GAP DISTRIBUTIONS AND HOMOGENEOUS DYNAMICS

JAYADEV S. ATHREYA

Abstract. We survey the use of dynamics of $SL(2, \mathbb{R})$-actions to understand gap distributions for various sequences of subsets of $[0, 1)$, particularly those arising from special trajectories of various two-dimensional dynamical systems. We state and prove an abstract theorem that gives a unified explanation for some of the examples we present.

1. Introduction

The study of the distribution of gaps in sequences is a subject that arises in many different contexts and has connections with many different areas of mathematics, including number theory, probability theory, and spectral analysis. In this paper, we study gap distributions from the perspective of dynamics and geometry, exploring examples connected with the dynamics of $SL(2, \mathbb{R})$-actions on moduli spaces of geometric objects, in particular the space of lattices and the space of translation surfaces.

The inspiration for this article is the a quote from the beautiful paper of Elkies-McMullen [8], referring to their explicit computation of the gap distribution of the sequence of fractional parts of $\sqrt{n}$, using the dynamics of the $SL(2, \mathbb{R})$-action on the space of affine unimodular lattices in $\mathbb{R}^2$.

"...the uniform distribution of lattices explains the exotic distribution of gaps."

Indeed, the main results of our paper, Theorem 2.2, Theorem 2.5, and Theorem 2.6, give unified explanations of several examples of ‘exotic’ gap distributions via uniform distribution on various moduli spaces of geometric objects.

1.1. Equidistribution, randomness, and gap distributions. Suppose that for each positive integer $k$, we are given a finite list of points $F(k) \subset [0, 1)$, where by a list, we mean a finite non-decreasing sequence of real numbers where $N_k$ denotes the number of terms in the $k^{th}$ sequence $F(k)$. We write

$$F(k) = \{ F_k^{(0)} \leq F_k^{(1)} \leq \ldots \leq F_k^{(N_k)} \},$$

and we assume $N_k \to \infty$ as $k \to \infty$. In many situations, we are interested in the ‘randomness’ of the sequence of lists $\{F(k)\}_{k=1}^\infty$. A first test of ‘randomness’ is whether the lists $F(k)$ uniformly distribute in $[0, 1)$, that is the

J.S.A. partially supported by NSF grant DMS 1069153.
measures $\Delta_k = \frac{1}{N_k} \sum_{j=0}^{N_k} \delta_{F_k^{(j)}}$ converge weak-* to Lebesgue measure, i.e., for any $0 \leq a \leq b \leq 1$,

\[
\lim_{k \to \infty} \Delta_k(a, b) = b - a.
\]

A more refined question (not necessarily dependent on (1.1)) is to examine the distribution of gaps for the sequences $F(k)$. That is, form the associated normalized gap sets

\[
G(k) := \left\{ N_k \left( F_k^{(i+1)} - F_k^{(i)} \right) : 0 \leq i < N_k \right\},
\]

and given $0 \leq a < b \leq \infty$, what is the behavior of

\[
\lim_{k \to \infty} \frac{|G(k) \cap (a, b)|}{N_k}?
\]

If the sequence $F(k)$ is ‘truly random’, that is, given by

\[
F(k) = \{X(0) \leq X(1) \leq \ldots \leq X(k)\},
\]

where the $\{X(i)\}$ are the order statistics generated by independent, identically distributed (i.i.d.) uniform $[0, 1)$ random variables $\{X_n\}_{n=0}^\infty$, it is an exercise in probability theory to show that the gap distribution converges to a Poisson process of intensity 1. Precisely, for any $t > 0$,

\[
\lim_{k \to \infty} \frac{|G(k) \cap (t, \infty)|}{N_k} = e^{-t}
\]

However, many sequences that arise ‘in nature’ satisfy an equidistribution property but do not have Poissonian gaps. Following [8], we call such gap distributions exotic. In this paper, we discuss in detail some examples of exotic gap distributions, which, moreover, can be calculated (or at least shown to exist) using methods arising from homogeneous dynamics, in particular dynamics of $SL(2, \mathbb{R})$ actions on appropriate moduli spaces. In particular, the results we discuss share a similar philosophy; the sets $F(k)$ are associated to sets of angles or slopes of a discrete set of vectors in $\mathbb{R}^2$, and the gap distribution is studied by appropriate linear renormalizations, which can be viewed as part of an $SL(2, \mathbb{R})$ action on an appropriate moduli space of geometric objects. The main novelty of this paper is the statement of three meta-theorems (Theorem 2.2, Theorem 2.5, and Theorem 2.6), which give unified explanations of some of these examples by linking them to uniform distribution on various moduli spaces and which we expect can be used for future applications.

1.2. Organization of the paper. This paper is organized as follows: in the remainder of this introduction we state results about our main (previously studied) examples: the Farey sequences $F(Q)$ (§1.4); slopes for lattice vectors (§1.4.1); and saddle connection directions for translation surfaces (§1.5). We also briefly discuss the space of affine lattices and $\{\{\sqrt{n}\}\}_{n \geq 1}$ in §1.6. In §2, we state the main results Theorem 2.2, Theorem 2.5, and Theorem 2.6. We describe how to use these results to explain our examples.
in §§3-4, and prove the theorems in §5. Finally, in §6 we pose some natural questions suggested by our approach.

1.3. **Acknowledgements.** Parts of this work are based on joint work with J. Chaika, Y. Cheung, and S. Lelievre. I thank them for their stimulating intellectual partnership. Discussions with F. Boca, M. Boshernitzan, A. Eskin, J. Marklof, and A. Zaharescu helped the author to see the common structural features in this circle of ideas, and without them this paper would not have been possible. In particular, the stunning paper [14] of J. Marklof and A. Strombergsson has inspired not only this paper but the papers [2, 3, 4]. I would also like to thank the Polish Academy of Sciences for their hospitality, and in particular Piotr Przytycki for arranging a visit to Warsaw where many of the ideas in this paper were clarified. Finally, thanks are due to the organizers of the ICM Satellite Conference on Geometry, Topology, and Dynamics in Negative Curvature for the opportunity to speak at the meeting and also to contribute to this volume. We would also like to thank the anonymous referee for their careful reading and many remarks which clarified the exposition of this paper.

1.4. **Farey sequences.** Consider the integer lattice $\mathbb{Z}^2$. If we imagine an observer sitting at the origin 0, the ‘visible’ points in $\mathbb{Z}^2$ correspond to the set of primitive vectors, that is, integer vectors which are not integer multiples of other integer vectors. If we consider slopes of vectors (as opposed to angles), it is natural to consider the set of vectors with slopes in $[0, 1]$. The set of slopes of (primitive) integer vectors with horizontal component at most $Q$ intersected with the interval $[0, 1]$ gives the Farey sequence of level $Q$. More simply, $\mathcal{F}(Q)$ consists of the set of fractions in between 0 and 1 with denominator at most $Q$. We write

$$\mathcal{F}(Q) := \left\{ \frac{0}{1} < \gamma_1 = \frac{1}{Q} < \gamma_2 < \ldots < \gamma_i = \frac{p_i}{q_i} < \ldots \gamma_N = \frac{1}{1} \right\}$$

Here, $N = N(Q) = \sum_{i=1}^{Q} \varphi(i)$ is the cardinality of $\mathcal{F}(Q)$. By the above discussion, these correspond to the slopes of primitive integer vectors $\left( \frac{q_i}{p_i} \right)$ in the (closed) triangle $T_Q$ with vertices at $(0,0)$, $(Q,0)$, and $(Q,Q)$. That is, it is bounded above by the line $\{y = x\}$, below by the $x$-axis, and on the right by the line $\{x = Q\}$. The triangle $T_4$ is shown in Figure 1.

The sequences $\mathcal{F}(Q)$ equidistribute in $[0, 1]$ (by, for example, Weyl’s criterion [22]). We denote by $\mathcal{G}(Q)$ the set of normalized gaps between Farey fractions, that is,

$$\mathcal{G}(Q) = \left\{ \frac{N(Q)}{q_i q_{i+1}} : 0 \leq i < N(Q) \right\}.$$
Figure 1. The triangle $T_4$. Primitive integer vectors are given by dashed lines, and are labeled by their slopes.

The limiting distribution for $G(Q)$ is given by the following beautiful theorem of R. R. Hall, and illustrated in Figures 2 and 3. Let

$$\Omega := \{(u, v) \in [0,1]^2 : u + v > 1\},$$

and for $0 \leq a, b < \infty$, let

$$A_{a,b} = \{(u, v) \in \Omega : b^{-1} < uv < a^{-1}\},$$

and set $\tilde{A}_{a,b} := A_{a^{2/3}, a^{2/3}b}$.

