal{A}_{1,q}) = -\nu(\mathcal{A}_{q,\infty}) = \mathcal{O}(1/q), \tag{3.9}$$

$$\nu(\hat{\mathcal{A}}^2) = \mathcal{O}\biggl(\sum_{m=1}^{p} \frac{m^2}{m^3}\biggr) + \mathcal{O}\biggl(\frac{p^2}{p^2}\biggr) = \mathcal{O}(\log p), \tag{3.10}$$

$$\nu(|\hat{\mathcal{A}}|^3) = \mathcal{O}\biggl(\sum_{m=1}^{p} \frac{m^3}{m^3}\biggr) + \mathcal{O}\biggl(\frac{p^3}{p^2}\biggr) = \mathcal{O}(p), \tag{3.11}$$

$$\nu(\hat{\mathcal{A}}^4) = \mathcal{O}\biggl(\sum_{m=1}^{p} \frac{m^4}{m^3}\biggr) + \mathcal{O}\biggl(\frac{p^4}{p^2}\biggr) = \mathcal{O}(p^2). \tag{3.12}$$

Also note that $\hat{\mathcal{A}} = \hat{\mathcal{A}}_{1,q} - \hat{\mathcal{A}}_{p,q} + \nu(\hat{\mathcal{A}})$, hence due to (3.9) and the correlation bound (3.5) we have

$$\bigl|\nu\bigl(\hat{\mathcal{A}} \cdot (\hat{\mathcal{A}} \circ \mathcal{F}^k)\bigr)\bigr| \le 4C\theta^k + C'/q^2 \tag{3.13}$$

for some constant $C' > 0$ and all $k \ge 1$.

4. Moment estimates

Here we begin the proof of the CLT for the truncated function, i.e., (3.8). Our truncations have removed all excessively large values of the function $\mathcal{A}$, so now we can apply probabilistic arguments.

We shall use Bernstein's classical method based on the "big small block" technique. That is, we partition the time interval $[0, n-1]$ into a sequence of alternating big intervals (blocks) of length $P = [n^a]$ and small blocks of length $Q = [n^b]$ for some $0 < b < a < 1$. The number of big blocks is $K = [n/(P+Q)] \sim n^{1-a}$. There may be a leftover block in the end, of length $L = n - KP - (K-1)Q < P + Q$.

We denote by Δ_k, $1 \le k \le K$, our big blocks and set

$$\mathcal{S}_P^{(k)} = \sum_{i \in \Delta_k} \hat{\mathcal{A}} \circ \mathcal{F}^i, \qquad \mathcal{S}_n' = \sum_{k=1}^{K} \mathcal{S}_P^{(k)}$$

and

$$\mathcal{S}_n'' = \mathcal{S}_n \hat{\mathcal{A}} - \mathcal{S}_n' = \sum_{i \in [0,n-1] \setminus \cup \Delta_k} \hat{\mathcal{A}} \circ \mathcal{F}^i.$$

The second sum $\mathcal{S}_n''$ contains no more than $n'' = KQ + P \le 2n^h$ terms, where $h = \max\{a, 1-a+b\} < 1$.

Just as in the proof of (3.6), we estimate

$$\nu\bigl([S_n'']^2\bigr) = \mathcal{O}(n'' \log n) = \mathcal{O}(n^h \log n)$$

where the main contribution comes from the "diagonal" terms

$$\nu\big[\hat{\mathcal{A}} \circ \mathcal{F}^i\big]^2 = \mathcal{O}(\log p) = \mathcal{O}(\log n),$$

as all the other terms sum up to $\mathcal{O}(n'')$ due to (3.13). Now by Chebyshev's inequality for any $\varepsilon > 0$

$$\nu\big(|\mathcal{S}_n''| \geq \varepsilon\sqrt{n \log n}\big) \leq \frac{\text{const} \cdot n^h \log n}{\varepsilon^2 n \log n} \to 0.$$

Hence we can neglect the contribution from the small blocks, as well as from the last leftover block.

Thus Theorem 4 is equivalent to

$$\frac{\mathcal{S}_n'}{\sqrt{n \log n}} \Rightarrow \mathcal{N}(0, \sigma_{\mathcal{A}}^2), \tag{4.1}$$

By the Lévy continuity theorem, it suffices to show that the characteristic function

$$\phi_n(t) = \nu\bigg(\exp\bigg(\frac{it\mathcal{S}_n'}{\sqrt{n \log n}}\bigg)\bigg) = \nu\bigg(\prod_{k=1}^{K} \exp\bigg(\frac{it\mathcal{S}_P^{(k)}}{\sqrt{n \log n}}\bigg)\bigg)$$

converges, pointwise, to that of the normal distribution $\mathcal{N}(0, \sigma_{\mathcal{A}}^2)$, i.e., to $\exp(-\frac{1}{2}\,\sigma_{\mathcal{A}}^2 t^2)$.

First we need to decorrelate the contributions from different big blocks, i.e., we will prove that

$$\phi_n(t) = \prod_{k=1}^{K} \nu\bigg(\exp\bigg(\frac{it\mathcal{S}_P^{(k)}}{\sqrt{n \log n}}\bigg)\bigg) + o(1) \tag{4.2}$$

This requires bounds on multiple correlations defined below.

Let f be a Hölder continuous function on $\mathcal{M}$ which may have discontinuity lines coinciding with those of $\mathcal{F}$. Consider the products

$$f^- = (f \circ \mathcal{F}^{-p_1}) \cdot (f \circ \mathcal{F}^{-p_2}) \cdots (f \circ \mathcal{F}^{-p_k})$$

for some $0 \leq p_1 < \cdots < p_k$ and

$$f^+ = (f \circ \mathcal{F}^{q_1}) \cdot (f \circ \mathcal{F}^{q_2}) \cdots (f \circ \mathcal{F}^{q_r})$$

for some $0 \leq q_1 < \cdots < q_r$. Note that f^- depends on the values of f taken in the *past*, and f^+ on the values of g taken in the *future*. The time interval between the future and the past is $p_1 + q_1$.

Lemma 4.1. *Suppose f is a Hölder continuous on each cell $\mathcal{M}_m$ with Hölder exponent α_f and Hölder norm $\mathcal{K}_f$. Then*

$$\big|\nu(f^- f^+) - \nu(f^-)\nu(f^+)\big| \leq B\theta^{|p_1+q_1|} \tag{4.3}$$

where $\theta = \theta(\alpha_f) \in (0,1)$, *and*

$$B = C_0 \mathcal{K}_f \|f\|_\infty^{k+r}, \tag{4.4}$$

where $C_0 = C_0(\mathcal{D}) > 0$ *is a constant.*

To prove (4.2) we apply Lemma 4.1 to the function $f = \exp(it\hat{\mathcal{A}}/\sqrt{n \log n})$. We obviously have $\|f\|_\infty = 1$; for the Hölder exponent we have $\alpha_f = \alpha_{\hat{\mathcal{A}}}$ and for the Hölder norm $\mathcal{K}_f = t\mathcal{K}_{\hat{\mathcal{A}}}/\sqrt{n \log n}$.

Now the proof of (4.2) goes on by splitting off one big block at a time, and the accumulated error in the end will be $\mathcal{O}\big(K\mathcal{K}_{\hat{\mathcal{A}}}\theta^Q/\sqrt{n \log n}\big)$. This is small enough due to the following lemma:

Lemma 4.2. *The induced function* $\mathcal{A}$ *is Hölder continuous on each cell* $\mathcal{M}_m$. *Its restriction to* $\mathcal{M}_m$ *has Hölder exponent* $\alpha_{\mathcal{A}} > 0$ *determined by* α_A *alone (i.e., independent of* m*) and Hölder norm* $\mathcal{K}_{\mathcal{A},m} = \mathcal{O}(m^d)$ *for some* $d > 0$.

Since we truncated our function $\mathcal{A}$ at the level $q = \sqrt{n}\log\log n$, we have $\mathcal{K}_{\hat{\mathcal{A}}} = \mathcal{O}(n^d)$, and since we chose $Q = n^b$ for some $b > 0$, the factor θ^Q will suppress $\mathcal{K}_{\hat{\mathcal{A}}}$ and K. This completes the proof of (4.2).

Due to the invariance of ν we can rewrite (4.2) as

$$\phi_n(t) = \left[\nu\left(\exp\left(\frac{it\mathcal{S}_P}{\sqrt{n \log n}}\right)\right)\right]^K + o(1) \tag{4.5}$$

where $\mathcal{S}_P = \mathcal{S}_P^{(1)}$ corresponds to the very first big block. Next we use Taylor expansion

$$\exp\left(\frac{it\mathcal{S}_P}{\sqrt{n \log n}}\right) = 1 + \frac{it\mathcal{S}_P}{\sqrt{n \log n}} - \frac{t^2\mathcal{S}_P^2}{2n \log n} + \mathcal{O}\left(\frac{|\mathcal{S}_P|^3}{(n \log n)^{3/2}}\right) \tag{4.6}$$

and then integrate it. For the linear term, we use (3.9) and get

$$\nu(\mathcal{S}_P) = \mathcal{O}(P/q). \tag{4.7}$$

For the quadratic term, we have

$$\nu(\mathcal{S}_P^2) = P\nu(\hat{\mathcal{A}}^2) + \mathcal{O}(P) = \mathcal{O}(P \log p). \tag{4.8}$$

Indeed, the main contribution comes from the "diagonal" terms, $\nu(\hat{\mathcal{A}}^2) = \mathcal{O}(\log p)$, as all the other terms sum up to $\mathcal{O}(P)$ due to (3.13). Moreover, $\nu(\hat{\mathcal{A}}^2) = \nu(\mathcal{A}_{1,p}^2) + \mathcal{O}(1)$, hence (4.8) can be rewritten as

$$\nu(\mathcal{S}_P^2) = P\nu(\mathcal{A}_{1,p}^2) + \mathcal{O}(P). \tag{4.9}$$

The value $\nu(\mathcal{A}_{1,p}^2)$ must be computed precisely, to the leading order:

Lemma 4.3. *We have* $\nu(\mathcal{A}_{1,p}^2) = 2\sigma_{\mathcal{A}}^2 \log p + \mathcal{O}(1)$.

