GENERICITY ON CURVES AND APPLICATIONS: PSEUDO-INTEGRABLE BILLIARDS, EATON LENSES AND GAP DISTRIBUTIONS

KRZYSZTOF FRĄCZEK, RONGGANG SHI, AND CORINNA ULCIGRAI

Abstract. In this paper we prove results on Birkhoff and Oseledets genericity along certain curves in the space of affine lattices and in moduli spaces of translation surfaces. We also prove applications of these results to dynamical billiards, mathematical physics and number theory. In the space of affine lattices $ASL_2(\mathbb{R})/ASL_2(\mathbb{Z})$, we prove that almost every point on a curve with some non-degeneracy assumptions is Birkhoff generic for the geodesic flow. This implies almost everywhere genericity for some curves in the locus of branched covers of the torus inside the stratum $\mathcal{H}(1,1)$ of translation surfaces. For these curves (and more in general curves which are well-approximated by horocycle arcs and satisfy almost everywhere Birkhoff genericity) we also prove that almost every point is Oseledets generic for the Kontsevitch-Zorich cocycle, generalizing a recent result by Chaika and Eskin. As applications, we first consider a class of pseudo-integrable billiards, billiards in ellipses with barriers, which was recently explored by Dragović and Radnović, and prove that for almost every parameter, the billiard flow is uniquely ergodic within the region of phase space in which it is trapped. We then consider any periodic array of Eaton retroreflector lenses, placed on vertices of a lattice, and prove that in almost every direction light rays are each confined to a band of finite width. This generalizes a phenomenon recently discovered by Frączek and Schmoll which could so far only be proved for random periodic configurations. Finally, a result on the gap distribution of fractional parts of the sequence of square roots of positive integers, which extends previous work by Elkies and McMullen, is also obtained.

1. Introduction

In this paper we prove three quite different results which answer recent open questions in dynamical systems, mathematical physics and number theory. These three results all turn out to be applications of two more results on genericity in homogeneous and Teichmüller dynamics, which constitute the heart of this paper. The three applications, which are explained in the following sections of the introduction, concern more precisely the chaotic properties (specifically ergodicity) of a recently discovered class of *pseudo-integrable billiards* (see § 1.1), the behaviour of *light rays in periodic arrays of Eaton lenses* (see § 1.2) and the *gap distribution* of the sequence of *fractional parts of square roots of positive integers* (see § 1.3). The common result they exploit, which is stated in § 2.1, concerns Birkhoff genericity (under the geodesic flow) for almost every point on certain curves in the space of affine lattices and in the moduli space of certain translation surfaces. Furthermore, for the application to Eaton lenses, we also need a result on Oseledets genericity (for the Kontsevich-Zorich cocycle) for almost every parameter describing certain

Date: August 18, 2015.
2000 *Mathematics Subject Classification.* Primary 37A17; Secondary 37A40, 37J35, 11K38.
Key words and phrases. homogeneous dynamics, translation surfaces, Kontsevich-Zorich cocycle, equidistribution, ergodic theorem, Oseledets theorem, Eaton lenses, gap distribution, pseudo-integrable billiards.

curves in the moduli space of translation surfaces, which is stated in § 2.3. The final section § 1.4 of this introduction provides an outline of how the three main applications are related to these two genericity results and describes the structure of the rest of the paper.

1.1. **Pseudo-integrable billiards in ellipses.** In this section we answer an open question on the ergodic properties of billiards in ellipses with barriers, a new class of pseudo-integrable billiards recently described by Dragović and Radnović in [9].

A (planar) *mathematical billiard* is a dynamical system in which a point-mass moves inside a *billiard table* $T \subset \mathbb{R}^2$, i.e. a bounded domain $T \subset \mathbb{R}^2$ whose boundary ∂T consists of finite number of smooth curves. A *billiard trajectory* is the trajectory described by the point-mass which moves freely inside the table on segments of straight lines and undergoes elastic collisions (angle of incidence equals to the angle of reflection) when it hits the boundary of the table. The billiard flow $\{b_t\}_{t\in\mathbb{R}}$ is defined on a subset of the phase space S^1T that consists of the points $(x, v) \in T \times S^1$, where $S^1 = \{v \in \mathbb{C} : |v| = 1\}$, such that if x belongs to the boundary of T then v is an inward unit tangent vector. For $t \in \mathbb{R}$ and (x, v) in the domain of $\{b_t\}_{t\in\mathbb{R}}$, b_t maps (x, v) to $b_t(x, v) = (x_t, v_t)$, where x_t is the point reached after time t by flowing at unit speed along the billiard trajectory starting at x in direction of the unit vector v and v_t is the unit tangent vector to the trajectory at x_t.

Billiards are sometimes divided into *convex*, *chaotic* and *polygonal* billiards (see for example the survey by Tabachnikov [35]). The billiard in T is called *integrable* if an open subset of S^1T is filled by invariant sets so that the billiard flow restricted to each such set is isomorphic to a linear flow on a two-dimensional torus (only convex or polygonal billiards can be integrable). The billiard system inside any ellipse is integrable. Each invariant set is determined by a confocal ellipse or hyperbola (called a *caustic*) and consists of all trajectories tangent the caustic. In *chaotic billiards*, such as the famous Sinai billiard (a square with a convex scatterer), no such invariant sets exist and the billiard flow exhibits strong chaotic properties and in particular is ergodic on the whole phase space (with respect to the billiard invariant measure, see [35]), i.e. there are no billiard flow invariant sets of positive measure. An intermediate behaviour is exhibited by *rational polygonal billiards* (when T is a polygon whose angles are rational multiples of π), sometimes referred to as *pseudo-integrable* billiards: for every direction $v \in S^1$ the billiard flow in direction v on S^1T is confined to an invariant surface in the phase space and the billiard flow restricted to this invariant surface is typically ergodic, but in contrast with integrable billiards, the invariant surface is not a torus but has higher genus. The study of rational polygonal billiards is intimately connected to the rich area of research in translation surfaces and Teichmüller dynamics.

Recently, Dragović and Radnović discovered a new class of pseudo-integrable billiards, see [9], given by *billiards in ellipses with barriers*, that we now describe. Let $0 < b < a$. Denote by $\{\mathcal{C}_\lambda : 0 < \lambda < a\}$ the family of confocal ellipses (for $\lambda \le b$ with $\lambda = b$ the set of two focal points) and hyperbolas (for $b < \lambda < a$)

$$\frac{x^2}{a-\lambda} + \frac{y^2}{b-\lambda} = 1.$$

Let us consider the billiard flow inside the ellipse $\mathcal{C}_0$ with one linear vertical obstacle of length $\sqrt{b} - \sqrt{b-\lambda_0}$, $0 < \lambda_0 < b$, which is positioned as shown in Figure 1. This billiard table is denoted by $\mathcal{D}_{\lambda_0}$.

Dragović and Radnović observed in [9] that the phase space of the billiard flow on $\mathcal{D}_{\lambda_0}$ splits into invariant subsets $\mathcal{S}_\lambda$, $0 < \lambda < a$ so that the ellipse $\mathcal{C}_\lambda$ for $0 < \lambda \le b$ or the hyperbola $\mathcal{C}_\lambda$ for $b < \lambda < a$ is a caustic of all billiard trajectories in $\mathcal{S}_\lambda$ (see Figure 1). Thus, the ellipse with a barrier display the typical trapping phenomenon

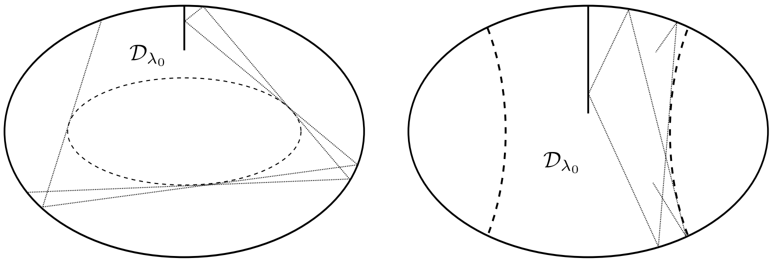

FIGURE 1.

of a pseudo-integrable billiard. In [9] they construct examples of regions where the flow is minimal but not uniquely ergodic, i.e. every trajectory is dense but not every trajectory is uniformly distributed. It is natural to expect that this behaviour is exceptional.

We answer affirmatively the natural conjecture (raised e.g. by Zorich) that for typical parameters $\lambda \in (0, a)$ of the caustic parameterizing the invariant region $\mathcal{S}_\lambda$ billiard trajectories are typically dense in the invariant subset and furthermore, the billiard flow restricted to $\mathcal{S}_\lambda$ is uniquely ergodic. Let us recall that a flow is *uniquely ergodic* if it admits a unique invariant probability measure, in which case it is also automatically ergodic with respect to this measure.

Theorem 1.1. *For almost every $\lambda \in (0, a)$ the billiard flow on $\mathcal{D}_{\lambda_0}$ restricted to $\mathcal{S}_\lambda$ is uniquely ergodic.*

This implies in particular that for typical parameters $\lambda \in (0, a)$, every billiard trajectory inside $\mathcal{S}_\lambda$ is dense and uniformly distributed. This result is proved in § 3 and will be deduced by the Birkhoff genericity results in § 2.1.

1.2. **Periodic systems of Eaton lenses.** In this section we study the behavior of light trajectories in a plane on which a lattice system of round retroreflector lenses of the same size (called *Eaton lenses*) is arranged. An *Eaton lens* is a circular lens which acts as a perfect retroreflector, i.e. so that each ray of light after passing through the Eaton lens is directed back toward its source, see Figure 2. Let $R > 0$ denote the radius of the lens. The refractive index (RI for short) in an Eaton lens varies from 1 (outside the lens) to infinity (at the center of the lens, where it is not defined) according to the formula $RI = \sqrt{2R/r - 1}$, where $0 < r \leq R$. As it was observed in [24], a light ray entering the Eaton lens at a point x_e moves (inside the lens) in an elliptic orbit whose focal point coincides with the center of the lens and then it leaves the lens at a point x_l so that the points x_e, x_l are the ends of the minor axis of the ellipse. Therefore, the direction of the light ray is reversed after passing through the lens. There is only one exception when the light ray hits the center of the lens and disappears. We adopt the convention that when the light ray hits the center, it turns back at the center and continues its motion backwards.

Denote by $L(\Lambda, R)$ the system of identical Eaton lenses of radius $R > 0$ arranged on the plane $\mathbb{R}^2$ so that their centers are placed at the points of a unimodular lattice $\Lambda \subset \mathbb{R}^2$, see Figure 3. We deal only with pairs (Λ, R) for which the lenses are pairwise disjoint and such pairs are called *admissible*. Admissibility is equivalent to $R < s(\Lambda)/2$, where $s(\Lambda)$ is the length of the shortest non-zero vector in Λ.

Frączek and Schmoll in [20] first observed the phenomenon that light rays in lattice systems of Eaton lenses are often trapped inside bands of finite width. Note that for every light ray there is a direction $\theta \in [0, 2\pi]$ such that the light ray flows in direction θ or $-\theta$ outside the lenses. Then θ is called the direction of the light ray. This direction is unique modulo π. Let us say that a direction θ is *trapped* if there

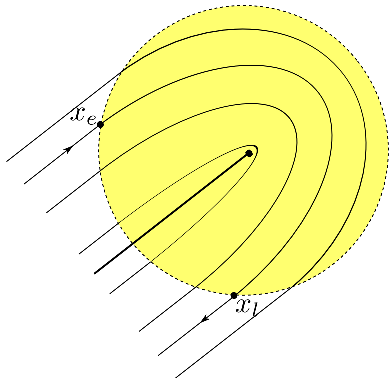

FIGURE 2. Eaton lens and a parallel family of light rays

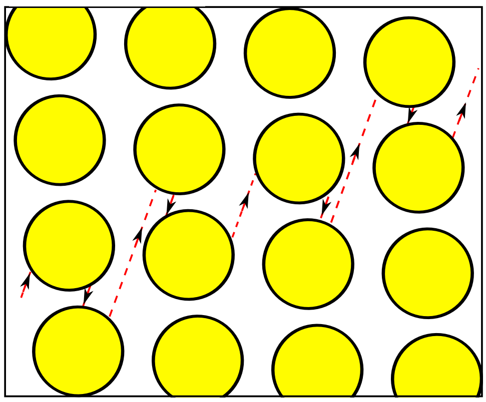

FIGURE 3. The system of lenses $L(\Lambda, R)$

exists $C(\theta) > 0$ and $v(\theta) \in S^1$ such that every light ray on $L(\Lambda, R)$ in direction θ is trapped in an infinite band of width $C(\theta) > 0$ parallel to the unit tangent vector $v(\theta)$.

Frączek and Schmoll considered *randomly chosen lattices* and proved that for every $0 < R < \sqrt{2\sqrt{3}}$ and for almost every R-admissible lattice Λ (with respect to the Haar measure on the space of lattices), light rays in the vertical direction are trapped. They also provided explicit examples of specific lattices and directions which are trapped. Their result, though, does not provide any information for the behaviour of typical light rays in a *fixed* admissible lattice configuration.

In this paper we answer this natural question (asked for example by Marklof and by the referee of [20]) by describing the behavior of light orbits on $L(\Lambda, R)$ in direction θ when an admissible pair (Λ, R) is fixed and the parameter θ varies.

Theorem 1.2. *Let (Λ, R) be an admissible pair. Then a.e. $\theta \in [0, 2\pi]$ is trapped.*

This result is proved in § 4. As in [20], we first reduce the system of Eaton lenses to a simpler model, a *system of flat lenses*, which can be unfolded and reduced to an infinite translation surface. For the definition and for more results for this related system, we refer the reader to § 4.

1.3. **Gap distribution of fractional parts of square roots.** Let $\mathbb{N} = \{1, 2, \ldots\}$. Consider a sequence $\{t_n\}_{n\in\mathbb{N}} \subset [0, 1]$ which is *equidistributed modulo one*, i.e. for any $0 \le a \le b \le 1$, $\#\{1 \le n \le N : t_n \in [a, b]\}/N$ tends to $|b-a|$ as $N \to \infty$. Given an initial block $\{t_1, \ldots, t_N\}$, the lengths of its complementary intervals, i.e. the connected components of $[0,1] \setminus \{t_1, \ldots, t_N\}$, are known as *gaps*. A natural question is to study the *gap distribution*, i.e. the limiting behavior of the sequence of gaps of $\{t_1, \ldots, t_N\}$, renormalized by their average length $1/N$, as $N \to \infty$ (see below for precise statements). In a celebrated paper, Elkies and McMullen considered the sequence $\{\sqrt{n} \mod 1\}_{n\in\mathbb{N}}$ of *fractional parts of square roots* and, establishing a connection with homogeneous dynamics, showed that the gap distribution of $\{\sqrt{n} \mod 1\}_{n\in\mathbb{N}}$ exists and is a non-standard distribution on $\mathbb{R}_{\geq 0}$ (in contrast to the Poissonian gap distribution displayed by random sequences).

More precisely, fix a real number $r \geq 1$ and let $0 = t_1 \leq t_2 \leq \ldots, \leq t_{\lfloor r \rfloor} < 1$ be increasingly ordered fractional parts of $\{\sqrt{1}, \sqrt{2}, \ldots, \sqrt{\lfloor r \rfloor}\}$, where $\lfloor r \rfloor$ is the smallest integer less than or equal to r. We set $t_{\lfloor r \rfloor + 1} = 1$ for convenience. The gap distribution of square roots of natural numbers describes the limit behaviors as $r \to \infty$ of renormalized consecutive gaps $\{(t_{n+1} - t_n)\lfloor r \rfloor\}_{n \leq \lfloor r \rfloor}$. Elkies and McMullen proved in [11] that there is a continuous probability density $F : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ of the form

$$F(s) = \begin{cases} 6/\pi^2 & t \in [0, 1/2], \\ F_2(t) & t \in [1/2, 2], \\ F_3(t) & t \in [2, \infty), \end{cases} \tag{1.1}$$

where F_2 and F_3 are explicit real analytic functions (we refer the reader to [11, Theorem 3.14] for explicit formulas) such that for every interval $[a, b] \subset \mathbb{R}_{\geq 0}$

$$\frac{1}{\lfloor r \rfloor} \#\{1 \leq n \leq \lfloor r \rfloor : (t_{n+1} - t_n)\lfloor r \rfloor \in [a, b]\} \to \int_a^b F(s)\, ds \qquad \text{as } r \to \infty.$$

An effective version of this result was recently obtained by Browning and Vinogradov [5].

We consider the distribution of normalized gaps containing a fixed $s \in [0, 1]$, which was suggested by Marklof. Let $k(r, s)$ be the largest positive integer k satisfying $t_k \leq s$. We define

$$L_r(s) = \lfloor r \rfloor (t_{k(r,s)+1} - t_{k(r,s)}). \tag{1.2}$$

For every $s \in [0, 1]$ which is not a fractional part of the square root of a natural number, $L_r(s)$ is the normalized length of the gap of the fractional parts of $\{\sqrt{1}, \ldots, \sqrt{\lfloor r \rfloor}\}$ which contains s. We are interested in the limit distribution of the sequence $\{L_{r_n}(s)\}_{n\in\mathbb{N}}$ of the normalized gaps containing s along geometric progressions $\{r_n = cq^n\}_{n\in\mathbb{N}}$ where $q > 1$ and $c \geq 1$. The reason we let r_n goes geometrically but not linearly is because from r to $r + 1$ we only add one number so at most one gap can change.

The main result of this part is the following.

Theorem 1.3. *Let $\{cq^n\}_{n\in\mathbb{N}}$ be any geometric progression with $c \geq 1$ and $q > 1$. Then for Lebesgue almost every $s \in [0, 1]$ the sequence $\{L_{cq^n}(s)\}_{n\in\mathbb{N}}$ converges in distribution to the probability measure $tF(t)\, dt$ on $\mathbb{R}_{\geq 0}$, i.e. for every interval $[a, b] \subset \mathbb{R}_{\geq 0}$*

$$\frac{1}{N} \#\{1 \leq n \leq N : L_{cq^n}(s) \in [a, b]\} \to \int_a^b tF(t)\, dt \qquad \textit{as } N \to \infty. \tag{1.3}$$

Theorem 1.3 raises a natural question, namely whether, for almost every $s \in [0, 1]$, (1.3) still hold if we replace the sequence $\{cq^n\}_{n\in\mathbb{N}}$ by $\{r_n(s)\}_{n\in\mathbb{N}}$ where $r_n(s)$ is the natural number such that the gap containing s is changing for the n-th time.

1.4. **Outline and structure of the paper.** We now give a brief explanation of how all these three applications rely on the same results (formulated in the next section § 2). At the end of this section we then explain the structure of the rest of the paper. It turns out that all three problems are related to *genericity along curves* in the space of affine lattices as follows.

In the paper [9] where Dragović and Radnović introduce *billiards in ellipses with barriers* as a new class of pseudo-integrable billiards, a key observation is that any given billiard trajectory in an ellipse with a barrier can be mapped, by a suitable change of coordinates, to a trajectory of a billiard in a rectangle with a barrier (see Proposition 3.1 in § 3). Billiards in *rational polygonal tables* (such as the rectangle with a barrier) have been successfully studied in the past decades through their connection with *Teichmüller dynamics.* A classical construction allows us to *unfold* the billiard to a translation surface (a surface with an almost everywhere Euclidean metric, see § 2.2) so that the billiard trajectory becomes a trajectory of the linear flow (i.e. a flat geodesic) on the surface and in virtue of a milestone result in Teichmüller dynamics, Masur's criterium (see Theorem 2.3 in § 2.2), unique ergodicity follows if one can show that the corresponding translation surface is *Birkhoff generic* for the Teichmüller flow (see § 2.1 for the definition).

Let us remark that a celebrated result by Kerkhoff, Masur and Smillie [26] from the '80s guarantees that in any rational billiard, the billiard flow in almost every direction is uniquely ergodic. Unfortunately, this result does not yield any information about trajectories in pseudo-integrable billiards. Indeed, as one changes the direction of the trajectory considered in the elliptical billiard (and hence its *caustic*), the *parameters* of the corresponding rectangular billiard table (such as lengths of sides and barrier) change too (while the slope of the image trajectories by the change of coordinates is fixed and equal to ± 1) and describe a one-parameter family, or *curve*, in the moduli space of translation surfaces. Finding suitable conditions on a curve of translation surfaces so that *almost every point on the curve* is Birkhoff generic is currently a widely open problem. In our setup, fortunately, the translation surfaces which are obtained by unfolding are all double covers of flat tori and thus the problem reduces to a homogeneous dynamics setup. The analogous result on genericity for almost every point on a curve in the space of affine lattices which we need is the first of our main results presented in the next section, see Theorem 2.1 and Corollary 2.2 in § 2.1.

One can reduce also the study of *systems of Eaton lenses* to the set up of translation surfaces. For a fixed direction of light rays, by replacing each Eaton lens by a flat lens (defined in § 4.1) and then taking a double cover, one can indeed reduce the behaviour of a light ray in the array to the study of a linear trajectory on an *infinite* translation surface (see § 4.1). Since the planar array of Eaton lenses is $\mathbb{Z}^2$ periodic, the infinite surface obtained is a periodic surface which is a $\mathbb{Z}^2$ cover of a genus two surface. The global behaviour of trajectories in the a $\mathbb{Z}^2$ cover turns out to be intimately related to Lyapunov exponents of the Teichmüller flow. In particular, as shown by Frączek and Schmoll in [20] (see also § 4.2), directions of bands which trap light rays are directions which correspond to negative Lyapunov exponents (in the plane spanned by the two homology classes which determine the cover). Thus, to establish that a direction is trapped one needs to prove that the corresponding genus two translation surface is *Oseledets generic* (see § 2.3). The main result by Frączek and Schmoll in [20] (mentioned in § 1.2, see also § 4.1), which deals with *random lattices*, follows from the standard Oseledets ergodic theorem (which is recalled in §2.3). To understand the behaviour for a *fixed lattice* and in particular to prove Theorem 1.2 one needs to know that almost every point on the curve of genus two translation surfaces obtained by the above reduction is

Birkhoff and Oseledets generic. *Birkhoff genericity* reduces as before to Birkhoff genericity for curves in the space of affine lattices, since the genus two surface turns out to be also in this case a double cover of a flat torus with a marked point. *Oseledets genericity* along the curve is based on our second main result, Theorem 2.5 in section 2.3. Let us remark that our two main results generalize in a special set up a recent work by Chaika and Eskin in [7], where they prove Birkhoff and Oseledets genericity for curves of translation surfaces which are *circles* (see § 2.3 for the precise formulation).

The seminal paper [11] by Elkies and McMullen on the *gap distribution of fractional parts of square roots* was the first to describe and to exploit the connection of this problem with homogeneous dynamics. In their paper it is shown that existence of the gap distribution follows from a result on equidistribution of certain curve in the space of affine lattices under the geodesic flow. Exploiting the same arguments, one can see in § 5 that our Theorem 1.3 can be derived from our result on Birkhoff genericity under the geodesic flow for the same curve (Theorem 2.1).

Structure of the paper. The rest of the paper is organized as follows. In the next section § 2 we recall background material and formulate the main results on which the applications are based. In § 2.1, we state Theorem 2.1 and Corollary 2.2 on Birkhoff genericity in the space of affine lattices, from which we derive in §2.4 the Birkhoff genericity result on translation surfaces (Theorem 2.10) used in the applications. We recall background material on translation surfaces and Teichmüller dynamics in §2.2 and §2.3. Theorem 2.5 on Oseledets genericity is stated in §2.3. Using the results in §2 we prove the three main applications in §3 (Theorem 1.1), §4 (Theorem 1.2) and §5 (Theorem 1.3). Section §6 is devoted to the proof of Theorem 2.1 on Birkhoff genericity, while section §7 contains the proof of Theorem 2.5 on Oseledets genericity. We refer the reader to the beginnings of sections §2, §6 and §7 for a more detailed outline of the content for each of these sections.

2. Results on genericity along curves

In this section we formulate the results on which the applications described in the introduction are based. Two of the main theorems in ergodic theory, the Birkhoff ergodic theorem (recalled at the beginning of section 2.1) and the Oseledets multiplicative ergodic theorem (see section 2.3), guarantee that given an ergodic measure preserving dynamical system on a probability space (Y, μ), almost every point in Y with respect to the measure μ is generic, in the sense that either the conclusion of the Birkhoff holds for $f \in L^1(Y)$ or the Oseledets theorem holds for a cocycle under suitable assumptions. If one considers a *curve* in the space Y, which has zero measure, a priori the conclusion of both Birkhoff and Oseledets theorems could fail for points in the curve. The main results in this paper concern two situations in which one can prove that almost every point along a certain curve is generic. The first result concerns curves in the space of affine lattices satisfying a non-degeneracy condition and is stated in § 2.1 after recalling Birkhoff ergodic theorem and the definition of the space of affine lattices. To state the second result, which concerns Oseledets genericity, we first recall background material on translation surfaces in § 2.2, and recall the definition of the Kontsevich-Zorich cocycle, a linear cocycle over the Teichmüller geodesic flow on the space of translation surfaces, which plays a fundamental role in Teichmüller dynamics. The main result stated in § 2.3 states that almost every point on certain curves in the moduli space of translation surfaces is Oseledets generic for the Kontsevich-Zorich cocycle. Finally, in § 2.4 we first introduce the space of translation surfaces of a special form, i.e. double covers of flat tori, which plays a key role in applications and describe its strict connection with the space of affine lattices. We then deduce from the Birkhoff genericity result in the space of affine lattices a result on unique ergodicity for curves of translation surfaces which are double covers of tori, which is then directly used in the applications.

Notation. We define here some notation which is used throughout the paper. The ring of $d \times d$ matrices $M_d(\mathbb{R})$ acts on $\mathbb{R}^d$ via linear transformations and this action will appear in the paper as hv for any $h \in M_d(\mathbb{R})$ and $v \in \mathbb{R}^d$. Let $\|\cdot\|$ denote the Euclidean norm on $\mathbb{R}^d$ as well as operator norm of matrices defined as

$$\|h\| = \sup_{v \in \mathbb{R}^d \setminus \{0\}} \frac{\|hv\|}{\|v\|}. \tag{2.1}$$

The identity matrix in $M_d(\mathbb{R})$ will be denoted by Id. We use $|\cdot|$ to denote the absolute value of real numbers or the Lebesgue measure of subsets of $\mathbb{R}$ according to the context. $\#S$ will denote cardinality of a set S.

2.1. **Birkhoff genericity along curves in $ASL_2(\mathbb{R})/ASL_2(\mathbb{Z})$.** We begin this section by first recalling the statement of Birkhoff ergodic theorem and defining the concept of Bikhoff genericity. Let Y be a locally compact, Hausdorff and second countable topological space. Let $\{\psi_t\}_{t\in\mathbb{R}}$ be a one-parameter topological flow on Y, i.e. $t \mapsto \psi_t$ is a homomorphism from $\mathbb{R}$ to the group of homeomorphisms of Y and $\mathbb{R} \times Y \ni (t, y) \mapsto \psi_t(y) \in Y$ is a continuous map .

Let μ be a $\{\psi_t\}_{t\in\mathbb{R}}$ invariant and ergodic probability measure on Y. The Birkhoff ergodic theorem says that, given a real valued measurable function f on Y with $\int_Y |f|\, d\mu < \infty$, one has

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(\psi_t(y))\, dt = \int_Y f\, d\mu \tag{2.2}$$

for μ almost every $y \in Y$. We say $y \in Y$ is *Birkhoff generic* with respect to (Y, μ, ψ_t) if (2.2) holds for every $f \in C_c(Y)$ where $C_c(Y)$ is the set of continuous compactly supported functions on Y.

We will be interested in Birkhoff genericity along curves in the space of affine lattices, that we now define. Let $SL_2(\mathbb{R})$ ($SL_2(\mathbb{Z})$) be the group of 2 by 2 matrices with real (integer) entries and determinant one. The quotient space $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ parameterizes the moduli space of two dimensional unimodular lattices. A (unimodular) *affine lattice* $\Lambda + v$ is determined by a lattice Λ in $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ and a point $v \in \mathbb{R}^2/\Lambda$ which describes the shift of new origin.

Let $G = ASL_2(\mathbb{R}) := SL_2(\mathbb{R}) \ltimes \mathbb{R}^2$, $\Gamma = ASL_2(\mathbb{Z}) := SL_2(\mathbb{Z}) \ltimes \mathbb{Z}^2$ and $X = G/\Gamma$. We represent elements of G by

$$(h, v) := \begin{pmatrix} h & v \\ 0 & 1 \end{pmatrix} \quad \text{where} \quad h \in SL_2(\mathbb{R}), v \in \mathbb{R}^2.$$

The multiplication for elements of G is given by

$$(h_1, v_1) \cdot (h_2, v_2) = (h_1h_2, h_1v_2 + v_1).$$

One can check that the space X parameterizes affine lattices: if the coset $h(SL_2(\mathbb{Z}))$ in $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ represents the unimodular lattice $\Lambda_h = h\mathbb{Z}^2$, then the coset $(h, v)\Gamma$ in G/Γ represents the unimodular affine lattice $\Lambda_h + v$.

Let us consider the following elements

$$a_t = \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad u(s_1, s_2, s_3) = \begin{pmatrix} 1 & s_1 & s_2 \\ 0 & 1 & s_3 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

The action of the diagonal group $\{a_t : t \in \mathbb{R}\}$ on X by left multiplication is known as *geodesic flow* on the space of affine lattices. It is well known that this action is ergodic with respect to the probability Haar measure μ_X on X.

We will consider curves in X of the form $u_\varphi(s)\Gamma$ where $\varphi : [0,1] \to \mathbb{R}$ is a C^1-function and $u_\varphi(s) := u(s, \varphi(s), 0)$. Let us remark that these curves have μ_X measure zero, hence the Birkhoff theorem does not yield information about genericity for points along these curves.

