LIMIT THEOREMS FOR PERTURBED PLANAR LORENTZ PROCESSES

DMITRY DOLGOPYAT, DOMOKOS SZÁSZ, TAMÁS VARJÚ

ABSTRACT. Modify the scatterer configuration of a planar, finite-horizon Lorentz process in a bounded domain. Sinai asked in 1981 whether for the diffusively scaled variant of the modified process convergence to Brownian motion still holds. The main result of the work answers Sinai's question in the affirmative. Other types of local perturbations are also investigated: finite horizon periodic Lorentz process in the half-strip or in the half-plane (in these models the local perturbation is the boundary condition) and finally finite horizon, periodic Lorentz process with a small, compactly supported external field in the strip. The corresponding limiting processes are Brownian motions with suitable boundary conditions and finally the skew Brownian motion on the line. The proofs combine Stroock-Varadhan's martingale method ([SV 71]) with those of our recent work ([DSzV 07]).

Subject classification: 37D50 billiards, 60F05 weak theorems

1. INTRODUCTION.

In this paper we consider systems which look like the periodic Lorentz process on a large part of the plane. Recall that the planar, periodic Lorentz process is the dynamics of a point particle moving in the plane with periodically situated, disjoint, convex scatterers removed. The motion of the particle is uniform with specular (i. e. optical) collisions at the scatterers. Throughout the paper we assume that the horizon is finite that is any ray intersects at least one scatterer (and then, in fact, infinitely many of them). We shall use the abbreviation FHLP for the finite horizon Lorentz process. The statistical properties of the periodic FHLP are well understood since it is nothing else than a $\mathbb{Z}^2$-extension of a finite horizon Sinai billiard given on the two-torus. Therefore the study of statistical properties of the FHLP is intertwined with those of Sinai billiards. Ergodicity

D.D was partially supported by NSF and IPST. D. Sz. and T. V. are grateful to Hungarian National Foundation for Scientific Research grants No. T 046187, K 71693, NK 63066 and TS 049835, and to Hungarian Science and Technology Foundation for grant No. A-9/03.

of Sinai billiards was established in [S 70], the Central Limit Theorem and stretched exponential decay of correlations are proven in [BS 81, BChS 91]. Exponential mixing was proven in [Y 98]. Other results include the local limit Theorem ([SzV 04]) and the almost sure invariance principle ([MN 07]). Finally, the fine recurrence properties of the FHLP studied in a companion paper [DSzV 07] also play an important role in our analysis.

Of course, an infinite periodic configuration is an idealization and it is interesting to understand how the theory of periodic FHLP can be used to study systems which look like a periodic Lorentz process apart from a compact part of the configuration space. The first result in this direction was obtained in [L 06] where the recurrence is proven for finite modifications of FHLP. In this paper we study limit theorems for local modifications of the FHLP.

Let us formulate our results.

For definiteness, denote by a) $Q = \cup_{i=1}^{\infty} O_i$ the configuration space of the Lorentz process, where the closed sets O_i are pairwise disjoint, strictly convex with $\mathcal{C}^3-$smooth boundaries; b) by $\Omega = Q \times S_+$ its phase space (where S_+ is the hemisphere of outgoing unit velocities); c) by $T : \Omega \to \Omega$ its discrete time mapping (the Poincaré section map) and finally d) by μ the f-invariant (infinite) Liouville-measure on Ω. If the scatterer configuration $\{O_i\}_i$ is $\mathbb{Z}^d$-periodic, then the corresponding dynamical system will be denoted by $(\Omega_{per} = Q_{per} \times S_+, T_{per}, \mu_{per})$ and it makes sense to factorize it by $\mathbb{Z}^d$ to obtain a Sinai billiard $(\Omega_0 = Q_0 \times S_+, T_0, \mu_0)$. The natural projection $\Omega \to Q$ (and analogously for Ω_{per} and for Ω_0) will be denoted by π_q.

In our first theorem $Q = Q_{per}$ outside a bounded domain. Select an initial point $x_0 = (q_0, v_0) \in \Omega$ according to a compactly supported probability measure $\mu_{(\mathbf{0})}$, absolutely continuous with respect to the Liouville measure μ. Then $\{T^n x_0 = (q_n, v_n) | n \in \mathbb{Z}\}$ is the Lorentz trajectory and the resulting configuration process $\{q_n | n \geq 0\}$ will be called *a finite modification of the FHLP*. (For simplicity we can assume that the unit is chosen so that $\mu_{(\mathbf{0})}$ is supported inside the unit torus and, moreover, $Q = Q_{per}$ outside the unit torus.)

Definition 1. *Assume $\{q_n \in \mathbb{R}^d | n \geq 0\}$ is a random trajectory. Then its diffusively scaled variant $\in C[0,1]$ (or $\in C[0,\infty]$) is defined as follows: for $N \in \mathbb{Z}_+$ denote $W_N(\frac{j}{N}) = \frac{q_j}{\sqrt{N}}$ $(0 \leq j \leq N$ or $j \in \mathbb{Z}_+)$ and define otherwise $W_N(t)(t \in [0,1]$ or $\mathbb{R}_+)$ as its piecewise linear, continuous extension.*

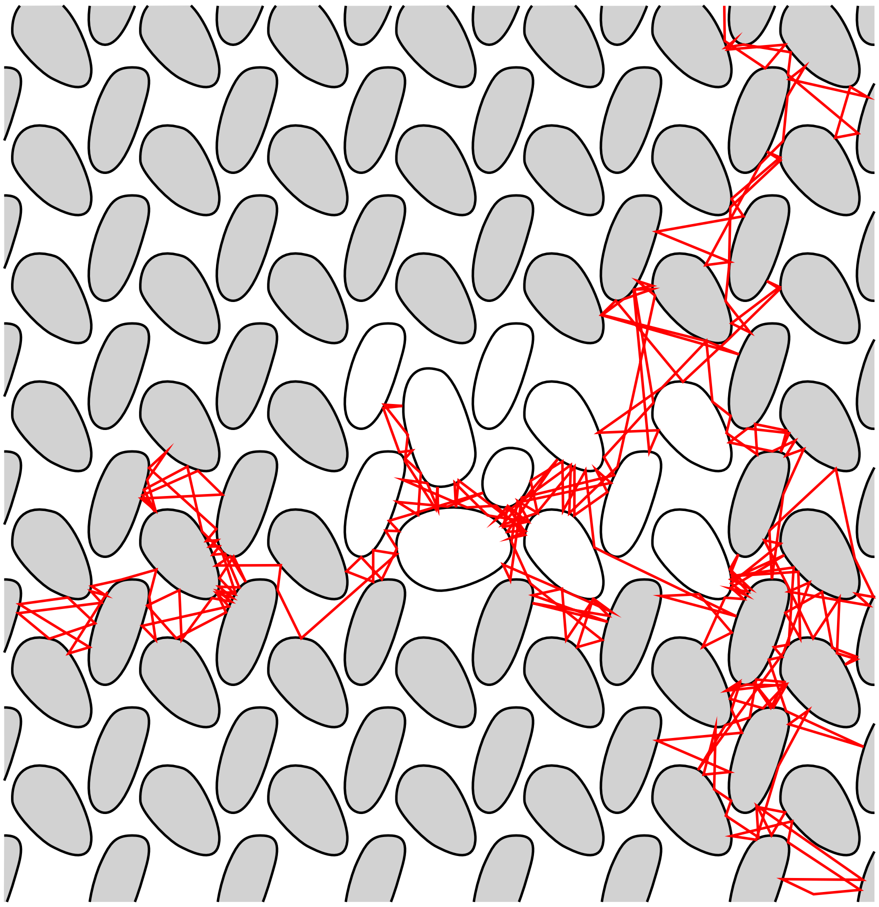

FIGURE 1. A 305 collisions trajectory segment in the finite modification of the FHLP (it is easily seen that, outside a bounded domain, the dark scatterers form a periodic configuration of scatterers)

By general theory, having shown the validity of our results in $C[0,1]$ (always weak convergence!) their truth in $C[0,\infty]$ is straightforward.

Theorem 1. *For finite modifications of the FHLP, as* $N \to \infty$, $W_N(t) \Rightarrow W_{\Sigma^2}(t)$ *(weak convergence in* $C[0,\infty]$*), where* $W_{\Sigma^2}(t)$ *is the Brownian Motion with the non-degenerate covariance matrix* Σ^2*. The limiting covariance matrix coincides with that for the unmodified periodic Lorentz process.*

The result of the previous theorem was conjectured by Sinai in 1981 (oral communication). (It had been tested for random walks with local impurities in [SzT 81].) There exist, however, other compelling mechanisms for local perturbations of periodic FHLP's and the aim of our forthcoming theorems is to treat some of them. Of course it is impossible to describe the most general perturbation. Our goal is to illustrate various difficulties appearing in treating various local perturbations and to introduce techniques to overcome these difficulties. Other applications of the techniques developed here can be found in [ChD 08]. For didactic reasons we formulate theorems in order of increasing difficulty of the proof. In particular in the proofs of Theorems 1 and 2 we use some estimates which are non-optimal and which are improved in the latter sections. However we do it in order to, first, make the proofs of Theorems 1 and 2 more accessible and, second, show what kind of estimates suffice for the proof of each result.

For the next two results we consider a FHLP in a horizontal strip $\mathbb{R} \times [0,1]$ (or in the half-strip $\mathbb{R}_+ \times [0,1]$). That is we study a periodic configuration of the disjoint convex scatteres in the strip such that any billiard trajectory intersects one of the obstacles. The notations we have introduced above have their natural analogues so for simplicity we do not repeat them here. By introducing coordinates (z_1, z_2) in the strip where $z_1 \in \mathbb{R}$ and $z_2 \in [0,1]$, $q_n = (z_{1n}, z_{2n})$ will be the position of the particle after n reflections.

In the next theorem we consider the half-strip $\mathbb{R}_+ \times [0,1]$. The specular reflection at the vertical boundary piece $z_1 = 0$ will play the role of the local perturbation (the result is valid independently of whether we permit some scatterers to intersect this piece or not only if we exclude the tangency of boundary pieces; in any case, apart from the 0-th cell the scatterer configuraqtion is periodic).

Theorem 2. *Consider a FHLP $\{z_{1,n}\}_{n\geq 0}$ in a halfstrip and let $W_N(t) \in \mathbb{R}_+$ be its diffusively scaled variant. Then, as $N \to \infty$, $W_N(t)$ converges weakly to a non-degenerate Brownian motion reflected at* 0.

Next we consider a particle in a whole strip in the presence of a compactly supported thermostatted field. Namely we assume that between the collisions the motion of the particle is given by

$$\dot{v} = E(q) - \frac{(E(q), v)}{(v, v)} v. \tag{1}$$

Theorem 3. *Consider a FHLP $\{z_{1,n}\}_{n\geq 0}$ in the strip in the presence of a small and compactly supported external field E and let $W_N(t) \in \mathbb{R}$ be its diffusively scaled variant. Then, as $N \to \infty$, $W_N(t)$ converges weakly to a skew Brownian motion.*

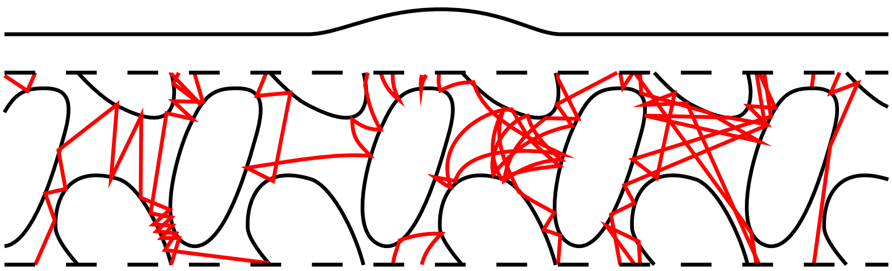

FIGURE 2. A 94 collisions trajectory segment of a FHLP in the presence of a compactly supported external field (whose strength is the function illustrated on top of the figure; observe that, outside a bounded interval, the orbits are linear)

Recall that the skew Brownian Motion is a process $\xi(t)$ such that $|\xi(t)|$ has the same distribution as the absolute value of usual Brownian Motion and its excursions are positive with probability p independently of each other. Thus for $p = 1$ ($p = 0$) we get reflected Brownian motion of $\mathbb{R}_+$ (respectively $\mathbb{R}_-$) and for $p = 1/2$ we have the standard Brownian Motion. The formal definition of the skew Brownian Motion is given in subsection 2.5 and its properties are described in [HSh 81].

The object of our last result is a FHLP in the half-plane $\{z_1 \geq 0\}$ with specular reflections at the vertical line $z_1 = 0$. In this case we delete all scatterers intersecting the vertical axis $z_1 = 0$, so for the resulting configuration space actually there are rays that do not intersect any scatterer (they are situated close to the vertical axis). Nevertheless, their existence — at least in the horizontal direction — only means a local perturbation and, as we will show, the limit is again a (reflected) Brownian motion.