Therefore, (4.9) takes form

$$\nu(\mathcal{S}_P^2) = 2P\sigma_{\mathcal{A}}^2 \log p + \mathcal{O}(P). \tag{4.10}$$

For the cubic term, we apply the Cauchy-Schwartz inequality:

$$\nu(|\mathcal{S}_P|^3) \le \left[\nu(\mathcal{S}_P^2)\nu(\mathcal{S}_P^4)\right]^{1/2}. \tag{4.11}$$

For the fourth moment we use expansion

$$\nu(\mathcal{S}_P^4) = \sum \nu\big(\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}\big) \tag{4.12}$$

where we denote $\hat{\mathcal{A}}_j = \hat{\mathcal{A}} \circ \mathcal{F}^j$, for brevity. We will consider ordered sets of indices, i.e., $0 \le j_1 \le j_2 \le j_3 \le j_4 < P$. We fix a large constant $C_1 \gg 1$ and divide the products $\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}$ into several types depending on the gaps between indices

$$D_1 = j_2 - j_1, \quad D_2 = j_3 - j_2, \quad D_3 = j_4 - j_3.$$

Case 1 (most significant): $|D_i| \le C_1 \log p$ for all $i = 1, 2, 3$. Then by the Hölder inequality and (3.12)

$$\left|\nu\big(\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}\big)\right| \le \nu(\hat{\mathcal{A}}^4) = \mathcal{O}(p^2),$$

thus the total contribution of such terms is $\mathcal{O}\big(Pp^2 \log^3 p\big)$.

Case 2 (of moderate significance): $|D_2| > C_1 \log p$ and $|D_i| \le C_1 \log p$ for $i = 1, 3$. We again apply Lemma 4.1 to the function $f = \hat{\mathcal{A}}$, then use Lemma 4.2 and the Hölder inequality:

$$\nu\big(\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}\big) = \left[\nu\big(\hat{\mathcal{A}}^2\big)\right]^2 + \mathcal{O}\big(\|\hat{\mathcal{A}}\|_\infty^4 n^d \theta^{C_1 \log p}\big)$$

It follows from (3.10) that the first term is $\mathcal{O}(\log^2 p)$, and if C_1 is large enough, the second term will be, say, $o(p^{-10})$. Hence the total contribution of all the above terms is $\mathcal{O}(P^2 \log^4 p)$.

Other cases (least significant): If $|D_1| > C_1 \log p$ and $|D_i| \le C_1 \log p$ for $i = 2, 3$, then the same argument gives, due to (3.9) and (3.11),

$$\nu\big(\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}\big) = \mathcal{O}\big(|\nu(\hat{\mathcal{A}})|\nu(|\hat{\mathcal{A}}|^3)\big) = \mathcal{O}(q^{-1}p),$$

(here and below we suppress correlations as they are just $o(p^{-10})$), so the total contribution of all these terms is $\mathcal{O}(P^2 q^{-1} p \log^2 p)$. If $|D_i| > C_1 \log p$ for $i = 1, 2$ and $|D_3| \le C_1 \log p$, then we get

$$\nu\big(\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}\big) = \mathcal{O}\big(|\nu(\hat{\mathcal{A}})|^2\nu(\hat{\mathcal{A}}^2)\big) = \mathcal{O}(q^{-2}\log p),$$

so the total contribution of all these terms is $\mathcal{O}(P^3 q^{-2} \log^2 p)$. Lastly, if $|D_i| > C_1 \log p$ for $i = 1, 2, 3$, then we get

$$\nu\big(\hat{\mathcal{A}}_{j_1}\hat{\mathcal{A}}_{j_2}\hat{\mathcal{A}}_{j_3}\hat{\mathcal{A}}_{j_4}\big) = \mathcal{O}\big(|\nu(\hat{\mathcal{A}})|^4\big) = \mathcal{O}(q^{-4}),$$

so the total contribution of all these terms is $\mathcal{O}(P^4 q^{-4})$.

Summarizing all the above cases gives an overall bound:

$$\nu(\mathcal{S}_P^4) = \mathcal{O}\big(Pp^2 \log^3 p\big). \tag{4.13}$$

Then (4.11) becomes, due to (3.10) and (4.13),

$$\nu(|\mathcal{S}_P|^3) = \mathcal{O}(Pp \log^2 p). \tag{4.14}$$

Now integrating (4.6) gives

$$\begin{aligned} \nu\left(\exp\left(\frac{it\mathcal{S}_P}{\sqrt{n \log n}}\right)\right) &= 1 - \frac{t^2 \sigma_{\mathcal{A}}^2 P}{2n} + \mathcal{O}\left(\frac{P}{n\sqrt{\log n}}\right) \\ &= \exp\left(-\frac{t^2 \sigma_{\mathcal{A}}^2 P}{2n} + \mathcal{O}\left(\frac{P}{n\sqrt{\log n}}\right)\right). \end{aligned} \tag{4.15}$$

Finally, (4.5) can be rewritten as

$$\phi_n(t) = \exp\left(-\frac{t^2 \sigma_{\mathcal{A}}^2 PK}{2n} + \mathcal{O}\left(\frac{PK}{n\sqrt{\log n}}\right)\right) + o(1) \tag{4.16}$$

which converges to $\exp(-\frac{1}{2}\,\sigma_{\mathcal{A}}^2 t^2)$, as desired. □

Remark. Now we can justify the need for the second truncation at level p. If we did not use it, our estimates on the first and second order terms in (4.6) would be still adequate, but the estimate on the third order term would not be satisfactory. Indeed, if we just replace p with q in (4.14), then the first error term in (4.16) would diverge.

5. Basic facts, Hölder norms, and correlations

In this section we begin our proofs of the model-specific facts stated as lemmas in the previous sections.

Dispersing billiards with cusps have been studied in [19], then with a mathematical rigor in [11]; see also [2] and [13]. Here we briefly summarize the basic facts; the reader is advised to check [11] for more details.

The map $\mathcal{F}\colon \mathcal{M} \to \mathcal{M}$ is uniformly hyperbolic, i.e., it expands unstable curves and contracts stable curves at an exponential rate. More precisely, if u is an unstable tangent vector at any point $x \in \mathcal{M}$, then $\|D_x\mathcal{F}^n(u)\| \geq c\Lambda^n \|u\|$ for some constants $c > 0$ and $\Lambda > 1$ and all $n \geq 1$. Similarly, if v is a stable tangent vector, then $\|D_x\mathcal{F}^{-n}(v)\| \geq c\Lambda^n \|v\|$ for all $n \geq 1$. There is no uniform upper bounds on the expansion and contraction rates, because those approach infinity near grazing (tangential) collisions.

The singularities of the original map $F\colon M \to M$ are made by trajectories hitting corner points (other than cusps) or experiencing grazing (tangential) collisions with $\partial\mathcal{D}$. The singularities of F lie on finitely many smooth compact curves. Those curves are stable in the sense that their tangent vectors belong to stable cones. Likewise, the singularities of F^{-1} are unstable curves.

The singularities of the induced map $\mathcal{F}$ are those of F plus the boundaries of the cells $\mathcal{M}_m$, $m \geq 1$. Those boundaries form a countable union of smooth compact stable curves that accumulate near the (unique) phase point whose trajectory runs directly into the cusp.

The structure of cells $\mathcal{M}_m$ and their boundaries are described in [11]. Each cell has length $\asymp m^{-7/3}$ in the unstable direction and length $\asymp m^{-2/3}$ in the stable direction. Its measure is $\mu(\mathcal{M}_m) \asymp m^{-7/3} \times m^{-2/3} = m^{-3}$. Incidentally, this is our Lemma 3.1.

The map $\mathcal{F} = F^m$ expands the cell $\mathcal{M}_m$ in the unstable direction by a factor $\asymp m^{5/3}$ and contracts it in the stable direction by a factor $\asymp m^{5/3}$, too. So the image $\mathcal{F}(\mathcal{M}_m)$ has 'unstable size' $\asymp m^{-2/3}$ and 'stable size' $\asymp m^{-7/3}$. The images accumulate near the (unique) phase point whose trajectory emerges directly from the cusp.

A characteristic feature of hyperbolic dynamics with singularities is the competition between hyperbolicity and the cutting by singularities. The former causes expansion of unstable curves, it makes them longer. The latter breaks unstable curves into pieces and thus produces shorter curves. One of the main results of [11] is a so called one-step expansion estimate [11, Eq. (5.1)] for the induced map $\mathcal{F}$, which guarantees that the expansion is stronger than the cutting by singularities, i.e., "on average" the unstable curves grow fast, at an exponential rate.

The one-step expansion estimate is a main tool in the subsequent analysis of statistical properties for the map $\mathcal{F}$. It basically implies the entire spectrum of standard facts: the growth lemmas, the coupling lemma for standard pairs and standard families, equidistribution estimates, exponential decay of correlations (including multiple correlations) for bounded Hölder continuous functions, limit theorems for the same type of functions, etc. All these facts with detailed proofs are presented in [10, Chapter 7] for general dispersing billiards (without cusps), but those proofs work for our map $\mathcal{F}$ almost verbatim (see [11, p. 749]). In particular, our Lemma 4.1 follows by a standard argument (see [10, Theorem 7.41]), so we will not repeat its proof here.

Proof of Lemma 4.2. Given $x, y \in \mathcal{M}_m$, we obviously have

$$\begin{aligned} |\mathcal{A}(x) - \mathcal{A}(y)| &\le \sum_{i=0}^{m-1} |A(F^i x) - A(F^i y)| \\ &\le \sum_{i=0}^{m-1} \mathcal{K}_A [\mathrm{dist}(F^i x, F^i y)]^{\alpha_A}. \end{aligned} \tag{5.1}$$

The images $F^i(\mathcal{M}_m)$, $i = 1, \dots, m-1$, keep stretching in the unstable direction and shrinking in the stable direction, as i increases (see [11, pp. 750–751]), thus we can assume that x, y lie on one unstable curve.