Theorem 2.1. *Let $\varphi : [0,1] \to \mathbb{R}$ be a C^1-function such that for any rational line $\mathcal{L}$ in $\mathbb{R}^2$ the Lebesgue measure of $\{s \in [0,1] : (s, \varphi(s)) \in \mathcal{L}\}$ is zero. Then for almost every $s \in [0,1]$ the coset $u_\varphi(s)\Gamma \in X$ is Birkhoff generic with respect to (X, μ_X, a_t).*

We can derive from Theorem 2.1 similar results for an arbitrary base point $x \in G/\Gamma'$, where Γ' is a lattice in G commensurable with Γ, i.e. $\Gamma \cap \Gamma'$ has finite index in both Γ and Γ'. We refer the reader to Corollary 6.5 for more details. For the applications, we will need a more general class of curves, for which equidistribution still holds and can be deduced from Theorem 2.1.

Corollary 2.2. *Suppose that $\psi : [0,1] \to G$ is a C^2-curve of the form*

$$\psi(s) = \left(\begin{pmatrix} h_{11}(s) & h_{12}(s) \\ h_{21}(s) & h_{22}(s) \end{pmatrix}, \begin{pmatrix} v_1(s) \\ v_2(s) \end{pmatrix} \right), \tag{2.4}$$

so that the determinant of the Wronskian matrix

$$M_\psi(s) = M_{h_{11},h_{12},v_1}(s) = \begin{pmatrix} h_{11}(s) & h_{12}(s) & v_1(s) \\ h'_{11}(s) & h'_{12}(s) & v'_1(s) \\ h''_{11}(s) & h''_{12}(s) & v''_1(s) \end{pmatrix} \tag{2.5}$$

is non-zero for a.e. $s \in [0,1]$. Then given $(h,v) \in G$ and a lattice Γ' of G commensurable with Γ one has for a.e. $s \in [0,1]$ the coset $\psi(s)(h,v)\Gamma'$ is Birkhoff generic with respect to $(G/\Gamma', \mu_{G/\Gamma'}, a_t)$ where $\mu_{G/\Gamma'}$ is the unique G-invariant probability measure on G/Γ'.

The proof of Theorem 2.1 will take almost all of § 6. The proof of Corollary 2.2 will be given at the end of § 6.2. It will be clear from § 6.2 that the assumptions on φ in Theorem 2.1 are necessary. The Birkhoff genericity result in the space of affine lattices has an immediate application to Birkhoff genericity for certain curves in a space of flat surfaces, which are branched covers of flat tori. In order to state this result (see § 2.4), let us first recall some basic background material on translation surfaces.

2.2. **Background material on translation surfaces.** A *translation surface* is a pair (M, ω) where M is an orientable compact surface and ω is a translation structure on M, that is a non-zero holomorphic 1-form also called an *Abelian differential.* Let us underline that the translation structure ω determines both a complex structure and an Abelian differential on M. Let $\Sigma_\omega \subset M$ denote the set of zeros of ω which are also the *singular points* of the translation structure. Let us consider the volume form $\nu_\omega = \frac{i}{2}\omega \wedge \overline{\omega} = \Re(\omega) \wedge \Im(\omega)$ which also will be treated as a volume measure. Since M is compact, $\nu_\omega(M)$ is finite and called the area of (M, ω).

Let M be a compact connected orientable surface and let $\Sigma \subset M$ be finite. Denote by $\mathrm{Diff}^+(M,\Sigma)$ the group of orientation-preserving diffeomorphisms of M that fix elements of Σ. Denote by $\mathrm{Diff}^+_0(M,\Sigma)$ the subgroup of elements $\mathrm{Diff}^+(M,\Sigma)$ which are isotopic to the identity. Let us denote by $\Gamma(M,\Sigma) := \mathrm{Diff}^+(M,\Sigma)/\mathrm{Diff}^+_0(M,\Sigma)$ the *mapping-class* group. We will denote by $\mathcal{T}(M,\Sigma)$ (respectively $\mathcal{T}_1(M,\Sigma)$) the *Teichmüller space of Abelian differentials* (respectively of unit area Abelian differentials), that is the space of orbits of the natural action of $\mathrm{Diff}^+_0(M,\Sigma)$ on the space of all Abelian differentials ω on M with $\Sigma_\omega = \Sigma$ (respectively, the ones with total area 1). We will denote by $\mathcal{M}(M,\Sigma)$ ($\mathcal{M}_1(M,\Sigma)$) the *moduli space* *of (unit area) Abelian differentials*, that is the space of orbits of the natural action of $\mathrm{Diff}^+(M,\Sigma)$ on the same space of (unit area) Abelian differentials. Thus $\mathcal{M}(M,\Sigma) = \mathcal{T}(M,\Sigma)/\Gamma(M,\Sigma)$ and $\mathcal{M}_1(M,\Sigma) = \mathcal{T}_1(M,\Sigma)/\Gamma(M,\Sigma)$.

The group $SL_2(\mathbb{R})$ acts naturally on the space of Abelian differentials as follows. Given a translation structure ω, consider the charts given by local primitives of the holomorphic 1-form. The new charts defined by postcomposition of this charts with an element of $SL_2(\mathbb{R})$ yield a new complex structure and a new differential which is Abelian with respect to this new complex structure, thus a new translation structure. We denote by $g\omega$ the translation structure on M obtained acting by $g \in SL_2(\mathbb{R})$ on a translation structure ω on M. Since the $SL_2(\mathbb{R})$ action commutes with that of $\mathrm{Diff}^+(M,\Sigma)$, it descends to action on $\mathcal{T}_1(M,\Sigma)$ and $\mathcal{M}_1(M,\Sigma)$. The *Teichmüller flow* is the restriction of this action to the diagonal subgroup $\{a_t = \mathrm{diag}(e^t, e^{-t}) : t \in \mathbb{R}\}$ of $SL_2(\mathbb{R})$ on $\mathcal{T}_1(M,\Sigma)$ and $\mathcal{M}_1(M,\Sigma)$. Here we slightly abuse the notation of a_t which has different meaning in (2.3). For $\theta \in \mathbb{R}$ we let

$$r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \in SL_2(\mathbb{R}).$$

We will deal also with the rotations $\{r_\theta : \theta \in \mathbb{R}\}$ that acts on a translation structure ω by $r_\theta\omega = e^{i\theta}\omega$.

Let $x_0 \in \mathcal{M}_1(M,\Sigma)$ and denote by $\mathcal{M} = \overline{SL_2(\mathbb{R})x_0}$ the closure of the $SL_2(\mathbb{R})$-orbit of x_0 in $\mathcal{M}_1(M,\Sigma)$. The celebrated result of Eskin, Mirzakhani and Mohammadi (see [14] and [15]) says that $\mathcal{M} \subset \mathcal{M}_1(M,\Sigma)$ is an affine $SL_2(\mathbb{R})$-invariant submanifold. Denote by $\mu_{\mathcal{M}}$ the corresponding affine $SL_2(\mathbb{R})$-invariant probability measure supported on $\mathcal{M}$. Recall that $\mu_{\mathcal{M}}$ is ergodic for the Teichmüller flow.

Recently Chaika and Eskin in [7] proved a finer result saying for a.e. $\theta \in [0,2\pi]$ the surface $r_\theta x_0$ is Birkhoff generic with respect to $(\mathcal{M}, \mu_{\mathcal{M}}, a_t)$. These results have applications to the dynamics on a translation surface, in virtue of Masur's ergodicity criterion, which constitutes one of the first and central results in Teichmüller dynamics.

On a given translation surface (M,ω), for every $\theta \in [0,2\pi]$ denote by $F_\theta = F_\theta^\omega$ the vector field in direction θ on $M \setminus \Sigma$, i.e. $\omega(F_\theta^\omega) = e^{i\theta}$. The corresponding directional flow $\{\psi_t^\theta\}_{t\in\mathbb{R}}$, also called a *translation flow*, on $M \setminus \Sigma$ preserves the volume measure ν_ω. We will use the notation $\{\psi_t^v\}_{t\in\mathbb{R}}$ for the *vertical flow* (corresponding to $\theta = \frac{\pi}{2}$). The flow $\{\psi_t^\theta\}_{t\in\mathbb{R}}$ is uniquely ergodic if the area is the unique invariant probability measure. Masur's ergodicity criterion relates unique ergodicity of the vertical flow on (M,ω) with the behaviour of the Teichmüller geodesic through (M,ω).

Theorem 2.3 (Masur's criterion, see [28])**.** *If $\{a_t\omega\}_{t\in\mathbb{R}}$ returns infinitely often to a compact set of $\mathcal{M}_1(M,\Sigma)$, then the vertical flow on (M,ω) is uniquely ergodic.*

Clearly, if (M,ω) is Birkhoff generic with respect to $(\mathcal{M}, \mu_{\mathcal{M}}, a_t)$, then the assumption of the theorem holds.

2.3. **Oseledets genericity for Kontsevich-Zorich cocycle.** Let us recall one of the definitions of the Kontsevich-Zorich cocycle, which is a fundamental tool in the study of translation surfaces and Teichmüller dynamics. Consider the projection $\pi : \mathcal{T}_1(M,\Sigma) \to \mathcal{M}_1(M,\Sigma)$ where $\mathcal{T}_1(M,\Sigma)$ and $\mathcal{M}_1(M,\Sigma)$ are respectively the Teichmüller and moduli space of translation surfaces introduced in the previous section. Let $\Delta \subset \mathcal{T}_1(M,\Sigma)$ be a Borel fundamental domain for the action of $\Gamma(M,\Sigma)$. For every $h \in SL_2(\mathbb{R})$ and $\tilde{x} \in \mathcal{T}_1(M,\Sigma)$ denote by $\psi_{h,\tilde{x}}$ the only element of $\Gamma(M,\Sigma)$ such that $\psi_{h,\tilde{x}}(h \cdot \tilde{x}) \in \Delta$.

Let us consider the cocycle $A : SL_2(\mathbb{R}) \times \mathcal{M}_1(M,\Sigma) \to GL(H_1(M,\mathbb{R}))$ given by

$$A(h,x)\zeta = (\psi_{h,\tilde{x}})_*\zeta \text{ for } \tilde{x} \in \Delta \text{ such that } \pi(\tilde{x}) = x. \tag{2.6}$$

Note that the cocycle A preserves $H_1(M,\mathbb{Z})$ and the non-degenerated symplectic structure on $H_1(M,\mathbb{R})$ given by the algebraic intersection form $\langle\,\cdot\,,\,\cdot\,\rangle$. Therefore, A can be considered as a cocycle taking values in $Sp(2g,\mathbb{Z})$. By the *Kontsevich-Zorich (KZ) cocycle* we mean the restriction of A to the diagonal subgroup $\{a_t : t\in\mathbb{R}\}$ of $SL_2(\mathbb{R})$, i.e.

$$A^{KZ} : \mathbb{R}\times\mathcal{M}_1(M,\Sigma)\to GL(H_1(M,\mathbb{R})), \qquad A^{KZ}(t,x)=A(a_t,x).$$

Let $\mu := \mu_{\mathcal{M}}$ be the affine $SL_2(\mathbb{R})$-invariant and ergodic probability measure on $\mathcal{M}=\overline{SL_2(\mathbb{R})x_0}$. By Moore's ergodicity theorem μ is ergodic for the subgroup $\{a_t\}$ action.

Suppose that $W\subset H_1(M,\mathbb{R})$ is a symplectic subspace (the symplectic form restricted to W is non-degenerated) of dimension $2d$. Moreover, assume that W is invariant for $SL_2(\mathbb{R})$ action on $\mathcal{M}$, i.e.

$$h\in SL_2(\mathbb{R}),\ x\in\mathcal{M}\Rightarrow A(h,x)W=W.$$

Then we can pass to the restricted cocycle $A_W^{KZ}:\mathbb{R}\times\mathcal{M}\to GL(W)$. Let

$$e^{2\lambda_1(t,x)}\geq\ldots\geq e^{2\lambda_d(t,x)}\geq e^{-2\lambda_d(t,x)}\geq\ldots\geq e^{-2\lambda_1(t,x)}$$

be the eigenvalues of $\big(A_W^{KZ}(t,x)\big)^{\mathrm{tr}}A_W^{KZ}(t,x)$. Oseledets multiplicative ergodic theorem says that there exist

$$\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_d\geq-\lambda_d\geq\ldots\geq-\lambda_2\geq-\lambda_1,$$

which are called *Lyapunov exponents*, such that for μ almost every $x\in\mathcal{M}$ and any $1\leq i\leq d$

$$\lim_{t\to\infty}\frac{1}{t}\lambda_i(t,x)=\lambda_i. \tag{2.7}$$

If (2.7) holds for every $1\leq i\leq n$ we say x is *Oseledets generic* with respect to $(\mathcal{M},\mu,a_t,A_W^{KZ})$, or simply Oseledets generic if the dynamical system and cocycle is understood.

Recently Chaika and Eskin in [7] considered *circles* in the space of translation surfaces, i.e. curves of the form $\{r_\theta x_0,\theta\in[0,2\pi]\}$ where x_0 is any given translation surface and proved that for a.e. $\theta\in[0,2\pi]$ the surface $r_\theta x_0$ is Oseledets generic. Their result turns out to be a fundamental tool for proving dynamical properties of directional flows on non-compact periodic translation surfaces, cf. [3], [8], [21] and [19] for the Ehrenfest wind tree model.

As we will see in § 4, the Chaika-Eskin theorem is however not sufficient to examine the behavior of light rays in lattice Eaton models as the angle of the rays changes. In such models we need to show that a.e. point is Birkhoff and Oseledets generic for curves different from those of the form $\theta\mapsto r_\theta x_0$.

We will prove Oseledets genericity for a class of curves which we will call *well approximated by horocycles*. Let $\mathcal{T}=\pi^{-1}(\mathcal{M})$ and consider a metric $d:\mathcal{T}\times\mathcal{T}\to\mathbb{R}_{\geq0}$ satisfying the following Γ-*invariance* and *growth* conditions:

$$d(\gamma(\tilde{x}),\gamma(\tilde{y}))=d(\tilde{x},\tilde{y})\ \text{for all}\ \gamma\in\Gamma(M,\Sigma); \tag{2.8}$$

$$d(\tilde{y},u(t)\cdot\tilde{y})\leq|t|\ \text{for every}\ t\in\mathbb{R}\quad\text{where}\ u(t)=\begin{pmatrix}1&t\\0&1\end{pmatrix}. \tag{2.9}$$

Note that the distance on $\mathcal{T}(M,\Sigma)$ defined in [2] and derived from a Finsler Γ-invariant metric on $\mathcal{T}(M,\Sigma)$ satisfies the above conditions.

We will prove an Oseledets type result along curves which satisfy the following technical assumption.

Definition 2.4. We say that $\varphi : I \to \mathcal{M}$ is *well approximated by horocycles* if it is of the form $\varphi = \pi \circ \widetilde{\varphi}$ with $\widetilde{\varphi} : I \to \mathcal{T}$ of class C^1 and there exist $C(\varphi) > 0$, $\rho \in \mathbb{N}$ and a metric d satisfying properties (2.8), (2.9) such that

$$d\big(a_t \cdot \widetilde{\varphi}(s + re^{-2t}), u(-r) \cdot a_t \cdot \widetilde{\varphi}(s)\big) \leq C(\varphi)|r|(1 + |r|^\rho)e^{-t} \tag{2.10}$$

for all $s \in I, r \in [-1, 1], t \geq 0$ with $s + re^{-2t} \in I$.

Theorem 2.5. *Let μ be the probability affine measure on an $SL_2(\mathbb{R})$-orbit closure $\mathcal{M}$. Let $\varphi : I \to \mathcal{M}$ be a curve which is well approximated by horocycles in the sense of Definition 2.4 and such that for a.e. $s \in I$ the point $\varphi(s)$ is Birkhoff generic with respect to $(\mathcal{M}, \mu, a_t)$. Suppose that the sum of positive Lyapunov exponents of the restricted KZ-cocycle $A_W^{KZ} : \mathbb{R} \times \mathcal{M} \to GL(W)$ is less than one. Then for a.e. $s \in I$ one has $\varphi(s)$ is Oseledets generic with respect to $(\mathcal{M}, \mu, a_t, A_W^{KZ})$.*

Remark 2.6. We remark that in our applications (in Section 4) the assumption that the sum of positive exponents is less than one in Theorem 2.5 is automatically satisfied. Indeed, we apply Theorem 2.5 to the KZ-cocycle A_W^{KZ} restricted to a two dimensional symplectic subspace W, which is symplectic orthogonal to the tautological bundle (the two dimensional $SL_2(\mathbb{R})$-invariant subbundle corresponding to extremal Lyapunov exponents 1 and -1). Thus, A_W^{KZ} has at most one positive Lyapunov exponent which is less than one.

Furthermore, in virtue of work in progress [12] announced recently by Eskin, Filip and Wright, the assumption that the sum of positive exponents is less than one could be removed from Theorem 2.5 whenever the subspace W meets a natural and verifiable condition. A more extensive discussion of this issue is postponed to the end of Section 7 (see §7.5). We decided to present and prove the above formulation of Theorem 2.5 (stated for invariant subspaces of dimension higher than two, but with the additional assumption on Lyapunov exponents) since the work [12] is not yet available and since we believe that also with this assumption it has further applications to billiards and systems of lenses (which will be investigated in forthcoming works).

2.4. **Genericity along curves in branched covers of a flat torus.** Let M be a compact connected orientable surface of genus 2 and let $\Sigma \subset M$ be a two-point subset. Denote by $\mathcal{M}^{dc} \subset \mathcal{M}(M, \Sigma)$ the (*moduli*) *space of double covers of flat tori* ramified over two distinguished points, i.e. elements of $\mathcal{M}^{dc}$ are represented by translation surfaces of the form $(M, q^*\omega_0)$, where (M_0, ω_0) is a flat torus, $q : M \to M_0$ is a double cover ramified over $q(\Sigma)$ and $q^*\omega_0$ is the pullback of the Abelian differential ω_0. If the area of ω_0 is 1 then the area of $q^*\omega_0$ is 2. Denote by $\mathcal{M}_2^{dc} \subset \mathcal{M}^{dc}$ the subspace of area 2 translation surfaces.

Consider a translation surface (M, ω) given in Figure 4: opposite sides of the two parallelograms are identified by parallel translations, while the sides of the slits are identified by parallel translations as indicated in Figure 4, i.e. between different copies. The surface (M, ω) has genus two and two conical singularities of cone angle 4π, which correspond to the endpoints of the slit. Moreover, $(M, \omega) \in \mathcal{M}^{dc}$.

The space $\mathcal{M}_2^{dc} \subset \mathcal{M}(M, \Sigma)$ is closed and $SL_2(\mathbb{R})$-invariant and under an assumption on irrationality of $(M, \omega) \in \mathcal{M}_2^{dc}$ its $SL_2(\mathbb{R})$-orbit closure is $\mathcal{M}_2^{dc}$. In virtue of Lemma 2.7 below the space $\mathcal{M}_2^{dc}$ and the dynamics on $\mathcal{M}_2^{dc}$ can be fully described in the language of homogeneous spaces studied in §2.1. Let μ_{dc} denote the affine $SL_2(\mathbb{R})$-invariant probability measure supported on $\mathcal{M}_2^{dc}$ (see previous section 2.2). Let G, Γ and X be as in § 2.1 and consider the lattice $\Gamma_2 := SL_2(\mathbb{Z}) \ltimes 2\mathbb{Z}^2$ which has index 4 in Γ. Denote by μ_2 the unique probability measure invariant under the left G-action on G/Γ_2. Let

$$G' := G \setminus \{(h, h\mathbb{Z}^2) : h \in SL_2(\mathbb{R})\} \quad \text{and} \quad X_2 := G'\Gamma_2 \subset G/\Gamma_2. \tag{2.11}$$

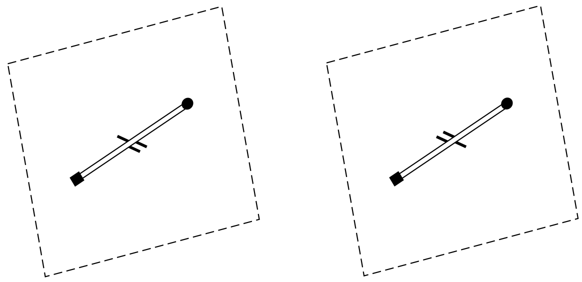

FIGURE 4. A translation surface (M, ω) in the space $\mathcal{M}^{dc}$ of double covers of tori.

We remark here that G' is an open subset of G and $\mu_2(X_2) = 1$.

Lemma 2.7. *There is a diffeomorphism $\Psi : \mathcal{M}_2^{dc} \to X_2$ such that:*

(i) Ψ commutes with $SL_2(\mathbb{R})$-action;
(ii) Ψ maps the natural affine measure μ_{dc} to μ_2.

Proof. Every translation surface (M, ω) in $\mathcal{M}_2^{dc}$ is uniquely determined by:

- the translation structure of the torus, i.e. $\mathbb{T}^2_\Lambda = \mathbb{R}^2/\Lambda$, where $\Lambda = h(\mathbb{Z}^2)$ for $h \in SL_2(\mathbb{R})$;
- two distinct marked points $\sigma_0, \sigma_1 \in \mathbb{T}^2_\Lambda$ so that $\sigma_0 = \mathbf{0} + \Lambda$;
- and finally the double cover ramified over $\Sigma := \{\sigma_0, \sigma_1\}$ is given by a relative homology class $\gamma \in H_1(\mathbb{T}^2_\Lambda, \Sigma, \mathbb{Z}/2\mathbb{Z}) \setminus H_1(\mathbb{T}^2_\Lambda, \mathbb{Z}/2\mathbb{Z})$.

Let $v_1 = h(1,0)^{\mathrm{tr}}$, $v_2 = h(0,1)^{\mathrm{tr}}$ be the vectors in $\mathbb{R}^2$ and choose $v \in \sigma_1 \subset \mathbb{R}^2$ (the coset σ_1 is treated as a subset in $\mathbb{R}^2$) in the parallelogram generated by v_1, v_2. Denote by $\zeta_0, \zeta_1, \zeta_2 \in H_1(\mathbb{T}^2_\Lambda, \Sigma, \mathbb{Z}/2\mathbb{Z})$ the homology classes given by the projections on $\mathbb{R}^2/\Lambda$ the vectors $\overrightarrow{\mathbf{0}v}$, v_1 and v_2 respectively. Since $\zeta_0, \zeta_1, \zeta_2$ establish a basis of $H_1(\mathbb{T}^2_\Lambda, \Sigma, \mathbb{Z}/2\mathbb{Z})$ and γ is not absolute we can decompose

$$\gamma = \zeta_0 + n_1\zeta_1 + n_2\zeta_2 \text{ with } n_1, n_2 \in \mathbb{Z}/2\mathbb{Z}.$$

Finally define the map $\Psi : \mathcal{M}_2^{dc} \to G/\Gamma_2$ by

$$\Psi(M, \omega) = (h, v + h(n_1, n_2)^{\mathrm{tr}}) \mod \Gamma_2.$$

This map is one-to-one and onto X_2 in (2.11), so we will identify $\mathcal{M}_2^{dc}$ with X_2. Let us consider the subgroup $SL_2(\mathbb{R}) \ltimes \{\mathbf{0}\} \subset G$ which we identify with $SL_2(\mathbb{R})$. This group defines a natural left action of $SL_2(\mathbb{R})$ on G/Γ_2 and X_2. Moreover, this action coincides with the standard $SL_2(\mathbb{R})$-action on $\mathcal{M}_2^{dc}$ via the correspondence between $\mathcal{M}_2^{dc}$ and X_2.

Since μ_{dc} is a smooth $SL_2(\mathbb{R})$-invariant and ergodic probability measure, the measure $\Psi_*\mu_{dc}$ is also $SL_2(\mathbb{R})$-invariant and ergodic. Moreover, $\Psi : \mathcal{M}_2^{dc} \to X_2$ is diffeomorphism and the subset $X_2 \subset G/\Gamma_2$ is open and has full μ_2 measure. According to Ratner's measure classification theorem every $SL_2(\mathbb{R})$-invariant and ergodic probability measure on G/Γ_2 is either μ_2 or supported on a proper closed sub-manifold. Therefore we have $\Psi_*\mu_{dc} = \mu_2$. □

Remark 2.8. By Lemma 2.7, Birkhoff genericity of (M, ω) in $(\mathcal{M}_2^{dc}, \mu_{dc}, a_t)$ is equivalent to the Birkhoff genericity of $\Psi(M, \omega)$ in $(G/\Gamma_2, \mu_2, a_t)$.

Remark 2.9. Let $\mathcal{T}_2(M, \Sigma)$ be the Teichmüller space of Abelian differentials with area 2. Denote by $\pi : \mathcal{T}_2(M, \Sigma) \to \mathcal{M}_2(M, \Sigma)$ the quotient map and let $\mathcal{T}_2^{dc} := \pi^{-1}(\mathcal{M}_2^{dc})$. Then we can lift the diffeomorphism $\Psi : \mathcal{M}_2^{dc} \to X_2$ to $\widetilde{\Psi} : \mathcal{T}_2^{dc} \to G'$ which also commutes with $SL_2(\mathbb{R})$ action. This gives a natural identification of the Teichmüller space $\mathcal{T}_2^{dc}$ of the double covers of flat tori with the subset G' of the group G. Furthermore, $\widetilde{\Psi}$ conjugates the action of $\Gamma(M,\Sigma)$ on $\mathcal{T}_2^{dc}$ with the right Γ_2-action on G'.

Let

$$W := \{\zeta \in H_1(M,\mathbb{R}) : \tau_*\zeta = -\zeta\},$$

where $\tau : M \to M$ is the only nontrivial element of the deck transformation group of any ramified double cover. The subspace W does not depend on the choice of ramified cover. Moreover, W is two dimensional symplectic and $SL_2(\mathbb{R})$-invariant. Therefore we can consider the restricted Kontsevich-Zorich cocycle $A_W^{KZ} : \mathbb{R} \times \mathcal{M}_2^{dc} \to GL(W)$. The following result, proved at the end of this section, gives an effective criterion for curves in $\mathcal{M}_2^{dc}$ to be almost everywhere Birkhoff and Oseledets (for A_W^{KZ}) generic.

Theorem 2.10. *Let $\varphi : I \to \mathcal{M}_2^{dc}$ be a C^2-curve such that $\Psi \circ \varphi(s) = \psi(s)g\Gamma_2$, where $g \in G$ and $\psi : I \to G$ is a C^2-curve such that $\det M_\psi(s) \neq 0$ for a.e. $s \in I$. Then $\varphi(s) \in \mathcal{M}_2^{dc}$ is Birkhoff and Oseledets generic for a.e. $s \in I$.*

In virtue of Masur's criterion for unique ergodicity (stated in § 2.2), we immediately get the following.

Corollary 2.11. *Let $\{(M_s,\omega_s), s \in I\}$ be a curve in the space of translation structures so that its image in $\mathcal{M}_2^{dc}$ satisfies the assumption of φ in Theorem 2.10. Then for a.e. $s \in I$ the vertical flow on (M_s,ω_s) is uniquely ergodic.*

We fix a right invariant Riemannian metric on G and use $d : G \times G \to \mathbb{R}_{\geq 0}$ to denote the corresponding metric. After rescaling we can assume that $d(Id, u(1)\cdot Id) = 1$. It is easy to see that the restriction of d to $G' \times G'$ meets (2.8)-(2.9). Since G' is identified with $\mathcal{T}_2^{dc}$ the transport of the metric d to $\mathcal{T}_2^{dc}$, which will be also denoted by d, meets (2.8)-(2.9) as well.

We now prove Theorem 2.10 using Theorem 2.1 and Theorem 2.5. The following lemma verifies that the curve given in Theorem 2.10 satisfies the nondegeneracy condition.

Lemma 2.12. *Let $\psi : I \to G$ (I is a compact interval) be a C^2-curve of the form $\psi(s) = (h(s), v(s)) \cdot (h_0, v_0)$. If*

$$h_{11}(s)h_{12}'(s) - h_{12}(s)h_{11}'(s) \neq 0 \quad \text{for every} \quad s \in I \tag{2.12}$$

then there is a C^2-diffeomorphism $\kappa : I_0 \to I$ from some closed interval I_0 such that the curve $\psi \circ \kappa : I_0 \to G$ satisfies the condition (2.10) *with respect to metric d given above.*

Proof. Let us consider the ordinary differential equation

$$\kappa'(s) = 1/(h_{11}(\kappa(s))h_{12}'(\kappa(s)) - h_{12}(\kappa(s))h_{11}'(\kappa(s)))$$

and let $\kappa : I_0 \to I$ be its solution. After changing the parameter we can pass to the case where $h_{11}(s)h_{12}'(s) - h_{12}(s)h_{11}'(s) = 1$ for every s and then we need to show that (2.10) holds.