Theorem 4. *Consider the diffusively scaled variant $W_N(t) \in \mathbb{R}_+ \times \mathbb{R}$ of a FHLP $\{q_n\}_{n\geq 0}$ in a halfplane $z_1 \geq 0$. Then, as $N \to \infty$, $W_N(t)$ converges weakly to a non-degenerate Brownian motion reflected at the z_2-axis.*

Theorem 5. *Theorems 1–4 remain valid for continuous time.*

2. PRELIMINARIES

In this section notions and theorems are collected, which later will be used or referred to. We also note that we will throughout use notions and results from our companion paper [DSzV 07]. For the correspondance of the notations of that work and of the present one, let us define the free flight vector $\kappa : \Omega \to \mathbb{R}^d$ as follows: for $x \in \Omega$ let $\kappa(x) = \pi_q T(x)) - \pi_q(x)$. Then for $x \in \Omega$ we define

$$S_n(x) = \sum_{k=0}^{n-1} \kappa(T^k(x)). \tag{2}$$

There is also a natural projection $\pi_{\Omega_{per}} : (\Omega_{per}, T_{per}, \mu_{per}) \to (\Omega_0, T_0, \mu_0)$. Consequently for $x \in \Omega_0$ we can also denote $\kappa(x) = \kappa(\pi^{-1}_{\Omega_{per}}(x))$ and

$$S_n(x) = \sum_{k=0}^{n-1} \kappa(T_0^k(x)). \tag{3}$$

For later reference we denote

$$K_2 = \max_{x \in \Omega} |\kappa(x)|. \tag{4}$$

2.1. **Hyperbolicity of the billiard map.** For definiteness, let $Q_0 = \cup_{i=1}^p O_i$ where the closed sets O_i are pairwise disjoint, strictly convex with $\mathcal{C}^3-$smooth boundaries. In Ω_0 it is convenient to use the product coordinates. Recall that

$$\Omega_0 = \{x = (q, v) | q \in Q_0, \langle v, n \rangle \geq 0\}$$

where $\langle \cdot, \cdot \rangle$ denotes scalar product, and n is the outer normal in the collision point. Traditionally for q one uses the arclength parameter and for the velocity the angle $\phi = \arccos \langle v, n \rangle \in [-\pi/2, \pi/2]$. In these coordinates the invariant measure is given by the density $\frac{1}{2l} \cos \phi \, dq \, d\phi$, where l is the overall perimeter of the scatterers. From our assumptions it follows that $0 < \min |\kappa| < \max |\kappa| < \infty$.

For our billiards there is a natural DT_0-invariant field $\mathcal{C}_x^u$ of unstable cones (and dually also a field $\mathcal{C}_x^s$ of stable ones) of the form $c_1 \leq \frac{d\phi}{dq} \leq c_2$ (or $-c_2 \leq \frac{d\phi}{dq} \leq -c_1$ respectively) where $0 < c_1 < c_2$ are suitable constants.

A connected smooth curve $\gamma \subset \Omega_0$ is called an *unstable curve* (or a *stable curve*) if at every point $x \in \gamma$ the tangent space $\mathcal{T}_x\gamma$ belongs to the unstable cone $\mathcal{C}_x^u$ (or the stable cone $\mathcal{C}_x^s$ respectively).

For an unstable curve γ (or a stable one) and for any $x \in \gamma$ denote by $\mathcal{J}_\gamma T_0^n(x) = ||D_x T_0^n(dx)||/||dx||$, $dx \in \mathcal{T}_x\gamma$ the *Jacobian* of the map f_0^n at the point x. Then the *hyperbolicity of the dynamics* means that there are constants $\Lambda > 1$ and $C > 0$ depending on the dynamics, only, such that for any unstable (or stable) curve γ and every $x \in \gamma$ and every $n \geq 1$ one has $\mathcal{J}_\gamma T_0^n(x) \geq C\Lambda^n$ (or $\mathcal{J}_\gamma T_0^{-n}(x) \geq C\Lambda^n$ respectively).

2.2. **Standard pairs.** Let us start with a heuristic introduction. Sinai's classical billiard philosophy ([S 70] reacts to the fact that dispersing billiards are hyperbolic (a nice property) but at the same time they are singular dynamical systems (an unpleasant property). Nevertheless smooth pieces of unstable (and of stable) invariant manifolds do exist for *expansion prevails partitioning*.

Though dispersing billiards are manifestly hyperbolic, they are not only singular but, added to that, close to the singularities the derivative of the map also explodes. This circumstance is the most unpleasant when one aims at proving the distortion estimates, basic for the techniques. To cope with this difficulty [BChS 91] introduced the idea of surrounding the singularities with a countable number of extremely narrow, so-called *homogeneity strips*, roughly parallel to the singularities. In these strips the derivative of the map can be large, but oscillates very little; this fact makes it possible to nevertheless establish the necessary distortion estimates. The boundaries of these homogeneity strips provide further singularities (causing further partitioning), the so-called *secondary* ones in contrast to the *primary singularities* (in our case only tangencies). The definition of homogeneity strips depends on a parameter denoted usually k_0. The larger k_0 is, the smaller the neighborhood of (primary) singularities is where one introduces the homogeneity strips. In certain bounds (e. g. in the growth lemmas) k_0 should be selected sufficiently large.

Let us now give precise definitions. For $k \geq k_0$ let

$$\mathbb{H}_k = \{(r,\phi) : \frac{\pi}{2} - k^{-2} < \phi < \frac{\pi}{2} - (k+1)^{-2}\},$$

$$\mathbb{H}_{-k} = \{(r,\phi) : \frac{\pi}{2} - k^{-2} < -\phi < \frac{\pi}{2} - (k+1)^{-2}\},$$

$$\mathbb{H}_0 = \{(r,\phi) : -(\frac{\pi}{2} - k_0^{-2}) < \phi < \frac{\pi}{2} - k_0^{-2}\}.$$

Take $L_1, L_2 \gg 1$ and $\theta < 1$ sufficiently close to 1.

An unstable curve is *weakly homogeneous* if it does not intersect any singularity (i. e. neither primary nor secondary one).

A weakly homogeneous unstable curve γ is *homogeneous* if it satisfies the distortion bound

$$\frac{\log J_\gamma T_0(x)}{\log J_\gamma T_0(y)} \leq L_1 \frac{d(x,y)}{\text{length}^{2/3}(\gamma)} \qquad x, y \in \gamma$$

and the curvature bound

$$\angle(\dot\gamma(x), \dot\gamma(y)) \leq L_1 \frac{d(x,y)}{\text{length}^{2/3}(\gamma)} \qquad x, y \in \gamma$$

We observe that if the $\mathcal{C}^2$ norm of γ is bounded and γ is *long* in the sense that either length$(\gamma) > \delta_0$ for some fixed constant δ_0 or γ crosses a whole homogeneity strip, then γ satisfies both the distortion and the curvature bounds.

Let $s^+(x, y)$ be the first time $T_0^s(x)$ and $T_0^s(y)$ are separated by a singularity.

A probability density ρ on a homogeneous unstable curve γ is called *a homogeneous density* if it satisfies the density bound

$$|\log \rho(x) - \log \rho(y)| \leq L_2 \theta^{s^+(x,y)}.$$

We will call the connected homogeneous components of an unstable (stable) curve the *H-components* of the curve. Given γ we let $\gamma_n(x)$ be the largest subcurve of $T_0^n \gamma$ containing $T_0^n x$ and such that $T_0^{-n} \gamma_n(x)$ does not contain singularities of T_0^n.

A *standard pair* is a pair $\ell = (\gamma, \rho)$ where γ is a homogeneous curve and ρ is a homogeneous density on γ.

Given a standard pair and a measurable $A : \Omega_0 \to \mathbb{R}$ we write

$$\mathbb{E}_\ell(A) = \int_\gamma A(x) dx$$

and length$(\ell) =$ length(γ).

In this work the precise definition of the standard pairs is not important but we shall take advantage of their invariance and equidistribution properties listed below and in subsection 2.3.

The fundamental tool used in our work is the so-called *growth lemma*. While hyperbolicity of Sinai billiards means that infinitesimal trajectories diverge exponentially fast, the growth lemma says that the exponential divergence also holds for most trajectories which are sufficiently close to each other.

We give two formulations of the growth lemma. The first and more classical one (statements (a) and (b) below) deals with curves while the second formulation (statements (c) and (d) below) deals with standard pairs. Let Ω denote the phase space of one of the systems appearing in Theorems 1-4.

Let γ be a homogeneous curve and for $n \geq 1$ and $x \in \gamma$ let $r_n(x)$ denote the distance of the point $T_0^n(x)$ from the nearest boundary point of the H-component $\gamma_n(x)$ containing $T_0^n(x)$.

Proposition 1. *(Growth lemma). If k_0 is sufficiently large, then*

(a) *there are constants $\beta_1 \in (0, 1)$ and $\beta_2 > 0$ such that for any $\varepsilon > 0$ and any $n \geq 1$*

$$\text{mes}_\ell(x : \ r_n(x) < \varepsilon) \leq (\beta_1 \Lambda)^n \text{mes}(x : \ r_0 < \varepsilon / \Lambda^n) + \beta_2 \varepsilon$$

(b) *there are constants $\beta_3, \beta_4 > 0$, such that if $n \geq \beta_3 |\log \text{length}(\gamma)|$, then for any $\varepsilon > 0$ and any $n \geq 1$ one has*

$$\text{mes}_\ell(x: \ r_n(x) < \varepsilon) \leq \beta_4 \varepsilon$$

(c) *If $\ell = (\gamma, \rho)$ is a standard pair, then*

$$\mathbb{E}_\ell(A \circ T_0^n) = \sum_\alpha c_{\alpha n} \mathbb{E}_{\ell_{\alpha n}}(A)$$

where $c_{\alpha n} > 0$, $\sum_\alpha c_{\alpha n} = 1$ and $\ell_{\alpha n} = (\gamma_{\alpha n}, \rho_{\alpha n})$ are standard pairs where $\gamma_{\alpha n} = \gamma_n(x_\alpha)$ for some $x_\alpha \in \gamma$ and $\rho_{\alpha n}$ is the pushforward of ρ up to a multiplicative factor.

(d) *If $n \geq \beta_3 |\log \text{length}(\ell)|$, then*

$$\sum_{\text{length}(\ell_{\alpha n}) < \varepsilon} c_{\alpha n} \leq \beta_4 \varepsilon.$$

(e) *For any $\beta_3 |\log(\text{length}(\ell)| \leq n_1 \leq n_2$ we have*

$$\mathbb{P}_\ell(\max_{j \in [n_1, n_2]} r_j(x) < \delta_0) \leq \text{Const}\beta_5^{n_2 - n_1}$$

for some $\beta_5 \in (0, 1)$.

Parts (a), (b). The restatement in terms of the standard pairs is taken from [ChD 07]. For part (e) see (e.g [ChD 07], Lemma 3.10).

In order to apply standard pairs to the problem at hand, observe that the Liouville measure can be decomposed as follows

$$\mu_0(A) = \int \mathbb{E}_{\ell_\alpha}(A) d\sigma(\alpha) \tag{5}$$

where σ is a factor measure such that

$$\sigma(\text{length}(\ell_\alpha) < \varepsilon) < \text{Const}\varepsilon. \tag{6}$$

We shall call measures satisfying (5) and (6) *admissible measueres*.

2.3. **Properties of standard pairs.** In the sequel we are still considering billiards (Ω_0, T_0, μ_0) and functions $A : \Omega_0 \to \mathbb{R}^d$, most frequently with $d = 2$. Let us introduce the space of functions (over (Ω_0, T_0, μ_0)) we are to consider. Take $\theta < 1$ close to 1. Let $s(x, y)$ be the smallest n such that either $T_0^n x$ and $T_0^n y$ or $T_0^{-n} x$ and $T_0^{-n} y$ are separated by a singularity. Define the dynamical Hölder space of functions $A : \Omega_0 \to \mathbb{R}$

$$\mathcal{H} = \{A : |A(x) - A(y)| < \text{Const}\theta^{s(x,y)}\}.$$

Let $A_n(x) = \sum_{j=0}^{n-1} A(T_0^j x)$.