**Theorem 1.1.** (R.R.Hall) Fix $0 \leq a < b < \infty$. Then

$$\lim_{Q \to \infty} \frac{|G(Q) \cap (a, b)|}{N(Q)} = 2|\tilde{A}_{a,b}|.$$

Differentiating the cumulative distribution function

$$F_{Hall}(t) := |\tilde{A}_{0,t}|,$$

one can compute the probability distribution function $P_{Hall}(t)$ so that

$$\int_{a,b} P_{Hall}(t) dt := 2|\tilde{A}_{a,b}|.$$

We call this distribution (and any scalings) Hall’s distribution. The graph of $P_{Hall}(t)$ is given in Figure 3 which is drawn from [6]. The points of non-differentiability $\frac{3}{\pi^2}$ and $\frac{12}{\pi^2}$ correspond to the transitions when the hyperbola $\{xy = \frac{3}{\pi^2}t^{-1}\}$ enters the region $\Omega$ ($t = \frac{3}{\pi^2}$) and when it hits the line $x + y = 1$ ($t = \frac{12}{\pi^2}$). In [3], we will, following [4], give a proof of Hall’s theorem inspired by the work of F. Boca, C. Cobeli, and A. Zaharescu [5]. They created a map $T : \Omega \to \Omega$, now known as the BCZ map, and used equidistribution properties of periodic orbits of this map to obtain many statistical results on $F(Q)$ and $G(Q)$. In [4], the author and Y. Cheung showed that these results could be obtained by studying the horocycle flow on the space $X_2 = SL(2, \mathbb{R})/SL(2, \mathbb{Z})$ of unimodular lattices in $\mathbb{R}^2$. 
Figure 2. The region $A_{a,b}$ in inside the region $\Omega$.

Figure 3. The limiting distribution of gaps for Farey fractions and, (appropriately rescaled) lattice slopes.

1.4.1. Geometry of Numbers. One can also study the behavior of an arbitrary unimodular lattice $\Lambda$. Let $\Lambda \subset \mathbb{R}^2$ be a unimodular lattice, and suppose $\Lambda$ does not have vertical vectors. Let $\{s_1 < s_2 < \cdots < s_n < \cdots\}$ denote the slopes of the vectors (written in increasing order) in the vertical strip $V_1 = \{(u,v)^T : u \in (0,1], v > 0\}$. Here, and below, we use $(u,v)^T$ to denote the column vector $\begin{pmatrix} u \\ v \end{pmatrix}$, as our matrices act on the left. Let

$$G_N(\Lambda) = \{s_{n+1} - s_n : 0 \leq n \leq N\}$$

denote the set of gaps in this sequence. Note that in this setting, we do not need to normalize, as our sequence is not contained in $[0,1)$. Then we have that the limiting distribution of $G_N$ is also given by Hall’s distribution. That is:

**Theorem 1.2.** [1] Let $0 \leq a \leq b \leq \infty$. Then

$$\lim_{N \to \infty} \frac{1}{N} |G_N(\Lambda) \cap (a,b)| = 2|A_{a,b}|.$$ 

1.5. Saddle Connections. We saw above that the Farey sequence could be interpreted geometrically as slopes of primitive integer vectors in $\mathbb{R}^2$. Primitive integer vectors also correspond to (parallel families) of closed geodesics.
on the torus $\mathbb{R}^2/\mathbb{Z}^2$, which can also be interpreted as closed billiard trajectories in the square $[0,1/2]^2$. A natural generalization would be to try and understand similar families of trajectories for higher-genus surfaces, and/or for billiards in more complex polygons. More precisely, let $P$ be a Euclidean polygon with angles in $\pi\mathbb{Q}$. The billiard dynamical system on $P$ is given the (frictionless) motion of a point mass at unit speed with elastic collisions with the sides, satisfying the law of geometric optics: angle of incidence = angle of reflection. A generalized diagonal for the polygon $P$ is a trajectory for the billiard flow that starts at one vertex of $P$ and ends at another vertex. Since the group $\Delta_P$ generated by reflections in the sides of $P$ is finite, the angle of a trajectory is well defined in $S^1 \cong S^1/\Delta_P$. The natural gap distribution question that arises in this context is:

**Question.** What is the limiting distribution of the gaps between angles of generalized diagonals (normalized in terms of the length)?

More generally, one can ask about the limiting distribution for gaps for saddle connections in the more general setting of translation surfaces. A translation surface is a pair $(M, \omega)$, where $M$ is a Riemann surface and $\omega$ a holomorphic 1-form. A saddle connection is a geodesic $\gamma$ in the flat metric induced by $\omega$, connecting two zeros of $\omega$. To each saddle connection $\gamma$ one can associate a holonomy vector $v_\gamma = \int_\gamma \omega \in \mathbb{C}$. The set of holonomy vectors $\Lambda_{sc}(\omega)$ is a discrete subset of $\mathbb{C} \cong \mathbb{R}^2$, and varies equivariantly under the natural $SL(2, \mathbb{R})$ action on the set of translation surfaces. Motivated by such concerns, and inspired by the work of Marklof-Strombergsson [14] (of which more below in §1.6), the author and J. Chaika [2] studied the gap distribution for saddle connection directions. The relationship between flat surfaces and billiards in polygons is given by a natural unfolding procedure, which associates to each (rational) polygon $P$ a translation surface $(X_P, \omega_P)$. The main result of [2] used the dynamics of the $SL(2, \mathbb{R})$ action on the moduli space $\Omega_g$ of genus $g$ translation surfaces to show that generically, a limiting distribution exists.

More precisely, given $R > 0$, let

$$F^\omega_R := \{ \arg(v) : v \in \Lambda_\omega \cap B(0, R) \}$$

denote the set of directions of saddle connections of length at most $R$. Masur [16] showed that the counting function $N(\omega, R) := |F^\omega(R)|$ grows quadratically in $R$ for any $\omega$. Denote the associated normalized gap set by $G^\omega(R)$.

**Theorem 1.3.** ([2, Theorem 1.1]) For almost every (with respect to Lebesgue measure on $\Omega_g$) translation surface $\omega$, there is a limiting distribution for the gap set $G^\omega(R)$. Moreover, this distribution has support at $0$, that is, for almost every $\omega \in \Omega_g$, and for any $\epsilon > 0,

$$\lim_{R \to \infty} \frac{|G^\omega(R) \cap (0, \epsilon)|}{N(\omega, R)} > 0.$$
1.5.1. **Lattice Surfaces.** The support at 0 in Theorem 1.3 is in contrast to the setting of the torus, where, as seen in Figure 3, there a gap between 0 and 3/π². This gap at 0 is, in some sense, due to the symmetry of the torus; if we think of the $SL(2,\mathbb{R})$ action on the moduli space $X_2$ of flat tori, the stabilizer of any point is (conjugate to) $SL(2,\mathbb{Z})$. More generally, it was shown in [2] that if $\omega$ is a lattice surface (i.e., the stabilizer of the flat surface $\omega$ under the $SL(2,\mathbb{R})$ action is a lattice) that the limiting distribution for gaps has no support at 0.

While it was in principle possible to compute the limiting distribution using the techniques in [2], the more geometric nature of the techniques in [4] and the use of horocycle flows on moduli spaces can be generalized to the setting of lattice surfaces to give a roadmap for explicitly calculating the limiting distribution of gaps. In joint work [3] with J. Chaika and S. Lelièvre, we proved Theorem 1.4 on the gap distribution for the golden $L$, which is a surface of genus 2 with one double zero, displayed in Figure 4.

**Theorem 1.4.** [3] There is an explicit limiting gap distribution for the set of slopes (equivalently, angles) for saddle connections on the golden $L$. The probability distribution function is differentiable except at a set of eight points.

**Remark:** The limiting and empirical distributions are shown in Figure 5, drawn from [3]. We refer the reader to [3] for the precise formulas for the limiting distribution.

1.6. **Visible affine lattice points.** Another natural generalization of the Farey sequence is to consider affine lattices, that is, translates of lattices by some fixed vector. We write

$$\Lambda = M\mathbb{Z}^2 + \mathbf{v},$$

where $M \in SL(2,\mathbb{R})$ and $\mathbf{v} \in \mathbb{R}^2$ (really $\mathbf{v}$ is well-defined up to the lattice $M\mathbb{Z}^2$, so we think of it as an element of the torus $\mathbb{R}^2/M\mathbb{Z}^2$). Marklof-Strombergsson [14] used dynamics on the space of affine lattices $\tilde{X}_2 =$...
The limiting and empirical distributions for gaps of saddle connection slopes on the golden L.