It was shown in [11, Eq. (4.5)] that unstable vectors u at points $x \in \mathcal{M}_m$ are expanded under $\mathcal{F} = F^m$ by a factor

$$\|D_x F^m(u)\| / \|u\| \asymp m \lambda_1 \lambda_{m-1}, \tag{5.2}$$

where

$$\lambda_1 = \|D_x F(u)\| / \|u\|, \qquad \lambda_{m-1} = \|D_x F^{m-1}(u)\| / \|D_x F^{m-2}(u)\|$$

are the one-step expansion factors at two“special” iterations at which the corresponding points $F(x)$ and $F^{m-1}(x)$ may come arbitrarily close to ∂M, i.e., experience almost grazing collisions. For this reason λ_1 and λ_{m-1} do not admit upper bounds, they may be arbitrarily large (see [11, p. 741]).

For those two iterations with unbounded expansion factors we can use the Hölder continuity (with exponent 1/2) of the original billiard map F, i.e.,

$$\mathrm{dist}(Fx, Fy) \le C_1 [\mathrm{dist}(x, y)]^{1/2}$$

for some $C_1 > 0$ (see, e.g., [10, Exercise 4.50]). Then due to (5.2) for all $i = 2, \dots, m-2$ we have

$$\mathrm{dist}(F^i x, F^i y) \le C_2 m \,\mathrm{dist}(Fx, Fy) \le C_1 C_2 m [\mathrm{dist}(x, y)]^{1/2}$$

for some $C_2 > 0$. Lastly, again by the Hölder continuity of F

$$\begin{aligned} \mathrm{dist}(F^{m-1} x, F^{m-1} y) &\le C_1 [\mathrm{dist}(F^{m-2} x, F^{m-2} y)]^{1/2} \\ &\le C_1^{3/2} C_2^{1/2} m^{1/2} [\mathrm{dist}(x, y)]^{1/4}. \end{aligned}$$

Adding it all up according to (5.1) gives

$$|\mathcal{A}(x) - \mathcal{A}(y)| \le \mathcal{K}_{\mathcal{A},m} [\mathrm{dist}(x, y)]^{\alpha_A/4}$$

with $\mathcal{K}_{\mathcal{A},m} = \mathcal{O}(m^2)$. Lemma 4.2 is proved. □

Proof of Lemma 3.2. Our argument is analogous to the proof of a similar correlation bound for the Lorentz gas with infinite horizon [8, Proposition 9.1].

The domain $\mathcal{F}(\mathcal{M}_m)$ can be foliated by unstable curves of length $\asymp m^{-2/3}$. Thus the conditional measure ν on $\mathcal{F}(\mathcal{M}_m)$ can be represented by a standard family $\mathcal{G}_m$ such that $Z(\mathcal{G}_m) = \mathcal{O}(m^{2/3})$; see [10, Section 7.4] for the definition and properties of standard families and the respective Z-function. We just remind the reader that given a standard family $\mathcal{G} = \{(W, \nu_W)\}$ of unstable curves $\{W\}$ with smooth probability measures $\{\nu_W\}$ on them, and a factor measure $\lambda_{\mathcal{G}}$ that defines a probability measure $\mu_{\mathcal{G}}$ on $\cup W$, the Z-function is defined by

$$Z(\mathcal{G}) \colon = \sup_{\varepsilon>0} \frac{\mu_{\mathcal{G}}(r_{\mathcal{G}} < \varepsilon)}{\varepsilon}$$

where $r_{\mathcal{G}}(x)$ denotes the distance from a point $x \in W \in \mathcal{G}$ to the nearer endpoint of W, i.e., $r_{\mathcal{G}}(x) = \operatorname{dist}(x, \partial W)$. If the curves $W \in \mathcal{G}$ have lengths $\asymp L$, then $Z(\mathcal{G}) \asymp 1/L$ (see [10, p. 171]). The images $\mathcal{G}_n = \mathcal{F}^n(\mathcal{G})$ are also standard families, and their Z-function satisfies

$$Z(\mathcal{G}_n) \leq c_1 \vartheta^n Z(\mathcal{G}) + c_2 \tag{5.3}$$

where $\vartheta \in (0,1)$ and $c_1, c_2 > 0$ are constants.

The further images $\mathcal{F}^n(\mathcal{M})$, $n \geq 1$, have the same property: the conditional measure ν on $\mathcal{F}^n(\mathcal{M}_m)$ can be represented by a standard family (for example, by $\mathcal{F}^{n-1}(\mathcal{G}_m)$) whose Z-function is $\mathcal{O}(m^{2/3})$ (in fact, the Z-function decreases exponentially under $\mathcal{F}$ due to (5.3)).

Now since the size of $\mathcal{M}_k$ in the instable direction is $k^{-7/3}$, we have

$$\nu\big(\mathcal{M}_k \cap \mathcal{F}^n(\mathcal{M}_m)\big) = \nu(\mathcal{M}_m) \cdot \mathcal{O}(m^{2/3} k^{-7/3}) = \mathcal{O}(m^{-7/3} k^{-7/3}) \tag{5.4}$$

for all $n \geq 1$ (this estimate was first derived in [13, p. 320]). Next we turn to the estimation of correlations $\nu\big(\hat{\mathcal{A}}_{p,q} \cdot (\hat{\mathcal{A}}_{p',q'} \circ \mathcal{F}^n)\big)$ that are involved in Lemma 3.2. For brevity, we denote $\mathcal{A}^{(1)} = \hat{\mathcal{A}}_{p,q}$ and $\mathcal{A}^{(2)} = \hat{\mathcal{A}}_{p',q'}$. Recall that $\big|\mathcal{A}^{(i)}|_{\mathcal{M}_m}\big| \leq cm$ for $i = 1, 2$ and some $c > 0$.

We truncate the functions $\mathcal{A}^{(i)}$ at two levels, $\mathbf{p} < \mathbf{q}$, which will be chosen later, i.e., we consider

$$\mathcal{A}^{(i)} = \mathcal{A}^{(i)}_{1,\mathbf{p}} + \mathcal{A}^{(i)}_{\mathbf{p},\mathbf{q}} + \mathcal{A}^{(i)}_{\mathbf{q},\infty}.$$

The functions $\mathcal{A}^{(i)}_{1,\mathbf{q}}$ are bounded (their ∞-norm is $C_i \colon = \|\mathcal{A}^{(i)}_{1,\mathbf{q}}\|_\infty = \mathcal{O}(\mathbf{q})$) and have Hölder norm $\mathcal{K}^{(i)} = \mathcal{O}(\mathbf{q}^d)$ by Lemma 4.2. Thus the standard correlation estimate [10, Theorem 7.37] (which is our Lemma 4.1 with $k = r = 1$, applied to two different functions) gives

$$\begin{aligned} \nu\big(\mathcal{A}^{(1)}_{1,\mathbf{q}} \cdot (\mathcal{A}^{(2)}_{1,\mathbf{q}} \circ \mathcal{F}^n)\big) &= \mathcal{O}\big((\mathcal{K}^{(1)} + \mathcal{K}^{(2)}) C_1 C_2 \theta^n\big) + \nu\big(\mathcal{A}^{(1)}_{1,\mathbf{q}}\big) \nu\big(\mathcal{A}^{(2)}_{1,\mathbf{q}}\big) \\ &= \mathcal{O}(\mathbf{q}^{d+2} \theta^n) + \mathcal{O}(\mathbf{q}^{-2}). \end{aligned} \tag{5.5}$$

Next,

$$\begin{aligned}
\bigl|\nu\bigl(\mathcal{A}^{(1)}_{1,\mathbf{p}}\cdot(\mathcal{A}^{(2)}_{\mathbf{q},\infty}\circ\mathcal{F}^n)\bigr)\bigr| &\le c^2\sum_{m=1}^{\mathbf{p}}\sum_{k=\mathbf{q}}^{\infty} mk\nu(\mathcal{M}_k\cap\mathcal{F}^n(\mathcal{M}_m))\\
&\le c^2\mathbf{p}\sum_{m=1}^{\mathbf{p}}\sum_{k=\mathbf{q}}^{\infty} k\nu(\mathcal{M}_k\cap\mathcal{F}^n(\mathcal{M}_m))\\
&\le c^2\mathbf{p}\sum_{k=\mathbf{q}}^{\infty} k\nu(\mathcal{M}_k)=\mathcal{O}(\mathbf{p}/\mathbf{q}), \qquad (5.6)
\end{aligned}$$

and a similar estimate holds for $\nu\bigl(\mathcal{A}^{(1)}_{\mathbf{q},\infty}\cdot(\mathcal{A}^{(2)}_{1,\mathbf{p}}\circ\mathcal{F}^n)\bigr)$. Lastly, by (5.4)

$$\begin{aligned}
\bigl|\nu\bigl(\mathcal{A}^{(1)}_{\mathbf{p},\infty}\cdot(\mathcal{A}^{(2)}_{\mathbf{p},\infty}\circ\mathcal{F}^n)\bigr)\bigr| &\le c^2\sum_{m=\mathbf{p}}^{\infty}\sum_{k=\mathbf{p}}^{\infty} mk\nu(\mathcal{M}_k\cap\mathcal{F}^n(\mathcal{M}_m))\\
&\le c^2\sum_{m=\mathbf{p}}^{\infty}\sum_{k=\mathbf{p}}^{\infty} m^{-4/3}k^{-4/3}\\
&=\mathcal{O}(\mathbf{p}^{-2/3}). \qquad (5.7)
\end{aligned}$$

Combining our estimates (5.5)–(5.7) gives

$$\nu\bigl(\mathcal{A}^{(1)}\cdot(\mathcal{A}^{(2)}\circ\mathcal{F}^n)\bigr)=\mathcal{O}\bigl(\mathbf{q}^{d+2}\theta^n+\mathbf{q}^{-2}+\mathbf{p}/\mathbf{q}+\mathbf{p}^{-2/3}\bigr)$$

(the shrewd reader shall notice that $\nu\bigl(\mathcal{A}^{(1)}_{\mathbf{p},\mathbf{q}}\cdot(\mathcal{A}^{(2)}_{\mathbf{p},\mathbf{q}}\circ\mathcal{F}^n)\bigr)$ is accounted for twice – once in (5.5) and once in (5.7) – but since (5.7) estimates absolute values, such a duplication cannot hurt).