Proposition 2. *Let ℓ be a standard pair, $A \in \mathcal{H}$ and, for statements (a), (b) and (d), take n such that $|\log \text{length}(\ell)| < n^{1/2-\delta}$. Then the following statements hold true:*

(a) There is a constant such that

$$\left|\mathbb{E}_\ell(A \circ T_0^n) - \int A d\mu_0\right| \leq \text{Const}\theta^n |\log \text{length}(\ell)|$$

(b) Let $A, B \in \mathcal{H}$ have zero mean. Then

$$\mathbb{E}_\ell(A_n B_n) = n\sigma_{A,B} + O(|\log^2 \text{length}(\ell)|)$$

where

$$\sigma_{A,B} = \sum_{j=-\infty}^{\infty} \int A(x)B(T_0^j x)d\mu_0(x).$$

(c) Let x be distributed according to ℓ and $w_n(t)$ be defined by

$$w_n\left(\frac{i}{n}\right) = \frac{S_i}{\sqrt{n}}$$

with linear interpolation in between. (S_i is the notation for partial sums of the mean free path from the Introduction). Then, as $n \to \infty$, w_n converges weakly (in $C([0,1] \to \mathbb{R}^2)$ to the 2 dimensional Brownian Motion with zero mean and covariance matrix D^2 given by

$$D^2 = \mu_0(\kappa_0 \otimes \kappa_0) + 2\sum_{j=1}^{\infty} \mu_0(\kappa_0 \otimes \kappa_n).$$

(d) There exists positive constants c_1, c_2 such that for every n and R satisfying $1 < R < n^{1/6-\delta}$ we have

$$\mathbb{P}_\ell(|A_n - n\int A d\mu_0| \geq R\sqrt{n}) \leq c_1 e^{-c_2 R^2}.$$

(e) There exists positive constants $\bar{c}_1, \bar{c}_2$ such that for every n and R satisfying $1 < R < n^{1/6-\delta}$ we have

$$\mathbb{P}_\ell(\max_{j\leq n} |A_j - j\int A d\mu_0| \geq R\sqrt{n}) \leq \bar{c}_1 e^{-\bar{c}_2 R^2}.$$

Parts (a) and (c) are proven in [Ch 99]. For part (b) see Lemma 5.12 of [ChD 07]. (The error estimate of part (b) is not stated explicitly in [ChD 07] but it can be easily deduced from the proof of Lemma 5.12.) Part (d) is proven in [ChD 07], Section A.4 for a particular A but the proof in the general case is exactly the same. Part (e) is proven in Appendix A of this paper.

2.4. **Tail of return times.** The study of the return time to a given scatterer for FHLP plays an important role in our analysis. In this subsection we present the estimates needed in our arguments. The proofs which are slight extensions of the results of [DSzV 07] are given in the Appendix B. Introduce the following notation: for a standard pair $\ell = (\gamma, \rho)$ let $[\ell]$ denote an index $m \in \mathbb{Z}^d$ such that $\pi_q \gamma$ intersects the m-th cell of the configuration space. (If this definition is not unique, then choose any index with this property.) Fix a small $\delta_0 > 0$.

Lemma 3. *(a) Consider planar FHLP. Fix a scatterer S and let Γ be a finite set of scatterers. Then there exist constants $C = C(\text{Card}(\Gamma)) > 0, k_0 = k_0(\text{Card}(\Gamma)), \xi$ such that for any standard pair ℓ such that $\pi_q \gamma \cap S \neq \emptyset$, $\text{length}(\ell) \geq \delta_0$ we have*

$$\mathbb{P}_\ell \left(q_j \notin (S \bigcup \Gamma) \textit{ for } j = k_0 \dots n \right) \geq \frac{C}{\log^\xi n}.$$

(b) For a FHLP in a strip or halfcylinder the following is true: for any standard pair ℓ such that $\text{length}(\ell) \geq \delta_0$ we have

$$\mathbb{P}_\ell \left(\text{Card}(j \leq n : q_j \in S) \leq k_0 \textit{ and } q_j \textit{ does not visit the vertical boundary} \right) \geq \frac{C}{\sqrt{n}}.$$

(c) For a FHLP in a strip or halfcylinder the following is true: for any standard pair ℓ such that $\text{length}(\ell) \geq \delta_0$ we have

$$\begin{aligned} (7) \quad & \mathbb{P}_\ell(\text{Card}(j \leq n : q_j \in S) \leq k_0, \quad max_{j \leq n} |q_j| > K\sqrt{n} \\ & \qquad \textit{and } q_j \textit{ does not visit the vertical boundary}) \leq \frac{C}{\sqrt{n} K^{100}}. \end{aligned}$$

2.5. **Martingale problems.** All limiting processes considered in this paper behave like the Brownian Motion with a specified boundary condition. Therefore these limiting processes are characterized by the fact that

$$\phi(W(t)) - \frac{1}{2} \int_0^t \sum_{ab=1,2} \sigma_{ab} D_{ab} \phi(W(s)) ds \tag{8}$$

is a martingale for a set of the functions dense in the domain of the generator of the corresponding process. Therefore, for showing the convergence of a sequence of stochastic processes to such a Brownian Motion, by general theory (cf. [SV 71], [SV 06]) it suffices to show that the limiting process $W(t)$ of any convergent subsequence of the processes in question (8) is a martingale for the suitable class of functions. In fact, these classes of functions are the following:

- BM in $\mathbb{R}^2$: C^2 functions of compact support (Theorem 1);
- BM in a halfline: C^2 functions of compact support satisfying $\frac{\partial \phi}{\phi x}(0) = 0$ (Theorem 2);
- skew BM: continuous functions of compact support which admit C^2 extensions to $(-\infty, 0]$ and $[0, \infty)$ such that

$$\phi'_+(0) = \mathbf{a}\phi'_-(0)$$

where $\mathbf{a}$ is the skewness parameter (Theorem 3). The meaning of the constant $\mathbf{a}$ is the following: if we start the skew Brownian Motion from 0 then $\mathbb{P}(W(t) > 0) = \frac{1}{\mathbf{a}+1}$;
- reflected BM in a halfplane $x_1 \geq 0$: C^2 functions of compact support satisfying $\frac{\partial \phi}{\partial x_1}(0, x_2) = 0$ (Theorem 4).

3. Proof of Theorem 1. Tightness.

Since any probability measure, absolutely continuous with respect to the Liouville measure is admissible in the sense of equations (5) and (6), it suffices to prove Theorem 1 in case the initial conditions are distributed according to some $\mathbb{P}_\ell$.

We begin the proof with the following result.

Lemma 4. *Let ℓ be a standard pair, and the initial point be distributed according to ℓ. Then $W_N(t)$ is tight in $C[0,1]$.*

Proof. It is sufficient to show that for any standard pair ℓ, for suitable constants C_1, C_2 and for $N \geq N_0$ sufficiently large, for any $n \geq 1$ one has

(9)
$$\max_{0 \geq m \geq 2^n - 1} \mathbb{P}_\ell \left(\left| W_N\left(\frac{m}{2^n}\right) - W_N\left(\frac{m+1}{2^n}\right) \right| \geq \frac{1}{2^{n/4}} \right) \leq C_1 \exp -(C_2 2^{n/4}).$$

Indeed, for any given $\varepsilon, \eta > 0$, by selecting n_0 to satisfy $\sum_{n \geq n_0} 2^{-n/4} < \varepsilon$ and $\sum_{n \geq n_0} C_1 \exp(-C_2 2^{n/4}) < \eta$ one can easily bound the modulus of continuity $\omega_{W_N}(\delta)$ for suitable $\delta \leq 2\sum_{n \geq n_0} 2^{-n}$ by using the convergence — uniform in $N \geq N_0$ — of the series

$$\sum_n \mathbb{P}_\ell \left(\max_{0 \leq m \leq 2^n - 1} \left| W_N\left(\frac{m}{2^n}\right) - W_N\left(\frac{m+1}{2^n}\right) \right| \geq \frac{1}{2^{n/4}} \right)$$

(cf. [B 68], Theorem 8.2). Further, the event in (9) can be rewritten as

$$|q_{m_2} - q_{m_1}| \geq \sqrt{\bar{m}} L \tag{10}$$

where $m_1 = \frac{mN}{2^n}$, $m_2 = \frac{(m+1)N}{2^n}$, $\bar{m} = m_2 - m_1 = \frac{N}{2^n}$, $L = 2^{n/4}$. (Important note: in the whole paper we pretend as if variables like $m_1, m_2, \bar{m}$, etc. were integers, though typically they are not; it is easy to see that the deviations are negligible whereas keeping track of the precise error terms would hinder perspicuity of ideas.)

Since

$$|q_{m_2} - q_{m_1}| \le K_2 \bar{m},$$

(10) is only possible if

$$\bar{m} > \frac{1}{K_2^2} 2^{n/2} \qquad \text{or in other words} \qquad N > K_2^{-2} 2^{3n/2}. \tag{11}$$

(Observe that, by the last inequality, for any given N the event in (9) can only hold for a finite number of n's, only.) Let τ be the first time $m_1 \le \tau \le m_2$ such that

$$|q_\tau| \ge \frac{L}{4}\sqrt{\bar{m}} + 1$$

(if there is no such time before m_2 we put $\tau = m_2$). Based upon our previous argument the very last inequality can only hold if $\tau \ge \frac{1}{4K_2^2 K_3} 2^{n/2}$. This inequality ensures that, however short the length of ℓ be, τ is arbitrary large if n_0 is large. Consequently, Propositions 1(d) and 2(d) will be applicable.

By the definition of τ and by (10), the event in (9) implies that

$$\sup_{0 \le k \le \bar{m}} |q_{\tau+k} - q_\tau| \ge \frac{L}{4}\sqrt{\bar{m}}. \tag{12}$$

Consider the Markov decomposition

$$\mathbb{E}_\ell (A \circ T^\tau) = \sum_\alpha c_\alpha \mathbb{E}_{\ell_\alpha}(A).$$

Since $\tau - m_1 \le \bar{m}$, Proposition 1(d) implies that for any $\bar{\delta}$ slightly larger than δ

$$\sum_{\log|\text{length}(\ell_\alpha)| > \bar{m}^{1/2-\delta}} c_\alpha < \text{Const.}\bar{m} \exp\left(-\bar{m}^{1/2-\delta}\right) < \text{Const.} \exp\left(-\bar{m}^{1/2-\bar{\delta}}\right).$$

Since (12) depends only on the unmodified part of the system, we can apply Proposition 2(e) to each α with $\log|\text{length}(\ell_\alpha)| \le \bar{m}^{1/2-\delta}$ to obtain

$$\mathbb{P}_\ell\left(\sup_{0 \le k \le \bar{m}} |q_{\tau+k} - q_\tau| \ge \frac{L}{4}\sqrt{m_2 - m_1} \right) \le C_1 e^{-C_2 2^{\frac{n}{4}}}$$

as claimed. Indeed, the condition of Proposition 2(e) is not directly applicable to the value $\frac{L}{4}$. However, for bounding

$$\mathbb{P}_\ell \left(\max_{k \le \bar{m}} |q_{\tau+k} - q_\tau| \ge \frac{L}{4}\sqrt{\bar{m}} \right)$$

it is sufficient to estimate a larger expression by also using (11) as follows

$$\mathbb{P}_\ell \left(\max_{k \le \bar{m}} |q_{\tau+k} - q_\tau| \ge \bar{m}^{\frac{1}{6}-\delta}\sqrt{\bar{m}} \right)$$

$$\le C_1 \left[\exp\left(-\bar{m}^{1/2-\bar{\delta}}\right) + \exp\left(-c_2 \left(\frac{N}{2^n}\right)^{\frac{1}{3}-2\delta} \right) \right] \le C_1 \exp\left(-C_2 2^{\frac{n}{4}}\right)$$

where $C_2 > 0$ is suitably small. The last inequality provides the sufficient bound.

□

4. PROOF OF THEOREM 1. MARTINGALE PROBLEM.

Here we finish the proof of Theorem 1. Recall that we are assuming that initial conditions are distributed according to some $\mathbb{P}_\ell$.

