Fine-scale distribution of roots of quadratic congruences*

Jens Marklof and Matthew Welsh

School of Mathematics, University of Bristol, Bristol BS8 1UG, U.K.

25 July 2022

Abstract

We establish limit laws for the distribution in small intervals of the roots of the quadratic congruence $\mu^2 \equiv D \bmod m$, with $D > 0$ square-free and $D \not\equiv 1 \bmod 4$. This is achieved by translating the problem to convergence of certain geodesic random line processes in the hyperbolic plane. This geometric interpretation allows us in particular to derive an explicit expression for the pair correlation density of the roots.

Contents

1 Introduction **2**
1.1 Previous work on roots . . . 4
1.2 Outline of paper . . . 6

2 Geodesic line processes **6**
2.1 Convergence of geodesic line processes . . . 9
2.2 Moments . . . 12
2.3 A one-dimensional point process . . . 16

3 Equidistribution along discrete orbits **17**
3.1 A Poincaré section for the horocycle flow . . . 17
3.2 Equidistribution of intersection points on a closed geodesic . . . 20

4 Conditioned geodesic line processes **21**
4.1 Convergence in distribution . . . 21
4.2 The intensity measure . . . 23
4.3 Moments . . . 33
4.4 Palm distribution and pair correlation . . . 34
4.5 Entry and return times . . . 35

5 Geodesics and roots of quadratic congruences **36**
5.1 Roots, ideals and binary quadratic forms . . . 36
5.2 Congruence subgroups . . . 40
5.3 Negative D . . . 43

*Research supported by EPSRC grant EP/S024948/1

1 Introduction

The object of this paper is the fine-scale distribution of solutions (the "roots") μ of the quadratic congruence $\mu^2 \equiv D \bmod m$, where D is a fixed non-zero integer and m runs through the positive integers. For each m, we consider the set of normalised roots in $\mathbb{T} = \mathbb{R}/\mathbb{Z}$,

$$Q_D(m) = \left\{ \frac{\mu}{m} \in \mathbb{T} : \mu^2 \equiv D \bmod m \right\}.$$

We order the points in $Q_D(m)$ in a specified (but arbitrary) finite sequence, and then construct an infinite sequence $(\xi_j)_{j=1}^{\infty}$ in $\mathbb{T}$ by first listing the points in $Q_D(1)$, then those in $Q_D(2)$, and so on. The first significant result in this setting is due to Hooley [14], who proved uniform distribution modulo one if D is not a perfect square. That is, for any $0 \le a < b \le 1$ we have that

$$\lim_{N\to\infty} \frac{1}{N} \#\big\{ j \le N : \xi_j \in [a,b) + \mathbb{Z} \big\} = b - a.$$

Hooley later extended this result to general polynomial congruences of degree two or higher [13]. His method has since been generalised and extended, see [18] and [31].

Having established uniform distribution, it is natural to ask for finer measures that capture the pseudo-random properties of the sequence. The simplest of these is the pair (or two-point) correlation measure $R_{2,N}$, which for a finite interval I is defined by

$$R_{2,N}(I) = \frac{1}{N} \#\big\{ i,j \le N : i \ne j,\ \xi_i - \xi_j \in N^{-1} I + \mathbb{Z} \big\}.$$

Our first principal result is the following.

Theorem 1.1. *Assume $D > 0$ is square-free and $D \not\equiv 1 \bmod 4$. Then there is an even and continuous function $w_D : \mathbb{R} \to \mathbb{R}_{\ge 0}$, such that for every finite interval I we have*

$$\lim_{N\to\infty} R_{2,N}(I) = \int_I w_D(v) \mathrm{d}v.$$

In other words, the pair correlation measure $R_{2,N}$ converges vaguely to a limit with bounded continuous density w_D. Our geometric approach will allow us to derive an explicit formula for w_D; three illustrations are given in figures 1–3. Compare this with the case of independent and uniformly distributed random variables in $\mathbb{T}$ where the limit density is $w(v) = 1$; or with the eigenphases of large random unitary matrices where the limiting pair correlation is $w(v) = 1 - (\sin(\pi v)/\pi v)^2$.

The conditions that D is squarefree and $D \not\equiv 1 \bmod 4$ in Theorem 1.1 ensure that the associated quadratic order $\mathbb{Z}[\sqrt{D}]$ is maximal, which will provide some simplifications in the proofs.

Our techniques are in fact strong enough to produce results for all higher-order correlation measures. Let ξ be a random variable in $\mathbb{T}$ which is distributed according to a Borel probability measure λ, and consider the random counting measure (random point process)

$$\Xi_{N,\lambda} = \sum_{j=1}^{N} \sum_{k\in\mathbb{Z}} \delta_{N(\xi_j - \xi + k)}$$

on $\mathbb{R}$. Here δ_x denotes the Dirac mass at $x \in \mathbb{R}$. For example, for a given interval $I \subset \mathbb{R}$ and integer k, we have

$$\mathbb{P}(\Xi_{N,\lambda}(I) = k) = \lambda(\{x \in \mathbb{T} : \mathcal{N}_I(x,N) = k\})$$

with

$$\mathcal{N}_I(x,N) = \#\{ j \le N : \xi_j \in x + N^{-1} I + \mathbb{Z} \}.$$

The scaling of the interval by N^{-1} ensures that we expect a finite number of points for typical x. The following theorem describes the precise distribution.

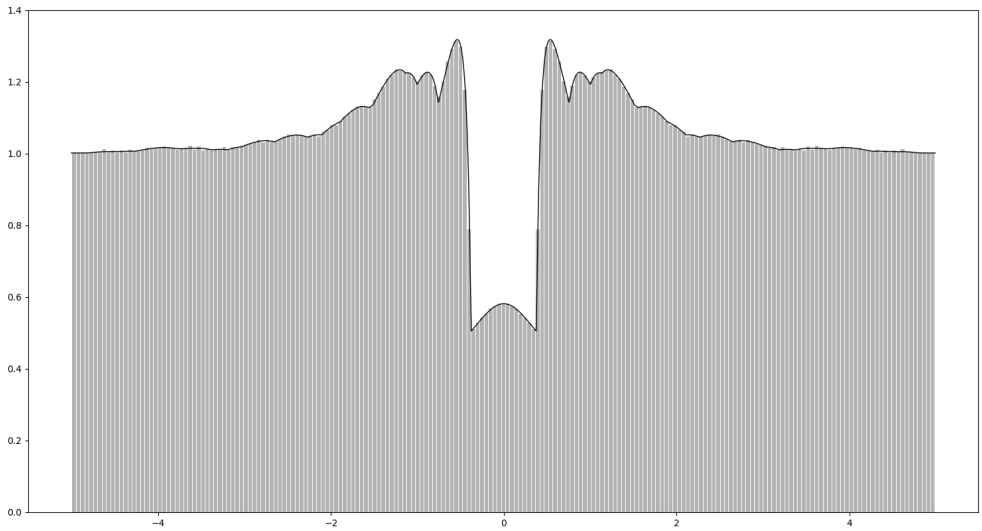

Figure 1: Pair correlation density w_2 compared to experiments with $N = 10^6$. The ring of integers $\mathbb{Z}[\sqrt{2}]$ has both class and narrow class number 1.

Theorem 1.2. *Assume $D > 0$ is square-free and $D \not\equiv 1 \bmod 4$. Then there exists a random point process Ξ depending only on D so that, for every Borel probability measure λ on $\mathbb{T}$ that is absolutely continuous with respect to the Lebesgue measure, we have convergence $\Xi_{N,\lambda} \to \Xi$ in distribution as $N \to \infty$.*

Specifically, for all $k_1, \dots, k_r \in \mathbb{Z}_{\geq 0}$ and finite intervals $I_1, \dots, I_r \subset \mathbb{R}$, we have that

$$\lim_{N\to\infty} \lambda\big(\{x \in \mathbb{T} : \mathcal{N}_{I_i}(x, N) = k_i \ \forall i\}\big) = \mathbb{P}\big(\Xi(I_i) = k_i \ \forall i\big)$$

and the limit is a continuous function of the endpoints of I_i.

We furthermore establish the uniform upper bound $\mathcal{N}_I(x, N) \ll \log N$, where the applied constant is independent of $x \in \mathbb{T}$ and N. This generalises Fouvry and Iwaniec's bound [9] in the case $D = -1$.

We provide an explicit description of the limit process Ξ in terms of tops of random geodesics in the hyperbolic plane, as well as an alternative description as entry times for a certain Poincaré section of the horocycle flow. A striking feature is that Ξ is independent of the choice of λ. Theorem 1.2 implies the convergence of various popular fine-scale statistics of the sequence $(\xi_j)_{j=1}^{\infty}$, including the gap distribution. Furthermore, we will prove the convergence of moments and higher correlation densities.

The above theorems in fact remain true (with different limit distributions) if we further restrict the roots to $m \equiv 0 \bmod n$, and $\mu \equiv \nu \bmod n$, for fixed $n > 0$ and $\nu \bmod n$ such that $\nu^2 \equiv D \bmod n$.

The results of this paper also hold for negative D. Here the proofs reduce to existing results [23, 26] on the fine-scale distribution of the real parts of hyperbolic lattice points, rather than the tops of geodesics. The studies of real parts of hyperbolic lattice points are closely related to the distribution of angles in hyperbolic lattices [3, 4, 17, 23, 27]. We also remark that results on counting the tops of geodesics can be found in [24, 25] as well as applications of a similar flavor. Their methods also apply in a general setting of negative curvature, where the notion of 'top' is naturally replaced by that of a common perpendicular between a fixed horosphere and geodesic.

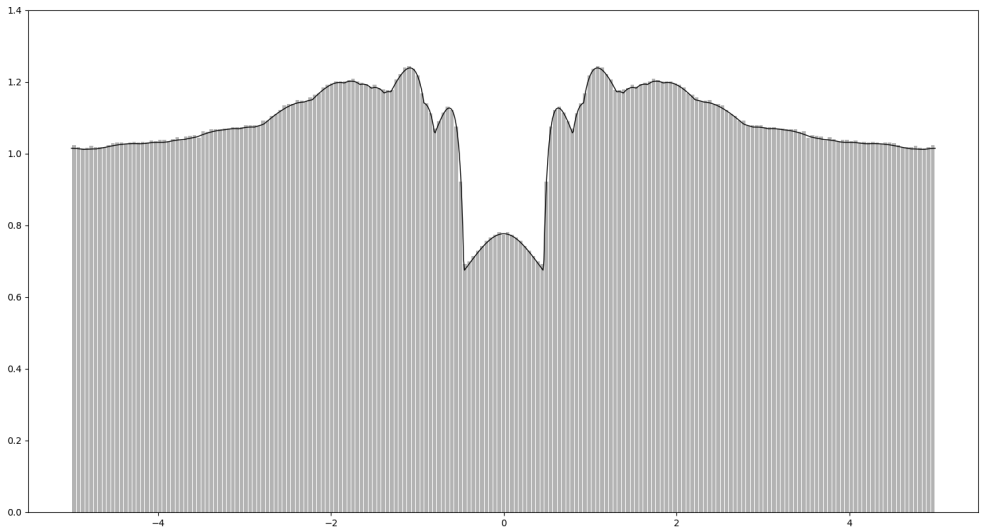

Figure 2: Pair correlation density w_3 compared to experiments with $N = 10^6$. The ring of integers $\mathbb{Z}[\sqrt{3}]$ has class number 1 and narrow class number 2.

1.1 Previous work on roots

In this section we review previous techniques used to study the sequence of roots $\frac{\mu}{m}$ and the results produced. As previously mentioned, Hooley [14] proved the equidistribution of this sequence. In fact, he produced a power-saving bound on the Weyl sums

$$\sum_{m \leq M} \sum_{\mu^2 \equiv D(m)} e\left(\frac{h\mu}{m}\right), \qquad e(x) := \mathrm{e}^{2\pi \mathrm{i} x}, \tag{1.1}$$

for non-zero integer h. These results are obtained by transforming (1.1) into a sum of Kloosterman sums, which is then estimated by an application of the Weil bound.

In [15], Iwaniec follows a similar strategy to bound the Weyl sum (1.1) with $D = -1$ but with various additional arithmetic constraints on m and μ, most notably the restriction to $m \equiv 0 \bmod n$. Iwaniec obtains enough uniformity in n so that a sieve can be applied, resulting in a proof that $l^2 + 1$ is infinitely often a product of at most two primes.

In light of the appearance of a sum of Kloosterman sums in these works, [14] and [15], one could in retrospect already predict the relevance of the spectral theory of SL(2) automorphic forms. However, the first application to analysing the sequence of $\frac{\mu}{m}$, found in the work of Bykovskiĭ, [5], proceeds by a more direct method and only applies to negative D. Roughly speaking, Bykovskiĭ's method exploits the classical connection between the roots $\mu \mod m$ and the $\mathrm{SL}(2, \mathbb{Z})$ orbits of Heegner points in the Poincaré upper half-plane $\mathbb{H}$. For example, the images of $\mathrm{i} \in \mathbb{H}$ under $\gamma \in \mathrm{SL}(2, \mathbb{Z})$ have the form

$$\gamma \mathrm{i} = \frac{\mu}{m} + \frac{\mathrm{i}}{m},$$

where μ satisfies $\mu^2 \equiv -1 \bmod m$ (see section 5.3 for more details). In this way, Bykovskiĭ turns the Weyl sum (1.1) into a finite sum of $\mathrm{SL}(2, \mathbb{Z})$-Poincaré series, which, after a spectral expansion, can be controlled using estimates for sums of Fourier coefficients of automorphic forms. These spectral techniques were extended substantially by Good [11], and similar problems were

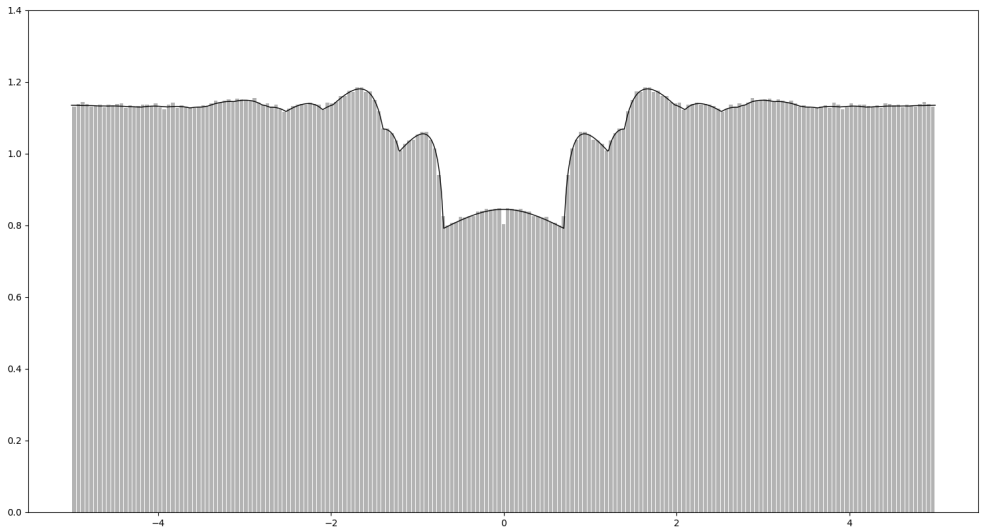

Figure 3: Pair correlation density w_{10} compared to experiments with $N = 10^6$. The ring of integers $\mathbb{Z}[\sqrt{10}]$ has both class and narrow class number 2.

considered in a very general context by Parkkonen and Paulin [24], [25], although no reference to roots of quadratic congruences are made.