Write $l := re^{-2t}$ and let us first estimate form above the norm

$$\begin{aligned}\|Id - &u(-r)\cdot a_t \cdot \psi(s)\cdot\psi(s+l)^{-1}\cdot a_{-t}\| \\ &= \|a_t \cdot (Id - u(-l)\cdot\psi(s)\cdot\psi(s+l)^{-1})\cdot a_{-t}\|.\end{aligned}$$

Let

$$\upsilon(l,s) := Id - u(-l)\cdot\psi(s)\cdot\psi(s+l)^{-1} \text{ and } \underline{\upsilon}(l,s) := Id - u(-l)\cdot h(s)\cdot h(s+l)^{-1}.$$

Since υ is of class C^2, $\upsilon(0,s) = 0$ and $\|\frac{\partial}{\partial l}\upsilon(l,s)\| \leq C\|\psi\|_{C^1}^3$ for some $C > 0$, by the mean value theorem,

$$\|\upsilon(l,s)\| \leq C\|\psi\|_{C^1}^3|l|.$$

Moreover,

$$\frac{\partial}{\partial l}\underline{\upsilon}(0,s) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - h(s)\cdot (h(s)^{-1})'$$

and $\|\frac{\partial^2}{\partial l^2}\upsilon(l,s)\| \leq C'\|\psi\|^2_{C^2}$ for some $C' > 0$. It follows that

$$\frac{\partial}{\partial l}\upsilon_{12}(0,s) = \frac{\partial}{\partial l}\underline{\upsilon}_{12}(0,s) = 1 - h'_{11}(s)h_{12}(s) + h_{11}(s)h'_{12}(s) = 0.$$

Then, by Taylor's formula,

$$|\upsilon_{12}(l,s)| \leq C'\|\psi\|^2_{C^2}|l|^2.$$

Therefore,

$$a_t \cdot \upsilon(l,s) \cdot a_{-t} = \begin{bmatrix} \upsilon_{11}(l,s) & e^{2t}\upsilon_{12}(l,s) & e^{t}\upsilon_{13}(l,s) \\ e^{-2t}\upsilon_{21}(l,s) & \upsilon_{22}(l,s) & e^{-t}\upsilon_{23}(l,s) \\ 0 & 0 & 0 \end{bmatrix}$$

with

$$\begin{gathered} |\upsilon_{11}(l,s)| \leq C\|\psi\|^3_{C^1}|r|e^{-2t},\ |e^{-2t}\upsilon_{21}(l,s)| \leq C\|\psi\|^3_{C^1}|r|e^{-4t} \\ |e^{2t}\upsilon_{12}(l,s)| \leq C'\|\psi\|^2_{C^2}|r|^2e^{-2t},\ |\upsilon_{22}(l,s)| \leq C\|\psi\|^3_{C^1}|r|e^{-2t} \\ |e^{t}\upsilon_{13}(l,s)| \leq C\|\psi\|^3_{C^1}|r|e^{-t},\ |e^{-t}\upsilon_{23}(l,s)| \leq C\|\psi\|^3_{C^1}|r|e^{-3t}. \end{gathered}$$

It follows that there exists $C(\psi) > 0$ such that

$$\|a_t \cdot \upsilon(l,s) \cdot a_{-t}\| \leq C(\psi)|r|(1+|r|)e^{-t} \leq 2C(\psi).$$

Moreover, there exists $C'' > 0$ such that for every $g \in G$ with $\|Id - g\| \leq 2C(\psi)$ we have $d(Id, g) \leq C''\|Id - g\|$. It follows that

$$\begin{aligned} &d\big(a_t\cdot\psi(s+re^{-2t}), u(-r)\cdot a_t\cdot\psi(s)\big) \\ &\quad = d\big(Id, u(-r)\cdot a_t\cdot\psi(s)\cdot\psi(s+re^{-2t})^{-1}\cdot a_{-t}\big) \\ &\quad \leq C''\|Id - u(-r)\cdot a_t\cdot\psi(s)\cdot\psi(s+re^{-2t})^{-1}\cdot a_{-t}\| \\ &\quad = C''\|a_t\cdot\upsilon(l,s)\cdot a_{-t}\| \leq C''C(\psi)|r|(1+|r|)e^{-t}. \end{aligned}$$

□

Proof of Theorem 2.10. We choose a countable collection $\{I_k\}_{k\in\mathbb{N}}$ of closed subinvervals of I so that $\bigcup_{k\in\mathbb{N}} I_k$ has full measure in I and (2.12) holds for every $s \in I_k$. We fix an interval I_k, it suffices to show that Theorem 2.10 holds for a.e. $s \in I_k$. The Birkhoff genericity of $\varphi(s)$ for a.e. $s \in I_k$ follows from Corollary 2.2 and Remark 2.8. By Lemma 2.12 and the correspondence between X_2 and $\mathcal{M}_2^{dc}$ the curve $\varphi|_{I_k}$ has a parameterization such that $\varphi|_{I_k} \circ \kappa$ is well approximated by horocycles and κ is a C^2-diffeomorphism. Since W is two dimensional and is the symplectic orthocomplement to the tautological bundle, using Theorem 2.5 together with Remark 2.6 we conclude that for a.e. $s \in I_k$, the element $\varphi(s) \in \mathcal{M}_2^{dc}$ is Oseledets generic which completes the proof. □

3. Ergodicity in elliptical billiards with barriers

In this section we prove Theorem 1.1 on unique ergodicity in elliptical billiards with barriers. We exploit the reduction of these family of billiards to polygonal billiards discovered by Dragović and Radnović in [9] and stated in § 3.1. The billiard flow on each $\mathcal{S}_\lambda$ is isomorphic to the vertical flow on a surface from $\mathcal{M}_2^{dc}$, which gives a smooth curve in $\mathcal{M}_2^{dc}$. In view of Corollary 2.11, it suffices to show that the curve satisfies a non-degeneracy property. In § 3.2, we prove the non-degeneracy property for the aforementioned curve which yields Theorem 1.1.

3.1. **Reduction to a family of polygonal billiards.** Let $\mathcal{D}_{\lambda_0}$ be the elliptic billiard table with a barrier described in § 1.1 and let $\mathcal{S}_\lambda$ for $0 < \lambda < a$ be one of its invariant regions. The fundamental observation, made by Dragović and Radnović in [9], is that the billiard flow on $\mathcal{S}_\lambda$ can be reduced to a polygonal billiard flow by a suitable change of coordinates. The polygonal billiard tables $\mathcal{P}_\lambda$ which are obtained after change of coordinates (given by elliptic integrals) are either a non-planar billiard table $\mathcal{P}_\lambda$ which is a cylinder with a vertical slit (for $0 < \lambda < b$), shown in Figure 5, or for $b < \lambda < a$, a rectangular billiard table with a vertical slit, shown in Figure 6. Billiard trajectories in $\mathcal{S}_\lambda$, which are tangent to the caustic $\mathcal{C}_\lambda$, are mapped to billiard trajectories in $\mathcal{P}_\lambda$ all in the *same* family of directions $\pm\pi/4$, $\pm 3\pi/4$. More precisely, we have the following.

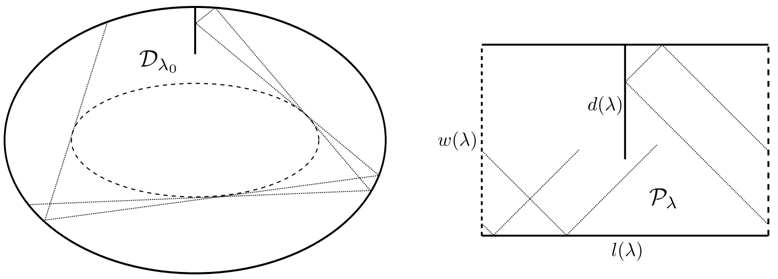

FIGURE 5.

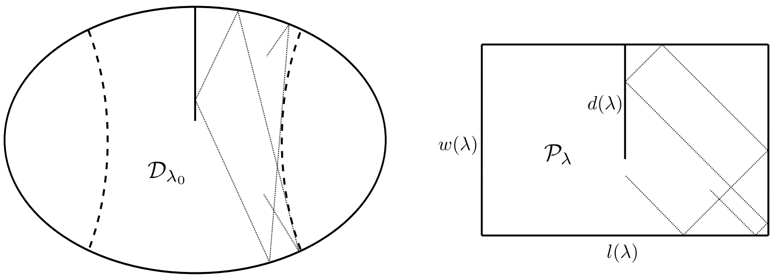

FIGURE 6.

Proposition 3.1 (c.f. [9])**.** *The billiard flow in $\mathcal{D}_{\lambda_0}$ restricted to the invariant region $\mathcal{S}_\lambda$ is isomorphic to the a billiard flow in directions $\pm\pi/4$, $\pm 3\pi/4$ inside a polygonal billiard $\mathcal{P}_\lambda$ as follows:*

(E) *If $\lambda_0 < \lambda < b$ then $\mathcal{P}_\lambda$ is the cylinder shown in Figure 5 with length*

$$l(\lambda) = 4\int_b^a \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}} = 4\int_{-\infty}^{\lambda} \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}}$$

and width

$$\begin{aligned} w(\lambda) &= \int_0^\lambda \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}} \\ &= \int_b^a \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}} - \int_{-\infty}^{0} \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}} \end{aligned}$$

and a linear vertical obstacle of length

$$d(\lambda) = \int_0^{\lambda_0} \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}}.$$

(E') *If $0 < \lambda \leq \lambda_0$ then $\mathcal{P}_\lambda$ is a rectangle such that its length $l(\lambda)$ and width $w(\lambda)$ are the same as in the previous case.*

(H) *If $b < \lambda < a$ then $\mathcal{P}_\lambda$ is the rectangle in Figure 6 with length*

$$l(\lambda) = 2\int_\lambda^a \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}} = 2\int_{-\infty}^b \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}}$$

and width

$$w(\lambda) = 2\int_0^b \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}}$$

and a linear vertical obstacle of length

$$d(\lambda) = \int_0^{\lambda_0} \frac{ds}{\sqrt{(a-s)(b-s)(\lambda-s)}}.$$

The billiards on $\mathcal{P}_\lambda$, $\lambda_0 < \lambda < b$, are *rational polygonal billiards*, i.e. the order of the group generated by reflections at the polygon sides is finite (in this case it is four). A standard procedure, known as *unfolding* or *Katok-Zemlyakov construction* (described in [18] and [25]), allows one to reduce a billiard in a *rational* polygon to a linear flow on a translation surface. In the case of the billiards $\mathcal{P}_\lambda$ in Proposition 3.1, the unfolding procedure consists in taking four copies of the billiard, one for each element of the group of reflections and gluing them. The resulting surface are shown in Figure 7.

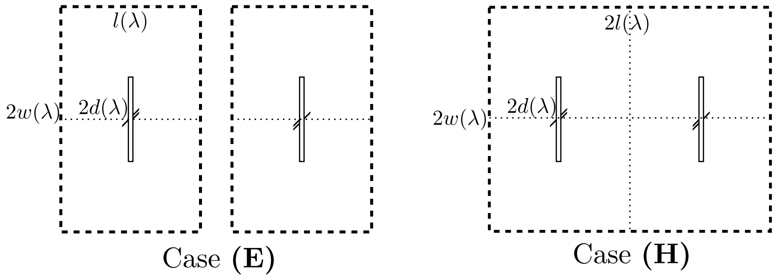

FIGURE 7. The surface $(M_\lambda, \omega_\lambda)$

Thus, the unfolding procedure reduces the family of billiard flows in Proposition 3.1 to a family of directional flows in a *fixed* direction $\pi/4$ on the translation surfaces $(M_\lambda, \omega_\lambda) \in \mathcal{M}^{dc}$, in Figure 7, parameterized by $\lambda_0 < \lambda < b$. The case **(E')**, where $0 < \lambda < \lambda_0$, should be treated separately, then $(M_\lambda, \omega_\lambda)$ is a torus. For $\lambda_0 < \lambda < b$ we will assume that $(M_\lambda, \omega_\lambda)$ is rescaled so that its area is two, and then rotated by $\pi/4$ so that the linear flow in direction $\pi/4$ is mapped to the vertical linear flow and still use the same notation $(M_\lambda, \omega_\lambda)$. Then these surfaces all belong to the moduli space $\mathcal{M}_2^{dc}$ of double covers of flat tori, described in 2.4.

Let us remark that given a *fixed* translation surface (M, ω), for almost every direction $\theta \in [0, 2\pi]$ the linear flow in direction θ is known to be uniquely ergodic by a celebrated result by Kerckhoff, Masur and Smillie, [26] (or, in virtue of the recent result by Chaika and Eskin, by Birkhoff genericity for a.e. point of the curve $r_\theta(M, \omega)$ and Masur's ergodicity criterion, see § 2.2). Proving unique ergodicity of the elliptical billiard flow on the invariant region $\mathcal{S}_\lambda$ for almost every $\lambda_0 < \lambda < b$, on the other hand, is equivalent to proving unique ergodicity of the vertical flow on $(M_\lambda, \omega_\lambda)$. By Masur's criterion it suffices to prove Birkhoff genericity for almost every point on the curve $\gamma := \{(M_\lambda, \omega_\lambda), \lambda \in (\lambda_0, b)\}$ (see also Remark 3.2). This curve γ is explicitly described in terms of the coordinates given by the identification of $\mathcal{M}_2^{cd}$ with the subset X_2 of the homogeneous space G/Γ_2 in Lemma 2.7 (see §3.2).

In § 3.2 we verify that γ satisfies the non-degeneracy assumptions of Corollary 2.11, to conclude that almost every point is Birkhoff generic.

Remark 3.2. In order to apply Masur's criterium, it is in principle enough to show that almost every point on the curve γ is *recurrent* under the Teichmüller flow (see Theorem 2.3). Let us remark though that recurrence under the diagonal flow in the homogeneous space G/Γ_2 (which is well-known) is *not sufficient*. Indeed, from Lemma 2.7 we know that the space $\mathcal{M}_2^{dc}$ is isomorphic to a *subset* $X_2 \subset G/\Gamma_2$ obtained by *removing* 2 closed $SL_2(\mathbb{R})$-orbits which correspond to the locus of flat tori where the marked point coincide with the origin of the lattice. One can hence construct a recurrent diagonal orbit in G/Γ_2 which approach H and lift it to a Teichmüller geodesics in $\mathcal{M}_2^{dc}$ which is *divergent* (since the separating slit endpoints collide towards a pinched surface).

On the other hand, to get recurrence in $\mathcal{M}_2^{dc}$ it is sufficient to prove *density* in the homogeneous setup (i.e. under the diagonal flow a_t for almost every point on the curve $\Psi\gamma$ in G/Γ_2, image by the isomorphism in Lemma 2.7). Density could in principle be deduced by equidistribution of the normalized volume measure on $a_t\Psi\gamma$ as $t \to \infty$ using similar arguments as in Elkies-McMullen [11], thus yielding a different strategy of proof. We here prove and exploit the stronger conclusion of almost everywhere Birkhoff genericity on γ, since the criterium used to prove it (Theorem 2.10) is also needed for the Eaton application in §4, for which Birkhoff genericity is needed in its full strength.

3.2. **Non-degeneracy of the curves in the space of double torus covers.**

In this section we conclude the proof of Theorem 1.1. Let us consider the curve

$$(\lambda_0, b) \cup (b, a) \ni \lambda \mapsto (M_\lambda, \omega_\lambda) \in \mathcal{M}_2^{dc}$$

described in § 3.1. In view of Corollary 2.11, it suffices to check the non-degeneracy assumptions for the curve $\lambda \mapsto \Psi(M_\lambda, \omega_\lambda)$, where $\Psi : \mathcal{M}_2^{dc} \to X_2 \subset G/\Gamma_2$ is the correspondence described in Lemma 2.7. Since $(M_\lambda, \omega_\lambda)$ is a the surface after rotation by $\pi/4$ one of the surfaces shown in Figure 7, its image is of the form $\psi(\lambda)\Gamma_2$, where

$$\psi(\lambda) = \left(\begin{pmatrix} \frac{l(\lambda)}{r(\lambda)} & -2\frac{w(\lambda)}{r(\lambda)} \\ \frac{l(\lambda)}{r(\lambda)} & 2\frac{w(\lambda)}{r(\lambda)} \end{pmatrix}, \begin{pmatrix} -2\frac{d(\lambda)}{r(\lambda)} \\ 2\frac{d(\lambda)}{r(\lambda)} \end{pmatrix} \right)$$

with $r(\lambda) = 2\sqrt{l(\lambda)w(\lambda)}$.

We will show that for $\lambda \in (\lambda_0, b) \cup (b, a)$ the determinant of $M_\psi(\lambda)$ (see (2.5)) is non-zero. The following result will help us to verify this requirement.

Lemma 3.3. *Let $f : (c_1, c_2) \cup (c_3, c_4) \to \mathbb{R}$ ($-\infty \le c_1 < c_2 < c_3 < c_4 \le \infty$) be a positive continuous function such that $\int_{c_1}^{c_2} f(s)\,ds$ and $\int_{c_3}^{c_4} f(s)\,ds$ are finite. Suppose that $\{A_i : 1 \le i \le k\}$ is a family of pairwise disjoint subintervals of $(c_1, c_2) \cup (c_3, c_4)$. Then for every $\lambda \in (c_2, c_3)$ we have*

$$\det\left[\int_{A_i} \frac{f(s)\,ds}{(\lambda - s)^{j-1}}\right]_{i,j=1,\ldots,k} \neq 0. \tag{3.1}$$

Proof. By the Vandermonde determinant formula, we have

$$\begin{aligned}\det\left[\int_{A_i} \frac{f(s)\,ds}{(\lambda - s)^{j-1}}\right]_{1\le i,j\le k} &= \int_{\prod_{i=1}^k A_i} \det\left[\frac{g(s_i)}{(\lambda - s_i)^{j-1}}\right]_{1\le i,j\le k} ds_1 \ldots ds_k \\ &= \int_{\prod_{i=1}^k A_i} \prod_{i=1}^k g(s_i) \prod_{1\le j<i\le k} \Big(\frac{1}{\lambda - s_i} - \frac{1}{\lambda - s_j}\Big)\, ds_1 \ldots ds_k.\end{aligned}$$

Since for $j < i$ the intervals A_i, $A_{i'}$ are disjoint and do not contain λ, the map

$$A_i \times A_j \to \mathbb{R} \qquad (s_i, s_j) \mapsto \frac{1}{(\lambda - s_i)} - \frac{1}{(\lambda - s_j)}$$

is of constant sign. Therefore, the integrated function is of constant sign as well (as the product of such functions). This gives the non-vanishing of the integral. □

Remark 3.4. Suppose that $a, b, c, r : I \to \mathbb{R}$ are C^2-functions such that r takes only non-zero values. Then $M_{ra,rb,rc}(s) = r(s)U(s)M_{a,b,c}(s)$, where $U(s)$ is a lower uni-triangular matrix. It follows that non-vanishing of the determinant of $M_{ra,rb,rc}(s)$ is inherited by the matrix $M_{a,b,c}(s)$. This observation will be used in the following lemma.

Lemma 3.5. *The determinant of $M_\psi(\lambda)$ is non-zero for every $\lambda \in (\lambda_0, b) \cup (b, a)$.*

Proof. **Case (E):** $\lambda \in (\lambda_0, b)$. Let us consider the C^∞-function

$$e : ((-\infty, \lambda_0) \cup (b, a)) \times (\lambda_0, b) \to \mathbb{R}_+, \quad e(s, \lambda) = \frac{1}{\sqrt{(a - s)(b - s)(\lambda - s)}}.$$

Then

$$l(\lambda) = 4 \int_b^a e(s, \lambda)\, ds, \qquad d(\lambda) = \int_0^{\lambda_0} e(s, \lambda)\, ds,$$
$$w(\lambda) = \frac{l(\lambda)}{4} - \widetilde{w}(\lambda) \quad \text{with} \quad \widetilde{w}(\lambda) := \int_{-\infty}^0 e(s, \lambda)\, ds$$

and

$$\frac{\partial e}{\partial \lambda}(s, \lambda) = -\frac{1}{2}\frac{e(s, \lambda)}{\lambda - s}, \qquad \frac{\partial^2 e}{\partial \lambda^2}(s, \lambda) = \frac{3}{4}\frac{e(s, \lambda)}{(\lambda - s)^2}.$$

Hence

$$l'(\lambda) = -2 \int_b^a \frac{e(s, \lambda)\, ds}{\lambda - s}, \qquad l''(\lambda) = 3 \int_b^a \frac{e(s, \lambda)\, ds}{(\lambda - s)^2},$$
$$\widetilde{w}'(\lambda) = -\frac{1}{2} \int_{-\infty}^0 \frac{e(s, \lambda)\, ds}{\lambda - s}, \qquad \widetilde{w}''(\lambda) = \frac{3}{4} \int_{-\infty}^0 \frac{e(s, \lambda)\, ds}{(\lambda - s)^2},$$
$$d'(\lambda) = -\frac{1}{2} \int_0^{\lambda_0} \frac{e(s, \lambda)\, ds}{\lambda - s}, \qquad d''(\lambda) = \frac{3}{4} \int_0^{\lambda_0} \frac{e(s, \lambda)\, ds}{(\lambda - s)^2}.$$

In view of Remark 3.4, we can consider the matrix $M_{l,\widetilde{w},d}(\lambda)$ instead of $M_\psi(\lambda)$. Let $A_1 = (a, b)$, $A_2 = (-\infty, 0)$, $A_3 = (0, \lambda_0)$. Since they are pairwise disjoint, by Lemma 3.3, we have

$$\det M_{l,\widetilde{w},d}(\lambda) = -\frac{3}{2} \det \left[\int_{A_i} \frac{e(s, \lambda)\, ds}{(\lambda - s)^{j-1}} \right]_{i,j=1,2,3} \neq 0.$$

Since $r(\lambda) > 0$, this completes the proof of the part **(E)**.

Case (H): $\lambda \in (b, a)$. Here we deal with $e : (-\infty, b) \times (b, a) \to \mathbb{R}$. Then

$$l(\lambda) = 2 \int_{-\infty}^b e(s, \lambda)\, ds, \quad w(\lambda) = 2 \int_0^b e(s, \lambda)\, ds, \quad d(\lambda) = \int_0^{\lambda_0} e(s, \lambda)\, ds$$

and

$$l'(\lambda) = -\int_{-\infty}^b \frac{e(s, \lambda)\, ds}{\lambda - s}, \qquad l''(\lambda) = \frac{3}{2} \int_{-\infty}^b \frac{e(s, \lambda)\, ds}{(\lambda - s)^2}$$
$$w'(\lambda) = -\int_0^b \frac{e(s, \lambda)\, ds}{\lambda - s}, \qquad w''(\lambda) = \frac{3}{2} \int_0^b \frac{e(s, \lambda)\, ds}{(\lambda - s)^2}$$
$$d'(\lambda) = -\frac{1}{2} \int_0^{\lambda_0} \frac{e(s, \lambda)\, ds}{\lambda - s}, \qquad d''(\lambda) = \frac{3}{4} \int_0^{\lambda_0} \frac{e(s, \lambda)\, ds}{(\lambda - s)^2}.$$

Again, in view of Remark 3.4, we deal with the matrix $M_{l,w,d}(\lambda)$ instead of $M_\psi(\lambda)$. Let

$$A_1 := (-\infty, 0),\ A_2 := (\lambda_0, b),\ A_3 := (0, \lambda_0).$$

Then

$$B_1 := (-\infty, b) = A_1 \cup A_2 \cup A_3,\ B_2 := (0, b) = A_2 \cup A_3,\ B_3 := (0, \lambda_0) = A_3.$$

Since A_1, A_2, A_3 are pairwise disjoint, by Lemma 3.3, we have

$$\begin{aligned}\det M_{l,w,d}(\lambda) &= -\frac{3}{2}\det\left[\int_{B_i}\frac{e(s,\lambda)\,ds}{(\lambda-s)^{j-1}}\right]_{i,j=1,2,3}\\ &= -\frac{3}{2}\det\left[\int_{A_i}\frac{e(s,\lambda)\,ds}{(\lambda-s)^{j-1}}\right]_{i,j=1,2,3} \neq 0.\end{aligned}$$

Since $r(\lambda) > 0$, this completes the proof. □

Proof of Theorem 1.1. As $\Psi(M_\lambda, \omega_\lambda) = \psi(\lambda)\Gamma_2$ for $\lambda \in (\lambda_0, b) \cup (b, a)$, in view of Corollary 2.11, for a.e. $\lambda \in (\lambda_0, a)$ the vertical flow on $(M_\lambda, \omega_\lambda) \in \mathcal{M}_2^{dc}$ is uniquely ergodic. Moreover, the billiard flow on $\mathcal{S}_\lambda$ is isomorphic (up to a linear rescaling of time) to the vertical flow on $(M_\lambda, \omega_\lambda)$ for every $\lambda \in (\lambda_0, b) \cup (b, a)$. This completes the proof for $\lambda \in (\lambda_0, a)$.

If $\lambda \in (0, \lambda_0)$ then, by Proposition 3.1, the billiard flow on $\mathcal{S}_\lambda$ is isomorphic to the directional flow in direction $\pi/4$ on the torus $\mathbb{R}^2/((2l(\lambda)\mathbb{Z}) \times (2w(\lambda)\mathbb{Z}))$. This flow is uniquely ergodic if and only if $w(\lambda)/l(\lambda)$ is irrational. The same argument as in the proof of the case **(E)** in Lemma 3.5 shows that $w'(\lambda)l(\lambda) - w(\lambda)l'(\lambda) \neq 0$ also for all $\lambda \in (0, \lambda_0)$. Therefore, the map $\lambda \mapsto \frac{w(\lambda)}{l(\lambda)}$ is strictly monotonic. It follows that $w(\lambda)/l(\lambda)$ is irrational for all but countably many parameters $\lambda \in (0, \lambda_0)$, which completes the proof. □

4. The beaviour of light rays in Eaton lenses systems

The application to the behaviour of light rays in periodic arrays of Eaton lenses exploits both the result on Birkhoff and Oseledets genericity (i.e. Theorem 2.10). The proof follows the arguments developed in the work by Frączek and Schmoll [20], which reduces the behaviour of rays (and specifically being trapped in a band) to a result on existence of Lyapunov exponents. Since Frączek and Schmoll in [20] were relying on the standard formulation of Oseledets genericity, they could only obtained a weaker result on random lattice configuration. The genericity results along curves that we prove in this paper, on the other hand, allows us to analyse the behaviour of any given admissible lattice configuration of lenses in almost every direction.

4.1. **Reduction to systems of flat lenses and periodic translation surfaces.** As in [20] we pass to a simpler model in which round lenses are replaced by their flat counterparts. By a flat lens of radius $R > 0$ perpendicular to the direction θ we mean simply any interval in $\mathbb{R}^2$ of length $2R$ perpendicular to vectors in direction θ. The light rays pass through the lens as follows: any light ray in direction θ or $\theta + \pi$ runs until hitting the flat lens and then is rotated by π around the center of the flat lens and runs in the opposite direction, see Figures 8. The system of flat lenses of radius R perpendicular to the direction θ whose centers are arranged at the points of the lattice Λ will be denoted by $F(\Lambda, R, \theta)$, see Figures 8. A simple observation (see [20] for details) shows that if a direction θ is trapped on $F(\Lambda, R, \theta)$ then it is also trapped on $L(\Lambda, R)$. Moreover, after rotation by $\pi/2 - \theta$ the system of flat lenses we can pass to vertically directed light rays. We take $v \in S^1$ to be the vertical unit vector. Denote by $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ the rotation (also the matrix of the rotation $r_\theta \in SL_2(\mathbb{R})$) by θ around the center of $\mathbb{R}^2$. Then

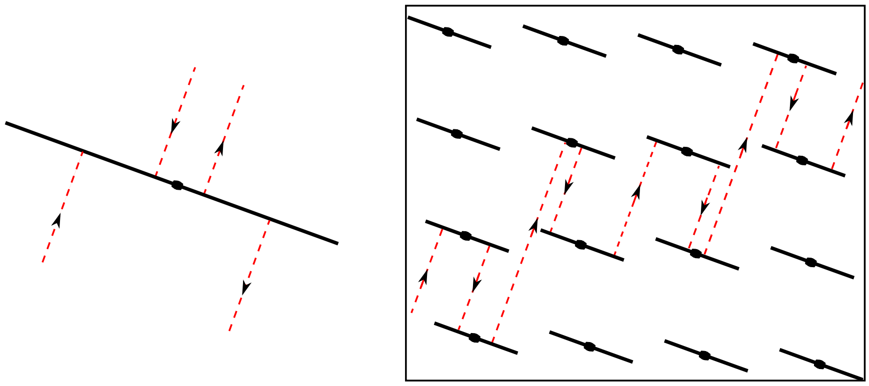

FIGURE 8. The system of lenses $F(\Lambda, R, \theta)$

$r_{\pi/2-\theta}F(\Lambda, R, \theta) = F(r_{\pi/2-\theta}\Lambda, R, v)$. In summary, Theorem 1.2 reduces to the following result.

Theorem 4.1. *Let* (Λ, R) *be an admissible pair. Then for a.e.* $\theta \in [0, 2\pi]$ *the vertical direction is trapped on* $F(r_\theta\Lambda, R, v)$*.*

In [20] (see Theorem 1.2), a weaker version of this result was proved saying that for every $R > 0$ and a.e. unimodular lattice Λ such that (Λ, R) is admissible the vertical direction is trapped on $F(\Lambda, R, v)$. Using a simple Fubini argument we also have the following seemingly stronger result which closely related to Theorem 1.2.

Corollary 4.2. *For every unimodular lattice* Λ*, a.e.* $0 < R < s(\Lambda)/2$ *and a.e.* $\theta \in [0, 2\pi]$ *the vertical direction on* $F(r_\theta\Lambda, R, v)$ *is trapped. In particular, for every unimodular lattice* Λ *and a.e.* $0 < R < s(\Lambda)/2$ *almost every direction* $\theta \in [0, 2\pi]$ *on* $L(\Lambda, R)$ *is trapped.*

Proof. Fix a unimodular lattice Λ_0 and $0 < R_0 < 1/\sqrt{2\sqrt{3}}$. Suppose contrary to our claim that there exists a set $A \subset (0, s(\Lambda_0)/2) \times [0, 2\pi]$ not of zero Lebesgue measure such that for every $(R, \theta) \in A$ the vertical direction is not trapped on $F(r_\theta\Lambda_0, R, v)$. For every $t \in R$ let $h(t) := \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$. Then for every admissible (Λ, R)

$$v \text{ is trapped on } F(\Lambda, R, v) \Leftrightarrow v \text{ is trapped on } a_sF(\Lambda, R, v) = F(a_s\Lambda, e^sR, v)$$

for all $s \in \mathbb{R}$ and

$$v \text{ is trapped on } F(\Lambda, R, v) \Leftrightarrow v \text{ is trapped on } F(h(t)\Lambda, R, v)$$

for all $t \in (-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. The two equivalences essentially follows from the fact that both operations does not change horizontal positions between flat lenses (up to rescaling in the first case). Therefore, $(R, \theta) \in A$ implies that v is not trapped on $F(h(t)a_{\log(R_0/R)}r_\theta\Lambda_0, R_0, v)$ for all $t \in (-\varepsilon, \varepsilon)$. Since the map

$$(t, s, \theta) \mapsto h(t)a_{\log(R_0/s)}r_\theta\Lambda_0$$

locally is a diffeomorphism between $\mathbb{R} \times \mathbb{R} \times [0, 2\pi]$ and the space of unimodular lattices $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$, the fact that $A \subset \mathbb{R}^2$ is not of measure zero together with a Fubini argument yield a set $\mathcal{A} \subset SL_2(\mathbb{R})/SL_2(\mathbb{Z})$ not of measure zero such that $\Lambda \in \mathcal{A}$ implies that v is not trapped on $F(\Lambda, R_0, v)$. This contradicts Theorem 1.2 in [20]. The second part of the corollary follows immediately from the first. $\square$

4.2. **Light behaviour vs Oseledets genericity.** In this section we relate the trapping phenomenon for light rays in almost all directions with Oseledets genericity along a curve of translation surfaces. We rely for this section on some basic steps of the proof of Theorem 4.1 which were developed in [20] and a technical result recently proved in [19].