Proof. Let ϕ be a smooth function of compact support. Denote $n = Nt$ and choose a small $\alpha > 0$. Let $L = N^\alpha$. Let $m_p = pL + z$ ($p \in \mathbb{Z}_+$)where z will be chosen later. Denote

$$\Delta_j = q_{j+1} - q_j.$$

We have

$$\phi\left(\frac{q_{m_{p+1}}}{\sqrt{N}}\right) - \phi\left(\frac{q_{m_p}}{\sqrt{N}}\right)$$

$$= \sum_{j=m_p+1}^{m_{p+1}} \frac{1}{\sqrt{N}} \left\langle D\phi\left(\frac{q_j}{\sqrt{N}}\right), \Delta_j \right\rangle + \frac{1}{2} \sum_{j=m_p+1}^{m_{p+1}} \frac{1}{N} \left\langle D^2\phi\left(\frac{q_j}{\sqrt{N}}\right) \Delta_j, \Delta_j \right\rangle + O(LN^{-3/2}).$$

Next for $m_p < j \le m_{p+1}$

$$D\phi\left(\frac{q_j}{\sqrt{N}}\right) = D\phi\left(\frac{q_{m_{p-1}}}{\sqrt{N}}\right) + \frac{1}{\sqrt{N}} \sum_{k=m_{p-1}+1}^{j} D^2\phi\left(\frac{q_{m_{p-1}}}{\sqrt{N}}\right) \Delta_k + O(L/N).$$

Hence

$$\phi\left(\frac{q_{m_{p+1}}}{\sqrt{N}}\right) - \phi\left(\frac{q_{m_p}}{\sqrt{N}}\right) = \sum_{j=m_p+1}^{m_{p+1}} \frac{1}{\sqrt{N}} \left\langle D\phi\left(\frac{q_{m_{p-1}}}{\sqrt{N}}\right), \Delta_j \right\rangle \tag{13}$$

$$+\frac{1}{N}\left[\frac{1}{2}\sum_{j=m_p+1}^{m_{p+1}}\left\langle D^2\phi\left(\frac{q_{m_{p-1}}}{\sqrt{N}}\right)\Delta_j,\Delta_j\right\rangle+\sum_{m_{p-1}<k<j}\left\langle D^2\phi\left(\frac{q_{m_{p-1}}}{\sqrt{N}}\right)\Delta_k,\Delta_j\right\rangle\right]$$

$$+O(L^2N^{-3/2}).$$

We now consider the Markov decomposition

$$\mathbb{E}_\ell(A\circ T^{m_p})=\sum_\alpha c_\alpha \mathbb{E}_{\ell_\alpha}(A\circ T^{m_{p-1}+m_p/2})=\mathcal{T}_1+\mathcal{T}_2$$

where $A=\phi\left(\frac{q_{m_1}}{\sqrt{N}}\right)-\phi\left(\frac{q_{m_0}}{\sqrt{N}}\right)$, $\mathcal{T}_1$ is the sum over α such that $|q_{m_{p-1}}|\geq KL$ and $\mathcal{T}_2$ is the sum over α such that $|q_{m_{p-1}}|<KL$. To estimate $\mathcal{T}_1$ split it $\mathcal{T}_1'+\mathcal{T}_1''$ where $\mathcal{T}_1'$ contains αs with length$(\ell_\alpha)>N^{-100}$. Since in any case the LHS of (13) is $O(L/\sqrt{N})$

$$\mathcal{T}_1''=O(L/N^{100.5})$$

by the Growth Lemma.

Decomposing $\mathcal{T}_1'=\mathcal{T}_{1a}+\mathcal{T}_{1b}+O(L^2N^{-3/2})$ according to the lines in (13) we get

$$\mathcal{T}_{1a}=O\left(\frac{L\theta^L}{\sqrt{N}}\right). \tag{14}$$

Indeed, $D\phi(\frac{q_{m_{p-1}}}{\sqrt{N}})$ varies little on each ℓ_α. Namely it can be approximated by a constant with error $O(\theta^L)$. Since Δ_0 has zero mean (14) follows by Proposition 2(a) (the factor of L comes since there are L terms).

To estimate $\mathcal{T}_{1b}$ we first observe that by the argument used to prove (14) we can bound for the contribution of each k,j to $\mathcal{T}_{1b}$ by $O\left(\frac{\theta^{j-k}}{N}\right)$. This shows that the total contribution of terms with $k<m_p$ is $O(\frac{1}{N})$. To estimate the contribution of the remaining terms we can use Proposition 2(b) to obtain

$$\mathcal{T}_{1b}=\frac{L}{2N}\sum_{st}D^2_{st}\phi\left(\frac{q_{m_{p-1}}}{\sqrt{N}}\right)\sigma^2_{st}.$$

Finally, since the summand in $\mathcal{T}_2$ is $O(L/\sqrt{N})$, we have

$$|\sum_p \mathcal{T}_2(p)|\leq\frac{\text{Const}L}{\sqrt{N}}\sum_p \mathbb{P}_\ell(|q_{m_{p-1}}|\leq KL).$$

The last sum can be rewritten as follows

$$\frac{\mathrm{Const}L}{\sqrt{N}}\sum_S \mathbb{E}_\ell(\mathrm{Card}(p:q_{m_{p-1}}\in S))$$

where the sum is taken over all scatterers within distance KL from the origin. Now we choose z so that the last sum is not more than its average over z, thus

$$\sum_S \mathbb{E}_\ell(\mathrm{Card}(p:q_{m_{p-1}}\in S))$$

$$\leq \frac{1}{L}\sum_S \mathbb{E}_\ell(\mathrm{Card}(j:q_j\in S)) \leq \mathrm{Const}L \max_S \mathbb{E}_\ell(\mathrm{Card}(j:q_j\in S))$$

since there are $O((KL)^2)$ scatterers within distance $O(L)$ from the origin.

Lemma 5. *There is a constant $\tilde{K}$ such that for all S*

$$\mathbb{E}_\ell(\mathrm{Card}(j\leq n:q_j\in S)) \leq \tilde{K}\log^{1+\xi} N$$

where ξ is the constant from Lemma 3.

Lemma 5 implies that

$$|\sum_p \mathbb{E}_\ell(\mathcal{T}_2(p))| \leq \mathrm{Const}\frac{L^2\log^{1+\xi} N}{\sqrt{N}} \to 0.$$

Thus if $W(t)$ is a limit point of $W_N(t)$, then taking the limit in (13) we get

$$\mathbb{E}_\ell\left(\phi(W(t))-\phi(W(0))-\frac{1}{2}\int_0^t\sum_{ab}D^2_{ab}\phi(W(s))\sigma^2_{ab}ds\right)=0. \tag{15}$$

A similar computation shows that if $\psi_1\dots\psi_m$ are smooth functions, then for any $s_1<s_2\dots s_m<t_1<t_2$ we have

$$\mathbb{E}_\ell\left(\left[\phi(W(t))-\phi(W(0))-\frac{1}{2}\int_{t_1}^{t_2}\sum_{ab}D^2_{ab}\phi(W(s))\sigma^2_{ab}ds\right]\prod_j\psi_j(W(s_j))\right)=0$$

proving Theorem 1. □

It remains to establish Lemma 5.

Proof of Lemma 5. Define two sequences m_1 and n_1 as follows. Let $m_0=0$ and let n_k be the first time after m_{k-1} such that $r_{n_k}(x)\geq\delta_0$.

Let m_k be the first time after n_k when $q_{m_k} \in S$. Then by the Growth Lemma we can find K so large that

$$\mathbb{P}_\ell(\max_{k \le N}(n_k - m_{k-1}) \ge K \log N) \le \frac{1}{N^{100}}. \tag{16}$$

Using Lemma 3(a) (with Γ being the modified part) we get inductively

$$\mathbb{P}_\ell(\max_{k \le b}(m_k - n_k) \le N) \le \left(1 - \frac{C}{\log^\xi N}\right)^b. \tag{17}$$

Let τ be the first time when $m_\tau - n_\tau > N$. (not visit S for n steps in a raw). Then (17) implies that

$$\mathbb{E}_\ell(\tau) \le \mathrm{Const} \log^\xi N.$$

Since

$$\mathrm{Card}(j \le n : q_j \in S) \le K\tau \log N + N \mathbb{1}_{\max_{k \le N}(n_k - m_{k-1}) \ge K \log N}$$

the lemma follows from (16). □

5. Proof of Theorem 2.

The proof of Theorem 2 is similar to the proof of Theorem 1 except that now Lemma 3(b) has to be used instead of Lemma 3(a). Accordingly the claim of Lemma 3 the equation (5) has to be replaced by

$$\mathbb{E}_\ell(\mathrm{Card}(j \le n : q_j \in S)) \le K\sqrt{N} \log N \tag{18}$$

which is much worse than (5). However now we want to establish (15) not for all functions but only for the functions in the domain of the reflected Brownian Motion, that is for functions satisfying $\phi'(0) = 0$. Accordingly $|\phi'(\frac{q_j}{\sqrt{N}})| \le \mathrm{Const}\frac{|q_j|}{\sqrt{N}}$. Important convention: for simplicity of notation in what follows q_j as an argument of the function ϕ will always denote the horizontal component of the vector q_j. Thus if $|q_{m_{p(j)}}| \le KL$ then

$$\left|\phi'\left(\frac{q_j}{\sqrt{N}}\right)\right| \le \mathrm{Const}\frac{L}{\sqrt{N}} \text{ and } \left|\phi\left(\frac{q_{m_{p+1}}}{\sqrt{N}}\right) - \phi\left(\frac{q_{m_p}}{\sqrt{N}}\right)\right| \le C\frac{L^2}{N}$$

Therefore we can estimate

$$|\sum_p \mathcal{T}_2(p)| \le O\left(\frac{L^2}{N}\right) \min_z \sum_{S:|q| \le KL} \mathbb{E}_\ell(\mathrm{Card}(p : q_{m_p} \in S)).$$

The last sum is less than its average over z, namely,

$$\frac{1}{L}\sum_{S:|q|\leq KL}\mathbb{E}_\ell(\text{Card}(j:q_j\in S))\leq\frac{1}{L}O(L)O(\sqrt{N}\log N)$$

where the second factor is due to the fact that there are $O(L)$ scatterers satisfying $|q|\leq KL$. Thus

$$\left|\sum_p \mathcal{T}_2(p)\right|=O\left(\frac{L^2\log N}{N}\right)$$

and the result follows.

6. First return maps. Statements

The difference between Theorems 1–2 and Theorems 3–4 is that for the former ones the terms with $|q_{m_p}|<KL$ can be estimated by their absolute values whereas for the later theorems this is not the case. For example, in Theorem 3 the skewness parameter **a** should be chosen carefully to make $\sum_p \mathcal{T}_2(p)\to 0$. Therefore the rough estimates like (18) are not enough for Theorems 3–4. Below we introduce some improvements based on a careful study of the first return maps. The proofs are given in Appendices C and D.

In the theorems below (Ω, f, μ) will either be the Lorentz process in the strip in the presence of an external field (Theorem 3) or the Lorentz process in the halfcylinder obtained by factorizing the Lorentz process in the half-plane over its group of (vertical) translational symmetries (Theorem 4).

For a fixed scatterer $S=\partial O$ let $T_{(S)}:S\times S_+\to S\times S_+$ be the first return map to the scatterer S. A different notation: $T_{[L]}$ be the first return map to $M_{[L]}:=\pi^{-1}(Q\cap\{|x|\leq L\})$.

Let $V_n(L)$ denote the number of visits of the Lorentz dynamics to $M_{[L]}$ up to time n.

Theorem 6. *(a) $T_{(S)}$ satisfies the assumptions of* [Y 98]. *In particular, $T_{(S)}$ is exponentially mixing.*

(b) There are constants $C,\bar{\alpha}$ such that for any S and for any $\delta>0$

$$\mathbb{E}_\ell\left(\text{Card}(j\leq n:q_j\in S, r_j(x)\leq\delta)\right)\leq C\left(\sqrt{n}\delta^{1/3}|\log\delta|^{\bar{\alpha}}+\log\text{length}(\ell)\right)$$

Theorem 7. *(a) $T_{[L]}$ has an SRB measure $\nu_{[L]}$ (for the cylinder case $\nu_{[L]}$ is the Liouville measure but in the presence of the field the existence of the SRB measure is a non-trivial statement). The mixing properties of $(T_{[L]},\nu_{[L]})$ can be summarized as follows.*

Let A be dynamical Holder function on $M_{[L]}$ such that

$$\int A d\nu_{[L]} = 0.$$

Then there are constants C, c, p such that for any standard pair ℓ and for any $n \geq C|\log \text{length}(\ell)|$ we have

$$\left|\mathbb{E}_\ell(A \circ T^n_{[L]})\right| \leq C||A||_{\mathcal{H}} \left(1 - \frac{c}{L^p}\right)^n.$$

(b) The family $\left\{\frac{V_n(L)}{L\sqrt{n}}\right\}$ is uniformly integrable (both in n and L).

(c) For any A such that $||A||_\infty \leq 1$, for any $n \geq \text{Const}||A||_{\mathcal{H}}$ and for any fixed $\delta > 0$ there exist positive constants C, c such that for abitrary $R < n^{1/6-\delta}$ we have

$$\mathbb{P}_\ell\left(\left|\sum_{j=0}^{n-1} A(T^j_{[L]}x) - n\nu_{[L]}(A)\right| \geq R\right) \leq Ce^{-c(R/L^p)^2}.$$

7. Proof of Theorem 3.

Now we describe the modifications needed to prove Theorem 3. In this case the domain of the generator consists of functions such that that ϕ is continuous, the one sided derivatives $\phi'_\pm(0)$ exist and

$$\phi'_+(0) = \mathbf{a}\phi'_-(0)$$

where $\mathbf{a}$ is the constant to be determined. Namely we want to choose $\mathbf{a}$ so that $\sum_p \mathcal{T}_2(p) \to 0$. Choose a large constant K_1 and denote $K_* = K_1 K_2$. where K_2 is defined by (4). Given j choose p so that $m_p \leq j < m_{p+1}$.

Let

$$\bar{\Delta}_j = \phi\left(\frac{q_{j+1}}{\sqrt{N}}\right) - \phi\left(\frac{q_j}{\sqrt{N}}\right).$$

We need to bound

$$\sum_j \mathbb{E}_\ell(\mathbb{1}_{|q_{m_{p(j)}}|<KL}\bar{\Delta}_j).$$

We split this sum into two parts.

(I) $j - m_{p(j)} < 2K_1 \log N$. There are two possibilities.