Now choosing $\mathbf{q}=\theta^{-n/(d+3)}$ and $\mathbf{p}=\mathbf{q}^{1/2}$ gives the desired exponential bound on correlations. □

6. Second moment calculation

Here we prove Lemma 4.3. This is the only place where we need a precise asymptotic formula, rather than just an estimate of the order of magnitude. This entails a detailed analysis of the 'high' cells $\mathcal{M}_m$ (where m is large) which are made by trajectories that go deep into the cusp and after exactly $m-1$ bounces off its walls exit it.

We use the results and notation of [11]. Let a cusp be made by two boundary components $\Theta_1,\Theta_2\subset\partial\mathcal{D}$. Choose the coordinate system as shown in Fig. 3, then the equations of Θ_1 and Θ_2 are, respectively, $y=f_1(x)$ and $y=-f_2(x)$, where f_i are convex C^3 functions, $f_i(x)>0$ for $x>0$, and $f_i(0)=f_i'(0)=0$ for $i=1,2$. We will use Taylor expansion for the functions f_i and their derivatives:

$$f_i(x)=\tfrac12 a_ix^2+\mathcal{O}(x^3),\qquad f_i'(x)=a_ix+\mathcal{O}(x^2),\qquad f_i''(x)=a_i+\mathcal{O}(x),$$

where $a_i = f_i''(0)$. Since the curvature of the boundary of dispersing billiards must not vanish, we have $a_i > 0$ for $i = 1, 2$.

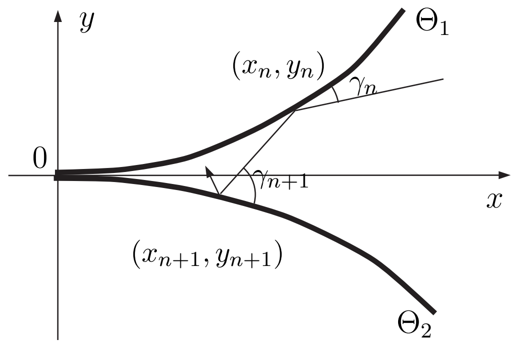

FIGURE 3. A cusp made by two curves, Θ_1 and Θ_2.

Consider a billiard trajectory entering the cusp and making a long series of N reflections there (so it belongs to $\mathcal{M}_{N+1}$). We denote reflection points by (x_n, y_n), where $y_n = f_1(x_n)$ or $y_n = -f_2(x_n)$ depending on which side of the cusp the trajectory hits. We also denote by $\gamma_n = \pi/2 - |\varphi_n|$ the angle made by the outgoing velocity vector with the line tangent to $\partial\mathcal{D}$ at the reflection point (x_n, y_n).

When the trajectory goes down the cusp, x_n decreases but γ_n grows. Then γ_n reaches $\pi/2$ and the trajectory turns back and starts climbing out of the cusp. During that period x_n grows back, but γ_n decreases. Denote by N_2 the deepest collision (closest to the vertex of the cusp), then

$$x_1 > x_2 > \cdots > x_{N_2} \leq x_{N_2+1} < x_{N_2+2} < \cdots < x_N.$$

It was shown in [11] that $N_2 = N/2 + \mathcal{O}(1)$. The following asymptotic formulas were also proven in [11]:

$$x_n \asymp n^{-1/3} N^{-2/3} \qquad \forall n = 1, \ldots, N_2 \tag{6.1}$$

Also, $\gamma_1 = \mathcal{O}(N^{-2/3})$ and

$$\gamma_n \asymp n x_n \asymp n^{2/3} N^{-2/3} \qquad \forall n = 2, \ldots, N_2. \tag{6.2}$$

During the exiting period ($N_2 \leq n \leq N$), we have, due to time reversal symmetry, $x_n \asymp (N-n)^{-1/3} N^{-2/3}$ and $\gamma_n \asymp (N-n)^{2/3} N^{-2/3}$, with the exception of $\gamma_N = \mathcal{O}(N^{-2/3})$. We also note that $\gamma_{N_2} = \pi/2 + \mathcal{O}(1/N)$.

The sequence (x_n, γ_n) satisfies certain recurrence equations (that follow from elementary geometry). If we assume that $y_n = f_1(x_n)$, and hence $y_{n+1} = -f_2(x_{n+1})$, then

$$\gamma_{n+1} = \gamma_n + \tan^{-1} f_1'(x_n) + \tan^{-1} f_2'(x_{n+1}) \tag{6.3}$$

and

$$x_{n+1} = x_n - \frac{f_1(x_n) + f_2(x_{n+1})}{\tan\bigl[\gamma_n + \tan^{-1} f_1'(x_n)\bigr]}. \tag{6.4}$$

If, on the other hand, $y_n = -f_2(x_n)$ (and hence $y_{n+1} = f_1(x_{n+1})$), then the above equations hold, but f_1 and f_2 must be interchanged. This is all proven in [11].

To motivate our further analysis we note that equations (6.3)–(6.4) can be approximated, to the leading order, by

$$\gamma_{n+1} - \gamma_n \approx 2\bar{a}x_n \qquad x_{n+1} - x_n \approx -\frac{\bar{a}x_n^2}{\tan\gamma_n} \tag{6.5}$$

where $\bar{a} = (a_1 + a_2)/2$. Now (6.5) can be regarded as discrete versions of two differential equations

$$\dot{\gamma} = 2\bar{a}x, \qquad \dot{x} = -\frac{\bar{a}x^2}{\tan\gamma}.$$

These equations were first derived (and solved) by Machta [19]. They have an integral $I = x^2 \sin\gamma$ (i.e., $\dot{I} = 0$). This suggests that the quantity $I_n = N^2 x_n^2 \sin\gamma_n$ should remain almost constant (the factor N^2 is included so that to make $I_n \asymp 1$).

Indeed, we have for all $n = 2, \dots, N-2$

$$I_{n+1} - I_n = \mathcal{O}(N^2 x_n^4/\gamma_n) \tag{6.6}$$

which follows by Taylor expansion of the functions involved in (6.3)–(6.4) and using the asymptotic formulas (6.1)–(6.2). (The largest error terms comes from the approximation of $\tan\bigl[\gamma_n + \tan^{-1} f_1'(x_n)\bigr]$ by $\tan\gamma_n$.) As a result, we have

$$I_{n+1} - I_n = \mathcal{O}\bigl(\max\{n^{-2}, (N-n)^{-2}\}\bigr),$$

hence

$$|I_n - I_{N_2}| = \mathcal{O}(n^{-1}). \tag{6.7}$$

Next we use an elliptic integral to introduce a new variable

$$s = \Phi(\gamma) := \int_0^{\gamma} \sqrt{\sin z}\, dz \tag{6.8}$$

and accordingly we put $s_n = \Phi(\gamma_n)$ for $n \le N_2$ (i.e., while γ_n keeps increasing). Then

$$s_{n+1} - s_n = \int_{\gamma_n}^{\gamma_{n+1}} \sqrt{\sin z}\, dz = \sqrt{\sin\gamma_n^*}\,(\gamma_{n+1} - \gamma_n)$$

for some $\gamma_n^* \in (\gamma_n, \gamma_{n+1})$. Again using Taylor expansion and (6.1)–(6.4) we obtain

$$\begin{aligned} s_{n+1} - s_n &= 2\bar{a}N^{-1}\sqrt{I_n + \mathcal{O}(1/n)} \\ &= 2\bar{a}N^{-1}\sqrt{I_{N_2}}\big(1 + \mathcal{O}(1/n)\big). \end{aligned} \tag{6.9}$$

Summing up from 1 to n gives

$$s_n = \frac{2\bar{a}n}{N}\sqrt{I_{N_2}} + \mathcal{O}\Big(\frac{\log n}{N}\Big). \tag{6.10}$$

In particular, for $n = N_2 = N/2 + \mathcal{O}(1)$ we get

$$\int_0^{\pi/2} \sqrt{\sin z}\, dz = \bar{a}\sqrt{I_{N_2}} + \mathcal{O}(N^{-1}\log N),$$

thus

$$I_{N_2} = \left[\frac{1}{\bar{a}}\int_0^{\pi/2}\sqrt{\sin z}\, dz\right]^2 + \mathcal{O}\left(\frac{\log N}{N}\right) \tag{6.11}$$

and (6.10) becomes, with notation $\kappa = \int_0^{\pi}\sqrt{\sin z}\, dz$,

$$s_n = \kappa n/N + \mathcal{O}\big(N^{-1}\log N\big). \tag{6.12}$$

We now estimate the sum $S_{N_2} = \sum_{n=1}^{N_2} A(r_n, \varphi_n)$, where (r_n, φ_n) are the standard coordinates of the reflection points (rather than (x_n, γ_n)). If the n'th collision occurs at the curve Θ_{i_n}, $i_n = 1, 2$, then $r_n = \bar{r}_{i_n} + r_{i_n}(x_n)$ where $\bar{r}_1 = r'_{i_1}$ and $\bar{r}_2 = r''_{i_2}$ are the r-coordinates of the vertex of the cusp, on the curves Θ_1 and Θ_2, respectively (see Section 1), and

$$r_i(x) := (-1)^{i_n+1}\int_0^x \sqrt{1 + [f_i'(x)]^2}\, dx.$$

Also, $\varphi_n = (-1)^{i_n}(\pi/2 - \gamma_n)$, which can be verified by direct inspection. Now

$$S_{N_2} = \sum_{n=1}^{N_2} A\Big(\bar{r}_{i_n} + r_{i_n}(x_n), (-1)^{i_n}(\pi/2 - \Phi^{-1}(s_n))\Big).$$