In [6], Duke, Friedlander, and Iwaniec extend Bykovskiĭ's approach to apply to the Weyl sums (1.1) restricted to $m \equiv 0 \bmod n$, analogously to Iwaniec's earlier extension of Hooley's work. Here Duke, Friedlander, and Iwaniec obtain enough uniformity in h and n to apply a sieve method and conclude the spectacular result that the sequence $\frac{\mu}{m}$ with m restricted to be prime is still equidistributed modulo one.

The first use of automorphic forms to understand the sequence of $\frac{\mu}{m}$ in the case with $D > 0$ came in the work of Hejhal, [12]. Hejhal approximates the Weyl sum (1.1) by an integral of a fixed automorphic form (related to the automorphic Green's function) over some closed geodesics in $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$. The results obtained in [12] only apply for select D and are not as strong as those previously mentioned, perhaps due to a lack of a clear geometric interpretation of the roots. Nevertheless, in [30] Toth manages to extend the results of Duke, Friedlander, and Iwaniec, [6] to the positive D setting. Here the geometric question is avoided by returning to the original methods of Hooley and Iwaniec, which transform the Weyl sum directly into a sum of Kloosterman sums.

It is also worth mentioning the approach to this equidistribution question outlined by Sarnak in [28]. Sarnak suggests relating the Weyl sum (1.1) to the Fourier coefficients of certain half integral weight automorphic forms by means of Salie sums. As shown by Duke, Friedlander, and Iwaniec in [7], this kind of method is capable of producing savings in the discriminant of the polynomial F itself. None of the previously mentioned methods, nor the new method we develop below, seem to be capable of this.

Other than equidistribution, the only statistics about the sequence $\frac{\mu}{m}$ that appear in the literature, as far as we are aware, are upper bounds for the number that can lie in very short intervals. For $D = -1$, Fouvry and Iwaniec [9] found that the number of $\frac{\mu}{m}$ with $m \leq M$ in an interval of length $\frac{1}{M}$ is $O(\log M)$. Our approach reproduces this bound for general D, cf. lemma 2.8. In [9] and in Friedlander and Iwaniec's breakthrough [10], this bound was a key input into

a sieve method for producing primes of the form $l^2 + p^2$ and $l^2 + k^4$, respectively.

1.2 Outline of paper

Let Γ be a discrete subgroup of the group of isometries G acting on the complex upper half-plane $\mathbb{H}$, and let $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_h$ be a finite collection of geodesics in $\mathbb{H}$ that project to closed geodesics on the hyperbolic surface $\Gamma\backslash\mathbb{H}$. In **section 2** we study the distribution of the geodesic lines $g^{-1}\gamma\boldsymbol{c}_l$ where γ runs through Γ and $l = 1, \ldots, h$. This becomes a geodesic random line process if g is taken to be a random isometry with respect to a suitable probability measure. In **section 2.1** we prove limit theorems for sequences of random geodesics that are randomly translated, followed by the proof of convergence of moments in **section 2.2**. Key ingredients of the proofs are the equidistribution of long horocycles, a Siegel-style mean value formula and no-escape-of-mass estimates to control the moments. A convenient parametrisation of a geodesic is given by its "top", i.e., the point on the geodesics with largest imaginary part. In **section 2.3** we project the tops of the random geodesics to the real line, which produces one-dimensional point processes. The convergence of these processes to a limit will eventually lead to the proof of theorem 1.1 for the roots of quadratic congruences.

The objective of **section 3** is to prepare the ground for geodesic line process conditioned so that at least one geodesic has a top in the imaginary axis. To this end, we introduce in **section 3.1** a Poincaré section of the horocycle flow, and prove various equidistribution results for the intersection points of horocycles with the section, with an interesting variant in **section 3.2** that considers the intersection points reduced to a closed geodesic.

These equidistribution results are then applied in **section 4** to establish limit theorems for the conditioned geodesic line processes. In analogy with section 2 we prove convergence in distribution in **section 4.1**, work out out a highly non-trivial Siegel-type volume formula in **section 4.2**, and prove convergence of moments in **section 4.3**. Projecting the tops to the real line as in section 2.3 yields a point process Ξ^0 which is distributed according the Palm distribution of Ξ. The corresponding limit theorems are stated in **section 4.4**. We give an alternative description of the processes Ξ and Ξ^0 in terms of the entry and return times for the horocycle flow in **section 4.5**.

The final **section 5** of this paper expresses the roots μ of quadratic congruences in terms of tops of geodesics and hence makes the connection with the distribution results established thus far. This is detailed in **section 5.1** for the setting described in the introduction, where the relevant discrete subgroup Γ is the modular group $\mathrm{SL}(2, \mathbb{Z})$. In **section 5.2** we extend the setting to allow additional restrictions $\mu \equiv \nu \bmod n$, which leads us to the congruence subgroups $\Gamma = \Gamma_0(n)$. In **section 5.3** we indicate how our results can be modified in the case of negative D, leading to random hyperbolic lattices, rather than geodesic line processes.

Acknowledgement. We would like to thank Zonglin Li for his helpful comments, particularly on section 5, and the referees for their careful reading and suggestions.

2 Geodesic line processes

We recall that $G = \mathrm{SL}(2, \mathbb{R})$ acts on the complex upper half-plane $\mathbb{H} = \{x + \mathrm{i}y : y > 0\}$ by fractional linear transformations and that $\mathrm{PSL}(2, \mathbb{R}) = G/\{\pm I\}$ is identified with the unit tangent bundle of $\mathbb{H}$ via

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left(\frac{a\mathrm{i} + b}{c\mathrm{i} + d}, \frac{\mathrm{i}}{(c\mathrm{i} + d)^2} \right),$$

where we have written the unit tangent bundle of $\mathbb{H}$ as (z, w) with $z \in \mathbb{H}$ and $w \in \mathbb{C}$, $|w| = \mathrm{Im}(z)$. This identification is nicely expressed in terms of the Iwasawa decomposition of G: if

$$g = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{\frac{1}{2}} & 0 \\ 0 & y^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \tag{2.1}$$

then $\pm g$ is identified with $(x + \mathrm{i}y, \mathrm{i}e^{-\mathrm{i}\theta}y)$. We note that multiplying g by $-I$ corresponds to changing θ in (2.1) to $\theta + 2\pi$, and so does not change the point $(x + \mathrm{i}y, \mathrm{i}e^{-\mathrm{i}\theta}y)$. In what follows we work with G directly rather than $G/\{\pm I\}$, and we only emphasise the difference when there is a chance of confusion.

Given $g \in G$, we set

$$\boldsymbol{c} = \boldsymbol{c}(g) = \left\{ g \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \mathrm{i} : t > 0 \right\} \subset \mathbb{H},$$

which defines, via the identification of $G/\{\pm I\}$ with the unit tangent bundle of $\mathbb{H}$, an oriented geodesic in $\mathbb{H}$. Every geodesic has two limit points on the boundary $\partial\mathbb{H} = \mathbb{R} \cup \{\infty\}$. We say that $\boldsymbol{c}$ has positive orientation if both endpoints are in $\mathbb{R}$ and the endpoint corresponding to $t \to 0$ is less than the endpoint corresponding to $t \to \infty$. This means that $\boldsymbol{c}$ is a semicircle traced from left to right. (Of course we could also assign an orientation to geodesics with one endpoint at ∞, but it is convenient not to include them in this definition.) For a positively oriented geodesic $\boldsymbol{c}$ we define $z_{\boldsymbol{c}} \in \mathbb{H}$, the top of the geodesic $\boldsymbol{c}$, to be the point in $\mathbb{H}$ on $\boldsymbol{c}$ with largest imaginary part. That is, the point on $\boldsymbol{c}$ closest (with respect to the hyperbolic metric) to a horocyclic neighborhood of the boundary point $\infty \in \partial\mathbb{H}$. We see that for every $z \in \mathbb{H}$ there is a unique positively oriented geodesic $\boldsymbol{c}$ such that $z = z_{\boldsymbol{c}}$. (Note that this would fail for geodesics with endpoint ∞). The aim is now study an infinite set of geodesics in $\mathbb{H}$ generated by a discrete subgroup Γ.

Let Γ be a lattice in G which is not co-compact, and we assume (without loss of generality) that Γ is scaled so that $\Gamma\backslash\mathbb{H}$ has a cusp of width one at ∞, and also that $-I \in \Gamma$. This means that

$$\Gamma_\infty = \Gamma \cap \left\{ \pm \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{R} \right\} = \left\{ \pm \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} : k \in \mathbb{Z} \right\}$$

is a subgroup of Γ; it is the stabiliser in Γ of $\infty \in \partial\mathbb{H}$.

In our applications to roots of quadratic congruences, we will work with the modular group $\Gamma = \mathrm{SL}(2, \mathbb{Z})$ and, more generally, the congruence subgroups

$$\Gamma_0(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) : c \equiv 0 \bmod n \right\}.$$

Consider the finite collection of positively oriented geodesics $\boldsymbol{c}_1 = \boldsymbol{c}(g_1), \ldots, \boldsymbol{c}_h = \boldsymbol{c}(g_h)$ with $g_1, \ldots, g_h \in G$ so that no two $\boldsymbol{c}_l$ are Γ-equivalent (i.e. $\Gamma\boldsymbol{c}_{l'} \neq \Gamma\boldsymbol{c}_l$ for $l' \neq l$) and that each $\boldsymbol{c}_l$ projects to a closed geodesic in $\Gamma\backslash\mathbb{H}$. This last assumption implies that the stabiliser $\Gamma_{\boldsymbol{c}_l}$ of $\boldsymbol{c}_l$ in Γ,

$$\Gamma_{\boldsymbol{c}_l} = \Gamma \cap \left\{ \pm g_l \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} g_l^{-1} : t > 0 \right\},$$

is conjugate to the infinite (projectively) cyclic group

$$g_l^{-1}\Gamma_{\boldsymbol{c}_l} g_l = \left\langle \pm \begin{pmatrix} \varepsilon_l & 0 \\ 0 & \varepsilon_l^{-1} \end{pmatrix} \right\rangle$$

for suitable $\varepsilon_l > 1$.

Let us now turn to the distribution of the infinite collection of geodesics

$$\bigcup_{l=1}^{h} \bigcup_{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_l}} \gamma\boldsymbol{c}_l$$

by studying their tops,

$$\mathcal{S} = \bigcup_{l=1}^{h} \bigcup_{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_l}} z_{\gamma \boldsymbol{c}_l}.$$

Note that in general $gz_{\boldsymbol{c}_l} \neq z_{g\boldsymbol{c}_l}$. However we have $\gamma z_{\boldsymbol{c}_l} = z_{\gamma \boldsymbol{c}_l}$ for $\gamma \in \Gamma_\infty$. Thus for such γ we have $\gamma \mathcal{S} = \mathcal{S}$, i.e. $\mathcal{S}$ is invariant under translations by 1, and we can therefore reduce the problem to $\mathcal{S}$ modulo one, with $\gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_{\boldsymbol{c}_l}$. Specifically, we are interested in the fine-scale statistics of the real parts of geodesic tops with imaginary part in the interval $[\alpha y, \beta y)$,

$$X_{\alpha,\beta}(y) = \biguplus_{l=1}^{h} X^l_{\alpha,\beta}(y) \tag{2.2}$$

with

$$X^l_{\alpha,\beta}(y) = \Big\{ \mathrm{Re}(z_{\gamma \boldsymbol{c}_l}) \bmod 1 : \gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_{\boldsymbol{c}_l},\ \alpha y \leq \mathrm{Im}(z_{\gamma \boldsymbol{c}_l}) < \beta y \Big\} \subset \mathbb{T},$$

in the limit $y \to 0$. Here $X_{\alpha,\beta}(y)$, $X^l_{\alpha,\beta}(y)$ are defined as multisets, i.e., we list points with their multiplicity. To quantify their distribution, define the number of points in small intervals by

$$\mathcal{N}_{I,\alpha,\beta}(x,y) = \#\Big(X_{\alpha,\beta}(y) \cap (x + yI + \mathbb{Z}) \Big), \tag{2.3}$$

where $I \subset \mathbb{R}$ is a fixed finite interval, and $1 \leq \alpha < \beta \leq \infty$. With

$$B_{I,\alpha,\beta} = \{ u + \mathrm{i}v \in \mathbb{H} : u \in I, \alpha \leq v < \beta \}, \tag{2.4}$$

we have

$$\mathcal{N}_{I,\alpha,\beta}(x,y) = \sum_{l=1}^{h} \#\Big\{ \gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_{\boldsymbol{c}_l} : z_{\gamma \boldsymbol{c}_l} \in x + yB_{I,\alpha,\beta} + \mathbb{Z} \Big\}.$$

In fact

$$\begin{aligned} \mathcal{N}_{I,\alpha,\beta}(x,y) &= \sum_{l=1}^{h} \#\Big\{ \gamma \in \Gamma / \Gamma_{\boldsymbol{c}_l} : z_{\gamma \boldsymbol{c}_l} \in x + yB_{I,\alpha,\beta} \Big\} \\ &= \sum_{l=1}^{h} \#\Big\{ \gamma \in \Gamma / \Gamma_{\boldsymbol{c}_l} : z_{\gamma \boldsymbol{c}_l} \in n(x)a(y)B_{I,\alpha,\beta} \Big\} \end{aligned}$$

with

$$n(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \quad a(y) = \begin{pmatrix} y^{\frac{1}{2}} & 0 \\ 0 & y^{-\frac{1}{2}} \end{pmatrix}.$$

Define for $g \in G$ and bounded $B \subset \mathbb{H}$ the counting function

$$\mathcal{N}_B(g) = \sum_{l=1}^{h} \#\{ \gamma \in \Gamma / \Gamma_{\boldsymbol{c}_l} : z_{g^{-1}\gamma \boldsymbol{c}_l} \in B \}.$$

Using the fact that $gz_{\boldsymbol{c}} = z_{g\boldsymbol{c}}$ for $g = n(x)a(y)$ we see that

$$\mathcal{N}_{I,\alpha,\beta}(x,y) = \mathcal{N}_B(n(x)a(y)). \tag{2.5}$$

where $B = B_{I,\alpha,\beta}$ is given by (2.4). We furthermore note that $\mathcal{N}_B(\gamma g) = \mathcal{N}_B(g)$ for all $\gamma \in \Gamma$. This motivates the definition of the following random line processes in $\mathbb{H}$.