(a) $|q_{m_p}| > 2K_* \log N$. The contribution of these terms is small which can be proved similarly to the treatment of $\mathcal{T}_1$ term in Theorem 1.

(b) $|q_{m_p}| \leq 2K_* \log N$. The treatment of these terms can be done similarly to the estimate of $\mathcal{T}_2$-terms in Theorem 2 yielding

$$O\left(\log N \times \frac{\log N}{\sqrt{N}} \times \frac{\sqrt{N}}{L}\right)$$

where the first factor appears because there are $O(\log N)$ scatterers in $\{|q| < 2K_1K_2 \log N)\}$ the second factor appears since $\bar{\Delta}_j = O(\frac{1}{\sqrt{N}})$ and for every p there are $\log N$ terms. The third factor is an average number of visits to each scatterer on the m_p subsequence (here we use Theorem 6 and choose z in the definition of p appropriately).

(II) $j - m_p \geq 2K_1 \log N$.

Let $j_k = K_1 2^k$. Define the following events

$$\bar{A}_j = \{|q_{j-K_1 \log N}| > K_* \log N\},$$
$$A_{jk} = \{|q_{j-j_k}| > j_k K_2 \text{ but } |q_{j-j_{k+1}}| \leq j_{k+1} K_2\},$$
$$A_j = \{|q_{j-K_1}| \leq K_*\}.$$

Note, that if K_2 is chosen to be larger than the maximal free-flight this is a complete system of events for $0 \leq k \leq \log_2 \log N$. Observe that since ϕ is not smooth at 0 we cannot use the Taylor decomposition if $|q_j| \leq L$ however we have

$$\bar{\Delta}_j = \phi'_-(0) \frac{\zeta(q_{j+1}, \mathbf{a}) - \zeta(q_j, \mathbf{a})}{\sqrt{N}} + O\left(\frac{1}{N}\right) \text{ if } |q_j| \leq K_*,$$

where

$$\zeta(q, \mathbf{a}) = \begin{cases} \mathbf{a}q & \text{if } q \geq 0 \\ q & \text{if } q < 0 \end{cases}$$

Now we split

$$\sum_j \mathbb{E}_\ell(\mathbb{1}_{|q_{m_{p(j)}}| < KL} \bar{\Delta}_j) = \sum_j \mathbb{E}_\ell(\mathbb{1}_{|q_{m_{p(j)}}| < KL} \mathbb{1}_{\bar{A}_j} \bar{\Delta}_j) +$$
$$\sum_{jk} \mathbb{E}_\ell(\mathbb{1}_{|q_{m_{p(j)}}| < KL} \mathbb{1}_{A_{jk}} \bar{\Delta}_j) + \sum_j \mathbb{E}_\ell(\mathbb{1}_{|q_{m_{p(j)}}| < KL} \mathbb{1}_{A_j} \bar{\Delta}_j)$$
$$= \sum_j \bar{\mathcal{E}}_j + \sum_{jk} \mathcal{E}_{jk} + \sum_j \mathcal{E}_j.$$

On the event $\bar{A}_j$ we surely avoid the perturbation for the whole $K_1 \log N$ trajectory segment. Hence for the first term we can apply the exponential mixing for Sinai billiards to get

$$\sum_j \bar{\mathcal{E}}_j = \sum_j O(\theta^{K_1 \log N}) = O(N^{-100})$$

provided that K_1 is large enough.

Next we estimate $\sum_j \mathcal{E}_{jk}$ for a given k. Consider a Markov decomposition

$$\mathbb{E}_\ell(A \circ T^{j-j_k}) = \sum_\alpha c_\alpha \mathbb{E}_{\ell_\alpha}(A).$$

Let

$$\mathcal{E}'_{jk} = \sum_{\text{length}(\ell_\alpha) > \theta^{j_k}} c_\alpha \mathbb{E}_{\ell_\alpha}(\mathbb{1}_{A_{jk}} \bar{\Delta}_j),$$

$$\mathcal{E}''_{jk} = \sum_{\text{length}(\ell_\alpha) \leq \theta^{j_k}} c_\alpha \mathbb{E}_{\ell_\alpha}(\mathbb{1}_{A_{jk}} \bar{\Delta}_j),$$

For terms in $\mathcal{E}'_{jk}$ we have

$$\mathbb{E}_{\ell_\alpha}(\Delta_j) = O(\theta^{j_k})$$

so

$$|\sum_j \mathcal{E}'_{jk}| \leq \text{Const} \left[\frac{\theta^{j_k}}{\sqrt{N}} + \frac{1}{N}\right] \sum_j \mathbb{P}_\ell(|q_{j-j_{k+1}}| \leq K_2 j_{k+1})$$

Now Theorem 6(b) tells us that the last sum is $O(j_k\sqrt{N})$. It follows that by choosing K_1 large we can make $\sum_{jk} \mathcal{E}'_{jk}$ as small as we wish. On the other hand

$$|\sum_{jk} \mathcal{E}''_{jk}| \leq \frac{\text{Const}}{\sqrt{N}} \sum_j \mathbb{P}_\ell(|q_{j-j_k}| \leq K_2 j_k \text{ and } r_{j-j_k}(x) \leq \theta^{j_k}).$$

Therefore Theorem 6(b) implies that by choosing K_1 large we can make $\sum_{jk} \mathcal{E}''_{jk}$ as small as we wish. Thus the main contribution to $\sum_p \mathcal{T}_2(p)$ comes from $\mathcal{E}_j$. In other words we proved that

$$\text{(19)} \qquad \mathbb{E}_\ell\left(\phi\left(\frac{q_n}{\sqrt{N}}\right) - \phi\left(\frac{q_0}{\sqrt{N}}\right) - \frac{1}{2N}\sum_j \phi''\left(\frac{q_n}{\sqrt{N}}\right)\sigma^2\right)$$

$$= \frac{\phi'_-(0)}{\sqrt{N}} \mathbb{E}_\ell\left(\sum_j \mathbb{1}_{A_j}\left(\zeta(q_{j+1}, \mathbf{a}) - \zeta(q_j, \mathbf{a})\right)\right) + o(1), \quad N \to \infty, K_1 \to \infty.$$

To estimate the last sum we consider the first return map $T_{[K_*]}$ to $|q| \leq K_*$. After reindexing we get

$$\frac{1}{\sqrt{N}} \sum_{j=1}^{V_n(K_*)} \hat{\zeta}(T^j_{[K_*]} x, \mathbf{a})$$

where $\hat{\zeta} = \zeta \circ T^{K_1+1} - \zeta \circ T^{K_1}$. Observe that

$$\frac{1}{V_n(K_*)} \sum_{j=1}^{V_n(K_*)} \hat{\zeta}(T^j_{[K_*]}x, a)$$

is uniformly bounded and it converges almost surely to $\int \hat{\zeta} d\nu_{[K_*]}$ (see Theorem 7(c)).

Denote $\hat{\nu}_{[K]} = \frac{\nu_{[K]}}{\nu_{[K]}(M_{[1]})}$ (this normalization is needed so that the restriction of $\hat{\nu}_{[\tilde{K}]}$ to $M_{[K]}$ equals $\hat{\nu}_{[K]}$ for $\tilde{K} \geq K$). Since $\{V_n(K)/\sqrt{N}K\}$ is uniformly integrable it follows that for large N

$$\left| \frac{1}{\sqrt{N}} \mathbb{E} \left(\sum_{j=1}^{V_n(K_*)} \hat{\zeta}(T^j_{[K_*]}x, a) \right) \right|$$

$$\leq 2 \int \hat{\zeta} d\nu_{[K_*]} \mathbb{E}_\ell \left(\frac{V_n(K_*)}{\sqrt{N}} \right) = 2 \int \hat{\zeta} d\hat{\nu}_{[K_*]} \mathbb{E}_\ell \left(\frac{V_n(K_*)}{K_* \sqrt{N}} \right) K_* \nu_{[K_*]}(M_{[1]}).$$

Lemma 6. *(a) There exists a limit*

$$\gamma(\mathbf{a}) = \lim_{K_* \to \infty} \int \hat{\zeta}(\cdot, \mathbf{a}) d\hat{\nu}_{[K_*]}.$$

Moreover there is $\eta > 0$ *such that* $\gamma(\mathbf{a}) - \int \hat{\zeta}(\cdot, \mathbf{a}) d\hat{\nu}_{[K_*]} = O(\theta^{K_*^\eta})$.

(b) There is a constant C *such that*

$$\nu_{[K_*]}(M_{[1]}) \leq \frac{C}{K_*}.$$

$\gamma(\mathbf{a})$ is an affine function of $\mathbf{a}$ because ζ is an affine function of $\mathbf{a}$. Thus we can choose $\mathbf{a}$ so that $\gamma(\mathbf{a}) = 0$. Then (19) gives

$$\mathbb{E}_\ell \left(\phi\left(\frac{q_n}{\sqrt{N}}\right) - \phi\left(\frac{q_0}{\sqrt{N}}\right) - \frac{1}{2N} \sum_j \phi''\left(\frac{q_n}{\sqrt{N}}\right) \sigma^2 \right) = o(1), \quad N \to \infty.$$

Hence any limit process will satisfy

$$\mathbb{E} \left(\phi(W(t)) - \phi(W(0)) - \frac{1}{2} \int_0^t \phi''(W(s)) \sigma^2 ds \right) = 0$$

and we are done as before. It remains to establish Lemma 6.

Proof. It suffices to show that

$$\int \zeta(T^{K_1}x, \mathbf{a}) d\hat{\nu}_{[K_*]} - \\ - \int \zeta(T^{K_1+1}x, \mathbf{a}) d\hat{\nu}_{[K_*+1]} = O\left(\theta^{K_*^\eta}\right).$$

We split the LHS into two parts

$$I = \int \zeta(T^{K_1}x, \mathbf{a})d\hat{\nu}_{[K_*]} - \int \zeta(T^{K_1+1}x, \mathbf{a})d\hat{\nu}_{[K_*]},$$

$$II = \int \zeta(T^{K_1+1}x, \mathbf{a})d\hat{\nu}_{[K_*]} - \int \zeta(T^{K_1+1}x, \mathbf{a})d\hat{\nu}_{[K_*+1]}.$$

To estimate I we observe that since $\hat{\nu}_{[K_*]}$ is $T_{[K_*]}$ invariant we have

$$I = \int \left[\zeta(T^{K_1}T_{[K_*]}x, \mathbf{a}) - \zeta(T^{K_1+1}x, \mathbf{a})\right] d\hat{\nu}_{[K_*]}.$$

But the integrand is different from zero only for points where $Tx \notin M_{[K_*]}$. Those points are near the boundary of $M_{[K_*]}$ and so by Proposition 2(d).

$$\nu_{[K_*]}(Tx \notin M_{[K_*]} \text{ but } \exists j \leq K_* + 1 \text{ such that } T^j x \text{ or } T^j(T_{[K_*]}x) \text{ visit the modified part})$$

$$\leq \text{Const}\theta^{K_*^\eta}$$

For the other orbits we can use the exponential mixing of Sinai billiards to show that

$$\int \mathbb{1}_{T_{[K_*]}x \neq Tx}\, \zeta(T^{K_1}T_{[K_*]}x)d\nu_{[K_*]}(x) \leq O\left(\theta^{K_*^\eta}\right) + O\left(\theta^{K_*}\right)$$

and

$$\int \mathbb{1}_{T_{[K_*]}x \neq Tx}\, \zeta(T^{K_1+1}x)d\nu_{[K_*]}(x) \leq O\left(\theta^{K_*^\eta}\right) + O\left(\theta^{K_*}\right)$$

Therefore

$$I = O\left(\theta^{K_*^\eta}\right) + O\left(\theta^{K_*}\right) = O\left(\theta^{K_*^\eta}\right).$$

Likewise

$$II = \int \zeta(T^{K_1+1}x, \mathbf{a})\mathbb{1}_{M_{[K_*+1]} - M_{[K_*]}}(x)d\hat{\nu}_{[K_*+1]} = O\left(\theta^{K_*^\eta}\right)$$

proving (a).

To prove (b) let $\mathbf{m}_{k_0}$ be the time of k_0-th return to $M_{[1]}$ (under $T_{[K_*]}$) where k_0 is the constant of Lemma 3. By parts (b) and (c) of Lemma 3 there are constants c and ϵ such that

$$\mathbb{P}(m_{k_0} \geq n) \geq \frac{c}{\sqrt{n}}$$

for $n \leq \epsilon K_*^2$. Hence

$$\nu_{[K_*]}(M_{[1]}) = \frac{k_0}{\mathbb{E}(m_{k_0})} = \frac{k_0}{\sum_{n=1}^{\infty} \mathbb{P}(m_{k_0} \geq n)} \leq \frac{k_0}{\sum_{n=1}^{\epsilon K_*^2} \mathbb{P}(m_{k_0} \geq n)} \leq \frac{C}{K_*}.$$

□

8. Proof of Theorem 4.

Here we explain the changes needed to prove Theorem 4. First, the proof of the tightness given in Section 3 has to be changed because here we modify the configuration along the line so the particle could 'slide' along this line for a long time. Thus while the tightness of $\frac{z_{1[Nt]}}{\sqrt{N}}$ can be proven as before a different argument is needed for $\frac{z_{2[Nt]}}{\sqrt{N}}$. We divide the proof into two lemmas. (For simplicity, in this sequel, notations of type q_j will denote the first component of the vector q_j, and notations of type Δ_{2j} will denote $\Delta_{2,j}$.)