$SL(2, \mathbb{R}) \ltimes \mathbb{R}^2 / SL(2, \mathbb{Z}) \ltimes \mathbb{Z}^2$ to study the gap distribution for the angles of visible affine lattice points. They in fact considered much more general problems, studying the distribution of visible affine lattice points in higher dimensions, but for the purposes of this paper, we focus on their two-dimensional results.

Consider the set of angles of lattice points of length at most $R$, that is,

$$F_\Lambda(R) := \{ \text{arg } w : w \in \Lambda \cap B(0, R) \}.$$

To calculate the associated gap distribution $P_\Lambda$, the key is to estimate the probability of finding multiple lattice points in ‘thinning’ wedges. Given $\sigma > 0, \theta \in [0, 2\pi)$ and $R > 0$ consider the wedge

$$A_\theta^R(\sigma) := \{ w \in \mathbb{R}^2 : w \in B(0, R), \text{arg}(w) \in (\theta - \sigma R^{-2}, \theta + \sigma R^2) \},$$

shown in Figure 6. Here, the factor of $R^{-2}$ corresponds to the normalizing factor $1/N$ above, since the cardinality of $F_\Lambda(R)$ is on the order of $R^2$. The gap distribution will be given by (the second derivative) of the limiting probability

$$p_{\Lambda,0}(\sigma) = \lim_{R \to \infty} \lambda(\theta : A_\theta^R(\sigma) \cap \Lambda = \emptyset)$$

that this wedge does not affine lattice points. This follows from the fact that if we let $P_\Lambda(t)$ denote the probability distribution function of the limiting gap distribution, we have

$$p_{\Lambda,0}(\sigma) = \sigma - \int_0^\sigma \int_0^t P_\Lambda(s)d\sigma dt$$

To compute $p_{\Lambda,0}(\sigma)$, note that rotating the region $A_\theta^R(\sigma)$ by the element

$$r_{-\theta} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix},$$

and scaling by

$$g_t = \begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{pmatrix},$$

(1.8)
Figure 6. The wedge $A_R^\theta(\sigma)$.

(with $t = 2 \log R$) we obtain (approximately) the triangle $T(\sigma)$ with vertices at $(0, 0)$ and $(1, \pm \sigma)$. Thus, the probability that (for a randomly chosen $\theta$) the region $A_R^\theta$ does not contain points from $\Lambda$ corresponds to the integral

$$\int_0^{2\pi} h_\sigma(g_\theta r_\theta \Lambda) d\theta,$$

where $h_\sigma : \tilde{X}_2 \mapsto \{0, 1\}$ is given by

$$h_\sigma(\Lambda) = \begin{cases} 1 & \Lambda \cap T(\sigma) \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

To understand the limits (as $R$, or equivalently $t \to \infty$) of the integrals $\int_0^{2\pi} h_\sigma(g_\theta r_\theta \Lambda) d\theta$, we need to apply tools from Ratner’s theory of unipotent flows. It can be shown that these integrals converge to $\int_{\tilde{X}_2} h_\sigma d\mu$, where the limit measure $\mu$ can be shown to be invariant under a unipotent subgroup $N \subset SL(2, \mathbb{R}) \ltimes \mathbb{R}^2$, following work of Shah [20]. Using Ratner’s Theorem [18], the possibilities for the measure $\mu$ are either:

- The Haar probability measure $\hat{\mu}_2$
- A probability measure supported on the set of torsion points $\tilde{X}_2[n]$, that is, the support of the measure is restricted to affine lattices $M\mathbb{Z}^2 + v$ whose translation vector $v$ satisfies $nv \in M\mathbb{Z}^2$.

Depending on the properties of the initial affine lattice $\Lambda$, torsion-supported measures can (and do) occur. As a consequence, the limiting gap distributions (which can all be explicitly computed) differ depending on the initial lattice, and in particular, whether the initial translation vector $v$ is a torsion point of the torus $\mathbb{R}^2/M\mathbb{Z}^2$ or not, and at what level [14, Theorem 1.3].

1.6.1. $\sqrt{n} \mod 1$. Extraordinarily, the gap distribution for generic affine lattices $\Lambda$ (that is, those for which the vector $v$ is not torsion, and so for which the limiting measure is Haar) coincides with the gap distribution for the sequence of fractional parts of $\{\sqrt{n}\}$. The reason they coincide is that both are coming from equidistribution of certain homogeneous trajectories.
with respect to the Haar measure on the space \( \tilde{X}_2 \). We strongly urge the reader to look at the beautiful paper [8] for further details on the connection of \( \sqrt{n} \mod 1 \) to homogeneous dynamics. Perhaps ironically, despite the inspiration for our paper being provided from their pithy quote, our theorems below do not seem to give a direct explanation of their result.

2. Meta-Theorems on Gap Distributions

As discussed in the introduction, the common thread that runs through these results is the creation of associated discrete sets of \( \mathbb{R}^2 \) and appropriate dynamical systems, which turns gap distribution questions into questions of equidistribution of orbits on certain moduli spaces. In this section, we state our main results, Theorem 2.2, Theorem 2.5, and Theorem 2.6, in a very general setting, to attempt to explain several of these phenomena simultaneously. We first fix notation and our general framework in §2.1 before stating Theorem 2.2 and Theorem 2.5 in §2.2 and Theorem 2.6 in §2.3.

2.1. Setup. Our setup is very similar to that of [1]. Let \( X \) be a locally compact metric space with a continuous \( SL(2,\mathbb{R}) \)-action, and a \( SL(2,\mathbb{R}) \)-equivariant assignment

\[
x \mapsto \Lambda_x
\]

that associates to each point \( x \in X \) a countable, discrete subset \( \Lambda_x \subset \mathbb{R}^2 \setminus \{0\} \). Equivariance, in this context, means that

\[
\Lambda_{gx} = g\Lambda_x
\]

for all \( g \in SL(2,\mathbb{R}), x \in X \). Our theorems will connect the gap distributions for slopes and angles in the set \( \Lambda_x \) to the orbit of the point \( x \) under various subgroups of the \( SL(2,\mathbb{R}) \) action.

2.2. Slope gaps. The key idea of our meta-theorems comes from a simple observation on the behavior of slopes under (vertical) shears (a corresponding discussion can be made with inverse slopes and horizontal shears). Let \( \mathbf{u} = (u_1, u_2)^T, \mathbf{v} = (v_1, v_2)^T \in \mathbb{R}^2 \). Let \( s_u = \frac{v_2}{u_1} \) and \( s_v = \frac{v_2}{v_1} \) denote their slopes. Let

\[
h_s = \begin{pmatrix} 1 & 0 \\ -s & 1 \end{pmatrix}.
\]

Observation. For any \( s \in \mathbb{R} \), the difference in slopes between \( h_s \mathbf{u} \) and \( h_s \mathbf{v} \) is the same as the difference in slopes between \( \mathbf{u} \) and \( \mathbf{v} \), that is,

\[
s_{h_s \mathbf{u}} - s_{h_s \mathbf{v}} = s_{\mathbf{u}} - s_{\mathbf{v}}.
\]

This follows from the (even simpler) observation that for any \( \mathbf{u} \in \mathbb{R}^2 \),

\[
s_{h_s \mathbf{u}} = s_{\mathbf{u}} - s.
\]
Thus, if we have a set of slopes \( \mathbb{S}(\Lambda) = \{s_1 < s_2 < \ldots s_n < \ldots \} \) of a family of vectors \( \Lambda \subset \mathbb{R}^2 \), the behavior of the associated gap set \( \mathbb{G}(\Lambda) = \{s_{i+1} - s_i : 1 \leq i \leq \infty \} \) is \textit{invariant} under the action of \( h_s \), that is
\[
\mathbb{G}(h_s \Lambda) = \mathbb{G}(\Lambda).
\]

In particular, by considering the flow \( h_s \) at the times \( s_1, s_2, \ldots \), we can recover the gap \( s_{i+1} - s_i \) by sampling the smallest positive element of \( \mathbb{S}(h_s \Lambda) \).

2.2.1. \textit{Minkowski properties, short vectors, and vertical trips.} For our theorem on slopes, we need two additional technical assumptions on the sets \( \Lambda_x \), a few definitions.

\textbf{Symmetry.} \( \Lambda_x \) is centrally symmetric, that is \( -\Lambda_x = \Lambda_x \).