Define

$$\Theta_{y,\lambda} = \sum_{l=1}^{h} \sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} \delta_{z_{(n(\xi)a(y))^{-1}\gamma\boldsymbol{c}_l}}$$

with the random variable ξ distributed according to the Borel probability measure λ on $\mathbb{T}$.

We also define

$$\Theta = \sum_{l=1}^{h} \sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} \delta_{z_{g^{-1}\gamma\boldsymbol{c}_l}}$$

with the random element g distributed with respect to Haar probability measure μ_Γ on $\Gamma\backslash G$ (μ_Γ is the Haar measure of G normalised so that $\mu_\Gamma(\Gamma\backslash G) = 1$). Since $\Theta_{y,\lambda}$ and Θ describe the distribution of random geodesics in $\mathbb{H}$, we refer to them as *geodesic line processes.* Of course, in the present parametrisation they can also be viewed as point processes describing the location of the geodesic tops in $\mathbb{H}$.

2.1 Convergence of geodesic line processes

We will now investigate the convergence of our random line processes, first in distribution, then (in the next section) the convergence of its moments.

Theorem 2.1. *For every Borel probability measure λ on $\mathbb{T}$ that is absolutely continuous with respect to the Lebesgue measure, we have convergence $\Theta_{y,\lambda} \to \Theta$ in distribution as $y \to 0$.*

In particular, for all $k_1, \dots, k_r \in \mathbb{Z}_{\geq 0}$, finite intervals I_i and $1 \leq \alpha_i < \beta_i \leq \infty$, we have that

$$\lim_{y\to 0} \lambda\big(\{x \in \mathbb{T} : \mathcal{N}_{I_i,\alpha_i,\beta_i}(x,y) = k_i \ \forall i\}\big) = \mathbb{P}\big(\Theta(B_{I_i,\alpha_i,\beta_i}) = k_i \ \forall i\big)$$

and the limit is a continuous function of α_i, β_i and the endpoints of I_i.

As an example of theorem 2.1, let us consider the special case of the void distribution, where $r = 1$ and $k_1 = 0$. Here, for $1 \leq \alpha < \beta \leq \infty$ we have that for any finite interval I

$$\lim_{y\to 0} \lambda\big(\{x \in \mathbb{T} : \mathcal{N}_{I,\alpha,\beta}(x,y) = 0\}\big) = \mathbb{P}\big(\Theta(B_{I,\alpha,\beta}) = 0\big),$$

with the formula

$$\mathbb{P}\big(\Theta(B) = 0\big) = \int_{\Gamma\backslash G} \Big(\prod_{l=1}^{h} \prod_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} \Big(1 - \chi_B(z_{g^{-1}\gamma\boldsymbol{c}_l})\Big)\Big) \mathrm{d}\mu_\Gamma(g),$$

where χ_B denotes the characteristic function of the set $B = B_{I,\alpha,\beta}$.

Our proof of theorem 2.1 follows similar lines as the arguments in section 4 of [23]. The main difference here is that the points $z_{\gamma\boldsymbol{c}_l}$ do not form a Γ-orbit in $\mathbb{H}$. Both here and in [23] the main ingredient for the proof is the equidistribution of long horocycles. The precise statement we use here can be found in theorem 5.6 of [22], which is based on ideas going back to [20] and [8].

We first need to compute the average of the functions $\mathcal{N}_B(g)$, which gives the intensity measure $\mathbb{E}\Theta$ of the process Θ. The analogue of this calculation in [23] is just a straightforward application of the unfolding technique, but here it is more involved due to the fact that the $z_{\gamma\boldsymbol{c}}$ are not an orbit of a point. This obstacle is overcome by an appropriate change of variables from the Iwasawa coordinates (2.1) which has a notably simple Jacobian as seen in the proof below. Set

$$\kappa_\Gamma = \frac{\ell}{2\pi \operatorname{vol}_{\mathbb{H}}(\Gamma\backslash\mathbb{H})}, \qquad \ell = 2\sum_{l=1}^{h} \log \varepsilon_l, \tag{2.6}$$

where $2\log\varepsilon_l$ and ℓ represent the individual and total lengths of the geodesics $\boldsymbol{c}_1, \dots, \boldsymbol{c}_h$ in $\Gamma\backslash\mathbb{H}$.

Proposition 2.2. *For any Borel set $B \subset \mathbb{H}$,*

$$\mathbb{E}\Theta(B) = \int_{\Gamma\backslash G} \mathcal{N}_B(g)\mathrm{d}\mu_\Gamma(g) = \kappa_\Gamma \operatorname{vol}_{\mathbb{H}}(B). \tag{2.7}$$

Proof. We have

$$\mathcal{N}_B(g) = \sum_{l=1}^{h} \sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} \chi_{S_B}(g^{-1}\gamma g_l),$$

where χ_{S_B} is the characteristic function of the set

$$S_B = \{g \in G : z_{g\boldsymbol{c}_0} \in B\}$$

with

$$\boldsymbol{c}_0 = \left\{ \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \mathrm{i} : t > 0 \right\}$$

defining an upward oriented geodesic. Unfolding shows that

$$\int_{\Gamma\backslash G} \mathcal{N}_B(g)\mathrm{d}\mu_\Gamma(g) = \sum_{l=1}^{h} \int_{\Gamma_{\boldsymbol{c}_l}\backslash G} \chi_{S_B}(g^{-1}g_l)\mathrm{d}\mu_\Gamma(g).$$

Changing variables $g \leftarrow g_l g$, the region of integration changes from $\Gamma_{\boldsymbol{c}_l}\backslash G$ to

$$(g_l^{-1}\Gamma_{\boldsymbol{c}_l} g_l)\backslash G = \left\langle \pm \begin{pmatrix} \varepsilon_l & 0 \\ 0 & \varepsilon_l^{-1} \end{pmatrix} \right\rangle \backslash G,$$

and so the integral becomes

$$\sum_{l=1}^{h} \int_{\mathcal{F}_l} \chi_{S_B}(g^{-1})\mathrm{d}\mu_\Gamma(g), \tag{2.8}$$

where $\mathcal{F}_l$ is a fundamental domain for the action of $\left\langle \pm \begin{pmatrix} \varepsilon_l & 0 \\ 0 & \varepsilon_l^{-1} \end{pmatrix} \right\rangle$ on G.

To compute this integral, we introduce a coordinate chart on G in which both S_B and a fundamental domain $\mathcal{F}_l$ are easy to describe. The region of G where this chart applies is the open set of g such that $g^{-1}\boldsymbol{c}_0$ is positively oriented. Clearly S_B is a subset of this region, as we only define $z_{\boldsymbol{c}}$ for positively oriented geodesics, and in what follows we only consider the subset of $\mathcal{F}_l$ in this region.

The choice of coordinates we use is explicitly

$$g^{-1} = n(u)a(v)k(\tfrac{\pi}{2})a(s)^{-1}, \tag{2.9}$$

where

$$k(\theta) = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}$$

and $u, v, s \in \mathbb{R}$ with $v, s > 0$. Since $k(\frac{\pi}{2})a(s)^{-1}\boldsymbol{c}_0$ is the positively oriented geodesic with top i, we see that in these coordinates $z_{g^{-1}\boldsymbol{c}_0} = u + \mathrm{i}v$. Moreover, we observe that in this coordinate chart the action of $\pm \begin{pmatrix} \varepsilon_l & 0 \\ 0 & \varepsilon_l^{-1} \end{pmatrix}$ is simply to multiply s by ε_l^2, and so

$$\mathcal{F}_l = \{a(s)k(-\tfrac{\pi}{2})a(v^{-1})n(u)^{-1} : 1 \le s < \varepsilon_l^2, v > 0, u \in \mathbb{R}\}$$

is a valid region for use in the integration (2.8).

We now express the Haar measure on G with respect to the coordinate chart (2.9). To this end, we write the matrix on the right in (2.9) in Iwasawa coordinates. We have

$$g^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2sv}}(u+v) & \frac{\sqrt{s}}{\sqrt{2v}}(u-v) \\ \frac{1}{\sqrt{2sv}} & \frac{\sqrt{s}}{\sqrt{2v}} \end{pmatrix},$$

and so we find that the Iwasawa coordinates (2.1) of g^{-1} are

$$\begin{aligned} x &= \left(\frac{1}{2sv}(u+v) + \frac{s}{2v}(u-v)\right)\left(\frac{1}{2sv} + \frac{s}{2sv}\right)^{-1} = u - \frac{s^2-1}{s^2+1}v \\ y &= \left(\frac{1}{2sv} + \frac{s}{2v}\right)^{-1} = \frac{2sv}{s^2+1} \\ \theta &= 2\arctan\left(s^{-1}\right). \end{aligned}$$

We note that the Jacobian matrix of this change of coordinates,

$$J = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial s} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial s} \\ \frac{\partial \theta}{\partial u} & \frac{\partial \theta}{\partial v} & \frac{\partial \theta}{\partial s} \end{pmatrix},$$

is upper triangular, and so we only need to compute the diagonal entries. We have

$$\frac{\partial x}{\partial u} = 1, \quad \frac{\partial y}{\partial v} = \frac{2s}{s^2+1}, \quad \frac{\partial \theta}{\partial s} = -\frac{2}{s^2+1},$$

and so

$$\mathrm{d}x\frac{\mathrm{d}y}{y^2}\mathrm{d}\theta = \mathrm{d}u\frac{\mathrm{d}v}{v^2}\frac{\mathrm{d}s}{s}. \tag{2.10}$$

Including the normalizing factor that makes μ_Γ a probability measure on $\Gamma\backslash G$, we have that the integral (2.8) is

$$\frac{1}{2\pi \mathrm{vol}_{\mathbb{H}}(\Gamma\backslash\mathbb{H})}\int_1^{\varepsilon_l^2}\iint\limits_{u+\mathrm{i}v\in B}\mathrm{d}u\frac{\mathrm{d}v}{v^2}\frac{\mathrm{d}s}{s},$$

and (2.7) follows. □

Alhtough not strictly necessary for our purposes, we note that we can extend proposition 2.2 to the following mean value formula.

Proposition 2.3. *For every measurable $f : \mathbb{H} \to \mathbb{R}_{\geq 0}$ (or $f \in L^1(\mathbb{H}, \mathrm{vol}_{\mathbb{H}})$),*

$$\int_{\Gamma\backslash G}\left(\sum_{l=1}^{h}\sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} f(z_{g^{-1}\gamma\boldsymbol{c}_l})\right)\mathrm{d}\mu_\Gamma(g) = \kappa_\Gamma\int_{\mathbb{H}} f(z)\mathrm{d}\mathrm{vol}_{\mathbb{H}}(z).$$

Proof. As $\mathrm{vol}_{\mathbb{H}}$ is a Borel measure, we may reduce to the case when f is the characteristic function of a Borel set, which is exactly the content of proposition 2.2. □

Following the technique in [23], we derive from proposition 2.2 the following estimates.

Lemma 2.4. *We have*

$$\mu_\Gamma\left(\{g \in \Gamma\backslash G : \mathcal{N}_B(g) \geq 1\}\right) \leq \kappa_\Gamma\,\mathrm{vol}_{\mathbb{H}}(B). \tag{2.11}$$

In particular, if $\mathrm{vol}_{\mathbb{H}}(\partial B) = 0$, then

$$\mu_\Gamma\left(\partial\{g \in \Gamma\backslash G : \mathcal{N}_B(g) = k\}\right) = 0. \tag{2.12}$$

Proof. Markov's inequality implies

$$\mu_\Gamma\left(\{g \in \Gamma\backslash G : \mathcal{N}_B(g) \geq 1\}\right) \leq \int_{\Gamma\backslash G} \mathcal{N}_B(g)\mathrm{d}\mu_\Gamma(g),$$

and so proposition 2.2 implies (2.11). We now observe that $g' \in \partial\{g \in \Gamma\backslash G : \mathcal{N}_B(g) = k\}$ implies $\mathcal{N}_{\partial B}(g') \geq 1$, and so (2.12) follows from (2.11). □

Proof of theorem 2.1. We now have all the ingredients to apply the equidistribution of long horocycles, theorem 5.6 of [22]. This theorem states in particular that, for λ absolutely continuous with respect to Lebesgue measure, if $S \subset \Gamma\backslash G$ has boundary of μ_Γ-measure 0, then

$$\lim_{y\to 0} \int_{\mathbb{T}} \chi_S(n(x)a(y))\mathrm{d}\lambda(x) = \mu_\Gamma(S),$$

where χ_S is the indicator function of S. Now (2.12) shows that the relevant set S for our theorem 2.1 does indeed have boundary of μ_Γ-measure zero. This furthermore implies that the limit distribution is continuous as stated. □

