1

Gap distributions and homogeneous dynamics

Jayadev S. Athreya[1]

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 7, Theorem 10, and Theorem 11, give unified explanations of several examples of 'exotic' gap

[1] J.S.A. partially supported by NSF CAREER grant DMS 1351853, NSF grant DMS 1069153, and grants DMS 1107452, 1107263, 1107367 "RNMS: GEometric structures And Representation varieties" (the GEAR Network).

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

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

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 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 $0 \leq a \leq b \leq 1$,

$$\lim_{k\to\infty} \Delta_k(a,b) = b - a. \tag{1}$$

A more refined question (not necessarily dependent on (1)) is to examine the distribution of *gaps* for the sequences $F(k)$. That is, form the associated *normalized gap sets*

$$G(k) := \left\{N_k\left(F_k^{(i+1)} - F_k^{(i)}\right) : 0 \leq i < N_k\right\}, \tag{2}$$

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

$$\lim_{k\to\infty} \frac{|G(k) \cap (a,b)|}{N_k}? \tag{3}$$

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 $t > 0$,

$$\lim_{k\to\infty} \frac{|G(k) \cap (t,\infty)|}{N_k} = e^{-t} \tag{4}$$

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 7, Theorem 10, and Theorem 11), 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 $\mathcal{F}(Q)$ (Section 1.4); slopes for lattice vectors (Section 1.4); and saddle connection directions for translation surfaces (Section 1.5). We also briefly discuss the space of affine lattices and $\{\{\sqrt{n}\}\}_{n\geq 1}$ in Section 1.6. In Section 2, we state the main results Theorem 7, Theorem 10, and Theorem 11. We describe how to use these results to explain our examples in Section 3-Section 4, and prove the theorems in Section 5. Finally, in Section 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.

Cambridge University Press
978-1-107-52900-7 - Geometry, Topology, and Dynamics in Negative Curvature
Edited by C. S. Aravinda, F. T. Farrell and J.- F. Lafont
Excerpt
More information

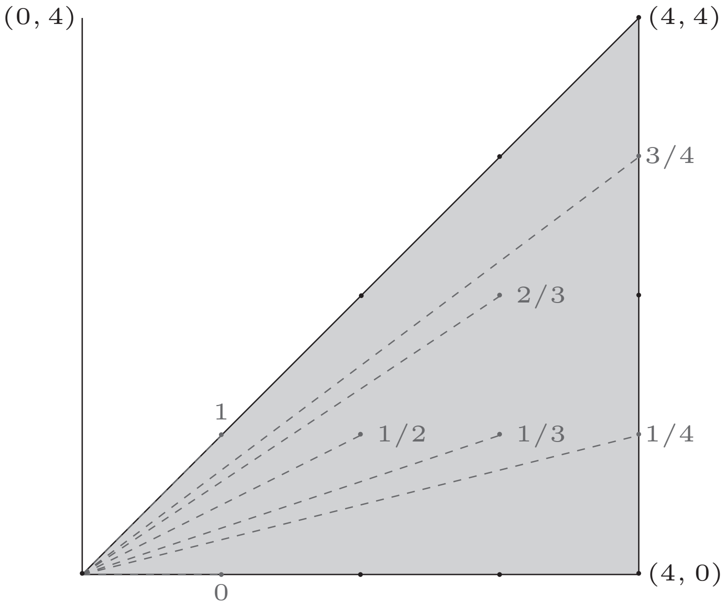

Figure 1 The triangle T_4. Primitive integer vectors are given by dashed lines, and are labeled by their slopes.

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

$$\mathcal{F}(Q) := \left\{\gamma_0 = \frac{0}{1} < \gamma_1 = \frac{1}{Q} < \gamma_2 \ldots < \gamma_i = \frac{p_i}{q_i} < \ldots \gamma_N = \frac{1}{1}\right\}$$

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 $\begin{pmatrix} q_i \\ p_i \end{pmatrix}$ 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.