First, recall that the function A is Hölder continuous with exponent α_A in the variables r and φ. It has the same Hölder continuity with respect to x, but in terms of s we have

$$\Big|A\big(r, \pi/2 - \Phi^{-1}(s')\big) - A\big(r, \pi/2 - \Phi^{-1}(s'')\big)\Big| = \mathcal{O}\big(|s' - s''|^{2\alpha_A/3}\big)$$

because $\Phi^{-1}(s) \sim s^{2/3}$ for small s, so our Hölder exponent reduces to $2\alpha_A/3$. Also note that the collisions at the curves Θ_1 and Θ_2 alternate, and the angle φ is negative when colliding at Θ_1 and positive when colliding at Θ_2 (see Fig. 3). Thus it is convenient to introduce

$$\bar{A}(\varphi) := \tfrac{1}{2}\left[A(r'_{i_1}, -\varphi) + A(r''_{i_2}, \varphi)\right].$$

Approximating an integral by Riemann sums gives

$$\begin{aligned} S_{N_2} &= \sum_{n=1}^{N_2} \bar{A}\big(\pi/2 - \Phi^{-1}(\kappa n/N)\big) \\ &\qquad + \mathcal{O}\Big(N(N^{-1}\log N)^{2\alpha_A/3} + \sum x_n^{\alpha_A}\Big) \\ &= \frac{N}{\kappa}\int_0^{\Phi^{-1}(\kappa/2)} \bar{A}\big(\pi/2 - \Phi^{-1}(s)\big)\, ds + \mathcal{O}\big(N^{1-\alpha_A/2}\big) \\ &= \frac{N}{\kappa}\int_0^{\pi/2} \bar{A}(\pi/2-\gamma)\sqrt{\sin\gamma}\, d\gamma + \mathcal{O}\big(N^{1-\alpha_A/2}\big) \\ &= \frac{N}{\kappa}\int_0^{\pi/2} \bar{A}(\varphi)\sqrt{\cos\varphi}\, d\varphi + \mathcal{O}\big(N^{1-\alpha_A/2}\big) \end{aligned} \tag{6.13}$$

By the time reversibility, the trajectory going out of the cusp during the period $N_2 \le n \le N$ has similar properties, but now the angle φ is positive when colliding at Θ_1 and negative when colliding at Θ_2. Thus

$$\sum_{n=N_2}^{N} A(r_n,\varphi_n) = \frac{N}{\kappa}\int_{-\pi/2}^{0} \bar{A}(\varphi)\sqrt{\cos\varphi}\, d\varphi + \mathcal{O}\big(N^{1-\alpha_A/2}\big), \tag{6.14}$$

Combining (6.13) and (6.14) gives

$$\sum_{n=1}^{N} A(r_n,\varphi_n) = \frac{N}{\kappa}\int_{-\pi/2}^{\pi/2} \bar{A}(\varphi)\sqrt{\cos\varphi}\, d\varphi + \mathcal{O}\big(N^{1-\alpha_A/2}\big) \tag{6.15}$$

where again $\kappa = \int_{-\pi/2}^{\pi/2}\sqrt{\cos\varphi}\, d\varphi$. This can be written as

$$\mathcal{A}|_{\mathcal{M}_{N+1}} = J_A N + \mathcal{O}\big(N^{1-\alpha_A/2}\big) \tag{6.16}$$

where

$$J_A = \frac{1}{2\kappa}\int_{-\pi/2}^{\pi/2}\big[A(r'_{i_1},\varphi) + A(r''_{i_2},\varphi)\big]\sqrt{\cos\varphi}\, d\varphi \tag{6.17}$$

For example, if A is a constant function ($A \equiv A_0$), then the left hand side of (6.16) is $(N+1)A_0$, and on the other hand $J_A = A_0$.

Next we turn to the proof of Lemma 4.3 per se. By (6.16)

$$\begin{aligned} \nu(\mathcal{A}_{1,p}^2) &= J_A^2 \sum_{m=1}^{p} \nu(\mathcal{M}_m) m^2 + \mathcal{O}(1) \\ &= 2J_A^2 \sum_{m=1}^{p} \nu(\mathcal{H}_m) m + \mathcal{O}(1) \end{aligned} \tag{6.18}$$

where $\mathcal{H}_m = \cup_{k=m}^{\infty}\mathcal{M}_k$ is the union of 'high' cells. Note that

$$\mu(\mathcal{H}_m) = \mu(\mathcal{H}'_m), \qquad \mathcal{H}'_m = \cup_{k=m}^{\infty} F^{[k/2]}(\mathcal{M}_k),$$

because the domains $F^{[k/2]}(\mathcal{M}_k)$ do not overlap. They consists of phase points deep in the cusp that are nearly half way in their excursions into the cusp. More precisely, if for $x \in \mathcal{H}'_m$ we denote by i^+ the number of forward collisions in the cusp before exiting it and by i^- the number of backward (past) collisions in the cusp before exiting it, then either $i^+ = i^-$ or $i^+ = i^- + 1$.

Also, the domain $\mathcal{H}'_m$ consists of two parts, one on Θ_1 (where the r-coordinates are near and above r'_{i_1}) and the other on Θ_2 (where the r-coordinates are near and below r''_{i_2}). On both parts the φ-coordinates are near zero.

More precisely, in our previous notation we have

$$I_{N_2} = N^2 x_{N_2}^2 \sin\gamma_{N_2} = N^2 x_{N_2}^2 \bigl(1 + \mathcal{O}(1/N^2)\bigr)$$

(because $\gamma_{N_2} = \pi/2 + \mathcal{O}(1/N)$), hence (6.11) implies $x_{N_2} = \kappa/(2\bar{a}N) + \mathcal{O}(N^{-2}\log N)$. Thus our domains have the range of x-coordinates

$$0 < x < \kappa/(2\bar{a}m) + \mathcal{O}(m^{-2}\log m)$$

The same bounds obviously hold for $|r - r'_{i_1}|$ and $|r - r''_{i_2}|$.

Now for each fixed x, the range of the φ-coordinate corresponds to the change of that coordinate during one iteration of F (indeed, at every iteration φ changes by $\asymp x$, while x only changes by $\mathcal{O}(x^2)$, cf. (6.3)–(6.4); besides, x is near it "stationary point" at iteration $N/2$, because it stops decreasing and starts increasing). So the range of φ can be estimated from (6.8)–(6.9): $\varphi \in [\varphi_1, \varphi_2]$ with

$$\varphi_2 - \varphi_1 = \kappa/N + \mathcal{O}(N^{-2}\log N) = 2\bar{a}x_{N_2} + \mathcal{O}(N^{-2}\log N).$$

Thus (remembering that $\mathcal{H}'_m$ consists of *two* parts)

$$\begin{aligned}\mu(\mathcal{H}'_m) &= 2c_\mu \int_0^{\frac{\kappa}{2\bar{a}m}+\mathcal{O}(\frac{\log m}{m^2})} \bigl[2\bar{a}r + \mathcal{O}(r^2|\log r|)\bigr]\,dr \\ &= \frac{c_\mu \kappa^2}{2\bar{a}m^2} + \mathcal{O}\Bigl(\frac{\log m}{m^3}\Bigr).\end{aligned}$$

We disregard the density $\sin\gamma$ of the measure μ because $\gamma_{N_2} = \pi/2 + \mathcal{O}(1/N)$. Now recall that $\nu(B) = \nu(\mathcal{R})\mu(B)$ for any set B. Therefore (6.18) becomes

$$\nu(\mathcal{A}_{1,p}^2) = \nu(\mathcal{R})c_\mu\kappa^2\bar{a}^{-1}J_A^2\log p + \mathcal{O}(1)$$

which completes the proof of Lemma 4.3. $\square$

7. Degenerate case

Here we prove the classical Central Limit Theorem (Theorem 3) for degenerate functions A, which are characterized by $\sigma_A^2 = 0$. We begin with Theorem 6 that deals with the induced map.

Proof of Theorem 6. As it follows from (6.16), we now have $\mathcal{A}|_{\mathcal{M}_m} = \mathcal{O}(m^{1-\alpha_A/2})$, because $J_A = 0$. Hence $\nu(\mathcal{A}^2) < \infty$. Moreover $\nu(|\mathcal{A}|^{2+\delta}) < \infty$ for some small $\delta > 0$. So the proof of the CLT will be much easier than it was in Sections 3–4 for the generic case $J_A \neq 0$.

Still, we will have to truncate $\mathcal{A}$ at least once, for two reasons: (i) the third and fourth moments of $\mathcal{A}$ may be infinite, and (ii) there is no uniform upper bound on the Hölder norm of $\mathcal{A}|_{\mathcal{M}_m}$, thus the multiple correlations estimate (Lemma 4.1) does not apply to $\mathcal{A}$.