Suppose that the pair (Λ, R) is admissible (Λ is unimodular) and let us consider flat lenses system $F(\Lambda, R, v)$. By an unfolding procedure similar to the one used in § 3, the system of flat lenses $F(\Lambda, R, v)$ can be reduced to a noncompact periodic translation surface, which is a $\mathbb{Z}^2$-cover of a compact translation surface $M(\Lambda, R) \in \mathcal{M}_2^{dc}$ obtained as follows (we refer the reader to [20] for further details). Take the translation torus $\mathbb{T}^2_\Lambda := \mathbb{R}^2/\Lambda$ with a horizontal interval (slit) $I \subset \mathbb{R}^2/\Lambda$ of length $2R$. Since (Λ, R) is admissible, I has no self-intersections. Next take two copies of such slitted torus and glue them together so that the bottom part of the slit on one torus is glued by translation to its top counterpart on the other torus, see Figure 9. Then

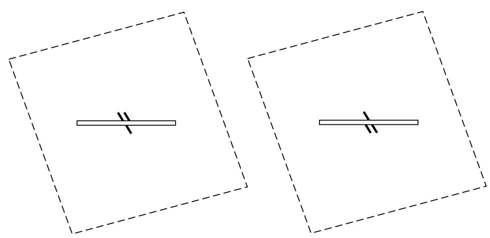

FIGURE 9. The surface $M(\Lambda, R)$

$$\Psi(M(\Lambda, R)) = (h, (2R, 0))\Gamma_2, \quad \text{where} \quad \Lambda = h(\mathbb{Z}^2). \tag{4.1}$$

It turns out that the behaviour of the vertical trajectories in $F(\Lambda, R, v)$ (and hence in $L(\Lambda, R)$) can be described by studying the asymptotic behaviour of the following homology classes. For every translation surface (M, ω) and $x \in M$ such that its positive vertical semi-orbit $\{\psi_t^v x : t \geq 0\}$ is well defined and for every $t > 0$ denote by $\sigma_t(x) \in H_1(M, \mathbb{Z})$ the homology class of the curve formed by the segment of the vertical orbit starting from x until time t closed up by the shortest curve returning to x. Then we have the following.

Proposition 4.3 (Theorem 3.2 in [20])**.** *Let $\tau : M(\Lambda, R) \to M(\Lambda, R)$ be the only nontrivial element of the deck transformation group of the double cover. Suppose that there is a non-zero homology class $\zeta \in H_1(M(\Lambda, R), \mathbb{R})$ and $C > 0$ such that*

$$\tau_*\zeta = -\zeta \text{ and } |\langle \sigma_t(x), \zeta \rangle| \leq C \text{ for all } x, t \text{ for which } \sigma_t(x) \text{ is defined.}$$

If the surface $M(\Lambda, R)$ has no vertical saddle connection, i.e. there is no vertical orbit segment that connect singular points, then the vertical direction on $F(\Lambda, R, v)$ is trapped.

Oseledets genericity plays a central role in verifying this type of assumptions in Proposition 4.3. The technical tool which allows us to check the assumptions in our specific context will be provided by the following general proposition recently proved by Frączek and Hubert in [19]. We should mention that the idea behind this result is not new (can be found also in [21] and [8]) and exploits the phenomenon of bounded deviation discovered by Zorich in [36, 37].

Proposition 4.4 ([19])**.** *Let $\mathcal{M} \subset \mathcal{M}_1(M, \Sigma)$ be an $SL_2(\mathbb{R})$-orbit closure. Suppose that $W \subset H_1(M, \mathbb{R})$ is a symplectic subspace which is $SL_2(\mathbb{R})$-invariant over $\mathcal{M}$. Suppose that the top Lyapunov exponent of the restricted cocycle $A_W^{KZ} : \mathbb{R} \times \mathcal{M} \to$*

$GL(W)$ is positive. If (M,ω) is Birkhoff and Oseledets generic with respect to $(\mathcal{M},\mu_{\mathcal{M}},a_t)$ then for every $\zeta\in E^-_\omega$ where

$$E^-_\omega=\Big\{v\in W:\lim_{t\to\infty}\frac{\log\|A^{KZ}_W(t,\omega)v\|}{t}<0\Big\},$$

there exists $C>0$ such that

$$|\langle\sigma_t(x),\zeta\rangle|\le C \text{ for all } x\in M, t>0 \text{ for which } \sigma_t(x) \text{ is defined.}$$

Proof of Theorems 4.1 and 1.2. Let (Λ,R) be a admissible pair and let $\Lambda=h(\mathbb{Z}^2)$ for some $h\in SL_2(\mathbb{R})$. As we have already mentioned in the beginning of § 4.1 that Theorem 1.2 reduces to Theorem 4.1, so we need to prove that for a.e. $\theta\in[0,2\pi]$ the vertical direction is trapped in $F(r_\theta\Lambda,R,v)$.

Let us consider the restricted Kontsevich-Zorich cocycle $A^{KZ}_W:\mathbb{R}\times\mathcal{M}^{dc}_2\to GL(W)$, where

$$W:=\{\zeta\in H_1(M,\mathbb{R}):\tau_*\zeta=-\zeta\}$$

and $\tau:M\to M$ is the only nontrivial element of the deck transformation group of the double cover. The subspace W is two dimensional symplectic and $SL_2(\mathbb{R})$-invariant. Let us consider the curve $[0,2\pi]\ni\theta\mapsto M(r_\theta\Lambda,R)\in\mathcal{M}^{dc}_2$. In view of (4.1),

$$\Psi(M(r_\theta\Lambda,R))=(r_\theta h,(2R,0))\Gamma_2.$$

Taking $\psi(\theta)=(r_\theta,(2R,0))$, we have that the determinant of

$$M_\psi(\theta)=\begin{bmatrix}\cos\theta & -\sin\theta & 2R\\ -\sin\theta & \cos\theta & 0\\ \cos\theta & \sin\theta & 0\end{bmatrix}$$

is non-zero. Therefore, by Theorem 2.10, for a.e. $\theta\in[0,2\pi]$ the surface $M(r_\theta\Lambda,R)$ is Birkhoff and Oseledets generic. Fix such $\theta\in[0,2\pi]$. Recall that all Lyapunov exponents for all invariant measures supported on the space of genus two surfaces are positive, see e.g. [4]. It follows that the top Lyapunov exponent of the reduced cocycle A^{KZ}_W is positive (in fact, is equal to 1/2). In view of Proposition 4.4, it follows that there exists non-zero $\zeta\in W$ and $C>0$ such that $|\langle\sigma_t(x),\zeta\rangle|\le C$ for all $x\in M(r_\theta\Lambda,R)$, $t>0$ for which $\sigma_t(x)$ is defined.

By Birkhoff genericity, the vertical flow on $M(r_\theta\Lambda,R)$ has no saddle connection. Therefore, by Proposition 4.3, the vertical direction is trapped on $F(r_\theta\Lambda,R,v)$. The result for Eaton model follows from the correspondence between trapped direction in $F(\Lambda,R,v)$ and $L(\Lambda,R)$. □

5. Gap distribution of square root of integers

In this section, we give the proof of Theorem 1.3, based on our Birkhoff genericity result for curves (Theorem 2.1). Our proof follows the same idea of the proof by Elkies and McMullen in [11].

Let us recall that a sequence of real numbers $\{t_n\}_{n\ge1}$ *converges in distribution* to a probability measure μ on $\mathbb{R}$ if for every $f\in C_c(\mathbb{R})$ we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}f(t_n)=\int_{\mathbb{R}}f(s)\,d\mu(s).$$

Convergence in distribution to μ is equivalent to

$$\lim_{N\to\infty}\frac{1}{N}\#\{1\le n\le N:t_n\in(-\infty,b]\}\to\mu((-\infty,b]) \tag{5.1}$$

for every $b\in\mathbb{R}$ with $\mu(\{b\})=0$. Let us first prove two simple Lemmas.

Lemma 5.1. *Let $\{t_n\}_{n\geq 1}, \{l_n\}_{n\geq 1}$ be two sequences of positive real numbers. Suppose that $\lim_{n\to\infty} \frac{t_n}{l_n} = 1$ and $\{t_n\}_{n\geq 1}$ converges in distribution to a probability measure μ, then so does $\{l_n\}_{n\geq 1}$.*

Proof. It suffices to show that for every $f \in C_c(\mathbb{R})$ one has

$$\lim_{n\to\infty} |f(t_n) - f(l_n)| = 0. \tag{5.2}$$

Given $\varepsilon > 0$, there exists $\delta > 0$ such that if $0 \leq s, t \leq \delta$, then $|f(s) - f(t)| < \varepsilon$. According to the assumption, there exits integer $N_1 > 0$ such that for $n \geq N_1$ either $t_n, l_n \leq \delta$ or $t_n, l_n \geq \frac{\delta}{2}$. Since f has compact support, there exists $N_2 > 0$ such that if $n \geq N_2$ and $t_n, l_n \geq \frac{\delta}{2}$ then $|f(t_n) - f(l_n)| < \varepsilon$. If follows that $|f(t_n) - f(l_n)| < \varepsilon$ provided that $n \geq \max\{N_1, N_2\}$. Therefore (5.2) holds and the proof is complete. □

Lemma 5.2. *Let $\{t_n\}, \{s_n\}, \{l_n\}$ be sequences of real numbers. Suppose that $s_n \leq t_n \leq l_n$ for every $n \geq 1$ and $\{s_n\}, \{l_n\}$ converge in distribution to a probability measure μ on $\mathbb{R}$, then so does $\{t_n\}$.*

Proof. The lemma follows from the usual squeeze lemma for sequences of real numbers and (5.1). □

Let us now turn to the proof of Theorem 1.3 using the strategy in [11]. Recall that $L_r(s)$ defined in (1.2) is the normalized gap containing s of fractional parts of integers up to r. We approximate $L_r(s)$ for $s > 0$ by the maximal area $L'_r(s)$ of $\triangle ABC$ (see Figure 10) with the following properties: B, C moves on the line $y = \sqrt{r} - s$, B, C are above and below the line $x = 2sy + s^2$ respectively, the interior of $\triangle ABC$ contains no points of $\mathbb{Z}^2$. We remark here that if the line segment of $x = 2sy + s^2$ contained in $\triangle ABC$ has a lattice point, then $L'_r(s) = 0$. For an affine lattice Λ we let $f(\Lambda)$ be the maximal area of triangles with the following properties: triangle's interior contains the line segment $\{0\} \times [0, 1]$ but no lattice points of Λ; triangle has one vertex $(0, 0)$ and the other vertices lie on the line $y = 1$.

Recall that $u(\cdot, \cdot, \cdot), a_t$ are defined in (2.3) and μ_X is the probability Haar measure on $X = ASL_2(\mathbb{R})/ASL_2(\mathbb{Z})$. It is noticed in [11] that

$$L'_r(s) = f(a_{\log\sqrt{r}} u(-2s, -s^2, s)\mathbb{Z}^2). \tag{5.3}$$

Let $F : [0, \infty) \to [0, \infty)$ be the piecewise analytic density function given by (1.1). It is proved in [11, §3.3] that

$$\int_0^l tF(t)\,dt = \mu_X(\Lambda \in X : f(\Lambda) \leq l).$$

Lemma 5.3. *Let $\{t_n\}_{n\in\mathbb{N}}$ be a sequence of natural numbers such that $\sum_{n=1}^{\infty} t_n^{-1} < \infty$. Then for Lebesgue almost every $s \in [0, 1]$ we have*

$$\lim_{n\to\infty} \frac{L_{t_n^2}(s)}{L'_{t_n^2}(s)} = 1.$$

Proof. Let $a, A > 1$ be integers. According to [11, Lemma 3.3 and 3.7] the set

$$\left\{\frac{1}{a-1} \leq s \leq 1 - \frac{1}{a-1} : \frac{2A+1}{2A+2} L'_{a^2}(s) \leq L_{a^2}(s) \leq \frac{2A+1}{2A} L'_{a^2}(s)\right\}$$

has Lebesgue measure

$$\geq 1 - \frac{(A+2)(A-1)+2}{a-1}.$$

The conclusion follows from the assumption $\sum_{n=1}^{\infty} t_n^{-1} < \infty$ and the Borel-Cantelli lemma.

□

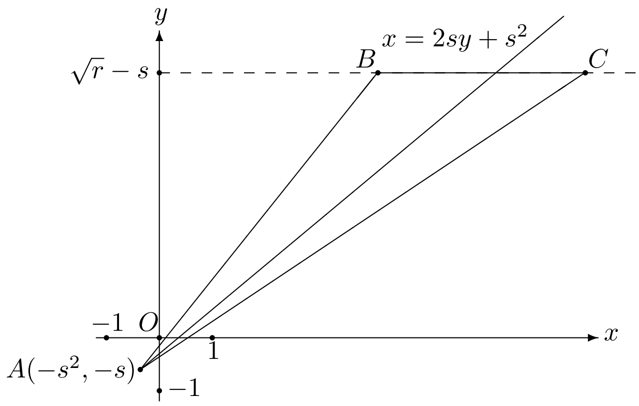

FIGURE 10.

Proof of Theorem 1.3. For every $r \geq 1$ and $s \in [0,1]$ it can be checked directly that

$$\frac{\lfloor r \rfloor}{(\lfloor \sqrt{r} \rfloor + 1)^2} L_{(\lfloor \sqrt{r} \rfloor+1)^2}(s) \leq L_r(s) \leq \frac{\lfloor r \rfloor}{\lfloor \sqrt{r} \rfloor^2} L_{\lfloor \sqrt{r} \rfloor^2}(s). \tag{5.4}$$

Note that coefficients of left and right hand sides of (5.4) converge to 1 as $r \to \infty$.

Let $r_n := cq^n$ and set

$$t_n(s) := L'_{(\lfloor \sqrt{r_n} \rfloor + 1)^2}(s), \; l_n(s) := L'_{\lfloor \sqrt{r_n} \rfloor^2}(s).$$

By Lemma 5.1, 5.2 and Lemma 5.3, it suffices to prove that for Lebesgue almost every $s \in [0,1]$ the sequences $\{t_n(s)\}_{n \in \mathbb{N}}$ and $\{l_n(s)\}_{n \in \mathbb{N}}$ converge in distribution to $\mu = tF(t)\,dt$ on $[0,\infty)$. By (5.3),

$$t_n(s) = f(a_{\log(\lfloor \sqrt{cq^n} \rfloor + 1)} u(-2s, -s^2, s)\mathbb{Z}^2) = f(a_{nt+t_0+c_n^-} u(-2s, -s^2, s)\mathbb{Z}^2)$$
$$l_n(s) = f(a_{\log \lfloor \sqrt{cq^n} \rfloor} u(-2s, -s^2, s)\mathbb{Z}^2) = f(a_{nt+t_0+c_n^+} u(-2s, -s^2, s)\mathbb{Z}^2)$$

where $t_0 = \log \sqrt{c}, t = \log \sqrt{q}$ and

$$c_n^- = \log \frac{\lfloor \sqrt{cq^n} \rfloor + 1}{\sqrt{cq^n}} \to 0, \quad c_n^+ = \log \frac{\lfloor \sqrt{cq^n} \rfloor}{\sqrt{cq^n}} \to 0 \quad \text{as } n \to \infty.$$

For every $l \geq 0$ let $E_l = \{\Lambda \in X : f(\Lambda) \leq l\}$ and let $\mathbb{1}_{E_l}$ be the indicator function of $E_l \subset X$. Using Corollary 2.2 and the fact that μ_X is the unique $\{a_t\}$-invariant probability measure on X with maximal entropy, it can be proved that for a.e. $s \in [0,1]$ and any $\varphi \in C_c(X)$ one has

$$\frac{1}{N} \sum_{n=1}^{N} \varphi(a_{(nt+t_0+c_n^{\pm})} u(-2s, -s^2, s)\Gamma) = \int_X \varphi \, d\mu_X.$$

It is proved in [11, Proposition 3.9] that $\mu_X(\partial E_l) = 0$. Therefore, for a.e. $s \in [0,1]$ we have

$$\frac{1}{N} \#\{1 \leq n \leq N : f(a_{(nt+t_0+c_n^{\pm})} u(-2s, -s^2, s)\mathbb{Z}^2) \leq l\}$$
$$= \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{E_l}(a_{(nt+t_0+c_n^{\pm})} u(-2s, -s^2, s)\Gamma) \to \mu_X(E_l) = \int_0^l tF(t)\,dt.$$

Therefore, for a.e. $s \in [0,1]$ the sequences $\{t_n(s)\}_{n \in \mathbb{N}}$ and $\{l_n(s)\}_{n \in \mathbb{N}}$ converge in distribution to μ, which completes the proof. $\square$

6. Proof of the Birkhoff genericity along curves

The aim of this section is to prove Theorem 2.1. We first setup some notations that will be used throughout this section. Let $G, \Gamma, X, a_t, u, u_\varphi$ be as in § 2.1 and φ be as in Theorem 2.1. For convenience we let $u(a) = u(a, 0, 0)$ and $u(a, b) = u(a, b, 0)$. The following notation will also be used during the proof:

$$H = \{(h, 0) : h \in SL_2(\mathbb{R})\} \leq G$$
$$U = \{u(t) : t \in \mathbb{R}\}$$
$$D = \{a_t : t \in \mathbb{R}\}$$
$$\rho : G \to SL_2(\mathbb{R}) \quad \text{which maps } (h, v) \to h.$$

We also denote by ρ the induced map of homogeneous spaces

$$\rho : G/\Gamma \to SL_2(\mathbb{R})/SL_2(\mathbb{Z}), \quad \rho(g\Gamma) = \rho(g)SL_2(\mathbb{Z}).$$

For every $s \in [0, 1]$ and $T > 0$ let $\mu_{s,T}$ be a probability measure on X given by

$$\mu_{s,T} = \frac{1}{T} \int_0^T \delta_{a_t u_\varphi(s)\Gamma} \, dt$$

where $\delta_{a_t u_\varphi(s)\Gamma}$ is the point mass measure on $a_t\, u_\varphi(s)\Gamma$. All the constants depending on φ will not be specified in this section.

6.1. **Outline of the proof of Theorem 2.1.**

The proof consists of two parts. We first consider any weak* limit of $\mu_{s,T}$ as $T \to \infty$ and prove that it is a measure on X invariant under the unipotent subgroup U (see section 6.2). It is easy to see that any weak* limit μ is also invariant under the group D. According to a result by Mozes in [29, Theorem 1], a finite measure on X invariant under the group DU is automatically H-invariant. Hence, one can apply the celebrated Ratner's Theorem (see [30]), which gives that any H-invariant and ergodic probability measure on X is homogeneous and supported on some orbit closure of H. The H-orbits are well known: each orbit is either closed or dense. To prove Theorem 2.1 we need a precise description of closed orbits, which is given at the beginning of §6.3. The main result to prove the second part is Proposition 6.4. In §6.3, assuming Proposition 6.4, we conclude the proof of Theorem 2.1. The proof of Proposition 6.4, which is rather long and technical, is postponed to the last two sections, i.e. §6.5, where a suitable mixed height function is constructed, and §6.6, where the height function is used to show that there is no accumulation of mass on closed orbits. Before the proof of Proposition 6.4, in § 6.4, we deduce from Theorem 2.1 the Birkhoff genericity result for more general curves, i.e. Corollary 2.2.

6.2. **Unipotent invariance.**

In this section we prove the unipotent invariance of every weak* limit of $\mu_{s,T}$ as $T \to \infty$. The methods of the proof are inspired by the work of Eskin-Chaika [7] and extends their technique to more general curves in our setup.

Proposition 6.1. *Let $\varphi : [0, 1] \to \mathbb{R}$ be a C^1-function. Then for a.e. $s \in [0, 1]$ any weak* limit μ of $\mu_{s,T}$ as $T \to \infty$ is invariant under the group U.*

Proof. For every function $f \in C_c^\infty(X)$ and every $r \in \mathbb{R}$ we let

$$f_r(t, s) = f(u(r)a_t u_\varphi(s)\Gamma) - f(a_t u_\varphi(s)\Gamma). \tag{6.1}$$

Fix a countable set $\mathcal{S} \subset C_c^\infty(X)$ which is dense in $C_c(X)$ with respect to sup-norm. Then it suffices to show that for every $f \in \mathcal{S}$ and $r \in \mathbb{R}$ we have

$$\frac{1}{T} \int_0^T f_r(t, s) \, dt \to 0 \quad \text{for a.e. } s \in [0, 1]. \tag{6.2}$$

By the proof of Lemma [7, Lemma 3.4] (cf. [33, Lemma 3.3]), to prove (6.2) it suffices to show that

$$\Big|\int_0^1 f_r(t,s)f_r(l,s)\,ds\Big| \le M(f,r)e^{-|l-t|}, \tag{6.3}$$

where $M(f,r)$ is a constant depending on f and r. Since in (6.3) the roles of l and t are symmetric we assume that $l>t$. Then for any

$$s\in I(s_0)=[s_0-e^{-l-t}, s_0+e^{-l-t}]\subset[0,1]$$

we have

$$\begin{aligned} a_tu_\varphi(s) &= a_tu(s-s_0,\varphi(s)-\varphi(s_0))a_{-t}\cdot a_tu_\varphi(s_0)\\ &= u\big(e^{2t}(s-s_0)), e^t\varphi'(\tau(s))(s-s_0)\big)\cdot a_tu_\varphi(s_0)\\ &= u(O(e^{-l+t}), O(e^{-l+t}))a_tu_\varphi(s_0), \end{aligned} \tag{6.4}$$

where $\tau(s)$ is a real number determined by the mean value theorem. Since $f:X\to\mathbb{R}$ is compactly supported there exists $C>0$ such that

$$|f(g_1\Gamma)-f(g_2\Gamma)|\le C\|(Id,0)-g_2g_1^{-1}\| \quad \text{for all} \quad g_1,g_2\in G, \tag{6.5}$$

where if $g_2g_1^{-1}=(h,v)$ then $\|(Id,0)-g_2g_1^{-1}\|=\|Id-h\|+\|v\|$. It follows from the definition of $f_r(t,s)$ in (6.1), (6.4) and the smoothness of f that

$$f_r(t,s)-f_r(t,s_0)=O_{f,r}(e^{-l+t}).$$

Therefore

$$\int_{I(s_0)} f_r(t,s)f_r(l,s)\,ds = f_r(t,s_0)\int_{I(s_0)} f_r(l,s)\,ds + |I(s_0)|O_{f,r}(e^{-l+t}). \tag{6.6}$$

On the other hand

$$\begin{aligned} u(r)a_lu_\varphi(s) &= a_lu(0,\varphi(s)-\varphi(s+re^{-2l}))a_{-l}\cdot a_lu_\varphi(s+re^{-2l})\\ &= u(0,-re^{-l}\varphi'(s+\tau(r)e^{-2l}))\cdot a_lu_\varphi(s+re^{-2l})\\ &= u(0,O_r(e^{-l+t}))a_lu_\varphi(s+re^{-2l}), \end{aligned}$$

where $\tau(r)$ is determined by the mean value theorem. According to (6.5), it follows that

$$f_r(l,s)=f(a_lu_\varphi(s+re^{-2l})\Gamma)-f(a_lu_\varphi(s)\Gamma)+O_{f,r}(e^{-l+t}),$$

and hence

$$\int_{I(s_0)} f_r(l,s)\,ds = O_{f,r}(e^{-l+t})|I(s_0)|. \tag{6.7}$$

In view of (6.6) and (6.7), we have

$$\int_{I(s_0)} f_r(t,s)f_r(l,s)\,ds = O_{f,r}(e^{-l+t})|I(s_0)|. \tag{6.8}$$

Now (6.3) follows by splitting $[0,1]=\bigcup_{1\le k\le m} I_k$ into interval $I_k=[s_{k-1},s_k]$ with $s_k-s_{k-1}=2e^{-l-t}$ for $1\le k<m$ and $s_m-s_{m-1}\le 2e^{-l-t}$, and then by applying (6.8) to interval I_k for $1\le k<m$. $\square$

6.3. **Proof of Theorem 2.1 assuming Proposition 6.4.** For every $h \in SL_2(\mathbb{R})$ the fiber
$$\rho^{-1}(h\, SL_2(\mathbb{Z})) \cong \mathbb{R}^2/\mathbb{Z}^2,$$
as $(h, v_1)\Gamma = (h, v_2)\Gamma$ if and only if $v_1 = v_2 + hv_0$ for some $v_0 \in \mathbb{Z}^2$. If the H-orbit of $x = (Id, v)\Gamma \in X$ is closed, then
$$\rho^{-1}(SL_2(\mathbb{Z})) \cap Hx = \{(Id, \gamma v)\Gamma : \gamma \in SL_2(\mathbb{Z})\}$$
is closed. In view of [23, Theorem 2], the $SL_2(\mathbb{Z})$-orbit of $v+\mathbb{Z}^2 \in \mathbb{R}^2/\mathbb{Z}^2$ is finite if $v \in \mathbb{Q}^2$ and dense otherwise. Therefore the closed H-orbits in X are exactly orbits of (Id, v) with $v \in \mathbb{Q}^2$.

For every $n \in \mathbb{N}$ and $i, j \in \mathbb{Z}$ we let
$$G[n]^{ij} = \{(h, h\,(i/n, j/n)) : h \in SL_2(\mathbb{R})\} \subset G.$$
Let $X[n]$ be the image of $\bigcup_{i,j\in\mathbb{Z}} G[n]^{ij} \subset G$ in X. Then $X[n]$ is a finite union of closed H-orbits and any closed H-orbit is contained in some $X[n]$. Therefore it suffices to show that for almost every $s \in [0,1]$ any weak* limit μ of $\mu_{s,T}$ as $T \to \infty$ puts no mass on $X[n]$.

Proposition 6.2. *Under the assumptions of Theorem 2.1, for a.e.* $s \in [0,1]$ *any weak* limit* μ *of* $\mu_{s,T}$ *as* $T \to \infty$ *is a probability measure on* X *and satisfies* $\mu(X[n]) = 0$ *for every positive integer* n*.*

Remark 6.3. Let
$$M_1 = 2 \sup_{s\in[0,1]} |\varphi'(s)| + 1. \tag{6.9}$$
According to the assumption for φ the set
$$\{s \in [0,1] : \varphi(s) = js/n + i/n \text{ for some } |j/n| \le M_1\}$$
has Lebesgue measure zero. Since this set is closed, there exist at most countably many open intervals $\{I_k\}$ such that:

- elements of $\{I_k\}$ are pairwise disjoint;
- $\bigcup I_k$ has full measure in $[0,1]$;
- for any $s \in I_k$ we have $\varphi(s) \neq js/n + i/n$ if $|j/n| \le M_1$.

Note that each I_k is a countable union of closed intervals. So it suffices to show that Proposition 6.2 holds for every closed interval I contained in some I_k.

Let K be a measurable subset of X, $T > 0$ and $s \in [0,1]$. The proportion of the trajectory $\{a_t u_\varphi(s)\Gamma : 0 \le t \le T\}$ in K is expressed by the function $\mathcal{A}_K^T : [0,1] \to [0,1]$ defined by
$$\mathcal{A}_K^T(s) := \frac{1}{T}\int_0^T \mathbb{1}_K(a_t u_\varphi(s)\Gamma)\, dt = \frac{1}{T}\int_0^T \delta_{a_t u_\varphi(s)\Gamma}(K)\, dt = \mu_{s,T}(K),$$
where $\mathbb{1}_K$ is the characteristic function of K. Proposition 6.2 follows from the following quantitative result.

Proposition 6.4. *Let* I *be a closed interval of* $(0,1)$ *and let* $n \in \mathbb{N}$*. Suppose that*
$$\inf\{|\varphi(s) - js/n + i/n| : s \in I, |j/n| \le M_1, i \in \mathbb{Z}\} = \sigma > 0. \tag{6.10}$$
Then for any $\varepsilon > 0$ *there exists a compact subset* K *of* $X \setminus X[n]$ *and* $\vartheta > 0$ *such that for any* $T > 0$
$$|\{s \in I : \mathcal{A}_K^T(s) \le 1 - \varepsilon\}| \le e^{-\vartheta T}|I|. \tag{6.11}$$

The proof of Proposition 6.4 will be given in § 6.5 and §6.6. Here we use it to derive Proposition 6.2 and Theorem 2.1.

Proof of Proposition 6.2. Fix $n \in \mathbb{N}$. By Remark 6.3, it suffices to show that the conclusion holds for almost every $s \in I$ whenever I is a closed interval as in Proposition 6.4. Given $0 < \varepsilon < 1$ we choose a compact subset $K_\varepsilon \subset X \setminus X[n]$ and $\vartheta > 0$ so that (6.11) holds. By taking $T = m$ for $m \in \mathbb{N}$ in (6.11) and using the Borel-Cantelli lemma we can find a subset I_ε of I with the following property: I_ε has full measure in I and for any $s \in I_\varepsilon$ any weak* limit μ of $\mu_{s,T}$ as $T \to \infty$ satisfies $\mu(X) \geq \mu(K_\varepsilon) \geq 1 - \varepsilon$. Therefore, $\mu(X[n]) < \varepsilon$. It follows that for any s in the full measure subset $\bigcap_{k\in\mathbb{N}} I_{\frac{1}{k}}$ of I we have $\mu(X) = 1$ and $\mu(X[n]) = 0$. □

Proof of Theorem 2.1. We claim that there exists a full measure subset I' of $[0, 1]$ such that for any $s \in I'$ any weak* limit μ of $\mu_{s,T}$ as $T \to \infty$ has the following properties:

(i) μ is a probability measure;
(ii) μ is invariant under the group U;
(iii) $\mu(X[n]) = 0$ for any positive integer n.

The claims (i) and (iii) follow from Proposition 6.2 and (ii) follows from Proposition 6.1.