2.2 Moments

We now turn to the moments of the counting functions $\mathcal{N}_{I,\alpha,\beta}(x,y)$. For $y > 0$, $\boldsymbol{s} = (s_1, \ldots, s_r) \in \mathbb{C}^r$, $\mathcal{I} = I_1 \times \cdots \times I_r$, $I_j \subset \mathbb{R}$ a finite interval, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)$ satisfying $1 \leq \alpha_j < \beta_j \leq \infty$, define the probabilities (cf. theorem 2.1)

$$P^\lambda_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}, y) = \lambda(\{x \in \mathbb{T} : \mathcal{N}_{I_j,\alpha_j,\beta_j}(x,y) = k_j \ \forall j\}),$$

$$P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}) = \mathbb{P}\big(\Theta(B_{I_i,\alpha_i,\beta_i}) = k_i \ \forall i\big),$$

where $\boldsymbol{k} = (k_1, \ldots, k_r)$ with k_j non-negative integers; the moments

$$M^\lambda_{\mathcal{I},\boldsymbol{\beta},\boldsymbol{\alpha}}(\boldsymbol{\kappa}, y) = \int_{\mathbb{T}} \prod_{1\leq j\leq r} \mathcal{N}_{I_j,\alpha_j,\beta_j}(x,y)^{\kappa_j}\mathrm{d}\lambda(x) = \sum_{k_1,\ldots,k_r\geq 0} k_1^{\kappa_1}\cdots k_r^{\kappa_r} P^\lambda_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}, y),$$

and

$$M_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa}) = \int_{\Gamma\backslash G} \prod_{1\leq j\leq r} \mathcal{N}_{I_j,\alpha_j,\beta_j}(g)^{\kappa_j}\mathrm{d}\mu_\Gamma(g) = \sum_{k_1,\ldots,k_r\geq 0} k_1^{\kappa_1}\cdots k_r^{\kappa_r} P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}),;$$

where $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_r) \in \mathbb{R}^r$ with $\kappa_j \geq 0$; and the moment generating functions

$$\begin{aligned} G^\lambda_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}, y) &= \int_{\mathbb{T}} \exp\left(\sum_{1\leq j\leq r} s_j \mathcal{N}_{I_j,\alpha_j,\beta_j}(x,y)\right)\mathrm{d}\lambda(x) \\ &= \sum_{k_1,\ldots,k_r\geq 0} \exp\left(\sum_{1\leq j\leq r} k_j s_j\right) P^\lambda_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}, y), \end{aligned}$$

and

$$\begin{aligned} G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}) &= \int_{\Gamma\backslash G} \exp\left(\sum_{1\leq j\leq r} s_j \mathcal{N}_{I_j,\alpha_j,\beta_j}(g)\right)\mathrm{d}\mu_\Gamma(g) \\ &= \sum_{k_1,\ldots,k_r\geq 0} \exp\left(\sum_{1\leq j\leq r} k_j s_j\right) P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}). \end{aligned}$$

Theorem 2.5. *Let λ be a probability measure on $\mathbb{T}$ with bounded density (with respect to Lebesgue measure), and $\mathcal{I}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ as above. Then there is a constant $c_0 > 0$ such that for $\boldsymbol{s} = (s_1, \dots, s_r) \in \mathbb{C}^r$ satisfying $\sum_j \max\{\mathrm{Re}(s_j), 0\} < c_0$ the function $G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s})$ is analytic and we have*

$$\lim_{y\to 0} G^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}, y) = G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}).$$

By a standard argument, theorem 2.5 implies the convergence as $y \to 0$ of the mixed moments $M^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa}, y)$.

Corollary 2.6. *Let λ be a probability measure on $\mathbb{T}$ with bounded density, and $\mathcal{I}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ as above. Then for all $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_r) \in \mathbb{R}^r$ with $\kappa_j \geq 0$, we have $M_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa})$ is finite and*

$$\lim_{y\to\infty} M^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa}, y) = M_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa}).$$

We also record the following corollary, which is simply an application of corollary 2.6 to $r = 1$ and $\kappa = 1$ together with the calculation in proposition 2.2.

For $0 \leq a < b \leq 1$ and $y > 0$, define the counting function for the distribution of the real parts of $z_{\gamma \boldsymbol{c}_l}$ modulo one,

$$N_{\alpha,\beta}(a, b, y) = \#\Big(X_{\alpha,\beta}(y) \cap ([a, b] + \mathbb{Z})\Big).$$

Corollary 2.7 (Equidistribution modulo one)**.** *For a finite interval $I \subset \mathbb{R}$, $1 \leq \alpha < \beta \leq \infty$ and probability measure λ with bounded density, we have*

$$\lim_{y\to 0} \int_{\mathbb{T}} \mathcal{N}_{I,\alpha,\beta}(x, y)\mathrm{d}\lambda(x) = \kappa_\Gamma \left(\frac{1}{\alpha} - \frac{1}{\beta}\right)\mathrm{length}(I) \tag{2.13}$$

and, for $0 \leq a < b \leq 1$,

$$\lim_{y\to 0} yN_{\alpha,\beta}(a, b, y) = \kappa_\Gamma \left(\frac{1}{\alpha} - \frac{1}{\beta}\right)(b - a). \tag{2.14}$$

Proof. As mentioned above, (2.13) follows from corollary 2.6 and proposition 2.2. Equation (2.14) follows from the observation that given any $\epsilon > 0$, we have for $0 < y < \epsilon$ and $I = [-\frac{1}{2}, \frac{1}{2}]$

$$\int_{a+\epsilon}^{b-\epsilon} \mathcal{N}_{I,\alpha,\beta}(x, y)\mathrm{d}x \leq yN_{\alpha,\beta}(a, b, y) \leq \int_{a-\epsilon}^{b+\epsilon} \mathcal{N}_{I,\alpha,\beta}(x, y)\mathrm{d}x.$$

Indeed, from (2.3), the left side (resp. right side) counts each point in $\xi \in X_{\alpha,\beta}(y)$ with $\xi \in [a, b] + \mathbb{Z}$ with weight at most (resp. at least) $y\mathrm{length}(I)$. In view of (2.13) (choose λ to be the uniform measure on the intervals $[a + \epsilon, b - \epsilon]$ and $[a - \epsilon, b + \epsilon]$, respectively), the limits of the left and right hand side exist for all $\epsilon > 0$. Taking $\epsilon \to 0$, we obtain the right hand side of (2.14). □

Our proof of theorem 2.5 follows the strategy of the proof of theorem 8 in [23]. As in the proof of theorem 2.1, our setting is somewhat complicated by the fact that our points $z_{\gamma \boldsymbol{c}_l}$ do not form a Γ-orbit of a point in $\mathbb{H}$. However there is a point in $w_l \in \mathbb{H}$ such that for any of our points $z_{\gamma \boldsymbol{c}_l}$, there is a $\gamma' \in \Gamma$ such that $z_{\gamma \boldsymbol{c}_l}$ is a bounded distance from $\gamma' w_l$ (and in fact $\gamma' \Gamma_{\boldsymbol{c}_l} = \gamma \Gamma_{\boldsymbol{c}_l}$). This fact gives us sufficient control over the points $z_{\gamma \boldsymbol{c}_l}$ for the purposes of proving theorem 2.5, as manifested in lemmas 2.8 and 2.9.

Lemma 2.8. *For every $\delta > 1$, there is a constant C_δ such that or any finite interval I, $0 \leq \alpha < \beta \leq \delta\alpha < \infty$, we have that*

(i) *for $x \in \mathbb{T}$, $y > 0$,*

$$\mathcal{N}_{I,\alpha,\beta}(x, y) < C_\delta \left(1 + \frac{\mathrm{length}(I)}{\alpha}\right); \tag{2.15}$$

(ii) *for* $g \in G$ *and* $B = B_{I,\alpha,\beta}$ *as in (2.4),*

$$\mathcal{N}_B(g) < C_\delta\Big(1 + \frac{\text{length}(I)}{\alpha}\Big); \tag{2.16}$$

(iii) *for* $x \in \mathbb{T}$, $0 < y \le \mathrm{e}^{-1}$,

$$\mathcal{N}_{I,1,\infty}(x,y) < C_2\big(1 + \text{length}(I)\big)\log y^{-1}; \tag{2.17}$$

(iv) *for* $0 \le a < b \le 1$, $0 < y \le 1$,

$$N_{1,\infty}(a,b,y) < 4C_2\, y^{-1}. \tag{2.18}$$

(The constant C_δ *only depends on* δ, Γ *and the geodesics* $\boldsymbol{c}_l$*.)*

Proof. Consider the number of hyperbolic lattice points in $B_{I,\alpha,\beta}$,

$$\widetilde{\mathcal{N}}_{I,\alpha,\beta}(g) = \sum_{l=1}^{h} \#\{\gamma \in \Gamma : g^{-1}\gamma g_l \mathrm{i} \in B_{I,\alpha,\beta}\}.$$

The set $g^{-1}\Gamma g_l$ is uniformly discrete, and the minimum distance between any two points is independent of g, since G acts by isometries. Thus there is a constant C_δ such that $\widetilde{\mathcal{N}}_{[0,1],1,\delta}(g) \le C_\delta$ for all $g \in G$. Note that

$$\widetilde{\mathcal{N}}_{[t,t+\alpha],\alpha,\beta}(g) \le \widetilde{\mathcal{N}}_{[t,t+\alpha],\alpha,\delta\alpha}(g) = \widetilde{\mathcal{N}}_{[0,1],1,\delta}(gn(t)a(\alpha)) \le C_\delta.$$

This bound is evidently uniform in t, α, β, g. A covering argument shows that

$$\widetilde{\mathcal{N}}_{I,\alpha,\beta}(g) \le C_\delta\Big(1 + \frac{\text{length}(I)}{\alpha}\Big).$$

We now show that this bound implies (2.16). Since $z_{\gamma\boldsymbol{c}_l}$ depends only on the coset $\gamma\Gamma_{\boldsymbol{c}_l}$, we claim that we can choose γ appropriately so that $\gamma g_l \mathrm{i}$ is at most a bounded hyperbolic distance from $z_{\gamma\boldsymbol{c}_l}$. Indeed, $\gamma g_l \mathrm{i}$ is on the geodesic $\gamma\boldsymbol{c}_l$, and $\gamma\Gamma_l\gamma^{-1}$ is generated by the hyperbolic motion along this geodesic moving points a distance of $2\log\varepsilon_l$. Hence γ can be chosen so that $\gamma g_l i$ is a distance at most $\log\varepsilon_l$ from $z_{\gamma\boldsymbol{c}_l}$. Therefore (2.16) follows by adjusting C_δ with a constant depending on ε_l. Next, (2.15) follows trivially from (2.16) in view of (2.5).

Again using the fact that there is a γ such that $\gamma g_l \mathrm{i}$ is a bounded distance from $z_{\gamma\boldsymbol{c}_l}$, it follows that there is a constant $Y > 0$ depending only on g_l and Γ such that $\mathrm{Im}(z_{\gamma\boldsymbol{c}_l}) \le Y$ for all γ (recall that since $\Gamma\backslash\mathbb{H}$ has a cusp at ∞, the maximum height of any orbit is bounded). We then have for $y < Y$ (the bound for $y \ge Y$ is obvious)

$$\begin{aligned}\mathcal{N}_{I,1,\infty}(x,y) &= \sum_{0\le j\le \log_2(Y/y)} \mathcal{N}_{I,2^j,2^{j+1}}(x,y)\\ &\le \sum_{0\le j\le \log_2(Y/y)} C_2\big(1+\text{length}(I)\big),\end{aligned}$$

using (2.15). This establishes (2.17).

The final bound (2.18) follows from a similar argument. For the upper bound we may take $[a,b] = [0,1]$. Then using periodicity in x, we have

$$\begin{aligned}N_{1,\infty}(0,1,y) &= \sum_{j=0}^{\infty} N_{1,2}(0,1,2^j y)\\ &= y^{-1}\sum_{j=0}^{\infty} 2^{-j}\int_0^1 \mathcal{N}_{[0,1],1,2}(x,2^j y)\mathrm{d}x,\end{aligned}$$

which in view of (2.15) is bounded by $4C_2\, y^{-1}$. This proves (2.18). □

Lemma 2.9. *Let λ be a measure with bounded density, $\mathcal{I}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ as above, and $\boldsymbol{k} = (k_1, \dots, k_r)$ with $k_j \geq 0$ integers. Then there exists $c_0 > 0$ such that*

$$\sup_{y>0} P^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}, y) \leq \sup_{y>0} \lambda\big(\{x \in \mathbb{T} : \mathcal{N}_{I_i,\alpha_i,\beta_i}(x, y) \geq k_i \ \forall i\}\big) \ll \exp\Big(-c_0 \max_{1\leq j\leq r} k_j\Big) \tag{2.19}$$

and

$$P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}) \leq \mathbb{P}\big(\Theta(B_{I_i,\alpha_i,\beta_i}) \geq k_i \ \forall i\big) \ll \exp\Big(-c_0 \max_{1\leq j\leq r} k_j\Big). \tag{2.20}$$

Proof. The inequalities "$\leq$" are trivial. As to the upper bounds "$\ll$", it is sufficient to establish them for $[\alpha_i, \beta_i) = [1, \infty)$. Let j be so that $k_j = \max_i k_i$. We may assume that $k_j \geq 1$, since (2.19) is otherwise evident. Now recalling that

$$\mathcal{N}_{I_j,1,\infty}(x, y) = \sum_{l=1}^{h} \#\{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_l} : \mathrm{Re}(z_{\gamma\boldsymbol{c}_l}) \in x + yI_j, \mathrm{Im}(z_{\gamma\boldsymbol{c}_l}) \geq y\}$$

and that by lemma 2.8

$$\sum_{l=1}^{h} \#\{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_l} : \mathrm{Re}(z_{\gamma\boldsymbol{c}_l}) \in x + yI_j, y \leq \mathrm{Im}(z_{\gamma\boldsymbol{c}_l}) \leq Yy\} \ll \log Y$$

for $Y > 1$ (note that this is not the same Y that appears in the proof of lemma 2.8), we conclude that for x such that $\mathcal{N}_{I_j,1,\infty}(x, y) \geq k_j$,

$$\sum_{l=1}^{h} \#\{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_l} : \mathrm{Re}(z_{\gamma\boldsymbol{c}_l}) \in x + yI_j, \mathrm{Im}(z_{\gamma\boldsymbol{c}_l}) \geq Cy\exp(c_0 k_j)\} \geq 1$$

for some positive C, c_0 independent of x and y. Setting $Y = C\exp(c_0 k_j)$ we have that

$$\{x \in \mathbb{T} : \mathcal{N}_{I_j,1,\infty}(x, y) \geq k_j\} \subset \{x \in \mathbb{T} : \mathcal{N}_{I_j,Y,\infty}(x, y) \geq 1\}. \tag{2.21}$$