\textbf{Minkowski Property.} For all \( x \in X \), there is a \( c = c(x) \) such that for any convex, centrally symmetric set \( K \subset \mathbb{R}^2 \) of volume at least \( c \),
\[
K \cap \Lambda_x \neq \emptyset.
\]

We leave it as an exercise to the reader that the minimal such constant \( c(x) \) is \textit{invariant} under the action of \( SL(2, \mathbb{R}) \).

\textbf{Vertically and horizontally short points.} We say that \( x \in X \) is \textit{vertical} (respectively \textit{horizontal}) if \( \Lambda_x \) contains a vertical (resp. horizontal) vector. Given a constant \( \eta \), we say that \( x \) is \( \eta \)-\textit{vertically short} (resp. \( \eta \)-\textit{horizontally short}) if \( \Lambda_x \) contains a vertical (resp. horizontal) vector of length at most \( \eta \).

\textbf{Vertical Strip Property.} For \( \eta > 0 \), let \( V_\eta := \{(u,v)^T : 0 < u \leq \eta, v > 0 \} \) denote the vertical strip of width \( \eta \) in \( \mathbb{R}^2 \). Suppose \( x \) is horizontal. Then \( \Lambda_x \cap V_\eta \) is non-empty for any \( \eta > 0 \).

\textbf{Exceptional points} We say that \( x \) is \( \eta \)-\textit{exceptional} if it is \( \frac{\eta}{4c(x)} \)-vertically short, that is, if it contains a vertical vector of length at most \( \frac{\eta}{4c(x)} \). We denote by \( X_\eta \) the set of non-\( \eta \)-exceptional points.

2.2.2. \textit{Horocycles and transversals.} Note that since \( h_s \) fixes vertical vectors, \( X_\eta \) is an \( h_s \)-invariant set for any \( \eta > 0 \). An important lemma, crucial to the statement of Theorem [2.2] is:

\textbf{Lemma 2.1.} Let \( \eta > 0 \), and let \( \Omega_\eta \subset X \) denote the set of \( \eta \)-\textit{horizontally short points}. Then \( \Omega_\eta \) is a transversal to the \( h_s \) action on the set \( X_\eta \). That is, for every \( x \in X_\eta \), the \( h_s \)-orbit of \( x \) intersects \( \Omega_\eta \) in a non-empty, discrete set of times.

We prove this lemma in \S 5. The crucial consequence of this lemma, for us, is the existence of the induced return map \( T_\eta : \Omega_\eta \to \Omega_\eta \) and return time function \( R_\eta : \Omega_\eta \to \mathbb{R}^+ \), given by
\[
R_\eta(x) = \min\{s > 0 : h_s x \in \Omega_\eta\}.
\]
and $T'(x) = h_{R(x)}x$. In particular, any finite $h_s$-invariant measure $\nu$ on $X_\eta$ can be decomposed as $d\nu = d\tilde{\nu}ds$, where $\tilde{\nu}$ is a $T_\eta$-invariant probability measure on $\Omega_\eta$ and we identify $X_\eta$ with the suspension space

$$\{(x, s) : x \in \Omega_\eta, 0 \leq s \leq R_\eta(x)\}/\sim,$$

where $(x, R_\eta(x)) \sim (T_\eta(x), 0)$.

2.2.3. Long horocycles. Define

$$S_\eta(x) := \{0 \leq s_1^\eta(x) < s_2^\eta(x) < \ldots < s_n^\eta(x) < \ldots\}$$

Let

$$G_N^\eta(x) := \{1 \leq i \leq N : s_{i+1}^\eta(x) - s_i^\eta(x)\}$$

denote the set of the first $N$ slope gaps. Here, we do not normalize, since the slopes $s_i^\eta$ are going to $\infty$ in all of our applications. Given $x \in X$, and $S > 0$, let $\sigma_{x,S}$ denote the Lebesgue probability measure $d\sigma_{x,S} = \frac{1}{S}ds$ on the segment $\{h_s x : 0 \leq s \leq S\}$.

**Theorem 2.2.** Let $t > 0$. Suppose $\sigma_{x,S} \to \nu$ as $S \to \infty$ (here, and below, convergence of measures is in the weak-$\star$ topology), where $\nu$ is a finite measure supported on $X_\eta$, and $d\nu = d\tilde{\nu}ds$ as above. Then for any $0 \leq a \leq b \leq \infty$, we have

$$\lim_{N \to \infty} \frac{|G_N^\eta(x) \cap [a, b]|}{N} = \tilde{\nu}(R_\eta^{-1}([a, b])).$$

That is, the slope gap distribution is given by the distribution of the return time function $R_\eta$.

This theorem has the following immediate corollary:

**Corollary 2.3.** Let $\nu$ be a ergodic, $h_s$-invariant probability measure supported on $X_\eta$. Then for $\nu$-almost every $x$,

$$\lim_{N \to \infty} \frac{|G_N^\eta(x) \cap [a, b]|}{N} = \tilde{\nu}(R_\eta^{-1}([a, b])).$$

2.2.4. Long closed horocycles. We can also state a theorem for $x$ with periodic $h_s$-orbits. Again, we need a definition and a preliminary (elementary) lemma. We say that $x \in X$ is $h_s$-periodic with period $s_0 > 0$ if $h_{s_0}x = x$, and $h_s x \neq x$ for all $0 < s < s_0$.

**Lemma 2.4.** Let $x$ be $h_s$-periodic. Then for every $\eta > 0$ there is an $N_0 = N_0(\eta)$ such that for any $N \geq N_0$,

$$G_N^\eta(x) = G_{N_0}^\eta(x).$$

That is, the set of gaps repeats.
We postpone the proof of the lemma to \[\text{(3.5)}\] Let \(g_t = \text{diag}(e^{t/2}, e^{-t/2})\) be as in \(\text{(1.10)}\). Then we have the conjugation relation
\[
(2.5) \quad g_t h_s g_{-t} = h_{se^{-t}}.
\]
Thus, if \(x\) is \(h_s\)-periodic with period \(s_0\), \(g_{-t} x\) is \(h_{s_0 e^{-t}}\)-periodic, since
\[
h_{s_0 e^{-t}} g_{-t} x = g_{-t} h_{s_0} x = g_{-t} x.
\]
Note also that if \(x \in X_{\eta}\), then \(g_{-t} x \in X_{e^t/2 \eta}\) for any \(t > 0\), since \(g_{-t}\) expands by \(e^{t/2}\) in the vertical direction. Let \(\rho_{x,T}\) denote the Lebesgue probability measure \(d \rho_{x,T} = \frac{ds}{e^{t s_0}}\) supported on the periodic orbit
\[
\{h_s g_{-T} x : 0 \leq s \leq e^T s_0\}.
\]

**Theorem 2.5.** Suppose \(\rho_{x,T} \to \nu\) as \(T \to \infty\), with \(\nu\) a finite measure supported on \(X_{\eta_0}\) for some \(\eta_0 > 0\) (in fact, it will be supported on any \(X_{\eta}\)), and \(d \nu = d \tilde{\nu} ds\), with \(\tilde{\nu}\) the \(T_{\eta_0}\)-invariant probability measure on \(X_{\eta_0}\). Then
\[
(2.6) \quad \lim_{\eta \to \infty} \frac{\left(\frac{2}{\eta_0}\right)^2 \mathbb{G}_{\eta_0}^N(x) \cap [a,b]}{N_0(\eta)} = \tilde{\nu}\left(R_{\eta_0}^{-1}([a,b])\right).
\]

To re-iterate, the right hand choice of this limit does not depend on the choice of \(\eta_0\), which can be chosen to be any positive number.

2.3. **Angle gaps and thinning wedges.** In this section, we work with angles, as opposed to slopes. We also do not need the assumptions of central symmetry and the Minkowski property, though they are present in our applications. Let \(x \in X\), and let
\[
\Theta_x(R) := \{0 < \theta_1 < \theta_2 < \ldots < \theta_N < 2\pi\}
\]
denote the set of angles of vectors in the set \(\Lambda_x \cap B(0,R)\), where \(B(0,R) = \{v \in \mathbb{R}^2 : \|v\|_2 \leq R\}\) is the ball of radius \(R\) in \(\mathbb{R}^2\), and \(N = N(R)\) is the cardinality of the set of distinct angles. Arguing as in \[\text{(1.6)}\] the gap distribution for \(\Theta_x(R)\) can be reduced to studying the limiting probability
\[
p_0(x,\sigma) = \lim_{R \to \infty} \lambda(\theta : A^\sigma_R(R) \cap \Lambda_x = \emptyset),
\]
where \(\lambda\) denotes the Lebesgue probability measure on \(S^1\), \(\sigma > 0\) is a fixed parameter, and
\[
(2.7) \quad A^\sigma_R(R) := \{v \in \mathbb{R}^2 : v \in B(0,R), \arg(v) \in (\theta - \sigma R^2, \theta + \sigma R^2)\},
\]
as shown in Figure 6. Let \(\lambda_{x,R}\) denote the Lebesgue probability measure \(d \lambda_{x,R} = \frac{d\theta}{2\pi}\) supported on the orbit \(\{g_t \theta x\}_{0 \leq \theta < 2\pi}\), with \(t = 2\ln R\).