Since μ is U-invariant, it is also H-invariant, by [29, Theorem 1]. According to Ratner's measure classification theorem any ergodic H-invariant probability measure on X is either μ_X or supported on some $X[n]$. Therefore claim (iii) implies that $\mu = \mu_X$. Since μ is an arbitrary weak* limit, it follows that $\mu_{s,T} \to \mu_X$ as $T \to \infty$ for any $s \in I'$. □

6.4. **Birkhoff genericity for more general curves.** We conclude this section by considering some extensions of Theorem 2.1. In particular, we give the proof of Corollary 2.2.

Corollary 6.5. *Let Γ' be a lattice in G commensurable with Γ. Let $\varphi : [0, 1] \to \mathbb{R}$ be a C^1-function, $h \in SL_2(\mathbb{R})$ and $v = (v_1, v_2)^{\mathrm{tr}} \in \mathbb{R}^2$. Suppose for any $(b, l)^{\mathrm{tr}} \in h\mathbb{Q}^2$ one has the Lebesgue measure of $\{s \in [0, 1] : \varphi(s) = (l - v_2)s + (b - v_1)\}$ is zero. Then for almost every $s \in [0, 1]$ the coset $u_\varphi(s)(h, v)\Gamma'$ is Birkhoff generic with respect to $(G/\Gamma', \mu_{G/\Gamma'}, a_t)$.*

Remark 6.6. Let us consider $h = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \in SL_2(\mathbb{R})$ with $h_{11} \neq 0$ and $v = (v_1, v_2)^{\mathrm{tr}} \in \mathbb{R}^2$. Then for any $x \in G/\Gamma'$ the forward trajectories of $(h, v)x$ and $u(\frac{h_{12}}{h_{11}}, \frac{v_1}{h_{11}})x$ with respect to the action of D are asymptotically parallel, i.e. there exists $g \in G$ such that the distance between $a_t(h, v)x$ and $ga_tu(\frac{h_{12}}{h_{11}}, \frac{v_1}{h_{11}})x$ tends to zero as $t \to \infty$. Therefore if one of them equidistributes so does the other. This observation and the description of closed H-orbits before Proposition 6.2 explains the assumption of φ in Theorem 2.1 and Corollary 6.5.

Lemma 6.7. *Let Γ' be a lattice in G commensurable with Γ. An element $(g, v)\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t) if and only if $(g, v)\Gamma'$ is Birkhoff generic with respect to $(G/\Gamma', \mu_{G/\Gamma'}, a_t)$.*

Proof. In the proof we will not use the fact that $\Gamma = ASL_2(\mathbb{Z})$, we only need to use commensurability of Γ and Γ', so without loss of generality we can assume that $\Gamma' \leq \Gamma$. Let μ and μ' be weak* limits of

$$\frac{1}{T}\int_0^T \delta_{a_t(h,v)\Gamma}\,dt \quad\text{and}\quad \frac{1}{T}\int_0^T \delta_{a_t(h,v)\Gamma'}\,dt \tag{6.12}$$

respectively along the same sequence $\{T_n\}$ with $T_n \to \infty$. It suffices to show that $\mu = \mu_{G/\Gamma}$ if and only if $\mu' = \mu_{G/\Gamma'}$.

Let us consider the natural projection $\pi : G/\Gamma' \to G/\Gamma$ which is a finite covering map. Since μ and μ' are weak* limits of (6.12) along the same sequence, it is easy to see that μ is a probability measure if and only if so is μ' and then $\mu = \pi_*(\mu')$.

Suppose that $\mu' = \mu_{G/\Gamma'}$. Since the map $\pi : G/\Gamma' \to G/\Gamma$ is G-equivalent, the measure μ is also G-invariant and probabilistic so it is equal to $\mu_{G/\Gamma}$.

The other direction (assuming $\mu = \mu_{G/\Gamma}$) is proved using entropy theory and we refer the readers to [10] for backgrounds. Since μ and μ' are D-invariant, π yields a factor map between $(G/\Gamma', a_1, \mu')$ and $(G/\Gamma, a_1, \mu)$. Therefore $h_{\mu'}(a_1) \geq h_\mu(a_1)$. The conclusion that $\mu' = \mu_{G/\Gamma'}$ follows from the following two facts, c.f. [10, §7]: (i) $h_{\mu_{G/\Gamma}}(a_1) = h_{\mu_{G/\Gamma'}}(a_1)$; (ii) $h_{\mu'}(a_1) \leq h_{\mu_{G/\Gamma'}}(a_1)$ and the equality holds if and only if $\mu' = \mu_{G/\Gamma'}$. □

Proof of Corollary 6.5. According to the observation in Remark 6.6 and Lemma 6.7 it suffices to show that for a.e. s the coset

$$u\left(\frac{h_{12} + h_{22}s}{h_{11} + h_{21}s}, \frac{v_2 s + v_1 + \varphi(s)}{h_{11} + h_{21}s}\right)\Gamma$$

is Birkhoff generic. By Theorem 2.1 it suffices to check that for any $\tilde{b}, \tilde{l} \in \mathbb{Q}$ the set

$$\{s \in [0,1] : v_2 s + v_1 + \varphi(s) = \tilde{l}(h_{12} + h_{22}s) + \tilde{b}(h_{11} + h_{21}s)\}$$

has Lebesgue measure zero. This follows from the assumption with $l = h_{21}\tilde{b} + h_{22}\tilde{l}$ and $b = h_{11}\tilde{b} + h_{12}\tilde{l}$. □

Proof of Corollary 2.2. According to Remark 6.6 and Lemma 6.7 it suffices to show that for a.e. $s \in [0,1]$ the coset

$$u\left(\frac{h_{12}(s)}{h_{11}(s)}, \frac{v_1(s)}{h_{11}(s)}\right)(h, v)\Gamma \tag{6.13}$$

is Birkhoff generic. First we show that the closed sets

$$\begin{aligned}
I_{l,b} &= \{s \in [0,1] : v_1(s) = l h_{12}(s) + b h_{11}(s)\}, \quad \text{where} \quad l, b \in \mathbb{R} \\
I_1 &= \{s \in [0,1] : h_{11}(s) = 0\} \quad \text{and} \\
I_2 &= \{s \in [0,1] : h_{11}(s) h'_{12}(s) - h_{12}(s) h'_{11}(s) = 0\}
\end{aligned}$$

have Lebesgue measure zero. Indeed, since h_{11}, h_{12}, v_1 are C^2-functions, it is easy to see that for any Lebesgue density point s from I_1, I_2 and $I_{l,b}$ respectively we have

$$h_{11}(s) = 0,\ h'_{11}(s) = 0,\ h''_{11}(s) = 0; \tag{6.14}$$

$$\begin{aligned}
h_{11}(s)h'_{12}(s) - h_{12}(s)h'_{11}(s) &= 0, \\
h_{11}(s)h''_{12}(s) - h_{12}(s)h''_{11}(s) &= (h_{11}h'_{12} - h_{12}h'_{11})'(s) = 0;
\end{aligned} \tag{6.15}$$

$$\begin{aligned}
v_1(s) &= l h_{12}(s) + b h_{11}(s),\ v'_1(s) = l h'_{12}(s) + b h'_{11}(s), \\
v''_1(s) &= l h''_{12}(s) + b h''_{11}(s)
\end{aligned} \tag{6.16}$$

respectively. Moreover, each of conditions (6.14), (6.15), (6.16) implies $\det M_\psi(s) = 0$. Since we assume $\det M_\psi(s) \neq 0$ almost every, the conclusion follows.

Since $[0,1] \setminus (I_1 \cup I_2)$ is open and its Lebesgue measure is one, we need to prove that for every closed interval $I \subset [0,1] \setminus (I_1 \cup I_2)$ the element (6.13) is Birkhoff generic in X for a.e. $s \in I$. For every such interval I the map $s \mapsto h_{12}(s)/h_{11}(s)$ yields a C^2-diffeomorphism between I and a compact interval J. Moreover, it gives a C^2-map $\varphi : J \to \mathbb{R}$ such that

$$\frac{v_1(s)}{h_{11}(s)} = \varphi\Big(\frac{h_{12}(s)}{h_{11}(s)}\Big) \quad \text{for all} \quad s \in I.$$

Since

$$u\left(\frac{h_{12}(s)}{h_{11}(s)}, \frac{v_1(s)}{h_{11}(s)}\right)(h,v)\Gamma = u\left(\frac{h_{12}(s)}{h_{11}(s)}, \varphi\Big(\frac{h_{12}(s)}{h_{11}(s)}\Big)\right)(h,v)\Gamma,$$

in view of Corollary 6.5, it suffices to show that for all real l,b the set $\{s \in J : \varphi(s) = ls + b\}$ has zero measure which directly follows from zero Lebesgue measure of the set $I_{l,b}$. This completes the proof. $\square$

6.5. **Height function.** The aim of this section is to show that for t sufficiently large there is a mixed height function on X, with respect to $X[n]$, satisfying certain contraction property along the orbits of a_t. Here mixed refers to the fact that we mix the height with respect to the cusp and $X[n]$. This height function will be applied in § 6.6 to prove the crucial Proposition 6.4. Throughout this section let M_1 be as in (6.9).

Lemma 6.8. *Let $n \in \mathbb{N}$ and let I be the closed interval as in Proposition 6.4. For t sufficiently large (depending on σ and I) there exists a measurable function $\beta_n : X \to [1,\infty]$ with the following properties:*

(i) there exists $b > 0$ (depending on σ, n, t) such that for any $m \in \mathbb{Z}_{\geq 0}$ and any interval $J \subset I$ with either $|J| \geq e^{-2mt}$ or $J = I$ one has

$$\int_J \beta_n(a_{(m+1)t}u_\varphi(s)\Gamma)\,ds < \frac{1}{2}\int_J \beta_n(a_{mt}u_\varphi(s)\Gamma)\,ds + b|J|; \tag{6.17}$$

(ii) for any $c > 0$ the set $\{x \in X : \beta_n(x) \leq c\}$ is compact;

(iii) for $x \in X$ one has $\beta_n(x) = \infty$ if and only if $x \in X[n]$;

(iv) for any $m \in \mathbb{Z}_{\geq 0}$, interval $J \subset I$ with $|J| \leq 2e^{-2mt}$ and any $s, \tilde{s} \in J$ one has

$$\beta_n(a_{mt}u_\varphi(\tilde{s})\Gamma) \leq 3\sigma^{-1}\beta_n(a_{mt}u_\varphi(s)\Gamma); \tag{6.18}$$

(v) for any $m \in \mathbb{Z}_{\geq 0}, s \in [0,1]$ and any $-t \leq \tau \leq t$ one has

$$\beta_n(a_\tau a_{mt}u_\varphi(s)) \leq e^t\beta_n(a_{mt}u_\varphi(s)). \tag{6.19}$$

Before proving the lemma we do some preparation. Let t be a positive real number which will be specified only in the proof of Lemma 6.8 (i).

Our mixed height function β_n, inspired by [15], combines the height with respect to the cusp and $X[n]$. The height of elements of X with respect to the cusp is measured by the continuous function $\alpha_0 : X \to [2^{-1/2}, \infty)$ where

$$\alpha_0((h,v)\Gamma) = \sup_{v_0 \in \mathbb{Z}^2\setminus\{0\}} \|hv_0\|^{-0.5}. \tag{6.20}$$

Lemma 6.9. *Let $\kappa : [-1,1] \to [0,\infty]$ be a measurable function. Suppose that there exists $c > 0$ such that*

$$|\{s \in [-1,1] : \kappa(s) < \varepsilon\}| \leq c\varepsilon \tag{6.21}$$

for every $\varepsilon > 0$. Then

$$\int_{-1}^{1} \frac{ds}{\kappa(s)^{0.5}} \leq 7c^{0.5}.$$

Proof. Let $I_0 = \{s \in [-1,1] : c\kappa(s) \geq 1\}$, $I_\infty = \{s \in [-1,1] : \kappa(s) = 0\}$ and

$$I_k = \{s \in [-1,1] : 2^{-k} \leq c\kappa(s) < 2^{-(k-1)}\}$$

for $k \in \mathbb{N}$. In view of (6.21), we have

$$|I_\infty| = 0, \quad |I_k| \leq 2^{-(k-1)} \quad \text{and} \quad |I_0| \leq 2.$$

Since $\frac{1}{\kappa(s)^{0.5}} \le c^{0.5}2^{0.5k}$ for $s \in I_k$, it follows that

$$\int_{-1}^{1} \frac{ds}{\kappa(s)^{0.5}} \le c^{0.5}\big(2 + \sum_{k=1}^{\infty} 2^{0.5k}2^{-(k-1)}\big) \le 7c^{0.5}.$$

□

We will use the following special case of [27, Lemma 3.3] to check (6.21).

Lemma 6.10. *Let $\kappa : [-1,1] \to [0,\infty]$ be a C^1-function. Suppose there exists $A_1, A_2 > 0$ such that for every $s \in [-1,1]$*

$$|\kappa(s)|, |\kappa'(s)| \le A_1 \quad \kappa'(s) \ge A_2, \tag{6.22}$$

then (6.21) holds for $c = \frac{24A_1}{A_2 \sup_{s\in[-1,1]} |\kappa(s)|}$.

An immediate consequence of above two lemmas is:

Corollary 6.11. *Let $a, l \in \mathbb{R}$ be such that $a^2 + l^2 > 0$. Then*

$$\int_{-1}^{1} \frac{ds}{|as+l|^{0.5}} < \frac{100}{(a^2+l^2)^{1/4}}. \tag{6.23}$$

Proof. If $2|a| \le |l|$, then $|as + l| \ge |l|/2 \ge (a^2+l^2)^{1/2}/4$ for all $s \in [-1,1]$, from which (6.23) follows. Now suppose $|2a| > |l|$. Using Lemma 6.10 for $\kappa(s) = as + l$ with $A_1 = |a| + |l|, A_2 = (a^2+l^2)^{1/2}/\sqrt{5}$ and $\sup_{s\in[-1,1]} |\kappa(s)| = |a| + |l|$ one has (6.21) holds for $c = 24\sqrt{5}/(a^2+l^2)^{1/2}$. Therefore, by Lemma 6.9,

$$\int_{-1}^{1} \frac{ds}{|as+l|^{0.5}} \le \frac{7\sqrt{24\sqrt{5}}}{(a^2+l^2)^{1/4}} < \frac{100}{(a^2+l^2)^{1/4}}.$$

□

This lemma allows us to get a linear inequality for the height function α_0.

Lemma 6.12. *For every $t \ge 20$ and every $x \in X$ one has*

$$\int_{-1}^{1} \alpha_0(a_t u(s)x)\, ds < \frac{1}{4}\alpha_0(x) + 2e^t. \tag{6.24}$$

Proof. We first note that for every $s \in [-1,1]$

$$\|\rho(a_t u(s))\| \le 2e^t. \tag{6.25}$$

Write $x = g\Gamma$ and recall that α_0 is defined according to shortest nonzero vectors of the lattice $\rho(g)\mathbb{Z}^2$. If $\alpha_0(x) > (2e^t)^{1/2}$ and $v \in \mathbb{R}^2$ is a shortest vector of $\rho(g)(\mathbb{Z}^2 \setminus \{0\})$, i.e. $\alpha_0(x) = \|v\|^{-0.5}$, then $\|\rho(a_t u(s))v\| < 1$ for every $s \in [-1,1]$, by (6.25). Hence $\rho(a_t u(s))v$ is a shortest vector of $\rho(a_t u(s)g)(\mathbb{Z}^2 \setminus \{0\})$ for any $s \in [-1,1]$. Write $v = (v_1, v_2)^{\mathrm{tr}}$ then

$$\alpha_0(a_t u(s)x)^2 = \|\rho(a_t u(s))v\|^{-1} = \|(e^t(v_2 s + v_1), e^{-t}v_2)\|^{-1} \le e^{-t}|v_2 s + v_1|^{-1}.$$

Using above inequality, Corollary 6.11 and assumption $t \ge 20$ one has

$$\int_{-1}^{1} \alpha_0(a_t u(s)x)\, ds \le 100e^{-0.5t}\|v\|^{-0.5} \le 100e^{-10}\|v\|^{0.5} < \frac{1}{4}\alpha_0(x)^{0.5}.$$

Therefore, in this case (6.24) holds. If $\alpha_0(x) \le (2e^t)^{1/2}$, then (6.25) implies $\alpha_0(a_t u(s)x) \le 2e^t$ for all $s \in [-1,1]$, from which (6.24) follows. □

Now we turn to the construction of the mixed height function β_n. There is a natural height function given in [6, §6] using Riemannian distance to $X[n]$ and this function satisfies a contraction property for the first return map to compact subsets. The height function used in [15] is much more complicated but it satisfies the contraction property without considering the first return map. One of the key observations for the height function in [15] is that the total number of pieces of $X[n]$ whose distance to $x \in X$ is comparable to $\alpha_0(x)^{-2}$ is finite. The following lemma can be interpreted as a simple version of this observation in our situation.

Lemma 6.13. *For every $(h, v) \in G$ there is at most one element $v_0 \in \frac{1}{n}\mathbb{Z}^2$ such that*

$$\|v - hv_0\| < \frac{1}{2n}\alpha_0((h, v)\Gamma)^{-2}. \tag{6.26}$$

The above lemma is clear from the definition of α_0. We will denote the unique element v_0 in this lemma by $\zeta_{h,v}$ if it exists. Otherwise we say $\zeta_{h,v}$ does not exist. If $\zeta_{h,v}$ exists then

$$\|v - h\zeta_{h,v}\| < \frac{1}{2n}2 \leq 1. \tag{6.27}$$

The height function $\alpha_n : X \to [1, \infty]$ with respect to the singular subspace $X[n]$ is defined by

$$\alpha_n((h, v)\Gamma) = \begin{cases} \|v - h\zeta_{h,v}\|^{-0.5} & \text{if } \zeta_{h,v} \text{ exists,} \\ 1 & \text{otherwise} \end{cases} \tag{6.28}$$

where we adopt the convention that $0^{-0.5} = \infty$. It can be checked directly that the definition of α_n does not depend on the choice of (h, v) in the coset $(h, v)\Gamma$.

Remark 6.14. By the definition of $\zeta_{h,v}$ and the continuity of α_0 the element $\zeta_{h,v}$ is locally constant if it exists. It follows that height function $\alpha_n : X \to [1, \infty]$ is lower semi-continuous.

Suppose that $a_{mt}u_\varphi(s) = (h, v)$ and $v_0 = (i/n, j/n)^{\mathrm{tr}} \in \frac{1}{n}\mathbb{Z}^2$. For $s \in [0, 1]$ we write

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1(s) \\ v_2(s) \end{pmatrix} = v - hv_0 = \begin{pmatrix} e^{mt}(\varphi(s) - \frac{i}{n} - \frac{js}{n}) \\ e^{-mt}\frac{j}{n} \end{pmatrix}. \tag{6.29}$$

Strictly speaking $v_i(s)$ above depends also on m and v_0. But usually when we use them m and v_0 are fixed so we omit this dependence for simplicity. During the proof of Lemma 6.8, we need to compare $v_i(s)$ with another $v_i(\tilde{s})$. So we express the latter in terms of former as follows:

$$\begin{aligned} \begin{pmatrix} v_1(\tilde{s}) \\ v_2(\tilde{s}) \end{pmatrix} &= \begin{pmatrix} e^{mt}\big(\varphi(\tilde{s}) - \varphi(s) - \frac{j(\tilde{s}-s)}{n}\big) + v_1 \\ v_2 \end{pmatrix} \\ &= \begin{pmatrix} e^{mt}\big(\varphi'(\hat{s}) - \frac{j}{n}\big)(\tilde{s} - s) + v_1 \\ v_2 \end{pmatrix} \end{aligned} \tag{6.30}$$

where $\hat{s}$ lies between s and $\tilde{s}$ and is determined by the mean value theorem.

In view of (6.10) there exists $\varepsilon = \varepsilon(\sigma) > 0$ such that $I_\varepsilon \subset [0, 1]$ where I_ε is the closed ε-neighborhood of I and

$$\inf\{|\varphi(s) - js/n + i/n| : s \in I_\varepsilon, |j/n| \leq M_1, i \in \mathbb{Z}\} \geq \sigma/2. \tag{6.31}$$

Lemma 6.15. *Let $(h, v) = a_{mt}u_\varphi(s)$ for some $m \in \mathbb{N}, s \in I_\varepsilon$ and $t > \log(2\sigma^{-1})$. If $\zeta_{h,v} = (i/n, j/n)^{\mathrm{tr}}$ exists then $|j/n| > M_1$.*

Proof. Assume the contrary, i.e. $|j/n| \le M_1$. Using (6.29), (6.31) and the assumption for t, we have

$$\|v - h\zeta_{h,v}\| \ge e^{mt}|\varphi(s) - \frac{i}{n} - \frac{js}{n}| > e^t\sigma/2 > 1,$$

contrary to (6.27). □

The mixed height function with respect to $X[n]$ is the function $\beta_n : X \to [1, \infty]$ defined by

$$\beta_n(x) = \alpha_n(x) + 8ne^t\alpha_0(x). \tag{6.32}$$

We will prove each property of Lemma 6.8 separately starting from the simplest.

Proof of (iii) in Lemma 6.8. Directly, by the definition of α_n and β_n, we have

$$\beta_n(x) = \infty \iff \alpha_n(x) = \infty \iff x \in X[n],$$

which gives (iii). □

Proof of (ii) in Lemma 6.8. By the definition of α_0, the set $\{x \in X : \alpha_0(x) \le c/8ne^t\}$ is a compact subset of X for every $c > 0$. As its subset, the set

$$\{x \in X : \beta_n(x) \le c\} \tag{6.33}$$

is relatively compact. Since α_0 is continuous and α_n is lower semi-continuous (Remark 6.14), the map β_n is also lower semi-continuous. Therefore the set (6.33) is closed and hence (ii) in Lemma 6.8 holds. □

Proof of (iv) in Lemma 6.8. We will show that (iv) holds for $t > \log 2\sigma^{-1}$. In case where $m = 0$ we have $\alpha_0(u_\varphi(s)\Gamma) = 1$ and $1 \le \alpha_n(u_\varphi(s)\Gamma) \le \sigma^{-1}$ for all $s \in I$. Therefore (6.18) holds.

Now assume $m \ge 1$. Let $a_{mt}u_\varphi(s) = (h, v)$, $a_{mt}u_\varphi(\tilde{s}) = (\tilde{h}, \tilde{v})$, $x = a_{mt}u_\varphi(s)\Gamma$ and $\tilde{x} = a_{mt}u_\varphi(\tilde{s})\Gamma$.

We claim that

$$\alpha_n(\tilde{x}) \le 6n\alpha_0(x) + 3\alpha_n(x). \tag{6.34}$$

If $\zeta_{\tilde{h},\tilde{v}}$ does not exist then (6.34) follows trivially. Otherwise suppose $\zeta_{\tilde{h},\tilde{v}} = v_0 = (i/n, j/n)^{\mathrm{tr}}$. By Lemma 6.15 we have $|j/n| > M_1$. To prove the claim we will use the notation of (6.29) and (6.30). Since $|\varphi'(\hat{s})| < M_1/2 < |j|/2n$ (see (6.9)), we have $|\varphi'(\hat{s}) - j/n| \le 3|j|/2n$. By the assumption $|\tilde{s} - s| \le 2e^{-2mt}$, it follows that

$$|e^{mt}(\varphi'(\hat{s}) - j/n)(\tilde{s} - s)| \le 3e^{-mt}|j|/n = 3|v_2|.$$

Hence by considering the cases where $|v_1| \ge 4|v_2|$ and $|v_1| < 4|v_2|$ separately one has

$$\begin{aligned}\|\tilde{v} - \tilde{h}v_0\| &\ge \max\{|v_2|, |v_1| - 3|v_2|\} \ge \frac{1}{4}\max\{|v_1|, |v_2|\}\\ &\ge \frac{1}{8}\sqrt{v_1^2 + v_2^2} = \frac{1}{8}\|v - hv_0\|.\end{aligned}$$

Therefore

$$\alpha_n(\tilde{x}) = \|\tilde{v} - \tilde{h}v_0\|^{-0.5} \le 3\|v - hv_0\|^{-0.5}. \tag{6.35}$$

If $\zeta_{h,v} = v_0$ then (6.34) follows from (6.35). If $\zeta_{h,v} \ne v_0$ (which means either $\zeta_{h,v}$ does not exist or it exists but is not equal to v_0), then it follows from the definition of $\zeta_{h,v}$ that $\|v - hv_0\|^{-0.5} \le 2n\alpha_0(x)$. Combine this with (6.35) we get (6.34).

We choose $v_0 \in \mathbb{Z}^2 \setminus \{0\}$ with $\alpha_0(\tilde{x}) = \|\tilde{h}v_0\|^{-0.5}$. Note that by assumption $\|h\tilde{h}^{-1}\| = \|u(e^{2mt}(s - \tilde{s}))\| \le 3$. So

$$\alpha_0(x) \ge \|hv_0\|^{-0.5} = \|h\tilde{h}^{-1} \cdot \tilde{h}v_0\|^{-0.5} \ge \frac{1}{2}\|\tilde{h}v_0\|^{-0.5} = \frac{1}{2}\alpha(\tilde{x}). \tag{6.36}$$

Therefore

$$\begin{aligned}\beta_n(\tilde{x}) &= \alpha_n(\tilde{x}) + 8ne^t\alpha_0(\tilde{x}) \\ &\le 3\alpha_n(x) + 6n\alpha_0(x) + 16ne^t\alpha_0(x) && \text{by (6.34), (6.36)} \\ &\le 3\beta_n(x).\end{aligned}$$

This completes the proof. □

Proof of (v) in Lemma 6.8. It is easy to see that

$$\alpha_0(a_\tau x) \le e^{0.5t}\alpha_0(x). \tag{6.37}$$

Let $a_{mt}u_\varphi(s) = (h, v)$ and $x = a_{mt}u_\varphi(s)\Gamma$. If $\zeta_{a_\tau h, a_\tau v}$ does not exist then (6.19) follows from (6.37). Otherwise take $\zeta_{a_\tau h, a_\tau v} = v_0$. We have

$$\begin{aligned}\alpha_n(a_\tau x) &= \|a_\tau v - a_\tau h v_0\|^{-0.5} \le e^{0.5t}\|v - hv_0\|^{-0.5} \\ &\le \begin{cases} 2ne^{0.5t}\alpha_0(x) & \text{if } v_0 \ne \zeta_{h,v} \\ e^{0.5t}\alpha_n(x) & \text{if } v_0 = \zeta_{h,v} \end{cases} \\ &\le e^{0.5t}2n\alpha_0(x) + e^{0.5t}\alpha_n(x).\end{aligned} \tag{6.38}$$

By (6.37) and (6.38) we have

$$\beta_n(a_\tau x) \le e^{0.5t}2n\alpha_0(x) + e^{0.5t}\alpha_n(x) + e^{0.5t}8ne^t\alpha_0(x) \le e^t\beta_n(x).$$

□

Proof of (i) in Lemma 6.8. We will show that for any $t, b > 0$ with

$$t \ge 30 + \log 2\sigma^{-1} + \log|I|^{-1} - \log\varepsilon \tag{6.39}$$

$$b \ge 32ne^{2t} + e^t\sigma^{-0.5} \tag{6.40}$$

property (i) of Lemma 6.8 holds. We note that (ii), (iii), (iv) and (v) of Lemma 6.8 hold according to the upper bound of t. So together with (i) the proof of Lemma 6.8 is complete.

In the remaining proof $t \ge 30$ will be used several times without being explicitly mentioned. The last term of the lower bound of t guarantees that $s + e^{-2mt}\tilde{s} \in I_\varepsilon$ for every $s \in I, \tilde{s} \in [-1, 1]$ and $m \in \mathbb{N}$. The third term of the upper bound of t implies that for $m \in \mathbb{N}$ we must have $|J| \ge e^{-2mt}$ according to the requirement of J. The second term of the upper bound of t shows that Lemma 6.15 holds.

In case where $m = 0$ we must have $J = I$ since otherwise $|J| \ge 1 > |I|$ which contradicts the assumption $J \subset I$. For every $s \in I$ one has

$$\begin{aligned}\beta_n(a_t u_\varphi(s)\Gamma) &\le e^t\beta_n(u_\varphi(s)\Gamma) && \text{by (v)} \\ &= e^t\alpha_n(u_\varphi(s)\Gamma) + 8ne^{2t}\alpha_0(u_\varphi(s)\Gamma).\end{aligned} \tag{6.41}$$

It follows directly from the definition (see (6.20)) that $\alpha_0(u_\varphi(s)\Gamma) = 1$. Also using (6.29) and the definition (see (6.28)) one has $\alpha_n(u_\varphi(s)\Gamma) \le \sigma^{-0.5}$. Therefore continuing (6.41) we have

$$\beta_n(a_t u_\varphi(s)\Gamma) \le e^t\sigma^{-0.5} + 8ne^{2t},$$

which in view of (6.40) implies (i) of Lemma 6.8.