By Markov's inequality, the λ measure of the set on the right of (2.21) is at most

$$\begin{aligned} &\int_{\mathbb{T}} \mathcal{N}_{I_j,Y,\infty}(x, y)\mathrm{d}\lambda(x) \ll \int_{\mathbb{T}} \mathcal{N}_{I_j,Y,\infty}(x, y)\mathrm{d}x \\ &\ll y \ \mathrm{length}(I_j) \sum_{l=1}^{h} \#\{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_j} : 0 \leq \mathrm{Re}(z_{\gamma\boldsymbol{c}_l}) \leq 1, \mathrm{Im}(z_{\gamma\boldsymbol{c}_l}) \geq yY\}, \end{aligned} \tag{2.22}$$

where the first inequality follows from λ having bounded density. It follows from lemma 2.8 that the quantity on the right of (2.22) is

$$\ll Y^{-1} \ll \exp(-c_0 k_j)$$

with the implied constant independent of y (but depending on I_j, the geodesics $\boldsymbol{c}_l$, etc.). This establishes (2.19), and (2.20) follows immediately from (2.19) and theorem 2.1. □

Proof of theorem 2.5. Having lemma 2.9, the proof of theorem 2.5 is easy to finish. To verify that $G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s})$ is analytic in the set of $\boldsymbol{s} = (s_1, \dots, s_r) \in \mathbb{C}^r$ satisfying $\sum_j \max\{\mathrm{Re}(s_j), 0\} < c_0$, we start by arranging

$$G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}) = \sum_{k\geq 0} \sum_{\substack{k_1,\dots,k_r\geq 0 \\ \max_j k_j = k}} \exp\left(\sum_{1\leq j\leq r} k_j s_j\right) P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}).$$

Applying (2.20) we find that

$$\begin{aligned}
&\sum_{k\geq K}\sum_{\substack{k_1,\dots,k_r\geq 0\\ \max_j k_j=k}}\left|\exp\left(\sum_{1\leq j\leq r}k_js_j\right)P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k})\right| \\
&\ll \sum_{k\geq K}(k+1)^r\exp\left(-\left(c_0-\sum_{1\leq j\leq r}\max\{\mathrm{Re}(s_j),0\}\right)k\right).
\end{aligned} \tag{2.23}$$

This bound clearly goes to 0 uniformly as $K\to\infty$ for $\boldsymbol{s}$ in compact subsets of the region $\sum_j\max\{\mathrm{Re}(s_j),0\}<c_0$, and so $G(\boldsymbol{s},\mathcal{I})$ is analytic in this region.

To verify that

$$\lim_{y\to 0}G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s},y)=G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}), \tag{2.24}$$

we recall that

$$G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s},y)=\sum_{k_1,\dots,k_r\geq 0}\exp\left(\sum_{1\leq j\leq r}k_js_j\right)P^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k},y).$$

Proceeding as above, we arrange this sum according to $k=\max_j k_j$, and split the sum into $k<K$ and $k\geq K$.

For the terms $k<K$, we apply theorem 2.1 to obtain

$$\begin{aligned}
&\lim_{y\to\infty}\sum_{0\leq k<K}\sum_{\substack{k_1,\dots,k_r\geq 0\\ \max_j k_j=k}}\exp\left(\sum_{1\leq j\leq r}k_js_j\right)P^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k},y) \\
&=\sum_{0\leq k<K}\sum_{\substack{k_1,\dots,k_r\geq 0\\ \max_j k_j=k}}\exp\left(\sum_{1\leq j\leq r}k_js_j\right)P_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k}).
\end{aligned}$$

Using (2.23), we see that this is equal to $G_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s})$ up to an error that goes to 0 as $K\to\infty$ since we are assuming $\sum_j\max\{\mathrm{Re}(s_j),0\}<c_0$. For the terms $k\geq K$, we apply (2.19) to obtain

$$\begin{aligned}
&\sum_{k\geq K}\sum_{\substack{k_1,\dots,k_r\geq 0\\ \max_j k_j=k}}|\exp\left(\sum_{1\leq j\leq r}k_js_j\right)P^{\lambda}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{k})| \\
&\ll \sum_{k\geq K}(k+1)^r\exp\left(-\left(c_0-\sum_{1\leq j\leq r}\max\{\mathrm{Re}(s_j),0\}\right)k\right).
\end{aligned}$$

As before, this goes to 0 as $K\to\infty$ since $\sum_j\max\{\mathrm{Re}(s_j),0\}<c_0$, and so we obtain (2.24). □

2.3 A one-dimensional point process

We will now restate theorem 2.1 (in the case $\alpha=1$, $\beta=\infty$) in a form closer to the setting of the introduction. The family of multisets

$$X_{1,\infty}(y)=\biguplus_{l=1}^{h}\left\{\,\mathrm{Re}(z_{\gamma\boldsymbol{c}_l})\bmod 1:\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_{\boldsymbol{c}_l},\ \mathrm{Im}(z_{\gamma\boldsymbol{c}})\geq y\right\}\subset\mathbb{T},$$

forms a nested subsequence as $y \to 0$. We list its elements (with multiplicity) as the sequence $(\xi_j)_{j=1}^{\infty}$, where we fix any ordering such that if $\xi_i \in X_{1,\infty}(y')$, $\xi_j \in X_{1,\infty}(y)$ with $i < j$, then $y' \geq y$. So in particular, the elements in $X_{1,\infty}(y)$ are listed as $\xi_1, \ldots, \xi_N$ with $N = N(y) = N_{1,\infty}(0, 1, y)$. As in the introduction, we are interested in the number $\mathcal{N}_I(x, N) = \mathcal{N}_{I_i,1,\infty}(x, N)$ of points in the interval $x + N^{-1}I + \mathbb{Z}$. We define the corresponding random point process

$$\Xi_{N,\lambda} = \sum_{j=1}^{N} \sum_{k \in \mathbb{Z}} \delta_{N(\xi_j - \xi + k)}$$

on $\mathbb{R}$, with $\xi \in \mathbb{T}$ distributed according to λ. Define the one-dimensional point process process Ξ by

$$\Xi(I) = \Theta(B_{\kappa_\Gamma^{-1} I,1,\infty})$$

with κ_Γ as in (2.6). With this scaling,

$$\mathbb{E}\Xi(I) = \mathbb{E}\Theta(B_{\kappa_\Gamma^{-1} I,1,\infty}) = \kappa_\Gamma \operatorname{vol}_{\mathbb{H}}(B_{\kappa_\Gamma^{-1} I,1,\infty}) = \operatorname{length}(I),$$

and hence the intensity measure of Ξ is the normalised Lebesgue measure on $\mathbb{R}$. Furthermore, since the distribution of Θ is invariant under translations $z \mapsto z + t$ of $\mathbb{H}$, the process Ξ is a (translation-) stationary point process in $\mathbb{R}$.

We have the following limit theorem.

Theorem 2.10. *For every Borel probability measure λ on $\mathbb{T}$ that is absolutely continuous with respect to the Lebesgue measure, we have convergence $\Xi_{N,\lambda} \to \Xi$ in distribution as $N \to \infty$.*

Specifically, for all $k_1, \ldots, k_r \in \mathbb{Z}_{\geq 0}$, finite intervals I_i, we have that

$$\lim_{N \to \infty} \lambda\big(\{x \in \mathbb{T} : \mathcal{N}_{I_i}(x, N) = k_i \ \forall i\}\big) = \mathbb{P}\big(\Xi(I_i) = k_i \ \forall i\big).$$

Proof. This follows directly from theorem 2.1. The only difference is the scaling by N rather than y^{-1}. The equivalence of the convergence in the two scalings follows from the asymptotics $yN(y) \to \kappa_\Gamma$ from (2.14) and the continuity of the limit distribution. □

3 Equidistribution along discrete orbits

Theorems 2.1 and 2.5 already give the existence of limits for many interesting fine-scale distributions of the the real parts of $z_{\gamma \boldsymbol{c}_l}$ mod 1. There are standard combinatorial arguments that allow the extension of the results to other statistics, such as gap distribution and pair correlation. See for instance [23], where the pair correlation is computed from second mixed moments. Here we will use a geometric approach, the surface of section method, from which the gap distribution and pair correlation follow more directly. This is similar to the approach in [1] for the statistics of Farey sequences, but with a different section, see section 4.5 for details. We refer the reader to [21] for more background on the connection between statistics of entry and return times for a given Poincaré section on one hand, and point processes and their Palm distributions on the other.

3.1 A Poincaré section for the horocycle flow

For $1 \leq \alpha < \beta \leq \infty$, define the two-dimensional section

$$S_{\alpha,\beta} = \{a(s)k(-\tfrac{\pi}{2})a(v^{-1}) : s > 0, \alpha \leq v < \beta\} \subset G.$$

We will show below that for $1 \le \alpha < \beta \le \infty$ the sets

$$\widetilde{S}^l_{\alpha,\beta} = \Gamma\backslash\Gamma g_l S_{\alpha,\beta} = \Gamma\backslash\Gamma g_l \{a(s)k(-\tfrac{\pi}{2})a(v^{-1}) : s \in [1, \varepsilon_l^2), \alpha \le v < \beta\}$$

and

$$\widetilde{S}_{\alpha,\beta} = \bigcup_{l=1}^{h} \widetilde{S}^l_{\alpha,\beta}$$

are measurable subsets of $\Gamma\backslash G$. Define the finite Borel measures $\nu_{\alpha,\beta}$, $\nu^l_{\alpha,\beta}$ on $\Gamma\backslash G$ by

$$\int_{\Gamma\backslash G} f(g)\,\mathrm{d}\nu^l_{\alpha,\beta}(g) = \frac{1}{2\pi \mathrm{vol}_{\mathbb{H}}(\Gamma\backslash\mathbb{H})} \int_\alpha^\beta \int_1^{\varepsilon_l^2} f(\Gamma g_l a(s) k(-\tfrac{\pi}{2}) a(v^{-1}))\, \frac{\mathrm{d}s}{s}\, \frac{\mathrm{d}v}{v^2},$$

and furthermore

$$\nu_{\alpha,\beta} = \sum_{l=1}^{h} \nu^l_{\alpha,\beta},$$

so that the support of $\nu_{\alpha,\beta}$ is $\widetilde{S}_{\alpha,\beta}$, and that of $\nu^l_{\alpha,\beta}$ is $\widetilde{S}^l_{\alpha,\beta}$.

Recall the definition of the multiset $X_{\alpha,\beta}(y)$ in (2.2).

Theorem 3.1. *For $1 \le \alpha < \beta \le \infty$ and $f : \mathbb{T} \times \Gamma\backslash G \to \mathbb{C}$ bounded continuous, we have*

$$\lim_{y\to 0} y \sum_{\xi \in X_{\alpha,\beta}(y)} f(\xi, \Gamma n(\xi) a(y)) = \int_{\mathbb{T}} \int_{\Gamma\backslash G} f(x, g) \mathrm{d}\nu_{\alpha,\beta}(g) \mathrm{d}x. \tag{3.1}$$

The plan for the proof of this statement is to show that $\widetilde{S}_{\alpha,\beta}$ is a Poincaré section for the horocycle flow $\Gamma g \mapsto \Gamma g n(t)$ on $\Gamma\backslash G$ such that the return times for the periodic orbit $\{\Gamma a(y) n(t) : t \in \mathbb{T}\}$ have the form $y^{-1}\mathrm{Re}(z_{\gamma_{\boldsymbol{c}_l}})$. We then introduce an ϵ-thickening of $\widetilde{S}_{\alpha,\beta}$ in the direction of the horocycle flow, and theorem 3.1 then follows by applying the equidistribution of long horocycles in $\Gamma\backslash G$.

Let us first investigate the structure of $\widetilde{S}^l_{\alpha,\beta}$. Recall $\{\Gamma g_l a(s) : s > 0\}$ is a closed geodesic in $\Gamma\backslash G$ for each l, and in particular the curves are not self-intersecting and are disjoint for different values of l. Therefore $\{\Gamma g_l a(s) : s > 0\} k(-\frac{\pi}{2}) a(v^{-1})$ are also closed curves in $\Gamma\backslash G$, which are not self-intersecting and are disjoint for different values of l. In fact, they do not intersect even after a slight thickening, as the following lemma shows.

Lemma 3.2. *For $1 \le \alpha < \beta < \infty$ and $\epsilon > 0$ define*

$$S^\epsilon_{\alpha,\beta} = \{a(s)k(-\tfrac{\pi}{2})a(v^{-1})n(u) : s > 0, \alpha \le v < \beta, |u| < \epsilon\} \subset G.$$

Then there is $\delta > 1$ sufficiently small (depending only on Γ and the g_l) and $\epsilon > 0$ sufficiently small (depending only on δ, Γ and the g_l), such that for $\beta < \delta\alpha$ and $\gamma \in \Gamma$ and $1 \le l_1, l_2 \le h$, we have that

$$\gamma g_{l_1} S^\epsilon_{\alpha,\beta} \cap g_{l_2} S^\epsilon_{\alpha,\beta} = \emptyset$$

unless $l_1 = l_2$ and $\gamma \in \Gamma_{\boldsymbol{c}_{l_1}}$.