**Theorem 2.6.** Suppose \(\lambda_{x,R} \to \nu\) as \(R \to \infty\), where \(\nu\) is a probability measure on \(X\). Then
\[
(2.8) \quad p_0(x,\sigma) = \nu(x \in X : \Lambda_x \cap T(\sigma) = \emptyset),
\]
where $T(\sigma)$ is the triangle with vertices at $(0,0)$ and $(1, \pm \sigma)$. Moreover, if we define

$$p_i(x, \sigma) := \lim_{R \to \infty} \lambda(\theta : |A_\theta^\sigma(R) \cap \Lambda_x| = i),$$

we have

$$(2.9) \quad p_i(x, \sigma) = \nu(x \in X : |\Lambda_x \cap T(\sigma)| = i).$$

3. Farey sequences

In this section, we show how to use Theorem 2.5 to prove Hall’s Theorem 1.1, and how to use Theorem 2.2 to prove the geometry of numbers result Theorem 1.2. We follow the exposition in [4], where it was shown how these theorems can be obtained from results on equidistribution of long orbits of the group $\{h_s\}$ on the space of unimodular lattices $X_2 = G/\Gamma$, where $G = SL(2, \mathbb{R})$ and $\Gamma = SL(2, \mathbb{Z})$. Following the notation in §2, our space $X$ is given by the space $X_2$, the assignment of a discrete set is given by $gSL(2, \mathbb{Z}) \mapsto \rightarrow g\mathbb{Z}_2^{prim}$, and we see immediately that the Minkowski and symmetry conditions of §2.2.1 are satisfied. It is also possible to verify the vertical strip condition explicitly (see [4]). We note that $\mathbb{Z}^2$ is $h_s$-periodic with period 1, since $h_1\mathbb{Z}^2 = \mathbb{Z}^2$. Given $\eta > 0$, the set of slopes in $V_\eta$ is the set of non-negative rational numbers with denominator at most $Q = \lfloor \eta \rfloor$. That is, we can write

$$S^\eta_N(\mathbb{Z}^2) := \{n + F(Q) : n \geq 0\},$$

where $n$ ranges over the non-negative integers. Thus, the gap set $G^\eta_N(\mathbb{Z}^2)$ corresponds to the set of gaps in the Farey sequence $F(Q)$, and has cardinality $N(Q)$. We recall that Sarnak [19] proved that the measures $\rho_{\mathbb{Z}^2, T}$ equidistribute with respect to the Haar measure probability $\mu_2$. That is,

$$\rho_{\mathbb{Z}^2, T} \rightarrow \mu_2.$$

Thus, to complete the proof of Theorem 1.1 (assuming Theorem 2.5), we need to describe the distribution of the return time function (with respect to the disintegrated measure) on the transversal

$$\Omega_\eta := \{\Lambda \in X_2 : \Lambda \text{ is } \eta - \text{horizontally short}\},$$

for some choice of $\eta > 0$.

3.1. Description of transversal and return map. The main result of [4] is to explicitly describe the transversal $\Omega_1$ (more generally, $\Omega_\eta$) and the associated return map and return time function. It was shown that $\Omega_\eta$ is in bijective correspondence with the set

$$\{(a, b) \in \mathbb{R}^2 : a, b \in (0, \eta], a + b > \eta\},$$

via the map

$$(a, b) \mapsto \Lambda_{a,b} = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \mathbb{Z}^2.$$

and that the roof function is given by

$$ R_\eta(a, b) = \frac{1}{ab}. $$

In particular, it does not depend on $\eta$. If we set $\eta = 1$, the return map is the BCZ map (named for its creators, Boca-Cobeli-Zaharescu [5]), given by

$$ T_1(a, b) = \left( b, -a + \left\lfloor \frac{1 + a}{b} \right\rfloor \right). $$

More generally,

$$ T_\eta(a, b) = \left( b, -a + \left\lfloor \frac{\eta + a}{b} \right\rfloor \right). $$

$T_\eta$ is piecewise linear, and all the linear maps have determinant 1, and so the Lebesgue probability measure $\frac{2}{\pi^2} dadb$ on $\Omega_b$ is $T_\eta$-invariant, and that $dadbds$ is, up to scaling, Haar measure on $X_2$ (which gives full measure to the set $X_{2,\eta}$ for any $\eta > 0$). The distribution of the roof function $R$ can be seen to be independent of $\eta$ (see §5.1.2 for an explanation of a general self-similarity phenomenon), so we can choose $\eta = 1$, obtaining the result of Theorem 1.1 up to the natural normalization factor $\frac{3}{\pi^2}$, which occurs since in the statement of Theorem 1.1 we normalize the gaps with a factor of $N(Q) \sim \frac{3}{\pi^2} Q^2$, as opposed to the factor $Q^2$ in the statement of Theorem 2.5. □

3.2. Proof of Theorem 1.2. The above computations are also crucial in the proof of Theorem 1.2 assuming Theorem 2.2. The other crucial step in the proof is Dani’s measure classification [7] result for $h_s$-invariant measures on $X_2$, which states that any ergodic $h_s$-invariant probability measure on $X_2$ is either supported on a periodic orbit or Haar measure. A lattice is $h_s$-periodic if and only if it has a vertical vector. To show that the existence of a vertical vector implies periodicity, note that such lattices can be written as

$$ \Lambda = \begin{pmatrix} 0 & -a^{-1} \\ a & b \end{pmatrix} Z^2, $$

and it is a direct verification to show

$$ h_a^2 \Lambda = \begin{pmatrix} 0 & -a^{-1} \\ a & b-a \end{pmatrix} Z^2 = \begin{pmatrix} 0 & -a^{-1} \\ a & b-a \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} Z^2 = \Lambda. $$

On the other hand, if $\Lambda$ is $h_s$ periodic, it must be divergent under $g_t$, which is equivalent to the existence of a vertical vector. This measure classification and characterization, combined with the Birkhoff ergodic theorem, yields:

Lemma 3.1. Let $\Lambda \in X_2$ be a lattice without vertical vectors. Then the Haar probability measure $\mu_2$. Thus, applying Theorem 2.2 we see that for any $\eta > 0$,

$$ \lim_{N \to \infty} \frac{1}{N} |G_N^\eta(\Lambda) \cap (a, b)| = \frac{2}{\eta^2} |\{(u, v) \in \Omega_\eta : R_\eta(u, v) \in (a, b)\}| = 2|A_{a, b}| $$
where $A_{a,b}$ is as in (1.5), and $|\cdot|$ is Lebesgue measure $dxdy$. The statement of Theorem 1.2 is the above statement for $\eta = 1$. □

4. Saddle connections

In this section, we sketch a proof of Theorem 1.3, following a similar strategy to the outline in §1.6, and attempting to indicate how it can be seen as a consequence of Theorem 2.6. We will particularly focus on the conclusion that for generic translation surfaces, the gap distribution has support at 0. Recall notation: $\Omega_g$ is the moduli space of holomorphic 1-forms on compact genus $g$ Riemann surfaces, and a saddle connection $\gamma$ on a surface $(M,\omega) \in \Omega_g$ is a geodesic (with respect to the flat metric given by the one-form) connecting two zeros of $\omega$ with no zeros in its interior. The holonomy vector $v_\gamma \subset \mathbb{C} \equiv \mathbb{R}^2$ associated to $\gamma$ is given by

$$v_\gamma = \int_\gamma \omega.$$ 

The set of holonomy vectors $\Lambda_\omega$ is a discrete subset of $\mathbb{R}^2$, and we are interested in the set of small gaps between directions of vectors. Precisely, letting $F^R_\sigma := \{\arg(v) : v \in \Lambda_\omega \cap B(0,R)\}$ denote the set of directions, we are interested in the normalized gap set $G_\omega(R)$, and in particular the size of the set $G_\omega(R) \cap (0,\epsilon)$ for $\epsilon > 0$. Note that having a small gap is equivalent to having two directions in a ‘thin wedge’.