In the rest of the proof we assume $m \ge 1$ and show that (6.17) holds for the interval $J \subset I$ with $|J| \ge e^{-2mt}$. It follows from the lower bound of $|J|$ that

$$\int_J \beta_n(a_{(m+1)t}u_\varphi(s)\Gamma)\,ds \le \int_J\int_{-1}^{1}\beta_n(a_{(m+1)t}u_\varphi(s + e^{-2mt}\tilde{s})\Gamma)\,d\tilde{s}ds. \tag{6.42}$$

Recall that β_n consists two summands α_0 and α_n and the former is well understood by Lemma 6.12. More precisely since $t \ge 30$ and

$$\alpha_0(a_{(m+1)t}u_\varphi(s + e^{-2mt}\tilde{s})\Gamma) = \alpha_0(a_t u(\tilde{s})a_{mt}u_\varphi(s)\Gamma),$$

Lemma 6.12 provides

$$\int_{-1}^{1} \alpha_0(a_{(m+1)t}u_\varphi(s+e^{-2mt}\tilde{s})\Gamma)\,ds \le \frac{1}{4}\alpha_0(a_{mt}u_\varphi(s)\Gamma) + 2e^t. \tag{6.43}$$

Our strategy for the term α_n is the following: for fixed s we will give an upper bound of the integral

$$\omega(s) := \int_{-1}^{1} \alpha_n(a_{(m+1)t}u_\varphi(s+e^{-2mt}\tilde{s})\Gamma)\,d\tilde{s} \tag{6.44}$$

using the data in the definition of $\beta_n(a_{mt}u_\varphi(s)\Gamma)$. This will be completed in (6.52). To simplify the notation we will not specify the dependence on s and set

$$\begin{aligned} &x = a_{mt}u(s,\varphi(s))\Gamma, \quad (h,v) = a_{mt}u(s,\varphi(s)),\\ &(h(\tilde{s}),v(\tilde{s})) = a_{(m+1)t}u_\varphi(s+e^{-2mt}\tilde{s}). \end{aligned}$$

We use $\zeta(\tilde{s})$ to denote $\zeta_{h(\tilde{s}),v(\tilde{s})}$ for simplicity. Let $v_0 = (i/n, j/n)^{\mathrm{tr}} \in \frac{1}{n}\mathbb{Z}^2$ and let the notation be as in (6.29) and (6.30). We set

$$\begin{aligned} \begin{pmatrix} w_1(\tilde{s}) \\ w_2(\tilde{s}) \end{pmatrix} :&= v(\tilde{s}) - h(\tilde{s})v_0 = \begin{pmatrix} v_1(s+\tilde{s}e^{-2mt}) \\ v_2(s+\tilde{s}e^{-2mt}) \end{pmatrix} \\ &= \begin{pmatrix} e^{(m+1)t}(\varphi(s+e^{-2mt}\tilde{s}) - \varphi(s)) - e^t e^{-mt}\frac{j}{n}\tilde{s} + e^t v_1 \\ e^{-t}v_2 \end{pmatrix}, \end{aligned} \tag{6.45}$$

where $v_1 = e^{mt}(\varphi(s) - \frac{i}{n} - \frac{js}{n})$ and $v_2 = e^{-mt}\frac{j}{n}$ as in (6.29). By mean value theorem there exists $\hat{s} \in [0,1]$ such that

$$w_1(\tilde{s}) = e^t\left[e^{-mt}\left(\varphi'(\hat{s}) - \frac{j}{n}\right)\tilde{s} + v_1\right]. \tag{6.46}$$

If $|j/n| > M_1$ which is always satisfied in below by Lemma 6.15 (recall that $m \ge 1$ and $t \ge \log 2\sigma^{-1}$), then we have $|\varphi'(\hat{s})| < M_1/2 < |j/2n|$, and hence $|j/2n| < |\varphi'(\hat{s}) - j/n| < 2|j/n|$. It follows that

$$|v_2|/2 = e^{-mt}|j|/(2n) \le |e^{-mt}(\varphi'(\hat{s}) - j/n)| \tag{6.47}$$

which will be used several times below for estimating $w_1(\tilde{s})$. To estimate (6.44) we consider two cases of v_0 (**Case A** and **Case B** below).

Case A: Assume $v_0 = \zeta(\tilde{s})$ where $\tilde{s} \in [-1,1]$ but $v_0 \ne \zeta_{h,v}$ which include the case where $\zeta_{h,v}$ does not exist. We will show that

$$\|v(\tilde{s}) - h(\tilde{s})v_0\|^{-0.5} \le ne^t\alpha_0(x). \tag{6.48}$$

If $|v_1| \ge 4|v_2|$, by (6.46) and (6.47) we have

$$\|v(\tilde{s}) - h(\tilde{s})v_0\| \ge |w_1(\tilde{s})| \ge e^t(|v_1| - 2|v_2|) \ge e^t\|v - hv_0\|/2\sqrt{2} \ge \frac{1}{n}\alpha_0^{-2}(x),$$

from which (6.48) follows. If $|v_1| < 4|v_2|$, we have

$$\|v(\tilde{s}) - h(\tilde{s})v_0\| \ge |w_2(\tilde{s})| \ge e^{-t}|v_2| \ge e^{-t}\|v - hv_0\|/4\sqrt{2} \ge \frac{1}{ne^{2t}}\alpha_0^{-2}(x),$$

from which (6.48) follows.

Case B: Assume $v_0 = \zeta_{h,v}$. We will show that

$$\int_{-1}^{1} \|v(\tilde{s}) - h(\tilde{s})v_0\|^{-0.5}\,d\tilde{s} < \frac{1}{4}\|v - hv_0\|^{-0.5}. \tag{6.49}$$

Since $\|v(\tilde{s}) - h(\tilde{s})v_0\| \ge |w_1(\tilde{s})|$, it suffices to estimate the lower bound of $w_1(\tilde{s})$. Here we consider two subcases (**Case B1** and **Case B2** below).

Case B1: $4|v_2| \le |v_1|$. Then, in view of (6.46) and (6.47), for every $\tilde{s} \in [-1,1]$ we have

$$|w_1(\tilde{s})| \ge e^t(|v_1| - 2|v_2|) \ge \frac{e^t}{2}|v_1| \ge \frac{e^t}{2\sqrt{2}}\sqrt{v_1^2 + v_2^2} > 64\|v - hv_0\|.$$

Therefore, (6.49) holds.

Case B2: $4|v_2| > |v_1|$. In this case we use Lemma 6.10 for the C^1-function $\kappa = w_1$. For this purpose we need upper and lower bound for $|w_1|$ and $|w_1'|$. By (6.45) one has

$$w_1'(\tilde{s}) = e^{(1-m)t}(\varphi'(s + e^{-2mt}\tilde{s}) - j/n) \tag{6.50}$$

whose sign is the same as j/n. By (6.47) we have

$$|w_1'(\tilde{s})| \le 2e^t|v_2| \le 2e^t\sqrt{v_1^2 + v_2^2} = 2e^t\|v - hv_0\|$$

and

$$|w_1'(\tilde{s})| \ge \frac{e^t}{2}|v_2| \ge \frac{e^t}{10}\sqrt{v_1^2 + v_2^2} = \frac{e^t}{10}\|v - hv_0\|.$$

In view of (6.46) and (6.47), we also have

$$|w_1(\tilde{s})| \le e^t(|v_1| + 2|v_2|) \le 3e^t\sqrt{v_1^2 + v_2^2} = 3e^t\|v - hv_0\|.$$

Moreover, (6.50) implies

$$\sup_{\tilde{s} \in [-1,1]} |w_1(\tilde{s})| \ge \inf_{\tilde{s} \in [-1,1]} |w_1'(\tilde{s})| \ge \frac{e^t}{10}\|v - hv_0\|. \tag{6.51}$$

Therefore, applying Lemma 6.10 to $\kappa = w_1$ with $A_1 = 3e^t\|v - hv_0\|, A_2 = \frac{e^t}{10}\|v - hv_0\|$ (here we use the notation of (6.22)), we have

$$\left|\{\tilde{s} \in [-1,1] : |w_1(\tilde{s})| \le \varepsilon\}\right| \le \frac{24 \cdot 30\varepsilon}{\sup_{\tilde{s} \in [-1,1]} |w_1(\tilde{s})|} \le \frac{7200\varepsilon}{e^t\|v - hv_0\|}$$

where in the last inequality we use (6.51). Therefore by Lemma 6.9 and $t \ge 30$, we get

$$\int_{-1}^{1} \|v(\tilde{s}) - h(\tilde{s})v_0\|^{-0.5} \le \int_{-1}^{1} \frac{ds}{|w_1(\tilde{s})|^{0.5}} \le \frac{600}{e^{0.5t}\|v - hv_0\|^{0.5}} < \frac{1}{4\|v - hv_0\|^{0.5}},$$

which gives (6.49).

Now we are ready to estimate $\omega(s)$ in (6.44). Assume $\zeta_{h,v}$ exists for the moment. Let

$$J_1 = \{\tilde{s} \in [-1,1] : \zeta(\tilde{s}) \text{ exists and } \zeta(\tilde{s}) \ne \zeta_{h,v}\},$$
$$J_2 = \{\tilde{s} \in [-1,1] : \zeta(\tilde{s}) \text{ exists and } \zeta(\tilde{s}) = \zeta_{h,v}\}.$$

By the definition of ω (see (6.44)) and α_n (see (6.28)), we have

$$\begin{aligned}
\omega(s) &\le 2 + \int_{J_1} \|v(\tilde{s}) - h(\tilde{s})\zeta(\tilde{s})\|^{-0.5}\, d\tilde{s} + \int_{J_2} \|v(\tilde{s}) - h(\tilde{s})\zeta_{h,v}\|^{-0.5}\, d\tilde{s} \\
&< 2 + 2ne^t\alpha_0(x) + \int_{-1}^{1} \|v(\tilde{s}) - h(\tilde{s})\zeta_{h,v}\|^{-0.5}\, d\tilde{s} && \text{by (6.48)} \\
&< 2 + 2ne^t\alpha_0(x) + \frac{1}{4}\|v - h\zeta_{h,v}\|^{-0.5} && \text{by (6.49)} \\
&= 2 + 2ne^t\alpha_0(x) + \frac{1}{4}\alpha_n(x).
\end{aligned} \tag{6.52}$$

It is not hard to see using (6.48) above that the final estimate of (6.52) still holds even though $\zeta_{h,v}$ does not exist. Therefore,

$$\begin{aligned}
&\int_J \beta_n(a_{(m+1)t}u_\varphi(s)\Gamma)\,ds \\
&\quad= \int_J\int_{-1}^{1}(\alpha_n + 8ne^t\alpha_0)(a_{(m+1)t}u_\varphi(s+e^{-2mt}\tilde{s}))\,d\tilde{s}\,ds && \text{by (6.42)}\\
&\quad< \int_J \big(\omega(s) + 2ne^t\alpha_0(a_{mt}u_\varphi(s)\Gamma) + 16ne^{2t}\big)\,ds && \text{by (6.43), (6.44)}\\
&\quad< \int_J \big(\frac{1}{4}\alpha_n + 4ne^t\alpha_0\big)(a_{mt}u_\varphi(s)\Gamma)\,ds + (16ne^{2t}+2)|J| && \text{by (6.52)}\\
&\quad\le \frac{1}{2}\int_J \beta_n(a_{mt}u_\varphi(s)\Gamma)\,ds + b|J| && \text{by (6.40),}
\end{aligned}$$

which completes the proof. □

6.6. **Proof of non-concentration on singular subspaces.** In this section we provide the proof of the crucial Proposition 6.4. We first describe a general strategy to derive quantitative results from an inequality similar to (6.17). We believe that it is of independent interest and can be used in other places.

Let Y be a topological space as in §2.1 and let $\phi : I_0 \to Y$ be a continuous map from a compact interval I_0 with positive length. Let $f : Y \to Y$ be a Borel measurable map. Let $\mathcal{I}_0 = \{I_0\}$ and for every $m \in \mathbb{N}$ let $\mathcal{I}_m$ be a partition of I_0 into at most countably many subintervals with positive length. We assume that $\{\mathcal{I}_m\}_{m\in\mathbb{Z}_{\ge 0}}$ is a filtration, i.e. $\mathcal{I}_{m+1}$ is a refinement of $\mathcal{I}_m$. We use $I_m(s)$ for $s \in I_0$ to denote the unique interval in $\mathcal{I}_m$ that contains s. Let $\beta : Y \to [1,\infty]$ be a measurable function such that there exists $0 < a < 1$ and $b > 0$ with the property that for any $m \in \mathbb{Z}_{\ge 0}$ and any $\tilde{s} \in I_0$

$$\int_{I_m(\tilde{s})} \beta(f^{m+1}\phi(s))\,ds < a\int_{I_m(\tilde{s})} \beta(f^m\phi(s))\,ds + b|I_m(\tilde{s})|. \tag{6.53}$$

We assume that β satisfy the following Lipschitz properties for some constant constant $M \ge 1$: for any $\tilde{s} \in I_0, m \in \mathbb{Z}_{\ge 0}$ and $s \in I_m(\tilde{s})$ one has

$$\beta(f^m\phi(\tilde{s})) \le M\beta(f^m\phi(\tilde{s})), \tag{6.54}$$

$$\beta(f^{m+1}\phi(\tilde{s})) \le M\beta(f^m\phi(\tilde{s})). \tag{6.55}$$

Assumptions (6.53) and (6.54) imply initial finiteness of β on $\phi(I_0)$, i.e. for sufficiently large positive number l

$$\phi(s) \in Y_l := \{y \in Y : \beta(y) \le l\} \quad \text{for all} \quad s \in I_0. \tag{6.56}$$

Let l be a positive number such that (6.56) holds and

$$c := \Big(a + \frac{b}{l}\Big) < 1. \tag{6.57}$$

Let $\omega_0 : I_0 \to \mathbb{Z}$ be the constant function $\omega_0(s) \equiv 0$. Next for every $r \in \mathbb{N}$ we define the r-th return time to Y_l inductively by

$$\omega_r(s) := \inf\{m > \omega_{r-1}(s) : f^m\phi(\tilde{s}) \in Y_l \text{ for some } \tilde{s} \in I_m(s)\}. \tag{6.58}$$

In the degenerate case where $\omega_{r-1}(s) = \infty$, we define $\omega_r(s) = \infty$. The reason we use this unnatural-looking definition is that the r-th return time function $\omega_r : I_0 \to \mathbb{Z}_{\ge 0} \cup \{\infty\}$ is locally constant with respect to the filtration $\{\mathcal{I}_m\}_{m\in\mathbb{Z}_{\ge 0}}$. That is, if $\omega_r(s) = m < \infty$, then $\omega_r(\tilde{s}) = m$ for all $\tilde{s} \in I_m(s)$. This kind of flattening (making a complicated function locally constant) allows us to calculate certain conditional expectations easily.

Lemma 6.16. *There exist $Q > 0$ and $\vartheta > 0$ such that for any integer $k \geq Q$, $r \in \mathbb{Z}_{\geq 0}$ and $\tilde{s} \in I_0$ with $m = \omega_r(\tilde{s}) < \infty$ one has the measure of the set*

$$J_{r,k}(\tilde{s}) := \big\{ s \in I_m(\tilde{s}) : \omega_{r+1}(s) - \omega_r(\tilde{s}) \geq k \big\} \tag{6.59}$$

is less than or equal to $e^{-\vartheta k} |I_m(\tilde{s})|$.

Remark 6.17. The above lemma says that for k sufficiently large the probability of $s \in I_m(\tilde{s})$ with the next return time to Y_l greater than or equal to k decays exponentially, i.e. less than or equal to $e^{-\vartheta k}$. Note that $\omega_0(\tilde{s}) \equiv 0$, so Lemma 6.16 implies that for a.e. $s \in I_0$ one has $\omega_r(s) < \infty$ for all $r \in \mathbb{N}$.

Proof. We will show that the lemma holds for

$$\vartheta = -\frac{1}{2}\log c \quad \text{and} \quad Q = \frac{2\log c - 2\log M}{\log c}.$$

Let $r, \tilde{s}$ and hence $m = \omega_r(\tilde{s})$ be fixed. For simplicity, set $J_k = J_{r,k}(\tilde{s})$ and let

$$A_k := \int_{J_{k+1}} \beta(f^{m+k}\phi(s))\, ds \leq \int_{J_k} \beta(f^{m+k}\phi(s))\, ds.$$

Note that the complement of J_k and hence J_k itself is a disjoint union of intervals in $\mathcal{I}_{m+k-1}$, so by (6.53) we have

$$A_k \leq a A_{k-1} + b|J_k|. \tag{6.60}$$

For all $s \in J_k$ we have $f^{m+k-1}\phi(s) \notin Y_l$, hence $\beta(f^{m+k-1}\phi(s)) > l$. Therefore

$$|J_k| \leq \frac{1}{l} \int_{J_k} \beta(f^{m+k-1}\phi(s))\, ds = \frac{A_{k-1}}{l}. \tag{6.61}$$

In view of (6.60) and (6.61), it follows that

$$A_k \leq \Big(a + \frac{b}{l}\Big) A_{k-1} \leq c A_{k-1}$$

and hence

$$A_k \leq c^k A_0. \tag{6.62}$$

According to (6.58), there is $s_0 \in I_m(\tilde{s})$ with $f^m\phi(s_0) \in Y_l$. Since $J_1 = I_m(\tilde{s})$, by Lipschitz property (6.54), we have

$$A_0 = \int_{I_m(\tilde{s})} \beta(f^m\phi(s))\, ds \leq M \int_{I_m(\tilde{s})} \beta(f^m\phi(s_0))\, ds \leq Ml|I_m(\tilde{s})|.$$

Therefore, by (6.61) and (6.62), we have

$$|J_k| \leq \frac{A_{k-1}}{l} \leq c^{k-1} M |I_m(\tilde{s})|.$$

It is easy to check that the conclusion holds for $k \geq Q$. $\square$

For a measurable subset $K \subset Y$ and $N \in \mathbb{N}$ the proportion of the trajectory $\{f^m\phi(s) : 0 \leq m \leq N-1\}$ in K is defined by

$$\mathcal{D}_K^N(s) := \frac{1}{N} \#\{0 \leq m \leq N-1 : f^m\phi(s) \in K\}.$$

Lemma 6.18. *Let $f : Y \to Y$, $\beta : Y \to [1, \infty]$ be measurable maps such that (6.53), (6.54) and (6.55) hold for a filtration $\{\mathcal{I}_m\}_{m\geq 0}$. Then for every $0 < \varepsilon_0 < 1$ there exist $0 < l_0 < \infty$ and $0 < c_0 < 1$ such that for $K_0 = Y_{l_0}$ and $N \in \mathbb{N}$*

$$\big|\big\{ s \in I_0 : \mathcal{D}_{K_0}^N(s) \leq 1 - \varepsilon_0 \big\}\big| \leq c_0^N |I_0|. \tag{6.63}$$

Proof. We choose $l > 0$ so that properties (6.57) and (6.56) hold. Let $\mathcal{A}_0 = \{\emptyset, I_0\}$ be the trivial σ-algebra on I_0. Inductively, let $\mathcal{A}_k$ for $k \geq 1$ be the σ-algebra on I_0 generated by $\mathcal{A}_i$ for $i < k$ and the sets

$$\{I_{\omega_k(s)}(s) : s \in I_0, \omega_k(s) < \infty\}.$$

It follows that $\{\mathcal{A}_k\}_{k\geq 0}$ is a filtration of σ-algebras on I_0 and ω_k is $\mathcal{A}_k$ measurable. Let I_0' be the subset of I_0 consisting of elements s such that $\omega_r(s) \neq \infty$ for any $r \in \mathbb{N}$. As noted in Remark 6.17 the set I_0' has full measure in I_0. By Lemma 6.16, there exist $Q > 0$ and $\vartheta > 0$ such that for every $k \geq Q, r \in \mathbb{Z}_{\geq 0}$ and $s \in I_0$ we get the exponential decay for the measure of the set $J_{r,k}(s)$ in (6.59). So all the conditions of [33, Lemma 3.1] are satisfied. It follows that there exist $Q_0 \in \mathbb{N}, 0 < c_0 < 1$ such that for every $N \in \mathbb{N}$ the measure of the set

$$J_N = \left\{ s \in I_0' : \frac{1}{N} \sum_{r=1}^{N} \mathbb{1}_{Q_0}(\omega_r(s) - \omega_{r-1}(s)) \geq \varepsilon_0 \right\}$$

is less than or equal to $c_0^N |I_0|$. Here $\mathbb{1}_{Q_0} : \mathbb{N} \to \mathbb{Z}_{\geq 0}$ is the truncation of the identity function defined by

$$\mathbb{1}_{Q_0}(k) = \begin{cases} k & \text{if } k \geq Q_0 \\ 0 & \text{otherwise.} \end{cases}$$

We claim that the lemma holds for this c_0 and $l_0 = lM^{Q_0}$. We now show that

$$\left\{ s \in I_0' : \mathcal{D}_{K_0}^N(s) \leq 1 - \varepsilon_0 \right\} \subset J_N, \tag{6.64}$$

which will complete the proof.

We fix $s \in I_0'$ with $\mathcal{D}_{K_0}^N(s) \leq 1 - \varepsilon_0$. By (6.56), $\phi(s) \in Y_l \subset Y_{l_0} = K_0$. Denote by $0 < n_1 < \cdots < n_k < N$ the sequence of consecutive times for which $f^{n_j}\phi(s) \notin K_0$, i.e

$$\beta(f^{n_j}\phi(s)) > M^{Q_0} l. \tag{6.65}$$

Since $\mathcal{D}_{K_0}^N(s) \leq 1 - \varepsilon_0$, we have $k/N \geq \varepsilon_0$. To prove (6.64) it suffices to find a subset R of $\{1, 2, \ldots, N\}$ such that

(6.66) $\omega_r(s) - \omega_{r-1}(s) \geq Q_0$ for every $r \in R$;

(6.67) for every $1 \leq j \leq k$ there exists $r \in R$ such that $\omega_{r-1}(s) < n_j < \omega_r(s)$.

Let

$$0 \leq m_1 = \max\{m < n_1 : m = \omega_r(s) \text{ for some } r \geq 0\},$$
$$\infty > m_1' = \min\{m > n_1 : m = \omega_r(s) \text{ for some } r \geq 0\}.$$

Then there exists positive integer $r \leq n_1 < N$ such that $m_1' = \omega_r(s)$ and $m_1 = \omega_{r-1}(s)$. We claim that $n_1 - m_1 \geq Q_0$. Indeed, otherwise, by (6.55) and (6.54) we have (note that $Q_0 \in \mathbb{N}$)

$$\beta(f^{n_1}\phi(s)) = \beta(f^{n_1-m_1} f^{m_1}\phi(s)) \leq M^{n_1-m_1}\beta(f^{m_1}\phi(s)) \leq M^{Q_0} l,$$

which contradicts (6.65). Therefore (6.66) holds for $r_1 = r$. If $n_k < m_1'$ then $R = \{r_1\}$ satisfies (6.67) and we are done. Otherwise we choose the smallest j such that $n_j > m_1'$. Then we can repeat the construction to find $r_2 = r$ with $r_1 < r \leq n_j$ so that $\omega_{r-1}(s) < n_j < \omega_r(s)$ and (6.66) holds. We continue this procedure until for $r_i = r$ we have $n_k < \omega_r(s)$. It follows directly from the construction that $R = \{r_1, \ldots, r_i\}$ satisfies (6.66) and (6.67). □

Let $\{\psi_t\}_{t\in\mathbb{R}}$ be a one-parameter flow on Y, see §2.1. Suppose that $f = \psi_\tau$ for some $\tau > 0$ and β satisfies the following Lipschitz property stronger than (6.55):

$$\beta(\psi_t f^m \phi(s)) \leq M\beta(f^m\phi(s)) \quad \text{for all } t \in [0, \tau], m \in \mathbb{Z}_{\geq 0} \text{ and } s \in I_0. \tag{6.68}$$

For any $K \subset Y$ and $T > 0$ denote by $\mathcal{A}_K^T : I_0 \to [0,1]$ the function

$$\mathcal{A}_K^T(s) := \frac{1}{T}\int_0^T \mathbb{1}_K(\psi_t\phi(s))\,dt.$$

Lemma 6.19. *Let $\{\psi_t\}_{t\in\mathbb{R}}$ be a continuous flow on Y and let $f = \psi_\tau$ for some $\tau > 0$. Let $\beta : Y \to [1,\infty]$ be a measurable function such that (6.53), (6.54) and (6.68) hold for a filtration $\{\mathcal{I}_m\}_{m\geq 0}$. Then for any $\varepsilon > 0$ there exist $0 < l_1 < \infty$ and $0 < c_1 < 1$ such that for $K = Y_{l_1}$ and every $T > 0$ one has*

$$\left|\{s \in I_0 : \mathcal{A}_K^T(s) \leq 1-\varepsilon\}\right| \leq c_1^T|I_0|. \tag{6.69}$$

Proof. By Lemma 6.18 applied to $f : Y \to Y$ and $\varepsilon_0 = \varepsilon/2$, we can find $0 < c_0 < 1$ and $l_0 > 0$ such that (6.63) holds for every positive integer N and $K_0 = Y_{l_0}$. By possibly enlarging l_0 we assume that (6.56) holds for $l = l_0$. We fix $n \in \mathbb{N}$ so that for $T \geq n\tau$ we have

$$\left(1-\frac{\tau}{T}\right)(1-\varepsilon_0) \geq 1-\varepsilon. \tag{6.70}$$

We claim that the lemma holds for $l_1 = l_0M^n$ and $c_1 = c_0^{1/2\tau}$.

For $T \leq n\tau$ it follows from (6.56) and (6.68) that (6.69) holds since the left hand side of it is always zero. Now assume $T > n\tau$. In view of (6.68), for every $m \in \mathbb{N}$,

$$f^m\phi(s) \in K_0 \implies \psi_t\phi(s) \in K \quad \text{for every} \quad t \in [m\tau, (m+1)\tau].$$

Therefore, in view of (6.70), we have

$$\mathcal{A}_K^T(s) \leq 1-\varepsilon \implies \mathcal{D}_{K_0}^{\lfloor T/\tau\rfloor}(w) \leq 1-\varepsilon_0.$$

It follow that

$$\left|\{s \in I_0 : \mathcal{A}_K^T(s) \leq 1-\varepsilon\}\right| \leq \left|\{s \in I_0 : \mathcal{D}_{K_0}^{\lfloor T/\tau\rfloor}(w) \leq 1-\varepsilon_0\}\right| \leq c_0^{\lfloor T/\tau\rfloor}|I_0| \leq c_1^T|I_0|.$$

$\square$

Proof of Proposition 6.4. We will apply Lemma 6.19 in the case where $Y = X$, $I_0 = I$ and $\psi_t = a_t$. We fix $t = \tau$ and $\beta = \beta_n$ such that Lemma 6.8 holds. We need to check that for this β and $f = \psi_\tau$ the assumptions of Lemma 6.19 hold. First we construct a filtration $\{\mathcal{I}_m\}_{m\in\mathbb{N}}$. Suppose we have already constructed $\mathcal{I}_{m-1}$. If $e^{-2mt} \leq |I|$, then we set $\mathcal{I}_m = \mathcal{I}_{m-1}$. Otherwise we divide each interval J of $\mathcal{I}_{m-1}$ consecutively into intervals of length e^{-2mt}, except for the last one which we allow to have length between e^{-2mt} and $2e^{-2mt}$. It is easy to see that (6.53), (6.54) and (6.68) follow from (i), (iv) and (v) of Lemma 6.8 respectively.

Therefore we can use Lemma 6.19 to find $0 < l < \infty$ and $\vartheta > 0$ so that (6.11) holds for $K = X_l := \{x \in X : \beta_n(x) \leq l\}$. Finally we note that (ii) and (iii) of Lemma 6.8 imply K is a compact subset of $X \setminus X[n]$. $\square$

7. Proof of the Oseledets genericity along curves

In this section, we prove Theorem 2.5. Throughout the section, we denote by μ the probability affine measure on an $SL_2(\mathbb{R})$-orbit closure $\mathcal{M}$. Suppose that $W \subset H_1(M,\mathbb{R})$ is a symplectic subspace which is invariant for the cocycle $A^{KZ} : SL_2(\mathbb{R}) \times \mathcal{M} \to GL(H_1(M,\mathbb{R}))$. We deal with the restricted cocycle $A_W^{KZ} : SL_2(\mathbb{R}) \times \mathcal{M} \to GL(W)$.

Assume that the sum of its positive Lyapunov exponents is less than 1. Then for every $1 \leq r \leq 2d$ ($2d$ is the dimension of W) the top Lyapunov exponent of the r-th exterior power cocycle $A^p : SL_2(\mathbb{R}) \times \mathcal{M} \to GL(\bigwedge^p W)$ is equal to the sum of r greatest Lyapunov exponents of $A_W^{KZ} : SL_2(\mathbb{R}) \times \mathcal{M} \to GL(W)$ (see e.g. [22] and [31]). Therefore, the proof Theorem 2.5 reduces to the following result:

Theorem 7.1. *Assume that the top Lyapunov exponent λ_{top} of $A^p : SL_2(\mathbb{R})\times\mathcal{M} \to GL(\bigwedge^p W)$ is less than 1. Suppose that $\varphi : I \to \mathcal{M}$ is a C^1-curve which is well approximated by horocycles (in the sense of Definition 2.4) and such that for a.e. $s \in I$ the element $\varphi(s) \in \mathcal{M}$ is Birkhoff generic with respect to $(\mathcal{M}, \mu, a_t)$. Then for a.e. $s \in I$ we have*

$$\lim_{t\to\infty} \frac{1}{t}\|A^p(a_t, \varphi(s))\| = \lambda_{top}.$$

Very rough outline of the proof. The proof of Theorem 7.1 is divided into several parts arranged in the following sections.

As it was observed by Chaika-Eskin in [7, the proof of Theorem 1.4] the vector bundle $\mathcal{M} \times \bigwedge^p W$ has an $SL_2(\mathbb{R})$-invariant (for the cocycle A^p) splitting $\bigoplus_{i=1}^k \mathcal{V}_i$ which is of class C^∞ and such that the restriction of the cocycle A^p to the subbundle $\mathcal{V}_i$ (denoted by $A_{\mathcal{V}_i}$) is strongly irreducible. This result is based on a recent result of Filip [16] on semisimplicity and rigidity of the Kontsevich-Zorich cocycle. For the definition of strong irreducibility we refer to [7].

In the first step of the proof we focus on the growth (along the curve φ) of any strongly irreducible restricted cocycle $A_{\mathcal{V}}$ where $\mathcal{V}$ is a smooth invariant subbundle. Let us denote by $\lambda_{\mathcal{V}}$ the top Lyapunov exponent of $A_{\mathcal{V}}$ with respect of the measure μ. By assumption, the image by a_t of the curve φ over small intervals is well approximated by horocycle arcs. Therefore, it is useful to control the growth of $A_{\mathcal{V}}$ restricted to such arcs. In § 7.1, using the idea and results of [7] we provide so called "good" subsets of $\mathcal{M}$ whose μ-measures are close to 1 and such that the growth of $A_{\mathcal{V}}$ is exponential with the exponent $\lambda_{\mathcal{V}}$ over short horocycle arcs starting from the "good" set.

In § 7.2, we give a tool for proving that for a.e. $s \in I$ the Teichmüller trajectory of $\varphi(s)$ visits the "good" set with the frequency close to 1. This is based on a simple application of the strong law of large numbers for weakly correlated random variables and the Birkhoff genericity of the curve.