Proof. Suppose that $\gamma \in \Gamma$ is such that

$$\gamma g_{l_1} a(s_1) k(-\tfrac{\pi}{2}) a(v_1)^{-1} n(u_1) = g_{l_2} a(s_2) k(-\tfrac{\pi}{2}) a(v_2)^{-1} n(u_2),$$

for some $s_1, s_2 > 0$, $\alpha \le v_1, v_2 < \beta$, and $|u_1|, |u_2| < \epsilon$. Replacing γ with $\gamma_2\gamma\gamma_1$ for suitable $\gamma_i \in \Gamma_{\boldsymbol{c}_i}$, we may assume furthermore that $1 \le s_1 < \varepsilon_{l_1}^2$, $1 \le s_2 < \varepsilon_{l_2}^2$. This implies that

$$g_{l_2}^{-1} \gamma g_{l_1} = \begin{pmatrix} s^{\frac{1}{2}} & 0 \\ 0 & s^{-\frac{1}{2}} \end{pmatrix} + O(\beta\epsilon + (\alpha^{-\frac{1}{2}}\beta^{\frac{1}{2}} - \alpha^{\frac{1}{2}}\beta^{-\frac{1}{2}})) \tag{3.2}$$

where $s = s_2/s_1$, and so $\varepsilon_{l_1}^{-2} < s < \varepsilon_{l_2}^2$. The condition that (3.2) holds for some $\varepsilon_{l_1}^{-2} < s < \varepsilon_{l_2}^2$ constrains γ to a fixed compact subset of G. The discreteness of Γ then implies that for $\beta/\alpha < \delta$ sufficiently close to 1 and $\beta\epsilon$ sufficiently small, if (3.2), then in fact

$$g_{l_2}^{-1}\gamma g_{l_1} = \begin{pmatrix} s^{\frac{1}{2}} & 0 \\ 0 & s^{-\frac{1}{2}} \end{pmatrix} \tag{3.3}$$

for some $\varepsilon_{l_1}^{-2} \leq s \leq \varepsilon_{l_2}^2$. Since the geodesics $\boldsymbol{c}_l$ are not Γ-equivalent, (3.3) implies $l_1 = l_2$, and so then $\gamma \in \Gamma_{\boldsymbol{c}_{l_1}}$. □

The lemma implies that $\widetilde{S}^l_{\alpha,\beta}$ are measurable subsets and $\nu_{\alpha,\beta}$ well defined as Borel measures. This observation will be key in the application of the surface-of-section method below.

Lemma 3.3. *Let $1 \leq \alpha < \beta < \delta\alpha < \infty$ with δ as in lemma 3.2. Then the map*

$$X_{\alpha,\beta}(y) \to \widetilde{S}_{\alpha,\beta}, \qquad \xi \mapsto \Gamma n(\xi)a(y) \tag{3.4}$$

is injective, and we have

$$X_{\alpha,\beta}(y) = \{\xi \in \mathbb{T} : \Gamma n(\xi)a(y) \in \widetilde{S}_{\alpha,\beta}\}.$$

Proof. Recall that $z_{\gamma\boldsymbol{c}_l}$ is the point on $\gamma\boldsymbol{c}_l$ where the tangent points directly to the right. The Iwasawa decomposition (2.1) hence tells us that

$$n(\mathrm{Re}(z_{\gamma\boldsymbol{c}_l}))a(\mathrm{Im}(z_{\gamma\boldsymbol{c}_l}))k(\tfrac{\pi}{2}) = \gamma g_l a(s) \tag{3.5}$$

for a unique $s > 0$. Therefore if $\alpha y \leq \mathrm{Im}(z_{\gamma\boldsymbol{c}_l}) < \beta y$, then

$$n(\mathrm{Re}(z_{\gamma\boldsymbol{c}_l}))a(y) = \gamma g_l a(s)k(-\tfrac{\pi}{2})a(v^{-1}) \in \gamma g_l S_{\alpha,\beta} \tag{3.6}$$

with $v = y^{-1}\mathrm{Im}(z_{\gamma\boldsymbol{c}_l})$, and so

$$\Gamma n(\mathrm{Re}(z_{\gamma\boldsymbol{c}_l}))a(y) \in \widetilde{S}^l_{\alpha,\beta} = \Gamma\backslash\Gamma g_l S_{\alpha,\beta}.$$

This shows that the map (3.4) is well defined. In view of lemma 3.2, the value of $v \in [\alpha, \beta)$ is in fact unique. Together with the uniqueness of $s \in [1, \varepsilon_l^2)$ this implies the injectivity of (3.4). □

Note that the previous results require $\beta < \infty$. The following result shows that restricting to finite intervals $[\alpha, \beta)$ only produces a small error.

Lemma 3.4. *For $\alpha > 0$ and $f : \mathbb{T} \times \Gamma\backslash G \to \mathbb{C}$ bounded, we have*

$$\limsup_{y\to 0} \left| y \sum_{\xi \in X_{\alpha,\infty}(y)} f(\xi, \Gamma n(\xi)a(y)) \right| \leq \kappa_\Gamma\, \alpha^{-1} \sup|f|.$$

Proof. Apply corollary 2.7, eq. (2.13). □

Proof of theorem 3.1. Since the sum over l is finite, we may assume without loss of generality that we only have one term, i.e., $h = 1$. We write $[\alpha, \beta)$ as a disjoint union of intervals $[\alpha_j, \alpha_{j+1})$, $j = 1, \ldots, J$, with $\alpha_1 = \alpha$ and $\alpha_{J+1} = \beta$. If $\beta = \infty$, we can choose α_J sufficiently large to make the contribution of the interval $[\alpha_J, \infty)$ to both sides of (3.1) as small as we wish; recall lemma 3.4. We may therefore assume without loss of generality that $\beta < \infty$. Assume furthermore $\alpha_{j+1} < \delta\alpha_j$ with $\delta > 1$ as in lemma 3.2. We will now calculate the contribution to (3.1) of each finite interval $[\alpha_j, \alpha_{j+1})$. For simplicity of notation, let us rename $[\alpha_j, \alpha_{j+1})$ as $[\alpha, \beta)$. We have

thus reduced the proof of the lemma to the statement for $1 \leq \alpha < \beta < \delta\alpha < \infty$ with $\delta > 1$ as in lemma 3.2.

What now remains to be shown is that for $f : \mathbb{T} \times \widetilde{S}^l_{\alpha,\beta} \to \mathbb{C}$ bounded continuous, we have

$$\lim_{y\to 0} y \sum_{\xi \in X^l_{\alpha,\beta}(y)} f(\xi, \Gamma n(\xi)a(y)) = \int_{\mathbb{T}} \int_{\Gamma\backslash G} f(x,g) \mathrm{d}\nu^l_{\alpha,\beta}(g)\mathrm{d}x.$$

To this end, we define for given $\epsilon > 0$ and f bounded continuous, the function $F_\epsilon : \mathbb{T} \times \Gamma\backslash G \to \mathbb{C}$ by

$$F_\epsilon(x, \Gamma g) = \begin{cases} f(x, \Gamma g_l a(s) k(-\frac{\pi}{2}) a(v^{-1})) & \text{if } g = g_l a(s) k(-\frac{\pi}{2}) a(v^{-1}) n(u) \in \Gamma g_l S^\epsilon_{\alpha,\beta} \\ 0 & \text{if } g \notin \Gamma g_l S^\epsilon_{\alpha,\beta}. \end{cases}$$

By lemma 3.2 and the assumptions on α, β, the function F_ϵ is well defined, is bounded and has compact support; it is continuous except at points with $u = \pm\epsilon$, which form a set of Haar measure zero, recall the parametrisation (2.10). The key observation is now that in view of (3.5)

$$\frac{1}{2\epsilon} \int_{\mathbb{T}} F_\epsilon(x, \Gamma n(x) a(y)) \mathrm{d}x = y \sum_{\xi \in X^l_{\alpha,\beta}(y)} f(\xi, \Gamma n(\xi) a(y)). \tag{3.7}$$

The equidistribution of closed horocycles (as stated in theorem 5.6 of [22]) implies that

$$\lim_{y\to 0} \frac{1}{2\epsilon} \int_{\mathbb{T}} F_\epsilon(x, \Gamma n(x) a(y)) \mathrm{d}x = \frac{1}{2\epsilon} \int_{\mathbb{T}\times\Gamma\backslash G} F_\epsilon(x,g)\, \mathrm{d}x\, \mathrm{d}\mu_\Gamma(g).$$

Using the formula (2.10) for μ_Γ with the correct normalisation factor, we work out that

$$\begin{aligned} &\frac{1}{2\epsilon} \int_{\mathbb{T}\times\Gamma\backslash G} F_\epsilon(x,g)\, \mathrm{d}x\, \mathrm{d}\mu_\Gamma(g) \\ &= \frac{1}{2\pi \mathrm{vol}_{\mathbb{H}}(\Gamma\backslash\mathbb{H})} \int_{\mathbb{T}} \int_\alpha^\beta \int_1^{\varepsilon_l^2} f(x, \Gamma g_l a(s) k(-\tfrac{\pi}{2}) a(v^{-1}))\, \frac{\mathrm{d}s}{s}\, \frac{\mathrm{d}v}{v^2}\, \mathrm{d}x. \end{aligned} \tag{3.8}$$

In conclusion, the right hand side of (3.7) converges to the right hand side of (3.8), as required. □

3.2 Equidistribution of intersection points on a closed geodesic

The following corollary of the discussion in the previous section may be of independent interest. It concerns the joint equidistribution of the points $\mathrm{Re}(z_{\gamma \boldsymbol{c}_l}) + \mathbb{Z} \in \mathbb{T}$ (with respect to Lebesgue measure) and the points $\Gamma z_{\gamma \boldsymbol{c}_l}$ on the closed geodesic $\Gamma\backslash\Gamma \boldsymbol{c}_l$ (with respect to the arc length measure). To be precise, for $\gamma \in \Gamma_\infty\backslash\Gamma/\Gamma_{\boldsymbol{c}_l}$, we define $t_l(\gamma)$ as the unique $t \in [0, 2\log\varepsilon_l)$ such that

$$\Gamma_{\boldsymbol{c}_l} \gamma^{-1} n(\mathrm{Re}(z_{\gamma \boldsymbol{c}_l})) a(\mathrm{Im}(z_{\gamma \boldsymbol{c}_l})) = \Gamma_{\boldsymbol{c}_l} g_l a(\mathrm{e}^t) k(-\tfrac{\pi}{2}) \tag{3.9}$$

holds; recall (3.6) and $2\log\varepsilon_l$ is the total length of $\boldsymbol{c}_l$. We remark that in the context of the roots $\mu^2 \equiv D \bmod m$, the natural harmonics on the closed geodesics are Hecke characters for $\mathbb{Q}(\sqrt{D})$.

Proposition 3.5. *For $1 \leq \alpha < \beta \leq \infty$ and $f : \mathbb{T} \times [0, 2\log\varepsilon_l] \to \mathbb{C}$ bounded continuous, we have*

$$\lim_{y\to\infty} y \sum_{\substack{\gamma \in \Gamma_\infty\backslash\Gamma/\Gamma_{\boldsymbol{c}_l} \\ \alpha y \leq \mathrm{Im}(z_{\gamma \boldsymbol{c}_l}) < \beta y}} f(\mathrm{Re}(z_{\gamma \boldsymbol{c}_l}), t_l(\gamma)) = \frac{\alpha^{-1} - \beta^{-1}}{2\pi \mathrm{vol}_{\mathbb{H}}(\Gamma\backslash\mathbb{H})} \int_0^{2\log\varepsilon_l} \int_{\mathbb{T}} f(x,t)\, \mathrm{d}x\, \mathrm{d}t. \tag{3.10}$$

Proof. As above, we may assume without loss of generality that $1 \leq \alpha < \beta < \delta\alpha < \infty$ with $\delta > 1$ as in lemma 3.2. We define $\tilde{f} : \mathbb{T} \times \tilde{S}^l_{\alpha,\beta} \to \mathbb{C}$ by

$$\tilde{f}(x, \Gamma g_l a(s) k(-\frac{\pi}{2}) a(v^{-1})) = \chi_{[\alpha,\beta)}(v) f(x, \log s).$$

This is well-defined by lemma 3.2, and we observe from (3.9) that

$$\lim_{y\to\infty} y \sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_{\boldsymbol{c}_l} \\ \alpha y \leq \mathrm{Im}(z_{\gamma \boldsymbol{c}_l}) < \beta y}} f(\mathrm{Re}(z_{\gamma \boldsymbol{c}_l}), t_l(\gamma)) = \lim_{y\to\infty} y \sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_{\boldsymbol{c}_l} \\ \alpha y \leq \mathrm{Im}(z_{\gamma \boldsymbol{c}_l}) < \beta y}} \tilde{f}(\mathrm{Re}(z_{\gamma \boldsymbol{c}_l}), \Gamma n(\mathrm{Re}(z_{\boldsymbol{c}_l})) a(y)).$$

Applying theorem 3.1, the right side becomes

$$\int_{\mathbb{T}} \int_{\Gamma\backslash G} \tilde{f}(x, g) \mathrm{d}\nu_{\alpha,\beta}(g) \mathrm{d}x = \frac{\alpha^{-1} - \beta^{-1}}{2\pi \mathrm{vol}_{\mathbb{H}}(\Gamma\backslash\mathbb{H})} \int_1^{\varepsilon_l^2} \int_{\mathbb{T}} f(x, \log s) \mathrm{d}x \frac{\mathrm{d}s}{s}.$$

Equation (3.10) follows after changing variables $t = \log s$. □

4 Conditioned geodesic line processes

In section 2.1 we considered the number of points $\mathcal{N}_{I,\alpha,\beta}(x, y)$ in small intervals (2.3), where x is distributed according to an absolutely continuous Borel probability measure on $\mathbb{T}$. In order to capture gap and nearest neighbour statistics, as well as pair and higher correlation measures, we now evaluate $\mathcal{N}_{I,\alpha,\beta}(\eta, y)$ for η randomly drawn from the sequence itself. That is, the random variable η is uniformly from the finite multiset $X_{\alpha,\beta}(y)$ as defined in (2.2), with fixed $1 \leq \alpha < \beta \leq \infty$. As in section 2.1, we define the corresponding geodesic line processes by

$$\Theta^0_{y,\alpha,\beta} = \sum_{l=1}^{h} \sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} \delta_{z_{(n(\eta)a(y))^{-1}\gamma \boldsymbol{c}_l}}$$

and

$$\Theta^0_{\alpha,\beta} = \sum_{l=1}^{h} \sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}} \delta_{z_{g^{-1}\gamma \boldsymbol{c}_l}}$$

with the random element g distributed with respect to the Borel probability measure $\widehat{\nu}_{\alpha,\beta}$ on $\Gamma\backslash G$, which is $\nu_{\alpha,\beta}$ normalised as a probability measure. That is,

$$\widehat{\nu}_{\alpha,\beta} = \kappa_\Gamma^{-1} \left(\frac{1}{\alpha} - \frac{1}{\beta} \right)^{-1} \nu_{\alpha,\beta}.$$

We may think of $\Theta^0_{y,\alpha,\beta}$ and $\Theta^0_{\alpha,\beta}$ as the processes $\Theta_{y,\lambda}$ resp. Θ (defined in section 2.1) conditioned to have a point on the line $\{z \in \mathbb{H} : \mathrm{Re}\, z = 0,\ \alpha \leq \mathrm{Im}\, z < \beta\}$.

4.1 Convergence in distribution

The key result of this section is the following analogue of theorem 2.1.