Let

$$\bar{z}_k = \sum_{j=0}^{k-1} \Delta_{2j} \mathbb{1}_{M^c_{[KL]}}(q_{j-L}), \quad \bar{\bar{z}}_k = \sum_{j=0}^{k-1} \Delta_{2j} \mathbb{1}_{M_{[KL]}}(q_{j-L}),$$

$$\bar{W}_N(t) = \frac{\bar{z}_{[Nt]}}{\sqrt{N}}, \quad \bar{\bar{W}}_N(t) = \frac{\bar{\bar{z}}_{[Nt]}}{\sqrt{N}}.$$

Lemma 7. *$\{\bar{W}_N(t)\}$ is tight.*

Proof. Consider the following function on Ω

$$A(x) = (q_1 - q_0)\mathbb{1}_{M^c_{[KL]}}(q_{-L}).$$

Taking the Markov decomposition

$$\mathbb{E}_\ell(A \circ f^n) = \sum_\alpha c_\alpha \mathbb{E}_{\ell_\alpha}(A \circ f^L) \tag{20}$$

applying Proposition 2(a) to the long components where $A \circ f^L \neq 0$ and using Proposition 1(b) to estimate the measure of short components we get

$$\mathbb{E}_\ell(A \circ f^n) = O(\theta^L). \tag{21}$$

Now arguing as in the proof of Proposition 6.1 of [ChD 07] we obtain the following bounds for $n_2 - n_1 > N^{3/7}$.

$$\mathbb{E}_\ell(\bar{z}_{n_1} - \bar{z}_{n_2}) = O(L),$$
$$\mathbb{E}_\ell((\bar{z}_{n_1} - \bar{z}_{n_2})^2) = O(|n_1 - n_2|),$$
$$\mathbb{E}_\ell((\bar{z}_{n_1} - \bar{z}_{n_2})^4) = O((n_2 - n_1)^2).$$

Indeed the proof of Proposition 6.1 in [ChD 07] relied only on the equidistibition lemma (Corollary 3.4 of [ChD 07]) and (21) is the analogue of such lemma in our situation. Now tightness can be derived from the last three estimates the same way Proposition 6.2 is derived from Proposition 6.1 in [ChD 07]. □

Lemma 8. $\max_k \frac{|\bar{\bar{z}}_k|}{\sqrt{N}}$ *converges to* 0 *in probability.*

Proof. Observe that

$$\bar{\bar{z}}_k = \sum_{j=1}^{V_k(KL)} \Delta_2(T^L T_{[KL]}^j x).$$

Hence Theorem 7(c) implies

$$\bar{\bar{z}}_k = c_L V_k(KL) + o_{\mathbb{P}}(L^p k^{1/4+\varepsilon})$$

where

$$c_L = \frac{1}{KL}\int_{M_{[KL]}} \Delta_2 \circ T^L d\mu,$$

μ is the Liouville measure and we write $\mathcal{A} = o_{\mathbb{P}}(\mathcal{B})$ if for any ε

$$\mathbb{P}(|\mathcal{A}| \geq \varepsilon|\mathcal{B}|)$$

tends to 0 faster than any power of N. Similarly to Lemma 6 we obtain that there exists the limit

$$\gamma = \lim_{L\to\infty} c_L KL \text{ and } \gamma - c_L KL = O(\theta^{L^\eta}).$$

Next

$$V_k(1) = \sum_{j=1}^{V_k(KL)} \mathbb{1}_{M_{[1]}}(T_{[KL]}^j x)$$

so using again Theorem 7(c) we get

$$V_k(KL) = KLV_k(1) + o_{\mathbb{P}}(L^p V_k(1)^{1/2+\varepsilon}).$$

Therefore

$$\max_k \left| \frac{\bar{\bar{z}}_k}{\sqrt{N}} - \gamma \frac{V_k(1)}{\sqrt{N}} \right| \to 0$$

so it remains to show that $\gamma = 0$. Let t_N be the first time when $V_t(1) = N$. Then the foregoing computation shows that

$$\mathbb{E}_\ell(\bar{\bar{z}}_{t_N}) = N(\gamma + o(1)).$$

Next we claim that

$$\mathbb{P}_\ell(t_N \geq N^{202}) = O\left(N^{-100}\right). \tag{22}$$

Indeed let $\bar{t}_1 \leq \bar{t}_2 \leq \cdots \leq \bar{t}_k$ be the consecutive visits to $M_{[1]}$ such that $r_{\bar{t}_j}(x) \geq \delta_0$. Applying Lemma 12 proven in Appendix C we prove by induction that

$$\mathbb{P}_\ell\left(\max_{j\leq k}(\bar{t}_j - \bar{t}_{j-1}) \geq n\right) = O\left(\frac{k}{\sqrt{n}} log^\alpha n\right). \tag{23}$$

(23) with $k = N$, $n = N^{202}$ easily implies (22). Since $\bar{\bar{z}}_{t_N}$ is always $O(N)$ we get

$$\mathbb{E}_\ell(\bar{\bar{z}}_N \mathbb{1}_{t_N \leq N^{202}}) = N(\gamma + o(1)).$$

On the other hand using (20) and (21) we see that

$$\mathbb{E}_\ell(\bar{z}_{\min(t_N, N^{202})}) = \sum_{n=0}^{N^{202}} O\left(\theta^L + \mathbb{P}_\ell(n - L \leq t_N \leq n)\right) = O\left(\theta^L N^{202} + L\right) = O(L).$$

Combining this with (22) we get

$$\mathbb{E}_\ell(\bar{z}_{t_N} \mathbb{1}_{t_N \leq N^{202}}) = O(L)$$

and so

$$\gamma = \lim_{m \to \infty} \frac{\mathbb{E}(z_{2,t_m} \mathbb{1}_{t_m < m^{202}})}{m}.$$

By the time reversal symmetry $\gamma = 0$. □

The second change comes in the estimate for the expectation of

$$\sum_j \frac{D_2 \phi(q_j / \sqrt{N})}{\sqrt{N}} \Delta_{2j} \mathbb{1}_{|q_{m_{p(j)}}| \leq L}.$$

Indeed we have

$$D_2 \phi(q_j / \sqrt{N}) \sim D_2 \phi(0, z_{2j} / \sqrt{N})$$

but as $z_{2j}/\sqrt{N}$ is not constant we can not factor it out like in the proof of Theorem 1. However we can divide the interval $[0, n]$ into intervals of length δN with small δ and use the tightness proven above to conclude that $D_2\phi(q_j/\sqrt{N})$ changes little on each interval so it can be factored out. The rest of the proof is similar to the proof of Theorem 3.

9. Continuous time.

Proof of Theorem 5. We shall show how to extend Theorem 1, other results are extended in a similar way. Let t_j be the time between j-th and $(j+1)$-st collisions and $\mathbf{L} = \mu_0(t_1)$ be the mean free path. Let $T_n = \sum_{j=0}^{n-1} t_j$ be the time of the n-th collision. Arguing as in the proof of Theorem 1 we show that the diffusively scaled version of $T_n - n\mathbf{L}$ converges to a Brownian Motion. In particular for any $\varepsilon > 0$ there exists $R > 0$ such that

$$\mathbb{P}_\ell(\max_{0 \leq k \leq n} |T_k - k\mathbf{L}| \geq R\sqrt{n}) \leq \varepsilon.$$

Thus the continuous time process is obtained from the discrete time process by the time change $s = t\mathbf{L}$. The result follows. □

APPENDIX A. MAXIMAL OSCILLATIONS.

Proof of Proposition 2(e). We apply the Reflection Principle (cf. the Lemma to Theorem 10.1 of [B 68]) to our situation. If the event in barckets hold, then let $\bar{j}$ be the first time $j \leq n$ such that

$$\left| A_j - j \int A d\mu_0 \right| \geq R\sqrt{n}.$$

Then Proposition 1 gives the decomposition

$$\mathbb{E}_\ell(A \circ T^{\bar{j}}) = \sum_\alpha c_\alpha \mathbb{E}_{\ell_\alpha}(A)$$

where

$$\sum_{\text{length}(\ell_\alpha) \leq \varepsilon} c_\alpha \leq \text{Const}.\varepsilon\, n.$$

Applying Proposition 2(d) to each α with $|\log \text{length}(\ell_\alpha)| < n^{1/2-\delta}$ we conclude that there are constants $\bar{C}_1, \bar{C}_2, \bar{C}_3$ and $\bar{\delta} > \delta$ such that

$$\mathbb{P}_\ell \left(\left| A_n - n \int A d\mu_0 \right| \geq (R - \bar{C}_3)\sqrt{n} \right) \geq$$

$$\geq \bar{C}_1 \mathbb{P}_\ell \left(\max_{j \leq n} \left| A_j - j \int A d\mu_0 \right| \geq R\sqrt{n} \right) - \bar{C}_2 \exp(-n^{\frac{1}{2}-\bar{\delta}}).$$

Therefore part (e) of Proposition 2 follows from part (d). □

APPENDIX B. RETURN TIMES.

Proof of Lemma 3. (a) Without loss of generality we can assume that S is in the 0-th cell. Take a standard pair ℓ with $\text{length}(\ell) \geq \delta_0$. It suffices to show that if R is sufficiently large and $d([\ell], (S \bigcup \Gamma)) \geq R$, then

$$\mathbb{P}_\ell \left(q_j \notin (S \bigcup \Gamma) \text{ for } j = 1 \dots n \right) \geq \frac{\text{Const}}{\log^\alpha n}. \tag{24}$$

We establish (24) in case $\text{Card}(\Gamma) = 1$, the general case is similar. For fixing our ideas we also assume that $d([\ell], S) \ll d([\ell], \Gamma)$, the other cases are easier. Take a sufficiently small $\varepsilon_0 > 0$. Let τ_1 be the first time τ such that either $|q_\tau| \geq R^{1+\varepsilon_0}$ or $|q_\tau| \leq R^{\varepsilon_0}$. It is proven in Sections 6 and 7 of [DSzV 07] that for any standard pair ℓ satisfying $\text{length}(\ell) \geq \delta_0$ and $[\ell] = R$ we have

$$\mathbb{P}_\ell \left(|q_{\tau_1}| \geq R^{1+\varepsilon_0} \text{ and } r_{\tau_1}(x) \geq R^{-100} \right) \geq \zeta \tag{25}$$

where $1-\zeta \asymp \varepsilon_0$, and thus ζ can be made as close to 1 as needed by choosing ε_0 small. Define τ_k as a first time τ after τ_{k-1} when either

$$|q_\tau| \geq R^{(1+\varepsilon_0)^k} \text{ or } |q_\tau| \leq R^{\varepsilon_0(1+\varepsilon_0)^{k-1}}.$$

Iterating (25) we get

$$\mathbb{P}_\ell\left(|q_{\tau_k}| \geq R^{(1+\varepsilon_0)^k} \text{ and } r_{\tau_k}(x) \geq R^{-100(1+\varepsilon_0)^{k-1}}\right) \geq \zeta^k. \tag{26}$$

Let $\bar{k}$ be the largest number such that

$$R^{(1+\varepsilon_0)^{\bar{k}}} < \frac{d(\Gamma,0)}{2}.$$

Applying (26) with $k=\bar{k}$ we see that the probability that the particle moves $(\frac{d(\Gamma,0)}{2})^{1/(1+\varepsilon_0)}$ away from the origin without visiting S is at least $c_1/\log(d(\Gamma,0))$.