We truncate $\mathcal{A}$ at a single level $q = \sqrt{n}\log\log n$, i.e., we replace $\mathcal{A}$ with $\mathcal{A}_{1,q}$. Due to (3.3), this truncation does not affect any limit laws. The formulas (3.9)–(3.12) are now replaced with

$$\nu(\mathcal{A}_{1,q}) = -\nu(\mathcal{A}_{q,\infty}) = \mathcal{O}\biggl(\sum_{m=q}^{\infty} \frac{m^{1-\alpha/2}}{m^3}\biggr) = \mathcal{O}\bigl(q^{-1-\alpha/2}\bigr), \tag{7.1}$$

$$\begin{aligned}\nu(\mathcal{A}_{1,q}^2) &= \nu(\mathcal{A}^2) - \nu(\mathcal{A}_{q,\infty}^2) \\ &= \nu(\mathcal{A}^2) + \mathcal{O}\biggl(\sum_{m=q}^{\infty} \frac{m^{2-\alpha}}{m^3}\biggr) = \nu(\mathcal{A}^2) + \mathcal{O}(q^{-\alpha}),\end{aligned} \tag{7.2}$$

$$\nu(|\mathcal{A}_{1,q}|^3) = \mathcal{O}\biggl(\sum_{m=1}^{q} \frac{m^{3-3\alpha/2}}{m^3}\biggr) = \mathcal{O}\bigl(q^{1-3\alpha/2}\bigr), \tag{7.3}$$

$$\nu(\mathcal{A}_{1,q}^4) = \mathcal{O}\biggl(\sum_{m=1}^{q} \frac{m^{4-2\alpha}}{m^3}\biggr) = \mathcal{O}\bigl(q^{2-2\alpha}\bigr). \tag{7.4}$$

where we denote $\alpha = \alpha_A$ for brevity. Next, because $\mathcal{A}_{1,q} = \hat{\mathcal{A}}_{1,q} + \nu(\mathcal{A}_{1,q})$ we have by Lemma 3.2 and (7.1)

$$\nu\bigl(\mathcal{A}_{1,q} \cdot (\mathcal{A}_{1,q} \circ \mathcal{F}^k)\bigr) = \mathcal{O}\bigl(\theta^k + q^{-2-\alpha}\bigr) \tag{7.5}$$

The estimate (7.5) remains valid if we replace either one of the $\mathcal{A}_{1,q}$'s (or both) with $\mathcal{A}_{q,\infty}$. We also have $\mathcal{A} = \mathcal{A}_{1,q} + \mathcal{A}_{q,\infty}$, thus

$$\nu\bigl(\mathcal{A} \cdot (\mathcal{A} \circ \mathcal{F}^k)\bigr) = \nu\bigl(\mathcal{A}_{1,q} \cdot (\mathcal{A}_{1,q} \circ \mathcal{F}^k)\bigr) + \chi_{n,k} \tag{7.6}$$

where the remainder term can be bounded as follows:

$$|\chi_{n,k}| = \mathcal{O}\bigl(\min\{\theta^k + q^{-2-\alpha}, q^{-\alpha/2}\}\bigr). \tag{7.7}$$

The first bound follows from the above modification of (7.5), and the second bound, $q^{-\alpha}$, comes just from the Cauchy-Schwartz inequality, because $\nu(\mathcal{A}^2_{q,\infty}) = \mathcal{O}(q^{-\alpha})$; see (7.2). We will use the first bound for $k > \log n$ and the second one for $k \le \log n$.

Now the analysis of Section 4 carries through with a few changes described below. First, naturally, $\sqrt{n \log n}$ is replaced everywhere with $\sqrt{n}$. The Taylor expansion (4.6) now reads

$$\exp\left(\frac{it\mathcal{S}_P}{\sqrt{n}}\right) = 1 + \frac{it\mathcal{S}_P}{\sqrt{n}} - \frac{t^2\mathcal{S}_P^2}{2n} + \mathcal{O}\left(\frac{|\mathcal{S}_P|^3}{n^{3/2}}\right). \tag{7.8}$$

For the linear term we have $\nu(\mathcal{S}_P) = \mathcal{O}\big(Pq^{-1-\alpha/2}\big)$ due to (7.1).

For the main, quadratic term we have (according to (1.4))

$$\nu(\mathcal{S}_P^2) = P\nu(\mathcal{A}^2_{1,q}) + 2\sum_{k=1}^{P-1}(P-k)\nu\big(\mathcal{A}_{1,q}\cdot(\mathcal{A}_{1,q}\circ\mathcal{F}^k)\big).$$

By using (7.2) and (7.5)–(7.7) we can replace $\mathcal{A}_{1,q}$ with $\mathcal{A}$ and get

$$\nu(\mathcal{S}_P^2) = P\nu(\mathcal{A}^2) + 2\sum_{k=1}^{P-1}(P-k)\nu\big(\mathcal{A}\cdot(\mathcal{A}\circ\mathcal{F}^k)\big) + \chi'_{n,P}$$

with $\chi'_{n,P} = \mathcal{O}(Pq^{-\alpha}\log n + P^2q^{-2-\alpha})$. Since $P \ll q^2$ we have $\chi'_{n,P} = \mathcal{O}(Pq^{-\alpha}\log n)$. Lastly, by Lemma 3.2

$$\nu(\mathcal{S}_P^2) = P\hat{\sigma}^2_{\mathcal{A}} + \mathcal{O}(1) + \chi'_{n,P},$$

where $\hat{\sigma}^2_{\mathcal{A}}$ is given by (2.16).

For the cubic term we apply (4.11) and then analyze the fourth order term as in Section 4 (except $\log p$ is now replaced with $\log q$). The most significant case gives a contribution of $\mathcal{O}(Pq^{2-2\alpha}\log^3 q)$, and the moderate significance case gives $\mathcal{O}(P^2\log^2 q)$ which can be neglected if we choose $P = [n^a]$ with $a < 1-\alpha$. Thus we get

$$\nu\big(|\mathcal{S}_P^3|\big) = \mathcal{O}(Pq^{1-\alpha}\log^2 q).$$

Now integrating (7.8) gives

$$\begin{aligned}\nu\left(\exp\left(\frac{it\mathcal{S}_P}{\sqrt{n}}\right)\right) &= 1 - \frac{t^2\hat{\sigma}^2_{\mathcal{A}}P}{2n} + \mathcal{O}\left(\frac{P}{n^{1+\alpha/4}}\right)\\ &= \exp\left(-\frac{t^2\hat{\sigma}^2_{\mathcal{A}}P}{2n} + \mathcal{O}\left(\frac{P}{n^{1+\alpha/4}}\right)\right). \end{aligned}\tag{7.9}$$

Finally, raising to the power K gives the desired result, just like in the end of Section 4. $\square$

Proof of Theorem 3. Our argument is very similar to the derivation of Theorem 1 from Theorem 4 in Section 2, so we only describe the differences.

The decomposition (2.7) remains valid, but (2.8) changes to

$$\frac{\mathcal{S}_{\mathbf{n}_1}\mathcal{A}}{\sqrt{n}} \Rightarrow \mathcal{N}(0, \hat{\sigma}_A^2),$$

which follows from Theorem 6 that we just proved. Now (2.9) become

$$\chi_1 = \frac{\mathcal{S}_{\mathbf{n}_2}\mathcal{A} - \mathcal{S}_{\mathbf{n}_1}\mathcal{A}}{\sqrt{n}}, \qquad \chi_2 = \frac{S_n A - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}}{\sqrt{n}}$$

which must both converge to zero in probability. The argument for χ_2 is the same as it was in Section 2, but our analysis of χ_1 requires more work. Indeed, since (2.10) remains valid, we now have to show that

$$\max_{1\leq j\leq C\sqrt{n\log n}} \frac{1}{\sqrt{n}} \left| \sum_{i=0}^{j} \mathcal{A}(\mathcal{F}^i x) \right| \tag{7.10}$$

converges to zero in probability (note that (7.10) replaces (2.11), while (2.12) should be replaced similarly and we omit it).

Since the number of iterations $C\sqrt{n\log n}$ in (7.10) exceeds the normalization factor $\sqrt{n}$, we cannot use the Birkhoff Ergodic Theorem anymore. Instead, we use specific features of our function $\mathcal{A}$ to prove an even stronger fact:

Proposition 7.1. *The function* $\max_{1\leq k\leq n} |\mathcal{S}_k\mathcal{A}|/n^{5/6}$ *converges to zero in probability, as* $n \to \infty$*.*

Proof. First, we truncate the function $\mathcal{A}$ replacing it with $\mathcal{A}_{1,q}$ as before. Note that $\mathcal{A}_{1,q} = \mathcal{O}(q^{1-\alpha/2}) = \mathcal{O}(n^{1/2-\alpha/4})$ (we drop less important logarithmic factors). Hence

$$\mathcal{S}_k\mathcal{A}_{1,q} - \mathcal{S}_{k'}\mathcal{A}_{1,q} = o(n^{5/6})$$

whenever

$$|k - k'| \leq \Delta\colon\ = [n^{1/3+\alpha/5}]$$

Thus it is enough to show that

$$\max_{1\leq j\leq n/\Delta} |\mathcal{S}_{j\Delta}\mathcal{A}_{1,q}|/n^{5/6} \tag{7.11}$$

converges to zero in probability. By Chebyshev's inequality

$$\nu\bigl(|\mathcal{S}_{j\Delta}\mathcal{A}_{1,q}| \geq \varepsilon n^{5/6}\bigr) \leq \frac{\mathrm{const}\cdot j\Delta}{\varepsilon^2 n^{5/3}}$$

because $\nu\bigl([\mathcal{S}_k\mathcal{A}_{1,q}]^2\bigr) = \mathcal{O}(k)$. Summing up over $j = 1, \dots, n/\Delta$ gives

$$\nu\Bigl(\max_{1\leq j\leq n/\Delta} |\mathcal{S}_{j\Delta}\mathcal{A}_{1,q}| \geq \varepsilon n^{5/6}\Bigr) \leq \frac{\mathrm{const}\cdot n^2/\Delta}{\varepsilon^2 n^{5/3}} \to 0.$$

This proves the proposition, which guarantees that (7.10) converges to zero in probability. □

The rest of the proof of Theorem 3 is the same as the proof of Theorem 1 in Section 2.

8. Weak Invariance Principle

Here we turn to the more specialized limit law, the WIP (Theorems 2 and 5). First we derive Theorem 2 from Theorem 5. Our argument is similar to [8, Section 3.1].

Proof of Theorem 2 from Theorem 5. To prove that a family of stochastic processes $W_N(s)$ weakly converges to a limit process $W(s)$, $0 < s < 1$, one needs to verify two conditions: (i) finite-dimensional distributions of $W_N(s)$ converge to those of $W(s)$, and (ii) the family $\{W_N(s)\}$ is tight.

The first condition means that the random vector $\{W_N(s_1), \dots, W_N(s_k)\}$ converges in distribution to $\{W(s_1), \dots, W(s_k)\}$ for every $k \geq 1$ and $0 \leq s_1 < \cdots < s_k \leq 1$. For $k = 1$ this is just Theorem 1 derived in Section 2, and our argument extends to $k > 1$ easily.

The tightness means that the family of probability measures $\{P_N\}$ on the space $C[0, 1]$ of continuous functions on $[0, 1]$ induced by the processes W_N have the following property: for any $\varepsilon > 0$ there exists a compact subset $K_\varepsilon \subset C[0, 1]$ such that $P_N(K_\varepsilon) > 1 - \varepsilon$ for all N. The compactness of K_ε means that the functions $\{F \in K_\varepsilon\}$ are uniformly bounded at $s = 0$ and equicontinuous on $[0, 1]$.