Theorem 4.1. *For $1 \leq \alpha < \beta \leq \infty$, $0 \leq a < b \leq 1$, we have convergence $\Theta^0_{y,\alpha,\beta} \to \Theta^0_{\alpha,\beta}$ in distribution as $y \to 0$.*

In particular, for all $k_1, \dots, k_r \in \mathbb{Z}_{\geq 0}$, finite intervals I_i, $0 \leq a < b \leq 1$, and $1 \leq \alpha_i < \beta_i \leq \infty$, we have that

$$\begin{aligned}&\lim_{y\to 0} y\#\big\{\xi \in X_{\alpha,\beta}(y) \cap ([a,b)+\mathbb{Z}) : \mathcal{N}_{I_i,\alpha_i,\beta_i}(\xi,y) = k_i \ \forall i\big\}\\ &= \kappa_\Gamma(b-a)\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B_{I_i,\alpha_i,\beta_i}) = k_i \ \forall i\big)\end{aligned}$$

and the limit is a continuous function of α_i, β_i and the endpoints of I_i.

In the special case of the void distribution (recall theorem 2.1) we have for instance that for any finite interval I (or finite unions of finite intervals),

$$\lim_{y\to 0} y\#\big\{\xi \in X_{\alpha,\beta}(y) : \mathcal{N}_{I,\alpha_1,\beta_1}(\xi,y) = 0\big\} = \kappa_\Gamma\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B_{I,\alpha_1,\beta_1}) = 0\big)$$

with the formula

$$\begin{aligned}\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B) = 0\big) &= \int_{\Gamma\backslash G}\left(\prod_{l=1}^{h}\prod_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}}\Big(1-\chi_B(z_{g^{-1}\gamma\boldsymbol{c}_l})\Big)\right)\mathrm{d}\widehat{\nu}_{\alpha,\beta}(g)\\ &= \kappa_\Gamma\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)\sum_{l_0=1}^{h}\int_\alpha^\beta\int_1^{\varepsilon_{l_0}^2}\left(\prod_{l=1}^{h}\prod_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}}\Big(1-\chi_{S_B}(a(v)k(\tfrac{\pi}{2})a(s)g_{l_0}^{-1}\gamma g_l)\Big)\right)\frac{\mathrm{d}s}{s}\,\frac{\mathrm{d}v}{v^2}.\end{aligned}$$

In particular, the choice $I = (0,t]$ yields the probability that the gap between consecutive elements is greater than t, and $I = [-t,0)\cup(0,t]$ yields the probability that the distance to the nearest neighbour is greater than t.

The proof of theorem 4.1 follows the same route as that of theorem 2.1, except that, rather than equidistribution of closed long horocycles, we now use the discrete averages established in theorem 3.1. We will not repeat the proof, other than to establish the following lemma which guarantees the regularity of the limit distribution.

Lemma 4.2. *For any finite interval I, $1 \leq \alpha < \beta \leq \infty$, $1 \leq \alpha_1 < \beta_1 \leq \infty$, we have*

$$\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B_{\partial I,\alpha_1,\beta_1}) \geq 1\big) = 0 \qquad \text{if } 0 \notin \partial I, \tag{4.1}$$

$$\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B_{\partial I,\alpha_1,\beta_1}) = 1\big) = 1 \qquad \text{if } 0 \in \partial I, \tag{4.2}$$

and

$$\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B_{I,\alpha_1,\alpha_1}) \geq 1\big) = 0. \tag{4.3}$$

Proof. To prove (4.3), it is sufficient to show that the set

$$\bigcup_{1\leq l_0\leq h}\bigcup_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}}\Big\{(s,v) : \operatorname{Im} z_{a(v)k(\frac{\pi}{2})a(s)g_{l_0}^{-1}\gamma\boldsymbol{c}_l} = \alpha_1\Big\} \subset \mathbb{R}^2_{>0} \tag{4.4}$$

has measure zero with respect to $\frac{\mathrm{d}s}{s}\,\frac{\mathrm{d}v}{v^2}$. Note that $z_{a(v)k(\frac{\pi}{2})a(s)g_{l_0}^{-1}\gamma\boldsymbol{c}_l} = v z_{k(\frac{\pi}{2})a(s)g_{l_0}^{-1}\gamma\boldsymbol{c}_l}$, which implies that the set $\{(s,v) : \operatorname{Im} z_{a(v)k(\frac{\pi}{2})a(s)g_{l_0}^{-1}\gamma\boldsymbol{c}_l} = \alpha_1\}$ has measure zero for every γ, l_0, and so does the countable union (4.4). This proves (4.3). Analogously, to prove (4.1), we need to show that

$$\bigcup_{1\leq l_0\leq h}\bigcup_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_l}}\Big\{(s,v) : \operatorname{Re} z_{g(s,v)\gamma\boldsymbol{c}_l} = \epsilon\Big\} \subset \mathbb{R}^2_{>0}, \tag{4.5}$$

where $g(s,v) = g(s,v,l_0) = a(v)k(\frac{\pi}{2})a(s)g_{l_0}^{-1}$, has measure zero for $\epsilon \neq 0$. This follows from the same argument as above. For (4.2), we need to investigate (4.5) in the case $\epsilon = 0$. Here we have

$$\Big\{(s,v) : \operatorname{Re} z_{g(s,v)\gamma \boldsymbol{c}_l} = 0\Big\} = \Big\{(s,v) : \operatorname{Re} z_{g(s,1)\gamma \boldsymbol{c}_l} = 0\Big\}. \tag{4.6}$$

If $l \neq l_0$, or if $l = l_0$ and $\gamma \neq \Gamma_{\boldsymbol{c}_l}$, there is a unique $s > 0$ such that $\operatorname{Re} z_{g(s,1)\gamma \boldsymbol{c}_l} = 0$, which shows (4.6) has measure zero in these cases. The remaining case is $l = l_0$, $\gamma = \Gamma_{\boldsymbol{c}_l}$, which yields $\operatorname{Re} z_{g(s,1)\gamma \boldsymbol{c}_l} = \operatorname{Re} z_{k(\frac{\pi}{2})\boldsymbol{c}_0} = 0$ for all v, s and hence corresponds to the probability one event. □

4.2 The intensity measure

We now turn to the intensity measure of the process $\Theta^0_{\alpha,\beta}$, defined by

$$\mathbb{E}\Theta^0_{\alpha,\beta}(B) = \int_{\Gamma\backslash G} \mathcal{N}_B(g)\mathrm{d}\widehat{\nu}_{\alpha,\beta}(g), \tag{4.7}$$

and prove an analogue of proposition 2.2. We first prove the following lemma that shows (4.7) is finite (and which will be useful to prove the convergence of moments in the next section).

Lemma 4.3. *Let $\mathcal{I}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ as in section 2.2, and $\boldsymbol{k} = (k_1, \ldots, k_r)$ with $k_j \geq 0$ integers. Then there exists $c_1 > 0$ such that*

$$\sup_{y>0} y\#\big\{\xi \in X_{\alpha_0,\beta_0}(y) : \mathcal{N}_{I_i,\alpha_i,\beta_i}(\xi, y) \geq k_i \ \forall i\big\} \ll \exp\Big(-c_1 \max_{1\leq j\leq r} k_j\Big) \tag{4.8}$$

and

$$\mathbb{P}\big(\Theta^0_{\alpha,\beta}(B_{I_i,\alpha_i,\beta_i}) \geq k_i \ \forall i\big) \ll \exp\Big(-c_1 \max_{1\leq j\leq r} k_j\Big). \tag{4.9}$$

Proof. We assume without loss of generality that $[\alpha_i, \beta_i) = [1, \infty)$ and label j so that $k_j = \max_{1\leq i\leq r} k_i$. Take λ to be the uniform probability measure on $[0,1]$, and $I_\epsilon = [-\frac{\epsilon}{2}, \frac{\epsilon}{2}]$ for some fixed $\epsilon > 0$. Then

$$\begin{aligned}
&\#\big\{\xi \in X_{\alpha_0,\beta_0}(y) : \mathcal{N}_{I_i,\alpha_i,\beta_i}(\xi, y) \geq k_i \ \forall i\big\} \\
&\leq \frac{1}{\epsilon}\sum_{k=1}^{\infty} k\ \lambda\big(\big\{x \in \mathbb{T} : \mathcal{N}_{I_\epsilon,1,\infty}(x,y) = k,\ \mathcal{N}_{I_i,1,\infty}(x,y) \geq k_i \ \forall i\big\}\big).
\end{aligned}$$

Lemma 2.9 tells us that this is bounded above by

$$\ll \sum_{k=1}^{\infty} k \exp(-c_0 \max\{k, k_j\}) \ll \exp\big(-\tfrac{1}{2}c_0 k_j\big).$$

This proves (4.8), which in turn, in view of theorem 4.1, implies the bound (4.9). □

We will now work out the expectation measure $\mathbb{E}\Theta^0_{\alpha,\beta}$. By linearity, it is sufficient to consider the case $\beta = \infty$ and test sets of the form $B = B_{I,\alpha_2,\infty}$.

Proposition 4.4. *Let $\alpha_1, \alpha_2 \in [1, \infty)$ and I a finite interval. Then*

$$\mathbb{E}\Theta^0_{\alpha_1,\infty}(B_{I,\alpha_2,\infty}) = \min\left\{1, \frac{\alpha_1}{\alpha_2}\right\} \delta_0(I) + \int_I W_{\alpha_1,\alpha_2}(v)\mathrm{d}v \tag{4.10}$$

where δ_0 denotes the delta mass at the orgin and $W_{\alpha_1,\alpha_2}:\mathbb{R}\to\mathbb{R}_{\geq 0}$ is the even and continuous function given by

$$W_{\alpha_1,\alpha_2}(v)=\frac{\alpha_1}{\ell v^2}\sum_{l_1,l_2=1}^{h}\sum_{\substack{\gamma\in\Gamma_{\boldsymbol{c}_{l_1}}\backslash\Gamma/\Gamma_{\boldsymbol{c}_{l_2}}\\ \gamma\boldsymbol{c}_{l_2}\neq\boldsymbol{c}_{l_1},\overline{\boldsymbol{c}_{l_1}}}} H_{\mathrm{sign}(g_{l_1}^{-1}\gamma g_{l_2}(0))}(q(\gamma,l_1,l_2),v\alpha_1^{-1},v\alpha_2^{-1}),$$

where $q(\gamma,l_1,l_2)=\frac{r+1}{r-1}$ with r the cross-ratio

$$r=\frac{((\gamma\boldsymbol{c}_{l_2})^+-\boldsymbol{c}_{l_1}^-)((\gamma\boldsymbol{c}_{l_2})^--\boldsymbol{c}_{l_1}^+)}{((\gamma\boldsymbol{c}_{l_2})^+-\boldsymbol{c}_{l_1}^+)((\gamma\boldsymbol{c}_{l_2})^--\boldsymbol{c}_{l_1}^-)},\tag{4.11}$$

and

$$\begin{aligned}&H_+(q,v_1,v_2)\\&\quad=\int_{-1}^{1}\chi_{[1,\infty)}\left(2v_1\frac{s+q}{s^2+2qs+1}\right)\chi_{[1,\infty)}\left(v_2\frac{1-s^2}{s^2+2qs+1}\right)\left|\frac{s^2+2qs+1}{(s+q)(s^2-1)}\right|\mathrm{d}s,\end{aligned}\tag{4.12}$$

$$\begin{aligned}H_-(q,v_1,v_2)=&\left(\int_{-\infty}^{-1}+\int_{1}^{\infty}\right)\\&\chi_{[1,\infty)}\left(2v_1\frac{s+q}{s^2+2qs+1}\right)\chi_{[1,\infty)}\left(v_2\frac{1-s^2}{s^2+2qs+1}\right)\left|\frac{s^2+2qs+1}{(s+q)(s^2-1)}\right|\mathrm{d}s.\end{aligned}\tag{4.13}$$

In the case $v_1=v_2=v$, we have explicitly

$$H_+(q,v,v)=\begin{cases}0 & \text{if } q<-1\\ 0 & \text{if } -1<q<1 \text{ and } v<\sqrt{2-2q}\\ h_q(s_1(q,v))-h_q(s_2(q,v)) & \text{if } -1<q<1 \text{ and } v>\sqrt{2-2q}\\ h_q(s_1(q,v))-h_q(-q+\sqrt{q^2-1}) & \text{if } q>1,\end{cases}\tag{4.14}$$

$$H_-(q,v,v)=\begin{cases}0 & \text{if } q<-1 \text{ and } |v|<\sqrt{2-2q}\\ h_q(s_1(q,v))-h_q(s_2(q,v)) & \text{if } q<-1 \text{ and } |v|>\sqrt{2-2q}\\ h_q(s_1(q,v))-h_q(s_2(q,v)) & \text{if } -1<q<1 \text{ and } v<-\sqrt{2-2q}\\ 0 & \text{if } -1<q<1 \text{ and } v>-\sqrt{2-2q}\\ h_q(-q-\sqrt{q^2-1})-h_q(s_2(q,v)) & \text{if } q>1,\end{cases}\tag{4.15}$$

with

$$h_q(s)=\log\frac{s+q}{1-s^2},$$

$$s_1(q,v)=\frac{-q+\sqrt{v^2+q^2-1}}{v+1},$$

and

$$s_2(q,v)=v-q-\sqrt{v^2+q^2-1}.$$

Proof. With the notation as in the proof of proposition 2.2, and $B=B_{I,\alpha_2,\infty}$,

$$\begin{aligned}&\int_{\Gamma\backslash G}\mathcal{N}_{I,\alpha_2,\infty}(g)\mathrm{d}\widehat{\nu}_{\alpha_1,\infty}(g)\\&=\frac{\alpha_1}{\ell}\sum_{l_1,l_2=1}^{h}\sum_{\gamma\in\Gamma/\Gamma_{\boldsymbol{c}_{l_2}}}\int_{\alpha_1}^{\infty}\int_{1}^{\varepsilon_{l_1}^2}\chi_{S_B}(a(v)k(\tfrac{\pi}{2})a(s)g_{l_1}^{-1}\gamma g_{l_2})\,\frac{\mathrm{d}s}{s}\,\frac{\mathrm{d}v}{v^2}.\end{aligned}\tag{4.16}$$

Recall that χ_{S_B} is the characteristic function of the set $S_B = \{g \in G : z_{g\boldsymbol{c}_0} \in B\}$. In view of lemma 4.3, the series is absolutely convergent.