Finally, in § 7.4, we collect the above mentioned results to prove Theorem 7.1. Since for a.e. $s \in I$ the Teichmüller trajectory of $\varphi(s)$ visits the "good" set with high frequency, we can study the growth of $A_{\mathcal{V}_i}$ for $1 \le i \le k$ by dividing the trajectory into steps where for most of these steps the trajectory are close to small horocycle arcs starting from the "good" set. For such points the growth of $A_{\mathcal{V}_i}$ is exponential with the exponent $\lambda_{\mathcal{V}_i}$. Therefore the growth of A^p is exponential with the exponent $\max_{1\le i\le k}\lambda_{\mathcal{V}_i} = \lambda_{top}$. Here we need an auxiliary result, presented in § 7.3, which use the fact that A^p is piecewise constant. In summary, the exponential growth for most steps finally yields the same growth along the whole trajectory starting from $\varphi(s)$.

7.1. **Existence of good sets.** Let $\mathcal{V}$ be a smooth $SL_2(\mathbb{R})$-invariant subbundle of $\mathcal{W} \times \bigwedge^p(W)$ such that $A_{\mathcal{V}}$ is strongly irreducible. In this section, roughly speaking, we prove for sufficiently large real number t the existence of a "good" open subset of $\mathcal{M}$ whose measure increases to 1 as $t \to \infty$ and for its element x the growth of the cocycle $A_{\mathcal{V}}(a_t u(r), x)$ is exponential with the exponent $\lambda_{\mathcal{V}}$ when r runs through some intervals of length proportional to e^{-2t}. This result helps to control the growth of $A_{\mathcal{V}}(a_t, \cdot)$ over short (of length $\approx e^{-2t}$) horocycle arcs starting from elements of the "good" set.

To prove the main result (Lemma 7.2 below) we use two results established by Chaika and Eskin in [7]. The first (Proposition 7.4) gives the existence of a "good" set for the Konstevich-Zorich cocycle generated by a random walk on $SL_2(\mathbb{R})$. The second (Proposition 7.3) allows us to show that for every horocycle arcs almost every orbit of the Teichmüller flow starting from such arc can be tracked sublinearly by a generic trajectory on $\mathcal{M}$ generated by the random walk. This sublinearity of the tracking guarantees that growth of $A_{\mathcal{V}}$ along the Teichmüller and random walk trajectories are the same. This enables us to pass from the random walk action on $\mathcal{M}$ to Teichmüller trajectories starting from a horocycle arc.

The main aim of this section is to prove the following result.

Lemma 7.2 (Good sets for horocycle arcs). *There exists $\lambda > 0$ such that for every $\varepsilon > 0$, $\delta > 0$ and $\sigma > 0$ there exists $L_0 \in \mathbb{N}$ such that for every natural $L \geq L_0$ there exists an open subset $E(\varepsilon, \sigma, L) = E_{\mathcal{V}}(\varepsilon, \sigma, L) \subset \mathcal{M}$ with $\mu(E(\varepsilon, \sigma, L)) > 1 - \delta$ such that for every $x \in E(\varepsilon, \sigma, L)$, for every interval $[a, b] \ni 0$ of length σ and every $v \in V(x)$ with $\|v\| = 1$ there exists a Borel set $R(v) = R(\varepsilon, L, x, v) \subset [a, b]$ with $Leb(R(v)) > (1-\varepsilon)\sigma$ such that for any $r_0 \in R(v)$ and $r \in [r_0 - \sigma e^{-2\lambda L}, r_0 + \sigma e^{-2\lambda L}]$ we have*

$$\lambda_{\mathcal{V}} - \varepsilon \leq \frac{1}{L\lambda} \log \|A_{\mathcal{V}}(a_{L\lambda} u(r), x)v\| \leq \frac{1}{L\lambda} \log \|A_{\mathcal{V}}(a_{L\lambda} u(r), x)\| \leq \lambda_{\mathcal{V}} + \varepsilon.$$

The proof of the lemma is based on the following two results established in [7]. Suppose that ϑ is an $SO(2, \mathbb{R})$-bi-invariant probability measure on $SL_2(\mathbb{R})$ which is compactly supported and absolutely continuous with respect to the Haar measure on $SL_2(\mathbb{R})$.

Proposition 7.3 (Sublinear tracking, see [7]). *There exists $\lambda = \lambda(\vartheta) > 0$ and a measurable map $\theta : SL_2(\mathbb{R})^{\mathbb{N}} \to [0, 2\pi]$ such that $\theta_*(\vartheta^{\mathbb{N}}) = \frac{1}{2\pi} Leb_{[0,2\pi]}$ and for $\vartheta^{\mathbb{N}}$-a.e. $\bar{g} = \{g_n\}_{n \in \mathbb{N}} \in SL_2(\mathbb{R})^{\mathbb{N}}$ we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \|a_{n\lambda} r_{\theta(\bar{g})} (g_n \cdots g_1)^{-1}\| = 0. \tag{7.1}$$

Proposition 7.4 (Good sets for random walks, see [7]). *For every $\varepsilon > 0$ and $\delta > 0$ there exists $L_0 \in \mathbb{N}$ such that for every $L \geq L_0$ there exists an open subset $E_{rw}(\varepsilon, L) \subset \mathcal{M}$ with $\mu(E_{rw}(\varepsilon, L)) > 1 - \delta$ such that for every $x \in E_{rw}(\varepsilon, L)$ and for every $v \in \mathcal{V}(x)$ with $\|v\| = 1$ there exists a measurable set $H(v) = H(\varepsilon, L, x, v) \subset SL_2(\mathbb{R})^L$ with $\vartheta^L(H(v)) > 1 - \varepsilon$ such that for any $(g_L, \ldots, g_1) \in H(v)$ we have*

$$\lambda_{\mathcal{V}} - \varepsilon \leq \frac{1}{L\lambda} \log \|A_{\mathcal{V}}(g_L \cdots g_1, x)v\| \leq \frac{1}{L\lambda} \log \|A_{\mathcal{V}}(g_1 \cdots g_L, x)\| \leq \lambda_{\mathcal{V}} + \varepsilon.$$

Remark 7.5. Recall that there exist $N \in \mathbb{N}$ and $C \geq 1$ such that

$$\|A_{\mathcal{V}}(g, x)\| \leq \|A^p(g, x)\| \leq C\|g\|^N \quad \text{for all} \quad x \in \mathcal{M},\ g \in SL_2(\mathbb{R}). \tag{7.2}$$

Since

$$A_{\mathcal{V}}(g_1 g_2^{-1}, g_2 x) A_{\mathcal{V}}(g_2, x) = A_{\mathcal{V}}(g_1, x)$$

for $g_1, g_2 \in SL_2(\mathbb{R})$ and $x \in \mathcal{M}$, it follows that

$$\big| \log \|A_{\mathcal{V}}(g_1, x)\| - \log \|A_{\mathcal{V}}(g_2, x)\| \big| \leq \log C + N \log \|g_1 g_2^{-1}\|. \tag{7.3}$$

Moreover, for every $v \in \mathcal{V}(x)$ with $\|v\| = 1$ we have

$$\big| \log \|A_{\mathcal{V}}(g_1, x)v\| - \log \|A_{\mathcal{V}}(g_2, x)v\| \big| \leq \log C + N \log \|g_1 g_2^{-1}\|. \tag{7.4}$$

Proof of Lemma 7.2. Let us consider measurable map $\tan : [0, 2\pi] \to \mathbb{R}$ and let

$$\kappa := \tan_*\Big(\frac{1}{2\pi} Leb_{[0,2\pi]}\Big) = \frac{dr}{\pi(1 + r^2)}.$$

Taking $\zeta : SL_2(R)^{\mathbb{N}} \to \mathbb{R}$ given by $\zeta = (-\tan) \circ \theta$ we have $\zeta_*(\vartheta^{\mathbb{N}}) = \kappa$. Note that, by Proposition 7.3,

$$\lim_{n\to\infty} \frac{1}{n\lambda} \log \|a_{n\lambda} u(\zeta(\bar{g}))(g_n \cdots g_1)^{-1}\| = 0 \tag{7.5}$$

for $\vartheta^{\mathbb{N}}$-a.e. $\bar{g} \in SL_2(\mathbb{R})^{\mathbb{N}}$. Indeed, for every $\theta \in [0, 2\pi] \setminus \{\pi/2, 3\pi/2\}$, $t > 0$ and $g \in SL_2(\mathbb{R})$ we have

$$\begin{aligned}\big|\log\|a_t u(\tan(\theta))g\| - \log\|a_t r_\theta g\|\big| &\leq \log\|a_t u(\tan(\theta)) r_{-\theta} a_{-t}\|\\ \leq \log\left\|\begin{pmatrix}\cos^{-1}\theta & 0\\ e^{-2t}\sin\theta & \cos\theta\end{pmatrix}\right\| &\leq \log\sqrt{\cos^{-2}\theta + 1}.\end{aligned}$$

Together with (7.1) and (7.3) this yields (7.5).

Now fix $\varepsilon > 0$, $\delta > 0$ and $\sigma > 0$. Let $\sigma_0 := \frac{\sigma}{\pi(\sigma^2+1)}$. Then $0 < \sigma_0 \leq 1/(2\pi)$. In view of Proposition 7.4, one can choose natural number

$$L_1 > \frac{2N(\sigma + \log C)}{\varepsilon\lambda}$$

so that for all $L \geq L_1$ we have $\mu(E_{rw}(\sigma_0\varepsilon/2, L)) > 1 - \delta$. Set

$$E(\varepsilon, \sigma, L) := E_{rw}(\sigma_0\varepsilon/2, L) \subset \mathcal{M}.$$

Next take

$$\varepsilon_1 := \frac{\varepsilon}{2N} - \frac{\sigma + \log C}{L_1\lambda} > 0$$

and for $n \geq L_1$ set

$$E'(n, \varepsilon_1) := \Big\{\bar{g} \in SL_2(\mathbb{R})^{\mathbb{N}} : \frac{1}{n\lambda}\log\|a_{n\lambda}u(\zeta(\bar{g}))(g_n \cdot \ldots \cdot g_1)^{-1}\| \leq \varepsilon_1\Big\}.$$

In view of (7.5), there exists $L_0 \geq L_1$ such that for $n \geq L_0$ we have $\vartheta^{\mathbb{N}}(E'(n, \varepsilon_1)) > 1 - \sigma_0\varepsilon/2$.

Take $L \geq L_0$ and fix $x \in E(\varepsilon, \sigma, L)$ and $v \in \mathcal{V}(x)$ with $\|v\| = 1$. Take the subset $H(\sigma_0\varepsilon/2, L, x, v) \subset SL_2(\mathbb{R})^L$ coming from Proposition 7.4 so that $\vartheta^L(H(\sigma_0\varepsilon/2, L, x, v)) > 1 - \sigma_0\varepsilon/2$. Setting

$$\widetilde{H}(v) := \{\bar{g} \in E'(L, \varepsilon_1) : (g_1, \ldots, g_L) \in H(\sigma_0\varepsilon/2, L, x, v)\}$$

we have $\vartheta^{\mathbb{N}}(\widetilde{H}(v)) > 1 - \sigma_0\varepsilon/2$. Then there exists a Borel subset $R \subset \mathbb{R}$ with $\kappa(R) > 1 - \sigma_0\varepsilon$ such that for every $r \in R$ we have $\zeta^{-1}(\{r\}) \cap \widetilde{H}(v) \neq \emptyset$.

Suppose that $r_0 \in R$ and take $\bar{g} \in \widetilde{H}(v)$ so that $\zeta(\bar{g}) = r_0$. Since $\bar{g} \in E'(L, \varepsilon_1)$, we have

$$\frac{1}{L\lambda}\log\|a_{L\lambda}u(r_0)(\bar{g}_{[1,L]})^{-1}\| \leq \varepsilon_1 \quad \text{with} \quad \bar{g}_{[1,L]} = g_L \cdots g_1.$$

If $|r - r_0| \leq \sigma e^{-2L\lambda}$ then

$$\begin{aligned}&\Big|\log\|a_{L\lambda}u(r)(\bar{g}_{[1,L]})^{-1}\| - \log\|a_{L\lambda}u(r_0)(\bar{g}_{[1,L]})^{-1}\|\Big|\\ &\quad\leq \log\|a_{L\lambda}u(r - r_0)a_{-L\lambda}\| = \log\|u(e^{2L\lambda}(r - r_0))\|\\ &\quad\leq \log(1 + e^{2L\lambda}|r - r_0|) \leq \log(1 + \sigma) \leq \sigma,\end{aligned}$$

so

$$\frac{1}{L\lambda}\log\|a_{L\lambda}u(r)(\bar{g}_{[1,L]})^{-1}\| \leq \varepsilon_1 + \frac{\sigma}{L\lambda}.$$

In view of (7.3) and (7.4), both

$$\frac{1}{L\lambda}\big|\log\|A_{\mathcal{V}}(a_{L\lambda}u(r), x)\| - \log\|A_{\mathcal{V}}(\bar{g}_{[1,L]}, x)\|\big|$$

$$\frac{1}{L\lambda}\big|\log\|A_{\mathcal{V}}(a_{L\lambda}u(r), x)v\| - \log\|A_{\mathcal{V}}(\bar{g}_{[1,L]}, x)v\|\big|$$

are bounded by

$$\frac{\log C}{L\lambda} + N\Big(\varepsilon_1 + \frac{\sigma}{L\lambda}\Big) \leq N\Big(\varepsilon_1 + \frac{\sigma + \log C}{L_1\lambda}\Big) = \varepsilon/2.$$

As $(g_1,\dots,g_L) \in H(\sigma_0\varepsilon/2, L, x, v)$ we have

$$\lambda_{\mathcal{V}} - \varepsilon/2 \le \frac{1}{L\lambda}\log\|A_{\mathcal{V}}(\bar g_{[1,L]},x)v\| \le \frac{1}{L\lambda}\log\|A_{\mathcal{V}}(\bar g_{[1,L]},x)\| \le \lambda_{\mathcal{V}} + \varepsilon/2.$$

It follows that, if $r\in R$ and $|r-r_0|\le \sigma e^{-2L\lambda}$ then

$$\lambda_{\mathcal{V}} - \varepsilon \le \frac{1}{L\lambda}\log\|A_{\mathcal{V}}(a_{L\lambda}u(r),x)v\| \le \frac{1}{L\lambda}\log\|A_{\mathcal{V}}(a_{L\lambda}u(r),x)\| \le \lambda_{\mathcal{V}} + \varepsilon.$$

Finally we need to show that for any interval $[a,b]\ni 0$ of length σ taking $R(\varepsilon,L,x,v) := R\cap[a,b]$ we have $Leb([a,b]\setminus R(\varepsilon,L,x,v)) < \sigma\varepsilon$. It follows directly from the fact that

$$\kappa([a,b]\setminus R(\varepsilon,L,x,v)) \le \kappa(\mathbb{R}\setminus R) < \sigma_0\varepsilon = \frac{\sigma}{\pi(\sigma^2+1)}\varepsilon$$

and the density of κ is $\frac{1}{\pi(r^2+1)} \ge \frac{1}{\pi(\sigma^2+1)}$ on $[a,b]$. This completes the proof. □

7.2. **Curve partitions and weak law of large numbers.** Let $U\subset\mathcal{M}$ be a "good" set for the time t. Then the interval I can be partitioned into intervals of length $\approx e^{-2tn}$ such that for most such intervals their images by $a_{nt}\circ\varphi : I\to\mathcal{M}$ are fully contained in U. This gives a sequence of partitions of I. In this section, applying the strong law of large numbers for weakly correlated random variables, we present an abstract setting for proving that for a.e. $s\in I$ the frequency $a_{nt}\varphi(s)$ in U is approximately $\mu(U)$.

Take $I=[0,1]$, $\varepsilon>0$, $0<\alpha,\sigma<1$. For every $n\ge 0$ let $d_n = \lfloor 1/(\sigma\alpha^n)\rfloor$ and denote by $\mathcal{I}_n$ the partition of the interval $[0, d_n\sigma\alpha^n]\subset I$ into d_n intervals of length $\sigma\alpha^n$. Denote by $\mathcal{F}_n$ the ring of sets generated by $\mathcal{I}_n$, i.e. each element of $\mathcal{F}_n$ is the union of some intervals from $\mathcal{I}_n$.

Assume that $1-\varepsilon < k\alpha < 1$ for some natural number k. Let us consider a sequence $\{A_n\}_{n\ge 0}$ of subsets of I such that for every $n\ge 0$ we have $A_n\in\mathcal{F}_{n+1}$ and for every interval $I_n\in\mathcal{I}_n$ the set $A_n\cap I_n$ is the union of exactly k intervals from $\mathcal{I}_{n+1}$. Then

$$k\alpha(1-\sigma\alpha^n) \le Leb(A_n) \le k\alpha.$$

Therefore, for $n\ge 0$ and $l\ge 1$ we have

$$|Leb(A_n\cap A_{n+l}) - Leb(A_n)Leb(A_{n+l})| \le (2+\sigma)\alpha^{l-1}.$$

Then, by the strong law of large numbers for weakly correlated random variables (see e.g. [32]),

$$\frac{1}{n}\sum_{j=0}^{n-1}(\mathbb{1}_{A_j}(s) - Leb(A_j)) \to 0 \text{ for a.e. } s\in I,$$

so

$$\lim_{n\to\infty}\frac{1}{n}\#\{0\le j\le n-1 : s\in A_j\} = k\alpha > 1-\varepsilon \text{ for a.e. } s\in I.$$

7.3. **Fundamental domains.** The Konstevich-Zorich cocycle is piecewise constant in terms of time and space. This fact helps to control the norm of A^p for fix time t and for relatively close points in $\mathcal{M}$. In this section we provide an effective tool for controlling the norm which will be useful in the proof of Theorem 7.1.

Let $\mathcal{T}$ be the Teichmüller space for $\mathcal{M}$ and let d be a metric on $\mathcal{T}$ satisfying (2.8) and (2.9). Denote by $\pi:\mathcal{T}\to\mathcal{M}$ the natural projection. Fix a fundamental domain $\Delta\subset\mathcal{T}$ for the action of the discrete group $\Gamma := \Gamma(M,\Sigma)$ on $\mathcal{T}$ so that

$$\mu(\pi(\partial\Delta)) = 0. \tag{7.6}$$

For every $\sigma>0$ let

$$\Delta_\sigma = \{\tilde x\in \operatorname{Int}\Delta : dist(\tilde x,\partial\Delta) > \sigma\}.$$

Then the set $\Delta_\sigma \subset \mathcal{T}$ is open and $\bigcup_{\sigma>0}\Delta_\sigma = \operatorname{Int}\Delta$. Therefore

$$\pi(\Delta_\sigma) \subset \mathcal{M} \text{ is open and } \lim_{\sigma\to 0}\mu(\pi(\Delta_\sigma)) = \mu(\pi(\operatorname{Int}\Delta)) = 1.$$

Lemma 7.6. *Let $\tilde{x},\tilde{y}\in\mathcal{T}$, $x=\pi(\tilde{x}), y=\pi(\tilde{y})\in\mathcal{M}$ be such that*

(i) $\tilde{x},\tilde{y}\in\Delta$; or,

(ii) $x\in\pi(\Delta_\sigma)$ and $d(\tilde{x},\tilde{y})\le\sigma$.

If $g\in SL_2(\mathbb{R})$ is such that $g\cdot x\in\pi(\Delta_\sigma)$ and $d(g\cdot\tilde{x},g\cdot\tilde{y})\le\sigma$ then $A^p(g,x)=A^p(g,y)$. Moreover, if $x\in\pi(\Delta_\sigma)$ and $d(\tilde{x},g\cdot\tilde{x})\le\sigma$ for some $g\in SL_2(\mathbb{R})$ then $A^p(g,x)=Id$.

Proof. Choose $\gamma\in\Gamma$ so that $\gamma(\tilde{x})\in\Delta$. By assumption (i) or (ii) ($d(\gamma(\tilde{x}),\gamma(\tilde{y}))=d(\tilde{x},\tilde{y})\le\sigma$), we also have $\gamma(\tilde{y})\in\Delta$. Let $\widehat{\gamma}\in\Gamma$ be such that $\widehat{\gamma}(g\cdot\tilde{x})\in\Delta$. Since $g\cdot x\in\pi(\Delta_\sigma)$, it follows that $\widehat{\gamma}(g\cdot\tilde{x})\in\Delta_\sigma$. Together with

$$d(\widehat{\gamma}(g\cdot\tilde{x}),\widehat{\gamma}(g\cdot\tilde{y}))=d(g\cdot\tilde{x},g\cdot\tilde{y})\le\sigma$$

this gives $\widehat{\gamma}(g\cdot\tilde{y})\in\Delta$. In summary

$$\gamma(\tilde{x}),\gamma(\tilde{y})\in\Delta \text{ and } (\widehat{\gamma}\circ\gamma^{-1})\big(\gamma(g\cdot\tilde{x})\big),(\widehat{\gamma}\circ\gamma^{-1})\big(\gamma(g\cdot\tilde{y})\big)\in\Delta.$$

It follows that $A^p(g,x)=\bigwedge^p((\widehat{\gamma}\circ\gamma^{-1})_*)=A^p(g,y)$.

Similarly, if $x\in\pi(\Delta_\sigma)$ and $d(\tilde{x},g\cdot\tilde{x})\le\sigma$ then $\gamma(\tilde{x}),\gamma(g\cdot\tilde{x})\in\Delta$. Therefore, $A^p(g,x)=\bigwedge^p((Id)_*)=Id$. □

7.4. **Oseledets genericity for almost every point on the curve.** In this last section, we conclude the proof: we exploit the Birkhoff generic behaviour along the curve and the existence of good sets to show that for most arcs which are quantitatively well approximated by a horocycle, thanks to the weak law of the large numbers, the time spent in each good set is large enough to guarantee typical Oseledets behaviour for almost every point in the curve.

Proof of Theorem 7.1. For any $\tilde{x}\in\mathcal{T}$ one can build a fundamental domain $\Delta(\tilde{x})\subset\mathcal{T}$ satisfying (7.6) so that $\tilde{x}\in\operatorname{Int}\Delta(\tilde{x})$. It follows that the interval I can be covered by a finite family of closed intervals such that for every such interval $J\subset I$ there is a fundamental domain $\Delta(J)\subset\mathcal{T}$ satisfying (7.6) so that $\widetilde{\varphi}(J)\in\operatorname{Int}\Delta(J)$. Therefore, without loss of generality we can assume that $I=[0,1]$ and $\widetilde{\varphi}(I)\in\operatorname{Int}\Delta$ for some fundamental domain $\Delta\subset\mathcal{T}$ satisfying (7.6).

Let us consider a C^∞ splitting $\bigoplus_{i=1}^k\mathcal{V}_i$ of $\mathcal{M}\times\bigwedge^pW$ such that each cocycle $A_{\mathcal{V}_i}$ is strongly irreducible. Then

$$\lambda_{top}=\max_{1\le i\le k}\lambda_{\mathcal{V}_i}.$$

For every $x\in\mathcal{M}$ let us consider the isomorphism

$$W\ni w\mapsto(w_1(x),\ldots,w_k(x))\in\mathcal{V}_1(x)\times\ldots\times\mathcal{V}_k(x)\ \text{ with }\ w=w_1(x)+\ldots+w_k(x).$$

Since φ is of class C^1 and each $\mathcal{V}_i$ is a C^∞-subbundle, there exists $C'\ge1$ such that

$$\sum_{1\le i\le k}\|w_i(\varphi(s))\|\le C'\|w\| \ \text{ for every }\ s\in I\ \text{ and }\ w\in W. \tag{7.7}$$

Moreover, for every $1\le i\le k$ we can choose a C^1-curves $I\ni s\mapsto v_i(s)\in\mathcal{V}_i(\varphi(s))$ over φ such that $\|v_i(s)\|=1$ for any $s\in I$. Then there exists $l>0$ such that

$$\|v_i(s)-v_i(s')\|\le l|s-s'| \ \text{ for all }\ s,s'\in I,\ 1\le i\le k. \tag{7.8}$$

Fix $0<\varepsilon<1$. Next choose $0<\sigma<1$ so that $\mu(\pi(\Delta_{2\sigma}))>1-\varepsilon/6$. By Lemma 7.2, there exists $L\in\mathbb{N}$ such that

$$L>\frac{8\max\{l\sigma,C'\}/\varepsilon+\log(C(\varphi)(1+\sigma^\rho))}{\lambda} \tag{7.9}$$

and

$$\mu(E_{\mathcal{V}_i}(\varepsilon/4k,\sigma,L)) > 1-\varepsilon/6k \quad \text{for} \quad 1\le i\le k,$$

where ρ is the natural factor derived from the well approximation of φ by horocycles (see (2.10)). Therefore the set

$$U_\varepsilon := \bigcap_{i=1}^k E_{\mathcal{V}_i}(\varepsilon/4k,\sigma,L)\cap\pi(\Delta_{2\sigma})\cap a_{-L\lambda}\pi(\Delta_{2\sigma})\subset\mathcal{M}$$

is open with $\mu(U_\varepsilon) > 1-\varepsilon/2$.

Let us consider the sequences $\{\mathcal{I}_n\}_{n\ge 0}$, $\{\mathcal{F}_n\}_{n\ge 0}$ of families of subsets in I described in § 7.2 for $\alpha := e^{-2L\lambda}$. Then the length of each interval from $\mathcal{I}_n$ is $\sigma e^{-2L\lambda n}$.

For $n\ge 0$ let us consider any interval $J=[m\sigma e^{-2L\lambda n},(m+1)\sigma e^{-2L\lambda n}]\in\mathcal{I}_n$ such that $a_{L\lambda n}\varphi(J)\cap U_\varepsilon\neq\emptyset$. Next fix $s_0=s_{0,n}=s_0(J)\in J$ with $a_{L\lambda n}\varphi(s_0)\in U_\varepsilon$. Then $a_{L\lambda n}\varphi(s_0)\in E_{\mathcal{V}_i}(\varepsilon/4,\sigma,L)\cap\pi(\Delta_{2\sigma})$ for $1\le i\le k$. Let

$$v_{n,i} := \frac{A_{\mathcal{V}_i}(a_{L\lambda n},\varphi(s_0))v_i(s_0)}{\|A_{\mathcal{V}_i}(a_{L\lambda n},\varphi(s_0))v_i(s_0)\|}\in\mathcal{V}_i(a_{L\lambda n}\varphi(s_0)) \quad \text{for} \quad 1\le i\le k. \tag{7.10}$$

Since $a_{L\lambda n}\varphi(s_0)\in E_{\mathcal{V}_i}(\varepsilon/4k,\sigma,L)$, by Lemma 7.2, there exists a Borel set

$$R(v_{n,i})\subset[m\sigma - s_0e^{2L\lambda n},(m+1)\sigma - s_0e^{2L\lambda n}] \text{ with } Leb(R(v_{n,i})) > (1-\frac{\varepsilon}{4k})\sigma$$

such that for any $r_0\in R(v_{n,i})$ and $r\in[r_0-\sigma e^{-2\lambda L},r_0+\sigma e^{-2\lambda L}]$ we have

$$\begin{aligned}\lambda_{\mathcal{V}_i}-\frac{\varepsilon}{4}&\le\frac{1}{L\lambda}\log\|A_{\mathcal{V}_i}(a_{L\lambda}u(r),a_{L\lambda n}\varphi(s_0))v_{n,i}\|\\&\le\frac{1}{L\lambda}\log\|A_{\mathcal{V}_i}(a_{L\lambda}u(r),a_{L\lambda n}\varphi(s_0))\|\le\lambda_{\mathcal{V}_i}+\frac{\varepsilon}{4}\end{aligned}$$

for $1\le i\le k$. Setting $\widetilde{R}_n := e^{-2L\lambda n}\bigcap_{i=1}^k R(v_{n,i})+s_0\subset J$ we have $Leb(\widetilde{R}_n) > (1-\varepsilon/4)|J|$ and for every $s_1\in\widetilde{R}_n$ if $|s-s_1|\le\sigma e^{-2L\lambda(n+1)}$ then for every $1\le i\le k$ we have

$$\begin{aligned}\lambda_{\mathcal{V}_i}-\frac{\varepsilon}{4}&\le\frac{1}{L\lambda}\log\|A_{\mathcal{V}_i}(a_{L\lambda}u(e^{2L\lambda n}(s-s_0)),a_{L\lambda n}\varphi(s_0))v_{n,i}\|\\&\le\frac{1}{L\lambda}\log\|A_{\mathcal{V}_i}(a_{L\lambda}u(e^{2L\lambda n}(s-s_0)),a_{L\lambda n}\varphi(s_0))\|\le\lambda_{\mathcal{V}_i}+\frac{\varepsilon}{4}.\end{aligned}\tag{7.11}$$

It follows that if for an interval $J'\in\mathcal{I}_{n+1}$ we have $J'\cap\widetilde{R}_n\neq\emptyset$ then (7.11) holds for every $s\in J'$. Since $Leb(\widetilde{R}_n) > (1-\varepsilon/4)|J|$, the number $\tilde{k}$ of such intervals included in J satisfies

$$(\tilde{k}+2)\sigma e^{-2L\lambda(n+1)}\ge Leb(\widetilde{R}_n) > (1-\varepsilon/4)\sigma e^{-2L\lambda n}.$$

Therefore, $\tilde{k}\ge k_0 := \lceil(1-\varepsilon/4)e^{2L\lambda}-2\rceil$ with

$$k_0e^{-2L\lambda}\ge 1-\frac{\varepsilon}{4}-2e^{-2L\lambda} > 1-\frac{\varepsilon}{2}.$$

Let $A_n(J)$ be the union of exactly k_0 intervals $J'\in\mathcal{I}_{n+1}$ with $J'\subset J$ and $J'\cap\widetilde{R}_n\neq\emptyset$. Hence (7.11) holds for every $s\in A_n(J)$. Finally set

$$A_n := \bigcup_{J\in\mathcal{I}_n} A_n(J)\in\mathcal{F}_{n+1}.$$

Since $k_0e^{-2L\lambda} > 1-\varepsilon/2$, by § 7.2,

$$\lim_{n\to\infty}\frac{1}{n}\#\{0\le j\le n-1 : s\in A_j\} = k_0\alpha = k_0e^{-2L\lambda} > 1-\varepsilon/2 \text{ for a.e. } s\in I.$$

On the other hand, $\varphi(s)\in\mathcal{M}$ is Birkhoff generic for a.e. $s\in I$. Therefore for a.e. $s\in I$ we have

$$\liminf_{n\to\infty}\frac{1}{n}\#\{0\le j\le n-1 : a_{L\lambda j}\varphi(s)\in U_\varepsilon\}\ge\mu(U_\varepsilon) > 1-\varepsilon/2.$$

It follows that

$$\liminf_{n\to\infty}\frac{1}{n}\#\{0\leq j\leq n-1: a_{L\lambda j}\varphi(s)\in U_\varepsilon, s\in A_j\}>1-\varepsilon \tag{7.12}$$

for a.e. $s\in I$. Therefore for every $s\in I$ satisfying (7.12) there exists $n_0=n_0(s)$ so that setting

$$D(s):=\{j\in\mathbb{N}: a_{L\lambda j}\varphi(s)\in U_\varepsilon, s\in A_j, j\geq 2\rho+2\}$$

we have $\#(D(s)\cap[0,n-1])>(1-\varepsilon)n$ for $n\geq n_0$.