For crossing the region where the particle can hit Γ we need a more delicate argument. To do so we define $\bar{\tau}_1$ as a first time τ after $\tau_{\bar{k}}$ such that

$$|q_{\bar{\tau}}| \geq d^{1+\varepsilon_0}(\Gamma,0)^{1+\varepsilon_0} \text{ or } |q_{\bar{\tau}}| \leq d^{1/(1+\varepsilon_0)^3(\Gamma,0)}.$$

Then by the argument of Lemma 6.1(a) of Section 6 of [DSzV 07] there exists a constant $c_2>0$ such that for any standard pair ℓ satisfying

$$|[l]| \geq \left(\frac{d(\Gamma,0)}{2}\right)^{1/(1+\varepsilon_0)} \text{ and length}(\ell) \geq d^{-100}(\Gamma,0) \tag{27}$$

we have

$$\mathbb{P}_\ell\left(|q_{\bar{\tau}_1}| \geq d^{1+\varepsilon_0}(\Gamma,0), \quad r_{\bar{\tau}_1}(x) \geq \delta_0 \text{ and } \bar{\tau}_1 - \tau_{\bar{k}} \leq d^{3(1+\varepsilon_0)}(\Gamma,0)\right) \geq c_2.$$

On the other hand, by Theorem 4 of [DSzV 07], for any standard pair satisfying (27)

$$\mathbb{P}_\ell\left(q_j \text{ visits } \Gamma \text{ before time } d^{3(1+\varepsilon_0)}(\Gamma,0)\right) \to 0 \text{ as } \varepsilon_0 \to 0, d(\Gamma,0) \to \infty.$$

Hence if ε_0 is sufficiently small, then we can arrange that for a suitable $c_3>0$

$$\mathbb{P}_\ell\left(|q_{\bar{\tau}_1}| \geq d^{1+\varepsilon_0}(\Gamma,0), \quad r_{\bar{\tau}_1}(x) \geq \delta_0 \text{ and } q_j \text{ does not visit } \Gamma \text{ before } \bar{\tau}_1\right) \geq c_3$$

Next let $\bar{\tau}_k$ be the first time τ after $\bar{\tau}_{k-1}$ such that either

$$|q_\tau| \geq d^{(1+\varepsilon_0)^k}(\Gamma,0) \text{ or } |q_\tau| \leq d(\Gamma,0).$$

The argument used to prove (26) shows that for any ℓ such that

$$|[l]| \geq d^{(1+\varepsilon_0)}(\Gamma,0), \quad \text{length}(\ell) > \delta_0$$

we have

$$\mathbb{P}_\ell(|q_{\bar{\tau}_k}| \geq R^{(1+\varepsilon_0)^k}) \geq \zeta^k.$$

Taking $\hat{k}$ such that

$$d^{(1+\varepsilon_0)^{\hat{k}}}(\Gamma, 0) = n$$

we get part (a).

Part (b) is proven in [DSzV 07] in case $[\ell] \in S$. To get the result in general, let τ^* be the first time the particle visits S and observe that by Theorem 11 of [DSzV 07] $\mathbb{P}_\ell(r_{\tau^*}(x) \geq \delta_0)$ is uniformly bounded from below so we can apply the result for $[\ell] \in S$.

The proof of part (c) is similar to the Proof of Lemma 11.1(c) of [DSzV 07]. □

Appendix C. First return to one scatterer.

Here we prove Theorem 6.

Let δ_0 be sufficiently small . Let $\tau_1 < \tau_2 < \dots \tau_k \dots$ be consecutive visits to S.

Lemma 9. *There are positive constants c_1, c_2 such that if ℓ is a standard pair,* length$(\ell) \geq \delta_0$ *then*

$$\mathbb{P}_\ell(\tau_1 < c_2 d^2(\ell, S), r_j(x) \geq \min(\delta_0, d^{-100}(q_j, S)) \text{ for } j \leq \tau_1) \geq c_1.$$

Proof. This follows from the proof of Theorem 10 of [DSzV 07]. □

Lemma 10. *There is a constant c_3 such that if ℓ is a standard pair such that* length$(\ell) \geq d^{-101}(\ell, S)$ *then*

$$\mathbb{P}_\ell(\tau_1 > n) \leq c_3 \frac{d(\ell, S) + 1}{\sqrt{n}}.$$

Proof. This follows from the proof of Theorem 8 of [DSzV 07]. □

Lemma 11. *There are constants $c_4, c_5 > 0, \theta_1 < 1$ such that for any standard pair the following holds. Let $\bar{n}$ be the first positive time when $r_{\tau_{\bar{n}}}(x) \geq \delta_0$ then*

$$\mathbb{P}_\ell(\bar{n} - c_4 |\log(\text{length}(\ell)| \geq n) \leq c_5 \theta_1^n.$$

Proof. We begin with the case when length$(\ell) \geq \delta_0$, and assume $\delta_0 < d^{-100}(x, S)$. Let k_1 be the smallest among the following numbers

- $c_2 d^2(\ell, S)$
- $\tau_1(x)$
- the first time k when $r_k(x) < \min(\delta_0, d^{-100}(q_k, S))$.

If $k_1 = \tau_1(x)$ we stop otherwise let m_1 be the first number after k_1 such that $r_{m_1}(x) \geq \delta_0$. Let k_2 be the smallest among the following numbers

- $m_1 + c_2 d^2(x_{m_1}, S)$
- $m_1 + \tau_1(x_{m_1})$
- the first time k after m_1 when $r_k(x) < \min(\delta_0, d^{-100}(q_k, S))$.

Continue this procedure until $q_{k_p} \in S$.

Observe that if δ_0 is small enough then our construction implies that $r_{k_p}(x) \geq \delta_0$. Also by Lemma 9

$$\mathbb{P}_\ell(p > n) < (1 - c_1)^n.$$

Next we claim that there are constants $c_6 > 0, \theta_2 < 1$ such that

$$\mathbb{P}_\ell(\text{Card}(i : k_j \leq i \leq m_j, q_i \in S) > n \mid p > j) \leq c_6 \theta_2^n \tag{28}$$

To derive (28) we distinguish two cases:

- $r_{k_j} \geq \exp(-\varepsilon |q_{k_j}|)$ where ε is sufficiently small. Since the orbit can not hit S during next $\frac{|q_{k_j}|}{K_2}$ iterations if
$$\text{Card}(i : k_j \leq i \leq m_j, q_i \in S) > n$$
then $m_j > n + \frac{|q_{k_j}|}{K_2}$ so the result follows from Proposition 1(e).
- $r_{k_j} < \exp(-\varepsilon |q_{k_j}|)$.
  - If $n > \bar{C} |\log r_{k_j}(x)|$, the result follows from Proposition 1(e).
  - If the opposite inequality
$$n \leq \bar{C} |\log r_{k_j}(x)| \tag{29}$$
holds, then for any $\delta > 0$ we have
$$\mathbb{P}_\ell(\delta \leq r_{k_j} < 2\delta, r_{k_j} < \exp(-\varepsilon |q_{k_j}|) \mid p > j) \leq$$
$$\leq \mathbb{P}_\ell(\delta \leq r_{k_j} < 2\delta, |q_{k_j}| < \frac{1}{\varepsilon} |\log \delta| \mid p > j)$$
Let t be the first time after m_{j-1} when $|q_t| \leq \frac{2}{\varepsilon} |\log \delta|$ (it can be m_{j-1} itself). By the definition of k_j on the event $\{|q_{k_j}| < \frac{1}{\varepsilon} |\log \delta|\}$ we have $r_t > (\frac{2}{\varepsilon} |\log \delta|)^{-100}$. Now Lemma 10 implies that for any standard pair ℓ such that $d(\ell, S) <$

$\frac{2}{\varepsilon}|\log\delta|$ and length$(\ell) > (\frac{2}{\varepsilon}|\log\delta|)^{-100}$ we have

$$\mathbb{P}_\ell(\tau_1 > n) \leq \text{Const}\frac{|\log\delta|}{\varepsilon\sqrt{n}}$$

and so by Growth Lemma for any n

$$\mathbb{P}_\ell(\min_{j\leq\tau_1} r_j(x) < \delta) \leq \text{Const}\frac{|\log\delta|}{\varepsilon\sqrt{n}} + n\delta.$$

Choosing $n = (|\log\delta|/\varepsilon\delta)^{2/3}$ we obtain:

$$\mathbb{P}_\ell(\delta \leq r_{k_j} < 2\delta, r_{k_j} < \exp(-\varepsilon|q_{k_j}|) \mid p > j) = O(\delta^{1/3}|\log\delta|^{2/3}).$$

This can be summed over $\delta_i = 1/2^i$. The desired bound follows from the fact that the largest possible δ satisfies $\delta \leq r_{k_j} \leq \exp(-n/\bar{C})$

Next denote $\phi_j(z) = \mathbb{E}_\ell(z^{N_j})$ where

$$N_j = \text{Card}(i : \exists\bar{j} \leq j,\ k_{\bar{j}} \leq i \leq m_{\bar{j}} \text{ and } q_i \in S).$$

We claim that there is a constant $\bar{c}$ such that for $|z| \leq \frac{1}{2}\left[1+\theta_2^{-1}\right]$ we have

$$|\phi_1(z)| \leq 1 + \bar{c}(|z|-1) \tag{30}$$

uniformly in ℓ. Indeed (28) shows that ϕ_1 is analytic and uniformly bounded in any disc of radius less than θ_2^{-1}. In particular $|\phi'| \leq \bar{c}$ for $|z| \leq \frac{1}{2}\left[1+\theta_2^{-1}\right]$. Combining this with the fact that $|\phi_1(z)| \leq 1$ if $|z| \leq 1$ we obtain the result. Now it is easy to show by induction that

$$\phi_j(z) \leq \sum_{m=1}^{j} (1-c_1)^m \left(1+\bar{c}(|z|-1)\right)^m.$$

Hence $\phi(z) = \lim_{j\to\infty}\phi_j(z)$ converges in some neighbourhood of 1 proving Lemma 11 if length$(\ell) \geq \delta_0$. In general case we define $k_0 = 0$ and m_0 to be the first time then $r_m(x) \geq \delta_0$ and argue as before. □

Lemma 11 implies the exponential mixing via the coupling algoritm of [Ch 06]. This proves Theorem 6(a).

Next we use this lemma to control the returns of short components. We need a preliminary result.

Lemma 12. *For any standard pair ℓ such that $[\ell] \in S$ and* $\text{length}(\ell) \geq \delta_0$ *we have*

$$\mathbb{P}_\ell(\tau_{\bar{n}} \geq n) \leq \frac{c_7 \log^\alpha n}{\sqrt{n}}.$$

Proof. We use an idea of [M 04]. By Lemma 11 we can choose a constant C such that $\mathbb{P}_\ell(\bar{n} \geq C \log n) \leq \frac{1}{n}$. Denote $\tau_0 = 0$. We need to show that

$$\mathbb{P}_\ell\left(\max_{1\leq j\leq C\log n}(\tau_j - \tau_{j-1}) \geq \frac{n}{C\log n}\right) \leq \frac{\text{Const}\log^\alpha n}{\sqrt{n}}.$$

To this end we show that for any $1 \leq j \leq C \log n$

$$\mathbb{P}_\ell\left(\tau_j - \tau_{j-1} \geq \frac{n}{C\log n}, \max_{1\leq l<j}(\tau_l - \tau_{l-1}) < \frac{n}{C\log n}\right)$$
$$\leq \frac{\text{Const}\log^{\alpha-1} n}{\sqrt{n}}.$$

The Growth Lemma (Proposition 1) implies that

$$\mathbb{P}_\ell\left(\max_{1\leq l<j}(\tau_l - \tau_{l-1}) < \frac{n}{C\log n} \text{ but } \min_{0\leq i\leq\tau_{j-1}} r_i(x) \leq \frac{1}{n^{100}}\right) \leq \frac{\text{Const}}{n^{99}}.$$

Hence if $\bar{m}_j$ is the first time m after τ_{j-1} such that $r_m(x) \geq \delta_0$, then there is a large constant c_8 such that

$$\mathbb{P}_\ell\left(\max_{1\leq l<j}(\tau_l - \tau_{l-1}) < \frac{n}{C\log n} \text{ but } \bar{m}_j - \tau_{j-1} > c_8 \log n\right) \leq \frac{\text{Const}}{n^{99}}.$$

(Here we were using that $r_{\tau_{j-1}}(x) > \frac{1}{n^{100}}$.) On the other hand

$$\mathbb{P}_\ell\left(\bar{m}_j - \tau_{j-1} \leq c_8 \log n \text{ but } \tau_j - \bar{m}_j > \frac{n}{C\log n} - c_8\log n\right) \leq \frac{\text{Const}(\log n)^{3/2}}{\sqrt{n}}$$

by Lemma 10. The result follows. □

Lemma 13.

$$(a) \quad \mathbb{P}_\ell(\exists i \leq \bar{n} : r_{\tau_i}(x) \leq \delta) \leq c_9 \delta^{1/3} |\log \delta|^\alpha.$$
$$(b) \quad \mathbb{E}_\ell(\text{Card}(i \leq \bar{n} : r_{\tau_i}(x) \leq \delta)) \leq c_{10} \delta^{1/3} |\log \delta|^{\bar{\alpha}}$$

where $\bar{\alpha} = \alpha + 1$.

Proof. (a) Let β be a parameter to be chosen later. We have

$$\mathbb{P}_\ell(\exists i \leq \bar{n} : r_{\tau_i}(x) \leq \delta) \leq \mathbb{P}_\ell(\tau_{\bar{n}} > \delta^{-\beta}) + \mathbb{P}_\ell(\tau_{\bar{n}} \leq \delta^{-\beta} \text{ but } \exists m \leq \tau_{\bar{n}} : r_m(x) \leq \delta)$$
$$\leq \text{Const}\left[\delta^{\beta/2}|\log\delta|^\alpha + \delta^{1-\beta}\right]$$

where the first term is estimated by Lemma 12 and the second term is estimated by the Growth Lemma. Choose $\beta = 2/3$. This prove (a).