4.1. $SL(2,\mathbb{R})$-action. A point $(M,\omega) \in \Omega_g$ determines (and is determined by) an atlas of charts from the surface punctured at the zeros of $\omega$ to the plane $\mathbb{C}$, given by integration of $\omega$. The transition maps for these charts are translations, and $\omega$ is given by $dz$ in these coordinates. This also gives a natural $SL(2,\mathbb{R})$ action on $\Omega_g$ via linear postcomposition with charts. There is ([15],[21]) a natural $SL(2,\mathbb{R})$-invariant probability measure on $\Omega_g$, which we denote $\mu_g$.

4.1.1. Equivariance. The assignment $(M,\omega) \mapsto \Lambda_\omega$ is an $SL(2,\mathbb{R})$-equivariant assignment, in the sense that for any $h \in SL(2,\mathbb{R})$,

$$\Lambda_{h\omega} = h\Lambda_\omega.$$ 

4.2. Renormalization. As above, we consider the wedges $A^\theta_R(\sigma)$. To understand the support of the distribution at 0, we consider probability that for a randomly chosen $\theta$ there are at least 2 points in the thinning wedge $A^\theta_R(\sigma)$. This will give us the probability that our gap set has gaps of size at most $\sigma$. As above, by renormalizing by $g_\theta$ and $r_\theta$, we can write this probability as

$$\int_{0}^{2\pi} h_{2,\sigma}(g_\theta r_\theta \Lambda) d\theta,$$
where \( h_{2,\sigma} : \Omega_g \mapsto \{0,1\} \) is given by

\[
(4.1) \quad h_{2,\sigma}(\omega) = \begin{cases} 
1 & |\Lambda_\omega \cap T(\sigma)| \geq 2 \\
0 & \text{otherwise}
\end{cases}
\]

4.3. **Equidistribution.** Since for \( g \geq 1 \) the space \( \Omega_g \) is not a homogeneous space, one cannot use Ratner-type technology to classify the possible limit measures for the measures \( \lambda_{\omega,R} \) supported on \( \{g_\theta r_\theta \omega\}_{0 \leq \theta < 2\pi} \). However, Eskin-Masur [9, Theorem 1.5] used a general ergodic result of Nevo to show that for \( \mu_g \)-almost every starting point \( \omega \), the measures converge (in an appropriately smoothed sense) to \( \mu_g \), that is, the main assumption of Theorem 2.6 is satisfied, with \( \nu = \mu_g \). It is worth noting that there is no known example of a point \( \omega \in \Omega_g \) for which there does not exist a measure \( \nu \) satisfying the hypothesis of Theorem 2.6. Using Nevo’s theorem, the author and J. Chaika [2] were able to conclude:

\[
\lim_{t \to \infty} \int_0^{2\pi} h_{2,\sigma}(g_\theta r_\theta \Lambda) d\theta = \int_{\Omega_g} h_{2,\sigma} d\mu_g.
\]

Finally, a measure estimate on the set of surfaces with simultaneous short saddle connections due to [17] (see also [10, Lemma 7.1]) suffices to show that the integral on the right hand side is positive for any \( \sigma > 0 \).

4.4. **Slopes.** For certain measures, Theorem 2.2 and Corollary 2.3 be applied in this setting, as the Minkowski condition is satisfied by [13, Theorem 1]. However, to give gap distributions for slopes of surfaces in the support of \( h_s \)-invariant measures on \( \Omega_g \), one still has to verify the vertical strip condition for the support of these measures, and then, identify the transversal, the roof function, and the associated measure \( \tilde{\nu} \) on the transversal in order to get explicit formulas. If we have a closed \( SL(2,\mathbb{R}) \) orbit, then the points in that orbit are lattice surfaces, where it is possible to carry out this program.

4.5. **Lattice Surfaces.** If \((M,\omega)\) is a lattice surface (for example the golden \( L \)), with stabilizer \( \Gamma \subset SL(2,\mathbb{R}) \), we restrict our parameter space \( X \) to from \( \Omega_g \) to the \( SL(2,\mathbb{R}) \)-orbit of \((M,\omega)\), which can be identified with \( SL(2,\mathbb{R})/\Gamma \) via

\[
g(M,\omega) \mapsto g\Gamma.
\]

It is possible (see [3]) to verify the vertical strip condition in this setting. Since the stabilizer \( \Gamma \) must be non-uniform, we can assume that \((M,\omega)\) is \( h_s \)-periodic, that is, that \( \Gamma \) contains an element \( h_{s_0} \) for some \( s_0 > 0 \). Applying Theorem 2.5, we obtain that the limiting distribution of the slope gaps for vectors in \( \Lambda_\omega \) (in the first quadrant) is given by the distribution of the return time function \( R \) for \( h_s \) to the transversal \( X_1 := \{g\Gamma \in X : \Lambda_{g_\omega} \text{ is 1-horizontally short}\} \). In [3], the transversal \( X_1 \), the return time function, and the return map were explicitly computed for
Figure 7. The transversal $X_1$ to the horocycle flow $h_s$ for the $SL(2, \mathbb{R})$-orbit of the golden $L$. The colors indicate different behaviors of the return map.

the translation surface $(M, \omega)$ associated to the golden $L$. The transversal, in this case, can be identified with the triangle with vertices at $(0,1)$, $(1,1)$, and $\left(1, \frac{1-\sqrt{5}}{2}\right)$, shown in Figure 7. In principal, with some effort, this program can be carried out for any lattice surface.

5. Abstract Homogeneous Dynamics

In this section, we prove our main results Theorem 2.2, Theorem 2.5, and Theorem 2.6. We prove Theorem 2.2 and Theorem 2.5 in §5.1 and Theorem 2.6 in §5.2.

5.1. Horocycle Flows. We recall notation: $X$ is our parameter space, equipped with an $SL(2, \mathbb{R})$-action, and given $x \in X$ we have the associated discrete set $\Lambda_x \subset \mathbb{R}^2$, satisfying:

**Equivariance:** For any $g \in SL(2, \mathbb{R})$, $\Lambda_{gx} = g\Lambda_x$.

**Symmetry:** $\Lambda_x$ is centrally symmetric, that is $-\Lambda_x = \Lambda_x$.

**Minkowski Property:** For all $x \in X$, there is a $c = c(x)$ such that for any convex, centrally symmetric set $K \subset \mathbb{R}^2$ of volume at least $c$,

$$K \cap \Lambda_x \neq \emptyset.$$

**Vertical Strip Property:** Fix $\eta > 0$. Let $V_\eta := \{(u, v)^T : 0 < u \leq \eta, v > 0\}$ denote the vertical strip in $\mathbb{R}^2$. Suppose $x$ is horizontal. Then $\Lambda_x \cap V_\eta$ is non-empty.
Recall that we say that \( x \in X \) is vertical (respectively horizontal) if \( \Lambda_x \) contains a vertical (resp. horizontal) vector, and that given a constant \( \eta \), we say that \( x \) is \( \eta \)-vertically short (resp. \( \eta \)-horizontally short) if \( \Lambda_x \) contains a vertical (resp. horizontal) vector of length at most \( \eta \). We say that \( x \) is \( \eta \)-exceptional if it is \( \eta \)-vertically short, that is, if it contains a vertical vector of length at most \( \frac{\eta}{4c(x)} \). We denote by \( X_\eta \) the set of non-\( \eta \)-exceptional points. Note that \( X_\eta \) and \( X \setminus X_\eta \) are both \( h_\omega \)-invariant sets.

5.1.1. Existence of return map. In this subsection, we prove Lemma 2.1 which gives the existence of a well-defined transversal \( \Omega_\eta \) and return map \( R_\eta \) for the flow \( h_\omega \) on the invariant set \( X_\eta \), for any fixed \( \eta > 0 \). Recall that

\[
\Omega_\eta = \{ x \in X : x \text{ is } \eta \text{-horizontally short} \},
\]

and

\[
V_\eta = \{ (u, v)^T \in \mathbb{R}^2 : u \in (0, \eta), v > 0 \}.
\]

Then we claim that for any \( x \in X_\eta \), the set \( S_\eta(x) \) of slopes of vectors of \( \Lambda_x \) in \( V_\eta \) defined in (2.2) corresponds precisely to the set of times

\[
\{ s > 0 : h_s x \in \Omega_\eta \},
\]

and that this set is non-empty. This claim follows from the observation that if \( (u, v)^T \in V_\eta \cap \Lambda_x \), then, setting \( s = \frac{u}{v} \) to be the slope of the vector, we have

\[
h_s(u, v)^T = (u, 0)^T,
\]

with \( 0 < u \leq \eta \), that is, \( h_s x \in \Omega_\eta \). This shows that the set of slopes \( S_\eta(x) \) is contained in the set of return times \( \{ s > 0 : h_s x \in \Omega_\eta \} \). For the reverse containment, suppose \( h_s x \in \Omega_\eta \). Then there is a vector \( (u, 0)^T \in \Lambda_{h_s x} \) with \( 0 < u \leq \eta \), so \( h_{-s}(u, 0)^T = (u, su) \in \Lambda_x \cap V_\eta \), so \( s \in S_\eta(x) \) as desired. To show that this set is discrete, note that if there was an accumulation point, say \( s_0 \), there would be a sequence of distinct vectors in \( \Lambda_x \cap V_\eta \) accumulating along the line of slope \( s_0 \), so by passing to a subsequence, we would obtain a sequence of distinct vectors in \( \Lambda_x \cap V_\eta \) converging to a point along the line segment \( \{ (u, su) : 0 < u \leq \eta \} \), a contradiction to the discreteness of \( \Lambda_x \).