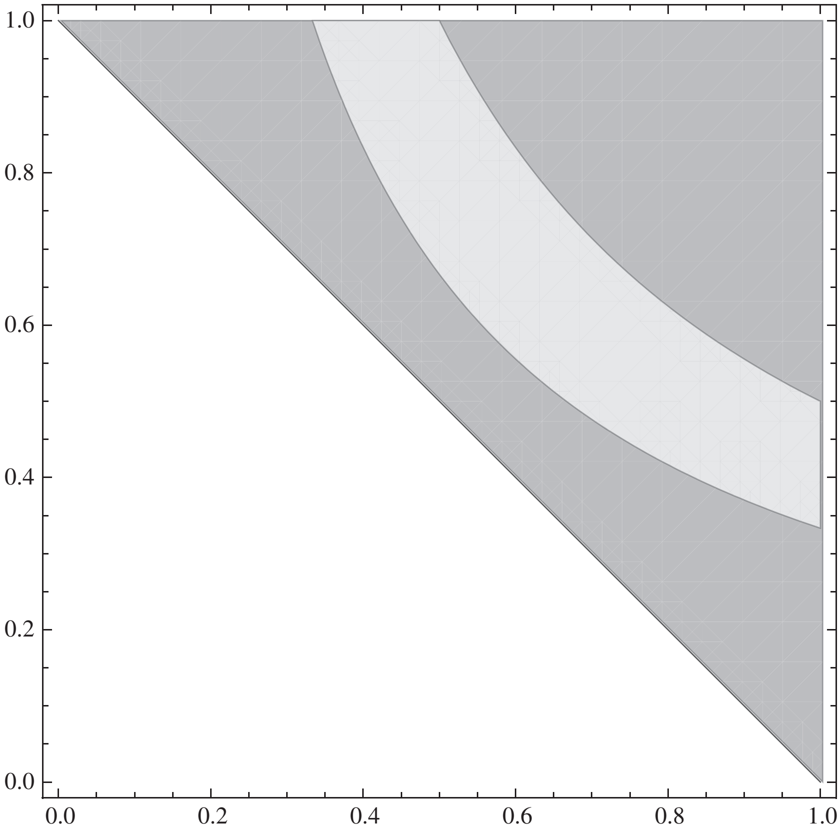

Figure 2 The region $A_{a,b}$ in inside the region Ω.

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,

$$\mathcal{G}(Q) = \left\{ N(Q)(\gamma_{i+1} - \gamma_i) = \frac{N(Q)}{q_i q_{i+1}} : 0 \leq i < N(Q) \right\}.$$

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

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

[2] Color versions of the Figures in this paper are available on the author's webpage.

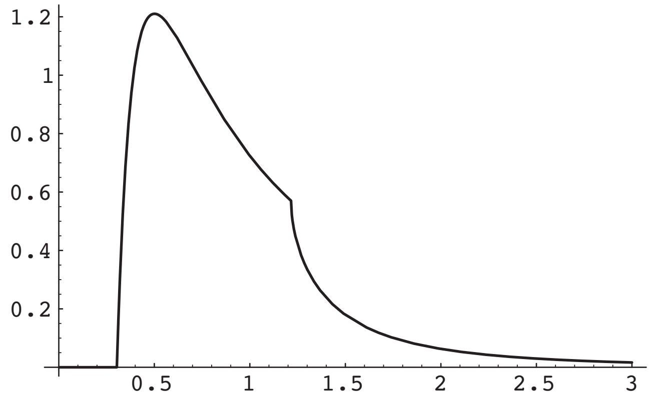

Figure 3 The limiting distribution of gaps for Farey fractions and, (appropriately rescaled) lattice slopes.

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

$$A_{a,b} = \left\{(u, v) \in \Omega : b^{-1} < uv < a^{-1}\right\}, \tag{5}$$

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

Theorem 1 *[11, R.R.Hall] Fix $0 \le a < b < \infty$. Then*

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

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 $\left\{xy = \frac{3}{\pi^2}t^{-1}\right\}$ enters the region Ω $\left(t = \frac{3}{\pi^2}\right)$ and when it hits the line $x + y = 1$ $\left(t = \frac{12}{\pi^2}\right)$. In Section 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 $\mathcal{F}(Q)$ and $\mathcal{G}(Q)$. 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 $\mathbb{R}^2$.

Geometry of Numbers

One can also study the behavior of an arbitrary unimodular lattice Λ. Let $\Lambda \subset \mathbb{R}^2$ be a unimodular lattice, and suppose Λ does not have vertical vectors. Let $\{s_1 < s_2 < \ldots < s_n < \ldots\}$ 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 2 *[4] Let* $0 \leq a < b \leq \infty$. *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 Δ_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 3 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, ω), where M is a Riemann surface and ω a holomorphic 1-form.