Now observe that by Lemma 11 it follows that

$$\mathbb{E}_\ell(\bar{n}1\!\!1_\Omega) \le \mathrm{Const} q |\log q|$$

for any set Ω such that $\mathbb{P}_\ell(\Omega) \le q$. Hence (b) follows from (a). □

We now prove Theorem 6(b). By Lemma 11 we can assume that $[\ell] \in S$ and that $\mathrm{length}(\ell) > \delta_0$. Let $0 = \bar{n}_0, \bar{n}_1, \bar{n}_2 \dots \bar{n}_k \dots$ be consequtive numbers such that $r_{\tau_{\bar{n}_k}}(x) \ge \delta_0$.

Using Lemma 3(b) it follows by induction that

$$\mathbb{P}_\ell(\max_{j \le k} \tau_{\bar{n}_j} - \tau_{\bar{n}_{j-1}} \le n) \le \left(1 - c_{11}/\sqrt{n}\right)^k.$$

Observe that if $m(n)$ is the first number such that $\tau_{\bar{n}_m} - \tau_{\bar{n}_{m-1}} \ge n$ then $V_n - V_n^{\delta_0} \le m$

In particular there is a constant c_{12} such that

$$\mathbb{P}_\ell(V_n - V_n^{\delta_0} \ge c_{12}\sqrt{n}) \le \frac{1}{2}.$$

On the other hand if $X_j = \mathrm{Card}\{\bar{n}_{j-1} < i \le \bar{n}_j,\ r_{\tau_i} < \delta\}$, then by Lemma 13

$$\mathbb{E}_\ell\left(\sum_{j=1}^{c_{12}\sqrt{n}} X_j\right) \le c_{13}\sqrt{n}\delta^{1/3}|\log\delta|^{\bar{\alpha}}.$$

Next let

$$\phi_n(\delta) = \max_{[\ell]\in S, \mathrm{length}(\ell) \ge \delta_0} \mathbb{E}_\ell(V_n^\delta).$$

Then the last two inequalities imply that

$$\phi_n(\delta) \le c_{13}\sqrt{n}\delta^{1/3}|\log\delta|^{\bar{\alpha}} + \frac{1}{2}\phi_n(\delta).$$

The result follows. □

APPENDIX D. SPECTRAL GAP FOR THE LARGE STRIP.

Proof of Theorem 7. Our proof is a generalization of the proof of the exponential mixing for Sinai billiards presented in [ChM 06]. The key technical tool is a so called coupling lemma. Let us present the statement of that result.

Lemma 14. *Given $\delta_0 > 0$ there exist $C > 0, \theta < 1, q > 0$ and $n \ge 1$ such that for any pair of standard pairs $\ell_1 = (\gamma_1, \rho_1), \ell_2 = (\gamma_2, \rho_2)$ supported on the same scatterer and such that* $\mathrm{length}(\ell_j) \ge \delta_0$, *there exist probability*

measures ν_1 and ν_2 and a constant $c \geq q$, and there exist families of standard pairs $\{\ell_{\beta j}\}_\beta$ and of positive constants $\{c_{\beta j}\}_\beta :\ j = 1, 2$, satisfying
(i)

$$\mathbb{E}_{\ell_j}(A \circ f^n) = c\nu_j(A) + \sum_{\beta j} c_{\beta j} \mathbb{E}_{\ell_{\beta j}}(A) \qquad j = 1, 2$$

with $c \geq q$;
(ii) Let λ denote the Lebesgue measure on [0,1]. There exist a measure preserving map $\pi : (\gamma_1 \times [0,1], \nu_1 \times \lambda) \to (\gamma_2 \times [0,1], \nu_2 \times \lambda)$ and constants $C > 0$ and $\theta < 1$ such that if $\pi(x_1, s_1) = (x_2, s_2)$ then

$$d(f^n x_1, f^n x_2) \leq C\theta^n \tag{31}$$

(iii) For each $\ell_{\beta j}$ we can define functions $n_{\beta j}$ so that

$$\mathbb{E}_{\ell_{\beta j}}(A \circ f^{n_{\beta j}}) = \sum_\alpha c_{\alpha\beta j} \mathbb{E}_{\ell_{\alpha\beta j}}(A),$$

where $\text{length}(\ell_{\alpha\beta j}) \geq \delta_0$ *and for each $m > 0$*

$$\sum_\beta c_{\beta j} \mathbb{P}_{\ell_{\beta j}}(n_{\beta j} \geq m)) \leq C\theta^m \qquad j = 1, 2.$$

In [DSzV 07] this lemma was formulated with (iii) replaced by

$$\widetilde{(iii)} \qquad \sum_{\beta : \text{length}(\ell_{\beta j}) \leq \rho} c_{\beta j} \leq \text{Const}(\delta_0)\rho.$$

For the Poincare map (iii) follows from $\widetilde{(iii)}$ and the Growth Lemma (Proposition 1 (b)). In our case (iii) follows from $\widetilde{(iii)}$ by combining Proposition 1 (b) and Lemma 11.

As in [ChM 06] we can deduce the exponential mixing by a repeated application of Lemma 14. More precisely we have the following statement.

Lemma 15. *Suppose that there are constants $\hat{\theta}$, n_L, c_L such that for any standard pairs ℓ_1 and ℓ_2 with* $\text{length}(\ell_j) \geq \delta_0$ *we have*

$$\mathbb{E}_{\ell_j}(A \circ f^{n_L}) = \sum_\alpha c_\alpha \mathbb{E}_{\ell_{\alpha j}}(A) + \sum_\beta c_{\beta j} \mathbb{E}_{\ell_{\beta j}}(A)$$

where $\sum_\alpha c_\alpha \geq c_L$ and $(\ell_{\alpha_1}, \ell_{\alpha 2})$ satisfy the conditions of Lemma 14, then for any $A \in \mathcal{H}$ such that $\nu_{[L]}(A) = 0$, for any standard pair ℓ and for any $n \geq C|\log \text{length}(\ell)|$ we have

$$|\mathbb{E}_\ell(A \circ T^n_{[L]})| \leq C(A) \left[\left(1 - c\frac{c_L}{n_L}\right)^n + \hat{\theta}^n \right].$$

Our goal is to verify the conditions of Lemma 15 with $n_L = c_1 L^{p_1}$, $c_L = c_2 L^{-p_2}$ (we do not pursue the optimal values of p_js).

Let ϵ be a small constant. We claim that the conditions of Lemma 15 are verified if γ_1, γ_2 belong to $\{|q - L/2| \leq \epsilon L\}$ with $\tilde{n}_L = \text{Const}(\epsilon L)^2$, $\tilde{c}_L = \bar{c}$ (the tildes mean that this values are only valid not for all curve but only for curves close to the middle of $M_{[L]}$). Indeed in the case of the non-modified Lorentz process this is proven in [DSzV 07]. However if ϵ is sufficiently small then we can make

$$\mathbb{P}_{\ell_j}(q_k \text{ visits the modified part before Const}(\epsilon L)^2)$$

as small as we wish due to Proposition 2(d). In particular we can make this probability smaller than $\bar{c}/4$ where $\bar{c}$ is the corresponding constant for the non-modified Lorentz process. This implies our claim.

Next we prove that there are constants c_3, c_4 such that for any standard pairs ℓ_1, ℓ_2 with $\text{length}(\ell_j) \geq \delta_0$ we have

$$\mathbb{P}_{\ell_j}(|q_{c_3 L^2} - L/2| \leq \epsilon L, r_{c_3 L^2}(x) \geq \delta_0) \geq \frac{c_4}{L}.$$

This is achieved in three steps. Let $\tilde{\tau}$ be the first time when

$$|q_{\tilde{\tau}} - L/2| \leq \frac{\epsilon L}{3}$$

then by Lemma 3(b) there exists c_5 such that

$$\mathbb{P}_{\ell_j}(\tilde{\tau} \leq c_3 L^2 \text{ and } q_k \text{ does not visit the modification for } k = 1 \dots \tilde{\tau}) \leq c_5/L.$$

Observe that $\tilde{\tau}$ can be significantly less than $c_3 L^2$ but by Proposition 2(c) there exists a constant c_6 such that

$$\mathbb{P}_{\ell_j}(|q_k - L/2| \leq \frac{2\epsilon L}{3} \text{ for } k = \tilde{\tau} \dots c_3 L^2) \geq c_6.$$

Finally we claim that there is a constant c_7 such that

$$\mathbb{P}_{\ell_j}(|q_k - L/2| \leq \epsilon L \text{ for } k = \tilde{\tau} \dots c_3 L^2 \text{ and } r_{c_3 L^2}(x) \geq \delta_0) \geq \mathbb{P}_{\ell_j}(|q_k - L/2| \leq \frac{2\epsilon L}{3} \text{ for } k = \tilde{\tau} \dots c_3 L^2)(1 - c_7 \delta_0). \quad (32)$$

Indeed if $|q_{c_3 L^2 - \sqrt{L} - L/2}| \leq 2\epsilon L/3$ then $|q_{c_3 L^2} - L/2| \leq \epsilon L$ so the result follows from the Growth Lemma (Proposition 1 (b)). Now part (a) of Theorem 7 follows from Lemma 15.

Part (b) of Theorem 7 follows from Theorem 6(b) while part (c) can be deduced from the exponential mixing by the method of [ChD 07], Section A4. □

REFERENCES

[B 68] P. Billingsley. *Convergence of Probability Measures*, pp. 253, Wiley, 1968.

[BChS 91] L. A. Bunimovich, N. I. Chernov and Ya. G. Sinai *Statistical properties of two dimensional dispersing billiards,* Russian Math. Surveys **46** 47–106, 1991.

[BS 81] L. A. Bunimovich and Ya. G. Sinai *Statistical properties of Lorentz gas with periodic configuration of scatterers,* Comm. Math. Phys. **78** 479–497, 1981.

[Ch 99] N. I. Chernov. *Decay of correlations and dispersing billiards,* Journal of Statistical Physics **94** 513–556, 1999.

[Ch 01] N. I. Chernov. *Sinai billiards under small external forces,* Ann. Henri Poincare **2** (2001) 197–236.

[Ch 06] N. I. Chernov. *Advanced statistical properties of dispersing billiards,* J. Stat. Phys. **122** 1061–1094, 2006.

[ChD 07] N. Chernov and D. Dolgopyat *Brownian Brownian Motion–1,* to appear in Memoirs AMS.

[ChD 08] N. Chernov and D. Dolgopyat *Galton board: limit theorems and recurrence,* preprint.

[ChM 06] N. Chernov and R. Markarian *Chaotic billiards,* Mathematical Surveys and Monographs, AMS, RI.

[DSzV 07] D. Dolgopyat, D. Szász and T. Varjú. *Recurrence properties of planar Lorentz process,* to appear in Duke Math J.

[HSh 81] J. M. Harrison and L. A. Shepp, *On skew Brownian motion,* Ann. Probab. **9** 309–313, 1981.

[L 06] M. Lenci *Typicality of recurrence for Lorentz processes*, Ergodic Theory Dynam. Systems **26** (2006) 799–820.

[M 04] R. Markarian *Billiards with polynomial decay of correlations,* Ergodic Theory Dynam. Systems **24** (2004) 177–197.

[MN 07] I. Melbourne and M. Nicol. *A vector-valued almost sure invariance principle for hyperbolic dynamical systems,* preprint.

[S 70] Sinai, Ya. G. Dynamical systems with elastic reflections: Ergodic properties of dispersing billiards, Russ. Math. Surv. **25** (1970), 137–189.

[SV 71] D. W. Stroock and S. R. S. Varadhan, *Diffusion processes with boundary conditions,* Comm. Pure Appl. Math. **24** 147–225, 1971.

[SV 06] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes,* Classics in Mathematics, Springer-Verlag, Berlin, 2006.

[SzT 81] D. Szász and A. Telcs. *Random walk in an inhomogeneous medium with local impurities*. J. of Statistical Physics. **26** 527–537, 1981.

[SzV 04] D. Szász and T. Varjú. *Local limit theorem for the Lorentz process and its recurrence in the plane* Ergodic Theory and Dynamical Systems **24** 257–278, 2004.

[Y 98] Young L.–S. *Statistical properties of dynamical systems with some hyperbolicity*, Ann. Math. **147** (1998) 585–650.

[Y 99] Young L.–S. *Recurrence times and rates of mixing,* Israel J. Math. **110** (1999) 153–188.

DMITRY DOLGOPYAT: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA,,
DOMOKOS SZÁSZ, TAMÁS VARJÚ: BUDAPEST UNIVERSITY OF TECHNOLOGY, MATHEMATICAL INSTITUTE, BUDAPEST, EGRY J. U. 1 HUNGARY H-1111,
E-mail address: dmitry@math.umd.edu, szasz@math.bme.hu, kanya@math.bme.hu