Finally, to show that this set is non-empty, we need to use the Minkowski and vertical strip properties. We first show that for any \( x \in X_\eta \), there is an \( s \in \mathbb{R} \) so that \( h_s x \in \Omega_\eta \). To show this, let

\[
A = [\frac{-\eta}{4}, \frac{\eta}{4}] \times \left[ \frac{-c}{4\eta}, \frac{c}{4\eta} \right],
\]

where \( c = c(x) \) is as in the Minkowski property. Then \( A \) is convex, centrally symmetric, and has volume \( c \). Since we assume \( x \in X_\eta \), \( \Lambda_x \) does not contain a vertical vector of length at most \( \frac{c}{4\eta} \), so we have a vector \( (v_1, v_2)^T \in A \cap \Lambda_x \), with \( v_1 \neq 0 \). Since \( \Lambda_x \) is centrally symmetric, we can assume \( v_1 > 0 \), and setting \( s = \frac{v_1}{v_2} \), we have \( (v_1, 0) \in h_s x \), so \( h_s x \in \Omega_\eta \) as desired.

Next, we will use the vertical strip property to show that for any \( \omega \in \Omega_\eta \), there is an \( s' > 0 \) so that \( h_{s'} \omega \in \Omega_\eta \). For \( \omega \in \Omega_\eta \), we can use the vertical...
strip property to find \((u, v)^T \in \Lambda_{\omega} \cap V_{\eta}\), and setting \(s' = \frac{v}{u}\), we have our result. This completes the proof of Lemma 2.1. □

5.1.2. Self-Similarity. We let \(R_{\eta} : \Omega_{\eta} \to \mathbb{R}^+\) be the return time function
\[
R_{\eta}(\omega) = \min\{s > 0 : h_s\omega \in \Omega_{\eta}\},
\]
and \(T_{\eta} : \Omega_{\eta} \to \Omega_{\eta}\) be the return map
\[
T_{\eta}(\omega) = h_{R_{\eta}(\omega)}\omega.
\]

Given two positive parameters \(\eta_1, \eta_2 > 0\), the maps \(T_{\eta_1}\) and \(T_{\eta_2}\) are related by the action of the diagonal subgroup \(g_t\), via the conjugation relation \(g_t h_s g_{-t} = h_{s e^{T/2}}\). Precisely, letting \(T = 2 \log \frac{\eta_1}{\eta_2}\), we have that \(g_T \Omega_{\eta_2} = \Omega_{\eta_1}\), since the horizontal vectors of length (at most) \(\eta_2\) get expanded by \(e^{T/2} = \frac{\eta_1}{\eta_2}\) to vectors of length (at most) \(\eta_1\), and the slopes of the vectors get scaled by \(e^{T} = \left(\frac{\eta_1}{\eta_2}\right)^2\). This yields the self-similarity relations
\[
T_{\eta_1}(g_T\omega) = g_T T_{\eta_2}(\omega),
\]
and
\[
R_{\eta_1}(g_T\omega) = \left(\frac{\eta_1}{\eta_2}\right)^2 R_{\eta_2}(\omega)
\]
for any \(\omega \in \Omega_{\eta_2}\).

5.1.3. Generic points. We now prove Theorem 2.2. Recall notation: \(\sigma_{x, S}\) denotes the Lebesgue probability measure \(\frac{ds}{S}\) on the orbit \(\{h_s x : 0 \leq s \leq S\}\). Our assumption is that
\[
\sigma_{x, S} \to \nu
\]
as \(S \to \infty\), where convergence is in the weak-⋆ topology (in fact, all of our convergence of measures will be in this topology, so we do not mention it again), and that \(\nu\) gives measure 0 to the set of vertical \(x\). We want to show (2.4), that is, for any \(0 \leq a \leq b \leq \infty\),
\[
\lim_{N \to \infty} \frac{|\mathcal{G}_N(\eta)(x) \cap [a, b]|}{N} = \tilde{\nu}\left(R_{\eta}^{-1}([a, b])\right),
\]
where
\[
\mathcal{G}_N(\eta)(x) = \{1 \leq i \leq N : s^\eta_{i+1}(x) - s^\eta_i(x)\}
\]
is the set of the first \(N\) gaps of \(\mathcal{S}_\eta(x)\), and \(\tilde{\nu}\) is the measure on \(\Omega_{\eta}\) such that \(d\nu = d\tilde{\nu}\). Let \(\omega_0 = h_{s^\eta_1(x)} \in \Omega_{\eta}\) be the first hitting point of \(\Omega_{\eta}\) for the orbit \(\{h_s x\}_{s > 0}\). Then the assumption \(\sigma_{x, S} \to \nu\) implies that the measures \(\sigma_{\omega_0, N}\) supported on the given by
\[
\sigma_{\omega_0, N} := \frac{1}{N} \sum_{i=0}^{N-1} \delta_{T_i(\omega_0)}
\]
satisfy
\[
\sigma_{\omega_0, N} \to \tilde{\nu}
\]
as \( N \to \infty \). Let \( \chi_{a,b} \) denote the indicator function of the set \( R_{\eta}^{-1}([a,b]) \). Then we claim

\[
\frac{|G_{N\eta}^\eta(x) \cap [a,b]|}{N} = \sigma_{\omega_0,N}(\chi_{a,b}),
\]

since

\[
R_{\eta}(T^i(\omega)) = s_{i+1}^\eta(x) - s_i^\eta(x),
\]

which follows from the proof of Lemma 2.1. That is,

**Observation.** The gap distribution is a Birkhoff sum for the roof function \( R_{\eta} \) along the orbit of the map \( T_{\eta} \).

Then, applying the convergence \( \sigma_{\omega_0,N} \to \tilde{\nu} \) to the bounded function \( \chi_{a,b} \), we have our result. The above argument also shows how Corollary 2.3 follows from the Birkhoff ergodic theorem applied to the system \( (\Omega_{\eta}, T_{\eta}, \tilde{\nu}) \) and the function \( \chi_{a,b} \). \( \square \)

5.1.4. **Periodic points.** To prove Theorem 2.5, we first need to prove Lemma 2.4, that is, we need to show if \( x \) is \( h_s \)-periodic, then the set of slope gaps eventually stabilizes. Let \( s_0 \) be the (minimal) period of \( x \), so \( h_s s_0 x = x \). Then

\[ h_{s_0} \Lambda x = \Lambda x. \]

Let

\[ M = \min\{n > 0 : s_0^n(x) \geq s_0\}. \]

Then

\[ S_{\eta}(h_{s_0} x) = \{s_{M+1}^\eta(x) - s_0 < s_{M+2}^\eta(x) - s_0 < \ldots < s_{M+1}^\eta(x) - s_0 < \ldots\}, \]

but \( h_{s_0} x = x \), so

\[ s_{M+i}^\eta(x) - s_0 = s_{i+1}^\eta(h_{s_0} x) = s_{i+1}^\eta(x), \]

and thus

\[ s_{i+1}^\eta(x) - s_i^\eta(x) = s_{M+i}^\eta(x) - s_{M+i-1}^\eta(x), \]

that is, the gaps have period \( M - 1 \). Setting \( N_0 = M - 1 \), we have, as desired, for any \( N \geq N_0 \),

\[ G_N^\eta(x) = G_{N_0}^\eta(x). \]