A saddle connection is a geodesic γ in the flat metric induced by ω, connecting two zeros of ω. To each saddle connection γ associate the holonomy vector $\mathbf{v}_\gamma = \int_\gamma \omega \in \mathbb{C}$. 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], 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, ω_P). The main result of [2] used the dynamics of the $SL(2, \mathbb{R})$ action on the moduli space Ω_g of genus g translation surfaces to show that generically, a limiting distribution exists.

More precisely, given $R > 0$, let

$$F_R^\omega := \{\arg(\mathbf{v}) : \mathbf{v} \in \Lambda_\omega \cap B(0, R)\} \tag{6}$$

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 ω. Denote the associated normalized gap set by $G^\omega(R)$.

Theorem 4 *([2, Theorem 1.1]) For almost every (with respect to Lebesgue measure on Ω_g) translation surface ω, 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 $\epsilon > 0$,*

$$\lim_{R\to\infty} \frac{|G^\omega(R) \cap (0, \epsilon)|}{N(\omega, R)} > 0. \tag{7}$$

Lattice Surfaces

The support at 0 in Theorem 4 is in contrast to the setting of the torus, where, as seen in Figure 3, there a gap between 0 and $3/\pi^2$. 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 ω is a *lattice surface* (i.e., the stabilizer of the flat surface ω under the $SL(2, \mathbb{R})$ action is a lattice) that the limiting distribution for gaps has no support at 0.

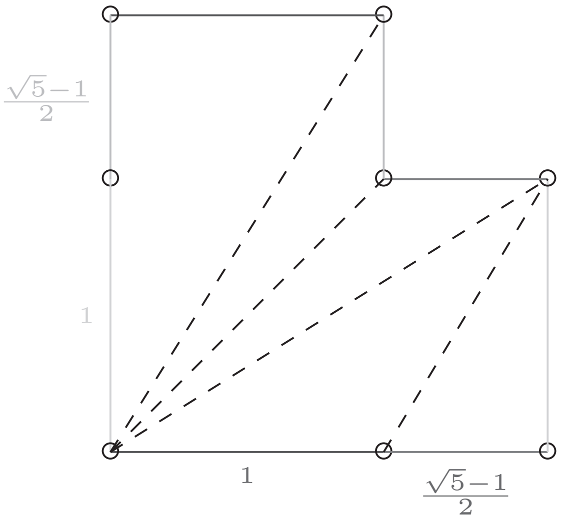

Figure 4 The Golden L. The long sides of the L each have length $\frac{1+\sqrt{5}}{2}$.

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. Lelievre, we proved Theorem 5 on the gap distribution for the golden L, which is a surface of genus 2 with one double zero, displayed in Figure 4.

Theorem 5 *[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

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

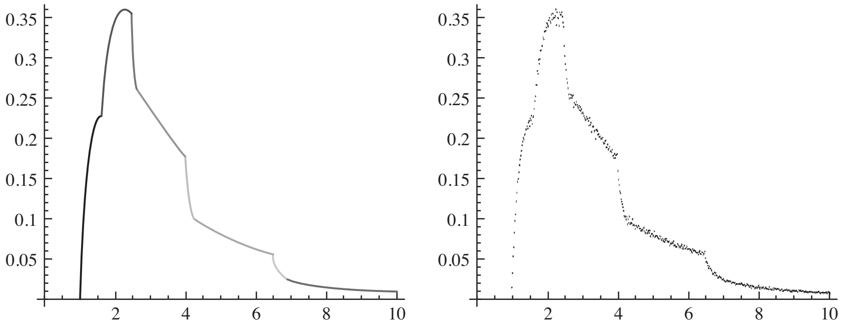

Figure 5 The limiting and empirical distributions for gaps of saddle connection slopes on the golden L.

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 $\tilde{X}_2 = SL(2,\mathbb{R}) \ltimes \mathbb{R}^2/SL(2,\mathbb{Z}) \ltimes \mathbb{Z}^2$ 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,

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

To calculate the associated gap distribution P_Λ, 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

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

shown in Figure 6. Here, the factor of R^{-2} corresponds to the normalizing factor $\frac{1}{N}$ 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

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

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