Take any $s\in I$ satisfying (7.12) and let us consider any $n\in D(s)$ greater than ρ. Choose $J\in\mathcal{I}_n$ containing s. Then $a_{L\lambda n}\varphi(J)\cap U_\varepsilon\neq\emptyset$. Set $s_0:=s_{0,n}=s_0(J)\in J$. Then

$$a_{L\lambda n}\varphi(s)\in U_\varepsilon\subset\pi(\Delta_{2\sigma}). \tag{7.13}$$

Since $s\in A_n\cap J=A_n(J)$, the inequalities (7.11) hold for s and $n\in D(s)$. Moreover, $a_{L\lambda n}\varphi(s)\in U_\varepsilon$ implies

$$a_{L\lambda n}\varphi(s), a_{L\lambda(n+1)}\varphi(s)\in\pi(\Delta_{2\sigma}).$$

The points s, s_0 belong to $J\in\mathcal{I}_n$, so $|e^{2L\lambda n}(s-s_0)|\leq\sigma$. In view of (2.10), it follows that for every $2(\rho+1)\leq n$ and $0\leq m\leq n+1$ we have

$$\begin{aligned}
&d\big(a_{L\lambda m}\cdot\widetilde{\varphi}(s), u(e^{2L\lambda m}(s_0-s))\cdot a_{L\lambda m}\cdot\widetilde{\varphi}(s_0)\big)\\
&\leq C(\varphi)|e^{2L\lambda m}(s_0-s)|(1+e^{2\rho L\lambda m}|s_0-s|^\rho)e^{-L\lambda m}\\
&\leq C(\varphi)\sigma e^{2L\lambda(m-n)}(1+\sigma^\rho e^{2\rho L\lambda(m-n)})e^{-L\lambda m}\\
&\leq C(\varphi)\sigma(1+\sigma^\rho)e^{L\lambda(m-2n+2\rho\max(0,m-n))}\leq C(\varphi)\sigma(1+\sigma^\rho)e^{-L\lambda}\leq\sigma.
\end{aligned}\tag{7.14}$$

By (7.14) applied to $m=n$ and $m=n+1$, we have

$$d(a_{L\lambda n}\cdot\widetilde{\varphi}(s), u(e^{2L\lambda n}(s_0-s))\cdot a_{L\lambda n}\cdot\widetilde{\varphi}(s_0))\leq\sigma$$

and

$$d(a_{L\lambda}\cdot a_{L\lambda n}\cdot\widetilde{\varphi}(s), a_{L\lambda}\cdot u(e^{2L\lambda n}(s_0-s))\cdot a_{L\lambda n}\cdot\widetilde{\varphi}(s_0))\leq\sigma.$$

Since $a_{L\lambda n}\varphi(s)$ and $a_{L\lambda}a_{L\lambda n}\varphi(s)$ belong to $\pi(\Delta_{2\sigma})$, by Lemma 7.6,

$$A^p(a_{L\lambda}, a_{L\lambda n}\varphi(s))=A^p(a_{L\lambda}, u(e^{2L\lambda n}(s-s_0))a_{L\lambda n}\varphi(s_0)). \tag{7.15}$$

In view of (2.9), we have

$$d\big(a_{L\lambda n}\widetilde{\varphi}(s_0), u(e^{2L\lambda n}(s_0-s))a_{L\lambda n}\widetilde{\varphi}(s_0)\big)\leq|e^{2L\lambda n}(s-s_0)|\leq\sigma. \tag{7.16}$$

Since $a_{L\lambda n}\varphi(s_0)\in\pi(\Delta_{2\sigma})$ (see (7.13)), by the second part of Lemma 7.6, it follows that

$$A^p(u(e^{2L\lambda n}(s_0-s)), a_{L\lambda n}\varphi(s_0))=Id.$$

Therefore, by (7.15),

$$\begin{aligned}
&A^p(a_{L\lambda}u(e^{2L\lambda n}(s_0-s)), a_{L\lambda n}\varphi(s_0))\\
&=A^p(a_{L\lambda}, u(e^{2L\lambda n}(s_0-s))a_{L\lambda n}\varphi(s_0))A^p(u(e^{2L\lambda n}(s_0-s)), a_{L\lambda n}\varphi(s_0))\\
&=A^p(a_{L\lambda}, a_{L\lambda n}\varphi(s)).
\end{aligned}\tag{7.17}$$

In view of (7.16) and (7.14), we also have

$$\begin{aligned}
d(a_{L\lambda n}\widetilde{\varphi}(s), a_{L\lambda n}\widetilde{\varphi}(s_0))\leq\,&d(a_{L\lambda n}\widetilde{\varphi}(s), u(e^{2L\lambda n}(s_0-s))a_{L\lambda n}\widetilde{\varphi}(s_0))\\
&+d(a_{L\lambda n}\widetilde{\varphi}(s_0), u(e^{2L\lambda n}(s_0-s))a_{L\lambda n}\widetilde{\varphi}(s_0))\leq 2\sigma.
\end{aligned}$$

Since $\tilde{\varphi}(s),\tilde{\varphi}(s_0)\in\Delta$ and $a_{L\lambda n}\varphi(s)\in\pi(\Delta_{2\sigma})$, by Lemma 7.6, it follows that

$$A^p(a_{L\lambda n},\varphi(s))=A^p(a_{L\lambda n},\varphi(s_0)). \tag{7.18}$$

In summary, (7.11) combined with (7.10), (7.17) and (7.18) yields for $1 \leq i \leq k$

$$\begin{aligned}\lambda_{\mathcal{V}_i} - \frac{\varepsilon}{4} &\leq \frac{1}{L\lambda} \log \frac{\|A^p(a_{L\lambda(n+1)}, \varphi(s))v_i(s_{0,n})\|}{\|A^p(a_{L\lambda n}, \varphi(s))v_i(s_{0,n})\|} \\ &\leq \frac{1}{L\lambda} \log \|A^p(a_{L\lambda}, a_{L\lambda n}\varphi(s))|_{\mathcal{V}_i(a_{L\lambda n}\varphi(s_{0,n}))}\| \leq \lambda_{\mathcal{V}_i} + \frac{\varepsilon}{4}\end{aligned} \tag{7.19}$$

for every $n \in D(s)$. Moreover, by (7.3),

$$\frac{1}{L\lambda}\big| \log \|A^p(a_{L\lambda}, a_{L\lambda n}\varphi(s))\|\big| \leq \frac{\log C + NL\lambda}{L\lambda} \leq C + N.$$

Therefore, by (7.7), for $n \geq n_0$ we have

$$\begin{aligned}&\frac{1}{L\lambda n} \log \|A^p(a_{L\lambda n}, \varphi(s))\| \leq \frac{1}{n}\sum_{j=0}^{n-1} \frac{1}{L\lambda} \log \|A^p(a_{L\lambda}, a_{L\lambda j}\varphi(s))\| \\ &\leq \frac{1}{n} \sum_{j\in D(s)\cap[0,n-1]} \frac{1}{L\lambda} \max_{1\leq i\leq k} \log C'\|A^p(a_{L\lambda}, a_{L\lambda j}\varphi(s))|_{\mathcal{V}_i(a_{L\lambda j}\varphi(s_{0,n}))}\| \\ &+ \frac{1}{n} \sum_{D(s)\cap[0,n-1]} \frac{1}{L\lambda} \log \|A^p(a_{L\lambda}, a_{L\lambda j}\varphi(s))\| \\ &\leq (\lambda_{top} + \frac{\varepsilon}{4} + \frac{C'}{L\lambda})\frac{\#(D(s)\cap[0,n-1])}{n} + \frac{\#([0,n-1]\setminus D(s))}{n}(C+N) \\ &\leq \lambda_{top} + \frac{\varepsilon}{2} + \varepsilon(C+N) \leq \lambda_{top} + \varepsilon(1+C+N).\end{aligned} \tag{7.20}$$

It follows that for a.e. $s \in I$ we have

$$\limsup_{n\to\infty} \frac{1}{L\lambda n} \log \|A^p(a_{L\lambda n}, \varphi(s))\| \leq \lambda_{top} + \varepsilon(1+C+N).$$

Let us consider the cocycle $((A^p)^{-1})^{tr}(g,x) = (A^p(g,x)^{-1})^{tr}$. The Lyapunov exponents of $((A^p)^{-1})^{tr}$ (with respect to μ) coincide with additive inverses of the Lyapunov exponents of A^p. As the last set is symmetric, the Lyapunov exponents of $((A^p)^{-1})^{tr}$ and A^p are the same. Moreover, $((A^p)^{-1})^{tr}$ meets all properties of A^p used in the current part of the proof, more precisely, $((A^p)^{-1})^{tr}$ is the composition of A with a group endomorphism so that the results of [16] can be applied. It follows that for a.e. $s \in I$ we have

$$\limsup_{n\to\infty} \frac{1}{L\lambda n} \log \|((A^p)^{-1})^{tr}(a_{L\lambda n}, \varphi(s))\| \leq \lambda_{top} + \varepsilon(1+C+N).$$

Since $\lambda_{top} < 1$ and $\varepsilon > 0$ is small, after more restrictive choice of s we have

$$\|A^p(a_{L\lambda m}, \varphi(s))\|, \|A^p(a_{L\lambda m}, \varphi(s))^{-1}\| \leq e^{(1-\varepsilon)L\lambda m}$$

for all m large enough. Let us recall that $s, s_0(J) \in J$ and the length of $J \in \mathcal{I}_n$ is equal to $\sigma e^{-2L\lambda n}$. Therefore, by (7.8), if $m \leq n+1$ then

$$\begin{aligned}&\Big|\log \frac{\|A_{\mathcal{V}_i}(a_{L\lambda m}, \varphi(s))v_i(s)\|}{\|A^p(a_{L\lambda m}, \varphi(s))v_i(s_{0,n})\|}\Big| \\ &\leq \|A^p(a_{L\lambda m}, \varphi(s))\|\|A^p(a_{L\lambda m}, \varphi(s))^{-1}\|\|v_i(s) - v_i(s_{0,n})\| \\ &\leq e^{(2-\varepsilon)L\lambda m} l\sigma e^{-2L\lambda n} = l\sigma e^{(2-\varepsilon(n+1))L\lambda} \leq l\sigma\end{aligned} \tag{7.21}$$

for any $n \in D(s)$. In view of (7.19) and (7.9), it follows that for $n \in D(s)$ we have

$$\frac{1}{L\lambda} \log \frac{\|A_{\mathcal{V}_i}(a_{L\lambda(n+1)}, \varphi(s))v_i(s)\|}{\|A_{\mathcal{V}_i}(a_{L\lambda n}, \varphi(s))v_i(s)\|} \geq \lambda_{\mathcal{V}_i} - \frac{\varepsilon}{4} - \frac{2l\sigma}{L\lambda} \geq \lambda_{\mathcal{V}_i} - \frac{\varepsilon}{2}.$$

Moreover, by (7.4), for every $n \geq 0$

$$\frac{1}{L\lambda}\Big|\log\frac{\|A_{\mathcal{V}_i}(a_{L\lambda(n+1)},\varphi(s))v_i(s)\|}{\|A_{\mathcal{V}_i}(a_{L\lambda n},\varphi(s))v_i(s)\|}\Big| \leq \frac{\log C+NL\lambda}{L\lambda} \leq C+N.$$

Choose a subbundle $\mathcal{V}_i$ so that $\lambda_{\mathcal{V}_i} = \lambda_{top}$. Then, for $n \geq n_0$

$$\begin{aligned}\frac{1}{L\lambda n}\log\|A^p(a_{L\lambda n},\varphi(s))\| &\geq \frac{1}{L\lambda n}\log\|A_{\mathcal{V}_i}(a_{L\lambda n},\varphi(s))v_i(s)\|\\ &= \frac{1}{n}\sum_{j=0}^{n-1}\frac{1}{L\lambda}\log\frac{\|A_{\mathcal{V}_i}(a_{L\lambda(j+1)},\varphi(s))v_i(s)\|}{\|A_{\mathcal{V}_i}(a_{L\lambda j},\varphi(s))v_i(s)\|}\\ &\geq (\lambda_{\mathcal{V}_i}-\frac{\varepsilon}{2})\frac{\#(D(s)\cap[0,n-1])}{n} - \frac{\#([0,n-1]\setminus D(s))}{n}(C+N)\\ &\geq (\lambda_{top}-\frac{\varepsilon}{2})(1-\varepsilon)-\varepsilon(C+N) \geq \lambda_{top}-\varepsilon(2+C+N).\end{aligned} \tag{7.22}$$

In view of (7.20) and (7.22), it follows that for a.e. $s \in I$

$$\begin{aligned}\lambda_{top}-\varepsilon(2+C+N) &\leq \liminf_{n\to\infty}\frac{1}{L\lambda n}\log\|A^p(a_{L\lambda n},\varphi(s))\|\\ &\leq \limsup_{n\to\infty}\frac{1}{L\lambda n}\log\|A^p(a_{L\lambda n},\varphi(s))\| \leq \lambda_{top}+\varepsilon(1+C+N).\end{aligned}$$

Suppose that $L\lambda n \leq t < L\lambda(n+1)$. Since $0 \leq t - L\lambda n < L\lambda$, by (7.3),

$$\big|\log\|A^p(a_t,\varphi(s))\| - \log\|A^p(a_{L\lambda n},\varphi(s))\|\big| \leq \log C + NL\lambda \leq L\lambda(C+N).$$

Therefore,

$$\begin{aligned}\frac{\log\|A^p(a_{L\lambda n},\varphi(s))\|}{L\lambda(n+1)} - \frac{C+N}{n+1} &\leq \frac{\log\|A^p(a_t,\varphi(s))\|}{t}\\ &\leq \frac{\log\|A^p(a_{L\lambda n},\varphi(s))\|}{L\lambda n}+\frac{C+N}{n}.\end{aligned}$$

It follows that for a.e. $s \in I$

$$\begin{aligned}\lambda_{top}-\varepsilon(2+C+N) &\leq \liminf_{t\to+\infty}\frac{1}{t}\log\|A^p(a_t,\varphi(s))\|\\ &\leq \limsup_{t\to+\infty}\frac{1}{t}\log\|A^p(a_t,\varphi(s))\| \leq \lambda_{top}+\varepsilon(1+C+N).\end{aligned}$$

Since $\varepsilon > 0$ can be chosen arbitrary small, we obtain

$$\lim_{t\to+\infty}\frac{1}{t}\log\|A^p(a_t,\varphi(s))\| = \lambda_{top} \text{ for a.e. } s \in I.$$

□

7.5. **A more general conditional version of Theorem 7.1.** In this section we explain how, assuming recent work in progress [12] by Eskin-Filip-Wright, one can remove the sum of Lyapunov exponents in Theorem 2.5 and obtain the more general formulation presented at the end of this section (see Theorem 7.9 at the very end).

Let us consider the splitting $\bigoplus_{i=1}^k \mathcal{V}_i$ of $\mathcal{M}\times\bigwedge^p W$ into strongly irreducible invariant subbundles $\mathcal{V}_i$, used in the proof of Theorem 7.1. Let us first of all remark that *if* $\mathcal{V}_i$ *are all locally constant* then the assumption $\lambda_{top} < 1$ can be removed from Theorem 7.1. Indeed, under this assumption, the proof of Theorem 7.1 becomes easier, i.e. the smooth curves $s \mapsto v_i(s)$ can be chosen to be constant and hence we do not need to use the inequality (7.21) which requires the assumption that $\lambda_{top} < 1$. We remark in addition that this is the only place in the proof where this assumption is used.

Filip in [16] showed that these subbundles locally vary polynomially in the period coordinates. As it was recently announced by Eskin (work in progress [12]), one can deduce from this property that $\mathcal{V}_i$ are indeed locally constant whenever the space W satisfies a condition described in the following paragraph.

Let $\mathcal{M} \subset \mathcal{M}_1(M,\Sigma)$ be an orbit closure. Since $\mathcal{M}$ is an affine submanifold its tangle bundle $T\mathcal{M}$, in period coordinates (see [1]) is determined by a subspace $T_{\mathbb{C}}\mathcal{M} \subset H^1(M,\Sigma,\mathbb{C})$ such that

$$T_{\mathbb{C}}\mathcal{M} = \mathbb{C} \otimes T_{\mathbb{R}}\mathcal{M} \quad \text{where} \quad T_{\mathbb{R}}\mathcal{M} \subset H^1(M,\Sigma,\mathbb{R}).$$

Let $p : H^1(M,\Sigma,\mathbb{R}) \to H^1(M,\mathbb{R})$ be the natural projection. Let us consider the space $p(T_{\mathbb{R}}\mathcal{M}) \subset H^1(M,\mathbb{R})$. Since $H^1(M,\mathbb{R})$ and $H_1(M,\mathbb{R})$ can be identified by the Poincaré duality, we will also denote by $p(T_{\mathbb{R}}\mathcal{M})$ the corresponding subspace of $H_1(M,\mathbb{R})$. Avila-Eskin-Möller, in [1], proved that $p(T_{\mathbb{R}}\mathcal{M})$ is an $SL_2(\mathbb{R})$-invariant symplectic subspace.

Theorem 7.7 ([12])**.** *If the space*

$$W \text{ is symplectic orthogonal to } p(T_{\mathbb{R}}\mathcal{M}) \tag{7.23}$$

then for every $p \geq 1$ the bundle $\mathcal{M} \times \bigwedge^p W$ has a splitting into locally constant strongly irreducible invariant subbundles.

Thus, the above remark combined with Theorem 7.7 yield the proof of Theorem 2.5 under condition (7.23) and without the assumption on the sum of positive Lyapunov exponents of A_W^{KZ}.

Let us remark that the key condition (7.23) holds for some classes of natural invariant subspaces when M is a ramified cover of another surface. Suppose that $q : M \to N$ is a covering map, ramified over a finite set Σ_q. Let us consider a stratum $\mathcal{M}(\alpha) \subset \mathcal{M}(N,\Sigma_q)$. Denote by $\widetilde{\mathcal{M}}(\alpha)$ the moduli space of all translation surfaces of the form $(M, q^*\nu)$ for $\nu \in \mathcal{M}(\alpha)$. Denote by D the deck group of the ramified cover $q : M \to N$. Let us consider the subspace $H_1^+(M,\mathbb{R}) \subset H_1(M,\mathbb{R})$ of all D-invariant elements. Then $H_1^+(M,\mathbb{R})$ is a symplectic $SL_2(\mathbb{R})$-invariant subspace which is naturally identified with $H_1(N,\mathbb{R})$. The space which plays a crucial role in billiard or lenses applications is its symplectic orthocomplement, namely

$$H_1^-(M,\mathbb{R}) := H_1^+(M,\mathbb{R})^{\perp}.$$

Also the space $H_1^-(M,\mathbb{R})$ is clearly symplectic and $SL_2(\mathbb{R})$-invariant.

Lemma 7.8. *If an orbit closure $\mathcal{M}$ is contained in $\widetilde{\mathcal{M}}(\alpha)$ then $p(T_{\mathbb{R}}\mathcal{M}) \subset H_1^+(M,\mathbb{R})$. Consequently, $H_1^-(M,\mathbb{R})$ is symplectic orthogonal to $p(T_{\mathbb{R}}\mathcal{M})$.*

Proof. Let us consider the induced action of the group D on $\mathcal{M}(M,\Sigma)$. By assumption, $\mathcal{M}$ is a subset of the set $\mathrm{Fix}_G\,\mathcal{M}(M,\Sigma)$ of fixed points of this action. Next, let us consider the induced action of D on the tangent bundle $T\mathcal{M}(M,\Sigma)$. Then

$$T\mathcal{M} \subset \mathrm{Fix}_D\, T\mathcal{M}(M,\Sigma).$$

Since each tangent space of $T\mathcal{M}(M,\Sigma)$ is identified with $H^1(M,\Sigma,\mathbb{C})$ and the identification is D-equivariant, we have

$$T_{\mathbb{C}}\mathcal{M} \subset \mathrm{Fix}_D\, H^1(M,\Sigma,\mathbb{C})$$

and hence

$$T_{\mathbb{R}}\mathcal{M} \subset \mathrm{Fix}_D\, H^1(M,\Sigma,\mathbb{R}).$$

It follows that

$$p(T_{\mathbb{R}}\mathcal{M}) \subset \mathrm{Fix}_D\, H^1(M,\mathbb{R}).$$

Finally, using the Poincaré duality we have $p(T_{\mathbb{R}}\mathcal{M}) \subset H_1^+(M,\mathbb{R})$. □

To conclude the present section, we can formulate the final result which we hope will be useful for further mathematical physics applications.

Theorem 7.9. *Let $\mathcal{M}$ be an $SL_2(\mathbb{R})$-orbit closure contained in $\widetilde{\mathcal{M}}(\alpha)$. Let $W \subset H_1^-(M,\mathbb{R})$ be a symplectic and $SL_2(\mathbb{R})$-invariant subspace. Suppose that $\varphi : I \to \mathcal{M}$ is a curve which is well approximated by horocycles and such that $\varphi(s)$ is Birkhoff generic with respect to $(\mathcal{M},\mu,a_t)$ for a.e. $s \in I$. Then one has that $\varphi(s)$ is Oseledets generic with respect to $(\mathcal{M},\mu,a_t,A_W^{KZ})$ for a.e. $s \in I$.*

Acknowledgments

We are indebted to J. Marklof, both for asking to R. Shi the question on gaps distributions answered in Theorem 1.3) and for showing to C. Ulcigrai the paper [9] by Dragović and Radnović (without knowing that he was suggesting two applications of the same result). We are grateful to J. Chaika and A. Eskin for explaining to us their work in [7] and patiently answering all our questions on it. We thank Y. Benoist, Y. Cheung, A.Mohammadi, A. Zorich and B. Weiss for useful discussions and comments.

We acknowledge both the *Max Planck Institute* in Bonn and the *Mathematical Science Research Institute* in Berkeley, CA for their hospitality when part of this collaborative work was done, respectively when K. Frączek and C. Ulcigrai were visitors during the program "*Dynamics and Numbers*" in Bonn and when R. Shi and C. Ulcigrai were research members of the program "*Geometric and Arithmetic Aspects of Homogeneous Dynamics*" at MSRI. We also thank the *University of Bristol* for hosting R. Shi for a seminar and research visit which started this collaboration.

K. Frączek is partially supported by the National Science Centre (Poland) grant 2014/13/B/ST1/03153. R. Shi is supported by by NSFC (11201388), NSFC (11271278). C. Ulcigrai is currently supported by the ERC Starting Grant ChaParDyn: the research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 335989. Part of this work was done under NSF 0932078 000 while R. Shi and C. Ulcigrai were visiting MSRI during Spring 2015.

References

[1] A. Avila, A. Eskin and M. Möller, *Symplectic and isometric SL(2,R)-invariant subbundles of the Hodge bundle*, to appear in J. Reine Angew. Math.

[2] A. Avila, S. Gouëzel and J.-Ch. Yoccoz, *Exponential mixing for the Teichmüller flow*, Publ. Math. Inst. Hautes Études Sci. 104 (2006), 143-211.

[3] A. Avila and P. Hubert, *Recurrence for the wind-tree model*, to appear in Ann. Inst. H. Poincaré Anal. Non Linéaire.

[4] M. Bainbridge, *Euler characteristics of Teichmüller curves in genus two*, Geom. Topol. 11 (2007), 1887-2073.

[5] T. Browning and I. Vinogradov, *Effective Ratner theorem for $SL(2,\mathbb{R}) \ltimes \mathbb{R}^2$ and gaps in $\sqrt{n}$ modulo* 1, `arXiv:1311.6387`.

[6] Y. Benoist and J.-F. Quint, *Stationary measures and invariant subsets of homogeneous spaces (II)*, J. Amer. Math. Soc. 26 (2013), 659-734.

[7] J. Chaika and A. Eskin, *Every flat surface is Birkhoff and Oseledets generic in almost every direction*, J. Mod. Dyn. 9 (2015), 1-23.

[8] V. Delecroix, P. Hubert and S. Lelièvre, *Diffusion for the periodic wind-tree model*, Ann. Sci. Éc. Norm. Supér. (4) 47 (2014), 1085-1110.

[9] V. Dragović and M. Radnović, *Pseudo-integrable billiards and arithmetic dynamics*, J. Mod. Dyn. 8 (2014), 109-132.

[10] M. Einsiedler and E. Lindenstrauss, *Diagonal actions on locally homogeneous spaces.* In: it Homogeneous Flows, Moduli Spaces and Arithmetic, Clay Math. Proc. 10, Amer. Math. Soc. (2010), 155-241.

[11] N.D. Elkies and C.T. McMullen, *Gaps in* $\sqrt{n}$ mod 1 *and ergodic theory*, Duke Math. J. 123(1) (2004), 95-139.

[12] A. Eskin, S. Filip and A. Wright, *The algebraic hull of the Kontsevich-Zorich cocycle*, in preparation.
[13] A. Eskin, G.A. Margulis and S. Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. (2) 147 (1998), 93-141.
[14] A. Eskin and M. Mirzakhani, *Invariant and stationary measures for the $SL_2(\mathbb{R})$ action on moduli space*, `arXiv:1302.3320`.
[15] A. Eskin, M. Mirzakhani and A. Mohammadi, *Isolation, equidistribution and orbit closures for the $SL(2,\mathbb{R})$ action on moduli space*, Ann. of Math. (2) 182 (2015), 673-721.
[16] S. Filip, *Semisimplicity and rigidity of the Kontsevich-Zorich cocycle*, `arXiv:1307.7314`.
[17] G. Forni, *Deviation of ergodic averages for area-preserving flows on surfaces of higher genus*, Ann. of Math. (2) 155 (2002), 1-103.
[18] R. Fox and R. Kershner, *Concerning the transitive properties of geodesics on a rational polyhedron*, Duke Math. J. 2 (1936), 147-150.
[19] K. Frączek and P. Hubert, *Recurrence and non-ergodicity in generalized wind-tree models*, `arXiv:1506.05884`.
[20] K. Frączek and M. Schmoll, *Directional localization of light rays in a periodic array of retro-reflector lenses*, Nonlinearity 27 (2014), 1689-1707.
[21] K. Frączek and C. Ulcigrai, *Non-ergodic $\mathbb{Z}$-periodic billiards and infinite translation surfaces*, Invent. Math. 197 (2014), 241-298.
[22] I.Ya. Goldsheid, G.A. Margulis, *Lyapunov exponents of a product of random matrices*, Russian Math. Surveys 44 (1989), 11-71.
[23] Y. Guivarc'h and A.N. Starkov, *Orbits of linear group actions, random walks on homogeneous spaces and toral automorphisms*, Ergodic Theory Dynam. Systems 24 (2004), 767-802.
[24] J.H. Hannay and T.M. Haeusserab, *Retroreflection by refraction*, J. Mod. Opt. 40 (1993), 1437-1442.
[25] A. Katok and A. Zemljakov, *Topological transitivity of billiards in polygons*, (Russian) Mat. Zametki 18 (1975), 291-300.
[26] S. Kerckhoff, H. Masur and J. Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. of Math. (2) 124 (1986), 293-311.
[27] D. Kleinbock and G.A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) 148 (1998), 339-360.
[28] H. Masur, *Hausdorff dimension of the set of nonergodic foliations of a quadratic differential*, Duke Math. J. 66 (1992), 387-442.
[29] S. Mozes, *Epimorphic subgroups and invariant measures*, Ergodic Theory Dynam. Systems 15 (1995), 1207-1210.
[30] M. Ratner, *On Raghunathan's measure conjecture*, Ann. of Math. (2) 134 (1991), 545-607.
[31] D. Ruelle, *Ergodic theory of differentiable dynamical systems*, Inst. Hautes Études Sci. Publ. Math. 50 (1979), 275-306.
[32] L. Russell, *Strong laws of large numbers for weakly correlated random variables*, Michigan Math. J. 35 (1988), 353-359.
[33] R. Shi, *Pointwise equidistribution for one parameter diagonalizable group action on homogeneous space*, `arXiv:1405.2067`.
[34] A.N. Starkov, *Dynamical systems on homogeneous spaces*, Translations of Mathematical Monographs, 190. American Mathematical Society, Providence, RI, 2000.
[35] S. Tabachnikov, *Geometry and Billiards*, Student Mathematical Library, vol. 30, American Mathematical Society, Providence, RI, 2005.
[36] A. Zorich, *Deviation for interval exchange transformations*, Ergodic Theory Dynam. Systems 17 (1997), 1477-1499.
[37] A. Zorich, *How do the leaves of a closed 1-form wind around a surface?*, Amer. Math. Soc. Transl. Ser. 2, 197 (1999), 135-178.

Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland
E-mail address: `fraczek@mat.umk.pl`

School of Mathematical Sciences, Xiamen University, Xiamen 361005, PR China
E-mail address: `ronggang@xmu.edu.cn`

School of Mathematics, University of Bristol, Howard House, Queens Ave, BS8 1SD, Bristol, United Kingdom
E-mail address: `corinna.ulcigrai@bris.ac.uk`