This completes the proof of Lemma 2.4. To complete the proof of Theorem 2.5 we note that fixing an arbitrary \( \eta_0 > 0 \), periodic orbits for \( h_s \) naturally correspond to periodic orbits for the return map \( T_{\eta_0} \). The above argument shows the periodic orbit of \( T_{\eta_0} \) associated to the orbit

\[ \{h_s g_{-T} x : 0 \leq s \leq e^T s_0\} \]

has length \( N_0(\eta) \), where \( T = 2 \log(\frac{\eta_0}{\eta}) \). By moving \( x \) by \( h_s \), and applying Lemma 2.1, we can assume \( x \in \Omega_{\eta_0} \). We can then define the probability measure \( \sigma_{x,\eta} \) supported on the periodic orbit for \( T_{\eta_0} \) on \( \Omega_{\eta_0} \) induced by the probability measure \( \rho_{x,T} = \frac{ds}{s_0 e^T} \) supported on the orbit \( \{h_s g_{-T} x : 0 \leq s \leq e^T s_0\} \), via

\[
\sigma_{x,\eta} := \frac{1}{N_0(\eta)} \sum_{i=0}^{N_0(\eta)-1} \delta_{T_{\eta_0}^i(x)}. \]
Recall from §5.1.2 that $g_T$ scales $V_{n}$ to $V_{n_0}$ and multiplies slopes (and gaps of slopes) of vectors by the factor $(\frac{n}{n_0})^2$. Thus we have

$$\left(\frac{n}{n_0}\right)^2 \mathcal{G}_{N_0(\eta)}^\eta (x) \cap [a, b] = \sigma_{x,\eta}(\chi_{a,b}),$$

and applying our convergence assumption, $\rho_{x,T} \rightarrow \nu$, which implies

$$\sigma_{x,\eta} \rightarrow \tilde{\nu},$$

we have our result.

5.2. Circle Limits. The proof of Theorem 2.6 is essentially already contained in [14, §9] and in [2, §4], which we discussed in §1.6 and §4. We sketch the argument once again: we define, for $x \in X$, $\sigma > 0$,

$$p_i(x, \sigma) := \lim_{R \rightarrow \infty} \lambda(\theta : |A_\sigma^\theta (R) \cap \Lambda_x| = i).$$

By the renormalization argument described in §1.6, we have, for $R >> 0$,

$$|\Lambda_x \cap A_\sigma^\theta (R)| \approx |g_{t \theta - \theta} \Lambda_x \cap T(\sigma)|,$$

where $t = 2 \log R$. Thus,

$$p_i(x, \sigma) \approx \lambda_{x,R}(\{y \in X : |\Lambda_y \cap T(\sigma)| = i\}).$$

Applying our convergence assumption $\lambda_{x,R} \rightarrow \nu$, we have our result.

6. Further Questions

We collect some further questions and speculations. We discuss the space of translation surfaces in §6.1, the space of affine lattices in §6.2, and speculate wildly about other problems in §6.3.

6.1. Translation surfaces. It would be intriguing to push the machinery developed in this note further than the set of lattice surfaces. In particular, it would be interesting to check which $SL(2, \mathbb{R})$ (or $h_s$)-invariant measures on $\Omega_g$ are supported on the set of surfaces satisfying the vertical strip condition, and then to attempt to explicitly identify the transversal and the return time function (and the associated distribution). It would be particularly nice to do this for the Masur-Veech measure on $\Omega_g$, and to understand if there were any ‘continuity’ properties of the gap distributions— namely, given a sequence of surfaces $\omega_n \rightarrow \omega$, do the gap distributions of saddle connection directions converge in any appropriate sense? Also, are there any surfaces $\omega \in \Omega_g$ which do have ‘truly random’ behavior, that is, an exponential distribution of gaps? We conjecture that this is not possible.
6.2. **Affine Lattices.** A nice test case, which in fact corresponds to an $SL(2, \mathbb{R})$-invariant subset of $\Omega_2$ (see [2] §6), is the collection of **affine unimodular lattices**, as discussed in §1.6. If we could give an explicit description of the (4-dimensional) transversal $\Omega_2 \subset \tilde{X}_2 = (SL(2, \mathbb{R}) \ltimes \mathbb{R}^2) / (SL(2, \mathbb{Z}) \ltimes \mathbb{Z}^2)$, and understand the return time function, we would be able to recover the (2-dimensional) results of [14], explicitly using unipotent flows.

6.3. **Other gap distributions.** As discussed in the introduction, although our paper was inspired by [8], our machinery does not seem to immediately give the gap distribution for the sequence $\{\sqrt{n}\}_{n \geq 0}$. It would be a nice application to formulate the results of loc. cit. in the language our machine. A more ambitious, but probably very difficult project would be to try and understand the apparent exponential distribution of the gaps in $\{n^\alpha\}_{n \geq 0}$ for other $0 < \alpha < 1$. Another sequence of great interest in applications (see [12] §9) would be the gaps in the sequence of squares of lengths of lattice vectors $|a + b\tau|^2 : a, b \in \mathbb{Z}$ for any fixed $\tau \in \mathbb{C}$, which are also conjectured to be exponential. In fact, we would be very excited to see any application of our technology which lead to a non-exotic (i.e., exponential) distribution of gaps.

**References**

[1] J. S. Athreya, *Cusp excursions on parameter spaces*, preprint, arXiv:1104.2797
[2] J. S. Athreya and J. Chaika, *On the distribution of gaps for saddle connection directions*, to appear, Geometric and Functional Analysis.
[3] J. S. Athreya, J. Chaika, and S. Lelievre, *The distribution of gaps for saddle connections on the golden L*, in preparation.
[4] J. S. Athreya and Y. Cheung, *A Poincaré section for horocycle flow on the space of lattices*, preprint.
[5] F. Boca, C. Cobeli, and A. Zaharescu, *A conjecture of R. R. Hall on Farey points*. J. Reine Angew. Math. 535 (2001), 207 - 236.
[6] F. P. Boca and A. Zaharescu, *Farey fractions and two-dimensional tori*, in *Noncommutative Geometry and Number Theory* (C. Consani, M. Marcolli, eds.), Aspects of Mathematics E37, Vieweg Verlag, Wiesbaden, 2006, pp. 57-77.
[7] S. G. Dani, *On uniformly distributed orbits of certain horocycle flows*. Ergodic Theory Dynamical Systems 2 (1982), no. 2, 139 - 158 (1983).
[8] N.D. Elkies and C.T. McMullen, *Gaps in $\sqrt{n}$ mod 1 and ergodic theory*. Duke Math. J. 123 (2004), 95-139.
[9] A. Eskin and H. Masur, *Asymptotic Formulas on Flat Surfaces*, Ergodic Theory and Dynam. Systems, v.21, 443-478, 2001.
[10] A. Eskin, H. Masur, and A. Zorich, *Moduli spaces of abelian differentials: the principal boundary, counting problems, and the Siegel-Veech constants*. Publ. Math. Inst. Hautes Etudes Sci. No. 97 (2003), 61–179.
[11] R. R. Hall, *A note on Farey series*. J. London Math. Soc. (2) 2 1970 139 - 148.
[12] C. Hodgson and H. Masai, *On the number of hyperbolic 3-manifolds of a given volume*, preprint, arXiv:1203.6555
[13] P. Hubert and T. Schmidt, *Diophantine approximation on Veech surfaces*, preprint, arXiv:1010.3475
[14] J. Marklof and A. Strombergsson, *Distribution of free path lengths in the periodic Lorentz gas and related lattice point problems*, preprint, arXiv:0705.4305
[15] H. Masur, *Interval exchange transformations and measured foliations*. Ann. of Math. (2) 115 (1982), no. 1, 169–200.

[16] H. Masur, *The growth rate of trajectories of a quadratic differential*. Ergodic Theory Dynam. Systems 10 (1990), no. 1, 151-176.

[17] H. Masur and J. Smillie, *Hausdorff dimension of sets of nonergodic measured foliations*. Ann. of Math. (2) 134 (1991), no. 3, 455–543.

[18] M. Ratner, *Invariant measures and orbit closures for unipotent actions on homogeneous spaces*. Geom. Funct. Anal. 4 (1994), no. 2, 236-257.

[19] P. Sarnak, *Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series*. Comm. Pure Appl. Math. 34 (1981), no. 6, 719 - 739.

[20] N. Shah, *Limit distributions of expanding translates of certain orbits on homogeneous spaces*. Proc. Indian Acad. Sci. (Math Sci) 106(2), (1996), pp. 105-125.

[21] W. Veech, *Gauss measures for transformations on the space of interval exchange maps*. Ann. of Math. (2) 115 (1982), no. 1, 201–242.

[22] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. 77(1916), 313-352.

E-mail address: jathreya@illinois.edu

Department of Mathematics, University of Illinois Urbana-Champaign, 1409 W. Green Street, Urbana, IL 61801