All our functions vanish at $s = 0$, hence we only need to worry about the equicontinuity. That is, we need to verify that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\nu\Bigg(\sup_{\substack{0<s<s'<1 \\ |s'-s|<\delta}} |W_N(s') - W_N(s)| > \varepsilon \Bigg) < \varepsilon$$

for all $N \geq N_0(\varepsilon)$. Since the function $W_N(s)$ is continuous and piecewise linear, we only need to compare its values at "breaking points", where ns and ns' are integers.

We denote $n = sN$ and use notation $\mathbf{n}_1$ and $\mathbf{n}_2$ introduced in the previous derivation of Theorem 1 from Theorem 4. Similarly we denote $n' = s'N$ and use the respective values $\mathbf{n}_1'$, $\mathbf{n}_2'$. We have the decomposition (2.7) for both n and n'. Now

$$\nu\left(\sup \left| \frac{\mathcal{S}_{\mathbf{n}_1'}\mathcal{A}}{\sqrt{N \log N}} - \frac{\mathcal{S}_{\mathbf{n}_1}\mathcal{A}}{\sqrt{N \log N}} \right| > \varepsilon \right) < \varepsilon$$

due to the tightness of the family $\{\mathcal{W}_N(s)\}$ (which follows from Theorem 5).

Next, due to the Birkhoff Ergodic Theorem, $\mathcal{S}_m\mathcal{R} = m\nu(\mathcal{R}) + o(m)$, hence $\sup|\mathbf{n}_2 - \mathbf{n}_1|/N$ converges to zero in probability, i.e., for any $\varepsilon > 0$ and $\delta > 0$ there exists N_0 such that for all $N > N_0$

$$\nu\biggl(\sup_{0<s<1} |\mathbf{n}_2 - \mathbf{n}_1| > \delta N\biggr) < \varepsilon.$$

Thus, again due to the tightness of the family $\{\mathcal{W}_N(s)\}$, we get

$$\nu\biggl(\sup\biggl|\frac{\mathcal{S}_{\mathbf{n}_2}}{\sqrt{N\log N}} - \frac{\mathcal{S}_{\mathbf{n}_1}}{\sqrt{N\log N}}\biggr| > \varepsilon\biggr) < \varepsilon.$$

A similar estimate obviously holds for $\mathbf{n}_2'$ and $\mathbf{n}_1'$.

Lastly, due to Lemma 3.1, we have

$$\nu\bigl(|S_nA - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}| > \varepsilon\sqrt{N\log N}\bigr) = \mathcal{O}\biggl(\sum_{\sqrt{N\log N}}^{\infty} \frac{1}{m^3}\biggr) = \mathcal{O}\biggl(\frac{1}{N\log N}\biggr),$$

hence

$$\nu\bigl(\sup_{0<n<N} |S_nA - \mathcal{S}_{\mathbf{n}_2}\mathcal{A}| > \varepsilon\sqrt{N\log N}\bigr) = \mathcal{O}(1/\log N) \to 0,$$

and a similar estimate holds for n' and $\mathbf{n}_2'$. This completes the proof of Theorem 2 from Theorem 5. $\square$

Note that at the last point of the proof we had to use Lemma 3.1, i.e., a specific power law for the measures of the cells $\mathcal{M}_m$. The derivation of Theorem 1 from Theorem 4 did not require such specifics.

Proof of Theorem 5. First we note that N now plays the role of n in the proof of the CLT (Theorem 4). In particular, the truncation levels p and q must be defined by

$$p = \frac{\sqrt{N}}{(\log N)^{\omega}} \quad \text{and} \quad q = \sqrt{N}\log\log N. \tag{8.1}$$

Recall that the function $\mathcal{A}$ was replaced by its truncated version $\hat{\mathcal{A}}$ defined by (3.4) based on the estimates (3.3) and (3.7). Now, since the WIP requires us to control the entire path $\{\mathcal{S}_n\mathcal{A}\}$, $0 \le n \le N$, not just its final state $\mathcal{S}_N\mathcal{A}$, the estimate (3.7) must be upgraded to

$$\nu\biggl(\max_{1\le n\le N} |\mathcal{S}_n\hat{\mathcal{A}}_{p,q}| \ge \varepsilon\sqrt{N\log N}\biggr) \to 0, \tag{8.2}$$

The proof of (8.2) resembles the reflection principle in the theory of random walks. We consider "bad" sets

$$\mathcal{B}_{n,\varepsilon} = \bigl\{|\mathcal{S}_n\hat{\mathcal{A}}_{p,q}| \ge \varepsilon\sqrt{N\log N}\bigr\}$$

and "new additions" $\mathcal{N}_{n,\varepsilon} = \mathcal{B}_{n,\varepsilon} \setminus \cup_{i=1}^{n-1} \mathcal{B}_{i,\varepsilon}$. Due to (3.7), $\nu(\mathcal{B}_{N,\varepsilon}) \to 0$, thus it is enough to show that

$$\nu\big(\cup_{n=1}^{N} \mathcal{B}_{n,\varepsilon} \setminus \mathcal{B}_{N,\varepsilon/2}\big) = \sum_{n=1}^{N} \nu(\mathcal{N}_{n,\varepsilon} \setminus \mathcal{B}_{N,\varepsilon/2}) \to 0. \tag{8.3}$$

Each point $x \in \mathcal{N}_{n,\varepsilon} \setminus \mathcal{B}_{N,\varepsilon/2}$ satisfies (simultaneously) three conditions:

$$|\mathcal{S}_n \hat{\mathcal{A}}_{p,q}| \geq \varepsilon \sqrt{N \log N} \tag{8.4}$$

$$\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^{n-1} \neq 0, \tag{8.5}$$

$$|\mathcal{S}_N \hat{\mathcal{A}}_{p,q} - \mathcal{S}_n \hat{\mathcal{A}}_{p,q}| \geq \tfrac{1}{2}\, \varepsilon \sqrt{N \log N}. \tag{8.6}$$

Since $\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^{n-1} \neq 0$, we have $\mathcal{F}^{n-1}(x) \in \mathcal{M}_m$ for some $p \leq m < q$, i.e., the point $\mathcal{F}^{n-1}(x)$ lies in a 'high' cell with index $m \sim \sqrt{n}$ (modulo less significant logarithmic factors; see (8.1)).

The image $\mathcal{F}(\mathcal{M}_m)$ intersects cells $\mathcal{M}_{m'}$ with $c_1\sqrt{m} < m' < c_2 m^2$ for some $c_1, c_2 > 0$, see [11, 13], but typical points $y \in \mathcal{M}_m$ are mapped into cells $\mathcal{M}_{m'}$ with $m' \sim m^{1/2}$. Their further images are typically mapped into cells $\mathcal{M}_{m''}$ with $m'' \sim m^{1/4}$, etc. Since $m^{1/2} \sim n^{1/4} \ll p$, we typically have $\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^{n+i} = 0$ for several small i's.

Precisely, there exist $a, b > 0$ such that for any large $C > 0$ there is a subset $\mathcal{M}_m^* \subset \mathcal{M}_m$ of measure

$$\nu(\mathcal{M}_m \setminus \mathcal{M}_m^*) < K m^{-a} \nu(\mathcal{M}_m),$$

where $K = K(C) > 0$ and such that for every $y \in \mathcal{M}_m^*$ the images $\mathcal{F}^i(y)$ for $i = 1, \dots, C \log m$ never appear in cells $\mathcal{M}_k$ with $k > m^{1-b}$. This was proved in [13, p. 320]. We will apply this fact to cells $\mathcal{M}_m$ with $p \leq m < q$, hence we can replace $C \log m$ with $C \log N$.

The points falling into $\mathcal{M}_m \setminus \mathcal{M}_m^*$ for $p \leq m < q$ make a set of a negligibly small measure:

$$\nu\big(\cup_{n=1}^{N} \mathcal{F}^{-n}\big[\cup_{m=p}^{q} (\mathcal{M}_m \setminus \mathcal{M}_m^*)\big]\big) = \mathcal{O}(N p^{-2-a}) \to 0.$$

Hence we can assume that whenever $\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^{n-1} \neq 0$, we have $\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^{n+i} = 0$ for all $i = 0, 1, \dots, C \log N$.

For $i > C \log N$, the correlations between $\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^j$, $j < n$, and $\hat{\mathcal{A}}_{p,q} \circ \mathcal{F}^{n+i}$ are small due to Lemmas 4.1 and 4.2. One can easily check that they are $< N^{-50}$ if $C > 0$ is large enough. In the following, we use shorthand notation

$$F_n = |\mathcal{S}_n \hat{\mathcal{A}}_{p,q}| \mathbf{1}_{\mathcal{N}_{n,\varepsilon}}, \qquad G_n = |\mathcal{S}_N \hat{\mathcal{A}}_{p,q} - \mathcal{S}_n \hat{\mathcal{A}}_{p,q}|$$

where $\mathbf{1}_B$ stands for the indicator of the set B, and we also denote $D_N = \varepsilon\sqrt{N\log N}$. Now we have

$$\begin{aligned}\nu(\mathcal{N}_{n,\varepsilon}\setminus\mathcal{B}_{N,\varepsilon/2}) &\le \nu\bigl(F_n \ge D_N \ \&\ G_n \ge \tfrac{1}{2}\,D_N\bigr)\\ &\le \nu\bigl(F_nG_n \ge \tfrac{1}{2}\,D_N^2\bigr)\\ &\le 4\nu(F_n^2G_n^2)/D_N^4\\ &\le 4\bigl[\nu(F_n^2)\nu(G_n^2)+\mathcal{O}(N^{-50})\bigr]/D_N^4.\end{aligned}$$

The term $\mathcal{O}(N^{-50})$ accounts for correlations and is negligibly small. The second moment estimate (3.6) can be easily adapted to $\nu(G_n^2) = \mathcal{O}(n\log\log N) \le cN\log\log N$ for some $c>0$. Also note that

$$F_n = |\mathcal{S}_n\hat{\mathcal{A}}_{p,q}|\mathbf{1}_{\mathcal{N}_{n,\varepsilon}} \le (D_N + \|\hat{\mathcal{A}}_{p,q}\|_\infty)\mathbf{1}_{\mathcal{N}_{n,\varepsilon}} \le 2D_N\mathbf{1}_{\mathcal{N}_{n,\varepsilon}},$$

because $\|\hat{\mathcal{A}}_{p,q}\|_\infty \le \|A\|_\infty q \ll D_N$ for large N's. Thus

$$\sum_{n=1}^{N}\nu(\mathcal{N}_{n,\varepsilon}\setminus\mathcal{B}_{N,\varepsilon/2}) \le 16cN\log\log N/D_N^2 + o(1) \to 0,$$

which completes our proof of (8.3) and that of (8.2).

Thus we again can replace the unbounded function $\mathcal{A}$ with its truncated version $\hat{\mathcal{A}}$. That is, Theorem 5 would follow if we prove that

$$\hat{\mathcal{W}}_N(s) = \frac{\mathcal{S}_{sN}\hat{\mathcal{A}}}{\sqrt{\sigma_{\mathcal{A}}^2 N\log N}}, \qquad 0<s<1, \tag{8.7}$$

converges, as $N\to\infty$, to the standard Brownian motion.

Our proof of (8.7) is analogous to that of a similar property of the Lorentz gas with infinite horizon [8, Section 11]. The proof consists of two parts: (i) the weak convergence of finite-dimensional distributions of $\hat{\mathcal{W}}_N(s)$ to those of the Brownian Motion, and (ii) the tightness, see below. To derive (i), by the Lévy continuity theorem it is enough to show that for any $0<s_1<\cdots<s_k\le 1$, any sequences

$$\frac{n_1}{N}\to s_1, \frac{n_2}{N}\to s_2, \dots, \frac{n_k}{N}\to s_k,$$

and any fixed $t_1, t_2, \dots, t_k$ we have

$$\nu\left(\exp\left(\frac{i\sum_{j=1}^k t_j\mathcal{S}_{n_j}\hat{\mathcal{A}}}{\sqrt{N\log N}}\right)\right) \to \prod_{j=1}^k \exp\left(-\frac{\sigma_{\mathcal{A}}^2(s_j-s_{j-1})^2T_j^2}{2}\right) \tag{8.8}$$

where $s_0=0$ and $T_j = \sum_{r=j}^k t_r$. This convergence can be proved by the same big small block technique as in Section 4: small blocks allow us to decorrelate the contributions from big blocks, and in particular the contributions from the intervals $s_j - s_{j-1}$, which implies (8.8).

It remains to show that the family of functions

$$\{\hat{\mathcal{W}}_N(s)\} = \{\mathcal{S}_{sN}\hat{\mathcal{A}}/\sqrt{N\log N}\}, \qquad 0 < s < 1$$

is tight. By the standard argument (see, e.g., [3, Chapter 2]) it is enough to show that there exists a sequence $\{\delta_k\}$ with $\sum_k \delta_k < \infty$ such that $\nu(\tilde{\mathcal{M}}_{K,n}) \to 0$ as $K \to \infty$ uniformly in N, where

$$\tilde{\mathcal{M}}_{K,N} = \bigl\{\exists j,k\colon j < 2^k \text{ and } \bigl|\hat{\mathcal{W}}_N\left(\tfrac{j+1}{2^k}\right) - \hat{\mathcal{W}}_N\left(\tfrac{j}{2^k}\right)\bigr| > K\delta_k\bigr\}. \tag{8.9}$$

Let $\delta_k = 1/k^2$. First we estimate the ν-measure of

$$\tilde{\mathcal{M}}_{K,N,k,j} = \bigl\{|\mathcal{S}_{n_1}\hat{\mathcal{A}} - \mathcal{S}_{n_2}\hat{\mathcal{A}}| \ge \tfrac{1}{k^2}\, K\sqrt{N\log N}\bigr\} \tag{8.10}$$

where $n_1 = [jN/2^k]$ and $n_2 = [(j+1)N/2^k]$. Recall that $\hat{\mathcal{A}} = \mathcal{O}(p)$, hence

$$\mathcal{S}_{n_2}\hat{\mathcal{A}} - \mathcal{S}_{n_1}\hat{\mathcal{A}} = \mathcal{O}\bigl(p(n_2 - n_1)\bigr) = \mathcal{O}\bigl(pN/2^k\bigr).$$

Thus the set (8.10) is empty if $2^k/k^2 > N$, in particular if $k > 100\log N$. For $k < 100\log N$, we use the fourth moment estimate (4.13) and the Markov inequality to get

$$\begin{aligned}
\nu(\tilde{\mathcal{M}}_{K,N,k,j}) &\le \frac{k^8\nu\bigl([\mathcal{S}_{n_2}\hat{\mathcal{A}} - \mathcal{S}_{n_1}\hat{\mathcal{A}}]^4\bigr)}{K^4N^2\log^2 N} \\
&= \mathcal{O}\Bigl(\frac{k^8(n_2 - n_1)p^2\log^3 N}{K^4N^2\log^2 N}\Bigr) \\
&= \mathcal{O}\Bigl(\frac{k^8}{K^4 2^k \log^{99} N}\Bigr).
\end{aligned}$$

Summing over $j = 0, \dots, 2^k - 1$ and then over $k \le 100\log N$ gives $\nu(\tilde{\mathcal{M}}_{K,N}) = \mathcal{O}(1/K^4) \to 0$ as $K \to \infty$, uniformly in N, which implies the tightness. This completes the proof of Theorem 5. □

Acknowledgment. P. Bálint and N. Chernov acknowledge the hospitality of the University of Maryland which they visited and where most of this work was done. P. Bálint was partially supported by the Bolyai scholarship of the Hungarian Academy of Sciences and Hungarian National Fund for Scientific Research (OTKA) grants F60206 and K71693. N. Chernov was partially supported by NSF grant DMS-0969187. D. Dolgopyat was partially supported by NSF grant DMS-0555743.

References

[1] Bálint P. and Gouëzel S., *Limit theorems in the stadium billiard*, Comm. Math. Phys. **263** (2006), 461–512.

[2] Bálint P. and Melbourne I., *Decay of correlations and invariance principles for dispersing billiards with cusps, and related planar billiard flows*, J. Stat. Phys. **133** (2008), 435–447.

[3] Billingsley P. *Convergence of probability measures,* John Wiley & Sons, Inc., New York-London-Sydney 1968 xii+253 pp.

[4] Bunimovich L. A. and Sinai Ya. G. *Statistical properties of Lorentz gas with periodic configuration of scatterers,* Comm. Math. Phys. **78** (1980/81), 479–497.

[5] Bunimovich L. A., Sinai Ya. G., and Chernov N. I. *Markov partitions for two-dimensional billiards*, Russ. Math. Surv. **45** (1990), 105–152.

[6] Bunimovich L. A., Sinai Ya. G., and Chernov N. I. *Statistical properties of two-dimensional hyperbolic billiards,* Russ. Math. Surv. **46** (1991), 47–106.

[7] Bunimovich L. A. & Spohn H., *Viscosity for a periodic two disk fluid: an existence proof*, Comm. Math. Phys. **176** (1996), 661–680.

[8] Chernov N. and Dolgopyat D. *Anomalous current in periodic Lorentz gases with infinite horizon,* Russ. Math. Surv., **64** (2009), 73–124.

[9] Chernov N. and Dolgopyat D. *Brownian Brownian Motion-1,* Memoirs AMS **198**, no 927, 2009, (193 pp).

[10] Chernov N. and Markarian R., *Chaotic Billiards*, Mathematical Surveys and Monographs, **127**, AMS, Providence, RI, 2006. (316 pp.)

[11] Chernov N. and Markarian R., *Dispersing billiards with cusps: slow decay of correlations*, Comm. Math. Phys., **270** (2007), 727–758.

[12] N. Chernov and H.-K. Zhang, *Billiards with polynomial mixing rates*, Nonlinearity, **18** (2005), 1527–1553.

[13] Chernov N. and Zhang H.-K., *Improved estimates for correlations in billiards.* Commun. Math. Phys., **277** (2008), 305–321.

[14] Eagleson G. K., *Some simple conditions for limit theorems to be mixing*, Teor. Verojatnost. i Primenen. **21** (1976), 653–660.

[15] Gallavotti G. and Ornstein D., *Billiards and Bernoulli schemes*, Comm. Math. Phys. **38** (1974), 83–101.

[16] Gouëzel S., *Statistical properties of a skew product with a curve of neutral points*, Ergod. Th. Dynam. Syst. **27**, (2007) 123–151.

[17] Ibragimov I. A. and Linnik Yu. V., *Independent and stationary sequences of random variables*, Wolters-Noordhoff, Gröningen, 1971.

[18] Leonov V. P., *On the dispersion of time-dependent means of a stationary stochastic process*, Th. Probab. Appl., **6** (1961), 87–93.

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

[20] Machta J., Zwanzig R., *Diffusion in a periodic Lorentz gas*, Phys. Rev. Lett. **50** (1983), 1959–1962.

[21] Melbourne I. and Török A. *Statistical limit theorems for suspension flows*, Israel J. Math. **144** (2004), 191–209.

[22] Sinai Ya. G., *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards*, Russ. Math. Surv. **25** (1970), 137–189.

[23] Szasz D. and Varju T. *Limit Laws and Recurrence for the Planar Lorentz Process with Infinite Horizon,* J. Statist. Phys. **129** (2007), 59–80.

[24] Young L.-S. *Statistical properties of dynamical systems with some hyperbolicity*, Ann. Math. **147** (1998) 585–650.

P. Bálint: Institute of Mathematics, Budapest University of Technology and Economics, Budapest, Hungary

N. Chernov: Department of Mathematics, University of Alabama at Birmingham, Birmingham, AL 35294

D. Dolgopyat: Department of Mathematics, University of Maryland, College Park, MD 20742