The contribution to (4.16) of terms with $l_1 = l_2 = l$ and $\gamma = \Gamma_{\boldsymbol{c}_{l_2}}$ is

$$\frac{\alpha_1}{\ell} \sum_{l=1}^{h} \int_{\alpha_1}^{\infty} \int_1^{\varepsilon_l^2} \chi_{S_B}(a(v)k(\tfrac{\pi}{2})a(s)) \, \frac{\mathrm{d}s}{s} \, \frac{\mathrm{d}v}{v^2} = \min\left\{1, \frac{\alpha_1}{\alpha_2}\right\} \delta_0(I),$$

since $a(v)k(\frac{\pi}{2})a(s) \in S_B$ if and only if $0 \in I$ and $v \geq \alpha_2$. This explains the first term on the right hand side of (4.10).

To evaluate the remaining terms, we break the sum over $\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_{l_2}}$ ($\gamma \neq \Gamma_{\boldsymbol{c}_{l_2}}$) into a sum over $\Gamma_{\boldsymbol{c}_{l_1}}$ and a sum over double cosets $\Gamma_{\boldsymbol{c}_{l_1}} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}}$. To this end, we first note that we have the partition

$$\Gamma_{\boldsymbol{c}_{l_1}} \gamma \Gamma_{\boldsymbol{c}_{l_2}} = \bigcup_{\gamma' \in \Gamma_{\boldsymbol{c}_{l_1}}} \gamma' \gamma \Gamma_{\boldsymbol{c}_{l_2}},$$

with a disjoint union, if and only if $\gamma^{-1}\Gamma_{\boldsymbol{c}_{l_1}}\gamma \cap \Gamma_{\boldsymbol{c}_{l_2}}$ is trivial. Now if there were an element in this intersection, then it would fix the endpoints of $\boldsymbol{c}_{l_2}$ and $\gamma^{-1}\boldsymbol{c}_{l_1}$, and so if this element were nontrivial we would have $\gamma \boldsymbol{c}_{l_2} = \boldsymbol{c}_{l_1}$ or $\gamma \boldsymbol{c}_{l_2} = \overline{\boldsymbol{c}_{l_1}}$, where $\overline{\boldsymbol{c}_{l_1}}$ is the geodesic $\boldsymbol{c}_{l_1}$ with the opposite orientation. In the first case we would have $l_1 = l_2 = l$ and $\gamma \in \Gamma_{\boldsymbol{c}_l}$, which we have already considered. In the second case $\gamma \boldsymbol{c}_{l_2} = \overline{\boldsymbol{c}_{l_1}}$, the positive orientation condition implicit in the definition of the $z_{\boldsymbol{c}}$ implies that these terms give no contribution.

For all other terms we can use the double coset decomposition to unfold the integral over s, and their total contribution is

$$\begin{aligned}
&\sum_{l_1,l_2=1}^{h} \sum_{\substack{\gamma \in \Gamma/\Gamma_{\boldsymbol{c}_{l_2}} \\ \gamma \notin \Gamma_{\boldsymbol{c}_{l_2}}}} \int_{\alpha_1}^{\infty} \int_1^{\varepsilon_{l_1}^2} \chi_{S_B}(a(v)k(\tfrac{\pi}{2})a(s)g_{l_1}^{-1}\gamma g_{l_2}) \, \frac{\mathrm{d}s}{s} \, \frac{\mathrm{d}v}{v^2} \\
&= \sum_{l_1,l_2=1}^{h} \sum_{\substack{\gamma \in \Gamma_{\boldsymbol{c}_{l_1}} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}} \\ \gamma \boldsymbol{c}_{l_2} \neq \boldsymbol{c}_{l_1}, \overline{\boldsymbol{c}_{l_1}}}} \int_{\alpha_1}^{\infty} \int_0^{\infty} \chi_{S_B}(a(v)k(\tfrac{\pi}{2})a(s)g_{l_1}^{-1}\gamma g_{l_2}) \, \frac{\mathrm{d}s}{s} \, \frac{\mathrm{d}v}{v^2} \qquad (4.17) \\
&= \sum_{l_1,l_2=1}^{h} \sum_{\substack{\gamma \in \Gamma_{\boldsymbol{c}_{l_1}} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}} \\ \gamma \boldsymbol{c}_{l_2} \neq \boldsymbol{c}_{l_1}, \overline{\boldsymbol{c}_{l_1}}}} \int_0^{\infty} \int_{\alpha_1}^{\infty} \chi_I(\mathrm{Re}(z_{g^{-1}\gamma \boldsymbol{c}_{l_2}})) \chi_{[\alpha_2,\infty)}(\mathrm{Im}(z_{g^{-1}\gamma \boldsymbol{c}_{l_2}})) \, \frac{\mathrm{d}v}{v^2} \, \frac{\mathrm{d}s}{s},
\end{aligned}$$

where $g = g_{l_1} a(s) k(-\frac{\pi}{2}) a(v)^{-1}$. We now replace $s \leftarrow s s_\gamma^{-1}$ with s_γ chosen so that the backward endpoint of the geodesic $a(s_\gamma) g_{l_1}^{-1} \gamma \boldsymbol{c}_{l_2}$ is ± 1, the $\pm$ determined by whether the image of 0 under $g_{l_1}^{-1}\gamma g_{l_2}$ is positive or negative. The forward endpoint of the geodesic $a(s_\gamma) g_{l_1}^{-1} \gamma \boldsymbol{c}_{l_2}$ is then $\pm r$, where $r = r(\gamma, l_1, l_2)$ is the cross ratio of the points $\boldsymbol{c}_{l_1}^-, \boldsymbol{c}_{l_1}^+, (\gamma \boldsymbol{c}_{l_2})^+, (\gamma \boldsymbol{c}_{l_2})^-$, where we use $\boldsymbol{c}^\pm$ to denote the forward and backward endpoints of a geodesic $\boldsymbol{c}$. We note that r clearly does not depend on the choice of representative of the double coset $\Gamma_{\boldsymbol{c}_{l_1}} \gamma \Gamma_{\boldsymbol{c}_{l_2}}$, but that on the other hand $r(\gamma_1) = r(\gamma_2)$ only implies $G_{\boldsymbol{c}_{l_1}} \gamma_1 G_{\boldsymbol{c}_{l_2}} = G_{\boldsymbol{c}_{l_1}} \gamma_2 G_{\boldsymbol{c}_{l_2}}$, not necessarily $\Gamma_{\boldsymbol{c}_{l_1}} \gamma_1 \Gamma_{\boldsymbol{c}_{l_2}} = \Gamma_{\boldsymbol{c}_{l_1}} \gamma_2 \Gamma_{\boldsymbol{c}_{l_2}}$. Here $G_{\boldsymbol{c}}$ is the stabilizer in G of a geodesic $\boldsymbol{c}$.

With this change in s, we now have for $g = g_{l_1} a(s) k(-\frac{\pi}{2}) a(v)^{-1}$,

$$\mathrm{Re}(z_{g^{-1}\boldsymbol{c}_{l_2}}) = \frac{v}{2} \left(\frac{-s \pm r}{s \pm r} - \frac{-s \pm 1}{s \pm 1} \right),$$

and

$$\mathrm{Im}(z_{g^{-1}\boldsymbol{c}_{l_2}}) = \frac{v}{2} \left(\frac{-s \pm r}{s \pm r} + \frac{-s \pm 1}{s \pm 1} \right).$$

We change variables

$$s \leftarrow \pm\frac{-s+1}{s+1}, \ \frac{\mathrm{d}s}{s} \leftarrow 2\frac{\mathrm{d}s}{|s^2-1|},$$

and the range of integration of s is $\int_{-1}^{1}$ in the $+$ case and $\int_{-\infty}^{-1}+\int_{1}^{\infty}$ in the $-$ case. We have in both cases that

$$\begin{aligned}\mathrm{Re}(z_{g^{-1}\gamma \boldsymbol{c}_{l_2}}) &= \frac{v}{2}\frac{(r+1)s+(r-1)}{(r-1)s+(r+1)}+s = \frac{v}{2}\frac{s^2+2qs+1}{s+q},\\ \mathrm{Im}(z_{g^{-1}\gamma \boldsymbol{c}_{l_2}}) &= \frac{v}{2}\frac{(r+1)s+(r-1)}{(r-1)s+(r+1)}-s = \frac{v}{2}\frac{1-s^2}{s+q},\end{aligned}$$

where $q=\frac{r+1}{r-1}$.

The integral in (4.17) becomes

$$2\int_{-1}^{1}\int_{\alpha_1}^{\infty}\chi_I\left(\frac{v}{2}\frac{s^2+2qs+1}{s+q}\right)\chi_{[\alpha_2,\infty)}\left(\frac{v}{2}\frac{1-s^2}{s+q}\right)\frac{\mathrm{d}v}{v^2}\frac{\mathrm{d}s}{1-s^2}$$

in the $+$ case and

$$2\left(\int_{-\infty}^{-1}+\int_{1}^{\infty}\right)\int_{\alpha_1}^{\infty}\chi_I\left(\frac{v}{2}\frac{s^2+2qs+1}{s+q}\right)\chi_{[\alpha_2,\infty)}\left(\frac{v}{2}\frac{1-s^2}{s+q}\right)\frac{\mathrm{d}v}{v^2}\frac{\mathrm{d}s}{s^2-1}$$

in the $-$ case. We change variables

$$v \leftarrow 2\frac{s+q}{s^2+2qs+1}v, \ \frac{\mathrm{d}v}{v^2} \leftarrow \frac{1}{2}\left|\frac{s^2+2qs+1}{s+q}\right|\frac{\mathrm{d}v}{v^2}$$

to obtain

$$\int_I H_{\pm}(q, v\alpha_1^{-1}, v\alpha_2^{-1})\frac{\mathrm{d}v}{v^2},$$

with $H_{\pm}(q,v_1,v_2)$ as in (4.12) and (4.13) as claimed. Note here that, as we have only positive terms in the above rearrangements, the series

$$\sum_{l_1,l_2=1}^{h}\sum_{\substack{\gamma\in\Gamma_{\boldsymbol{c}_{l_1}}\backslash\Gamma/\Gamma_{\boldsymbol{c}_{l_2}}\\ \gamma\boldsymbol{c}_{l_2}\neq\boldsymbol{c}_{l_1},\overline{\boldsymbol{c}_{l_1}}}}\int_I H_{\mathrm{sign}(g_{l_1}^{-1}\gamma g_{l_2}(0))}(q(\gamma,l_1,l_2), v\alpha_1^{-1}, v\alpha_2^{-1})\,\frac{\mathrm{d}v}{v^2},$$

is still absolutely convergent. Thus for given α_1,α_2, the series

$$\sum_{l_1,l_2=1}^{h}\sum_{\substack{\gamma\in\Gamma_{\boldsymbol{c}_{l_1}}\backslash\Gamma/\Gamma_{\boldsymbol{c}_{l_2}}\\ \gamma\boldsymbol{c}_{l_2}\neq\boldsymbol{c}_{l_1},\overline{\boldsymbol{c}_{l_1}}}}\frac{1}{v^2}\,H_{\mathrm{sign}(g_{l_1}^{-1}\gamma g_{l_2}(0))}(q(\gamma,l_1,l_2), v\alpha_1^{-1}, v\alpha_2^{-1}), \tag{4.18}$$

converges absolutely for almost every v. This implies, by the monotonicity

$$H_{\pm}(q, v'\alpha_1^{-1}, v'\alpha_2^{-1}) \le H_{\pm}(q, v\alpha_1^{-1}, v\alpha_2^{-1}) \quad \text{if } 0<|v'|\le|v|,$$

that the series (4.18) converges for all v' uniformly on compacta not containing the origin. Lemmas 4.5 and 4.6 below establish the uniform convergence of (4.18) for v in a neighbourhood of 0, and so (4.18) is continuous as a function of v as each term in the series is continuous.

We now specialise to the case $\alpha_1 = \alpha_2 = 1$. We note that in the range of integration of $H_+(q,v)$, $1-s^2 \geq 0$, and so the condition

$$v\frac{1-s^2}{s^2+2qs+1} \geq 1 \tag{4.19}$$

forces v and $s^2+2qs+1$ to have the same sign, and so the condition

$$2v\frac{s+q}{s^2+2qs+1} \geq 1 \tag{4.20}$$

forces $s+q \geq 0$. Similarly, since $1-s^2 \leq 0$ in the range of integration of $H_-(q,v)$, v and $s^2+2qs+1$ have opposite signs, and so $s+q \leq 0$ in this case.

We split into three cases depending on whether $1 < q$, $-1 < q < 1$, or $q < -1$ (note that q is finite and not ± 1 since r, being a cross-ratio, is none of 0, 1, ∞). In the case when $q > 1$, $s+q \geq 0$ does not affect the range of integration for H_+, $\int_{-1}^{1}$, but $s+q \leq 0$ restricts the range of integration for H_- to $\int_{-\infty}^{-q}$. In the case when $-1 < q < 1$, the range of integration for H_+ is restricted to $\int_{-q}^{1}$, while that of H_- is restricted to $\int_{-\infty}^{-1}$. Finally, in the case when $q < -1$, the condition $s+q \geq 0$ forces $H_+ = 0$, while $s+q \leq 0$ restricts the range of integration for H_- to $\int_{-\infty}^{-1} + \int_{1}^{-q}$.

We note that if $|q| < 1$, then $s^2+2qs+1 > 0$ for all s, while if $|q| > 1$, $s^2+2qs+1$ has roots at $s = -q \pm \sqrt{q^2-1}$. We note that when $q < -1$, $0 < -q-\sqrt{q^2-1} < 1 < -q+\sqrt{q^2-1}$, while if $q > 1$, then $-q-\sqrt{q^2-1} < -1 < -q+\sqrt{q^2-1} < 0$. In the cases $q > 1$ and $q < -1$, we subdivide the range of integration to keep track of the sign of $s^2+2qs+1$. In the case $q > 1$ we write $H_+ = I_1^+ + I_2^+$ with I_1^+, I_2^+ having ranges of integration $\int_{-1}^{-q+\sqrt{q^2-1}}$, $\int_{-q+\sqrt{q^2-1}}^{1}$ respectively, and we write $H_- = I_1^- + I_2^-$ with I_1^-, I_2^- having ranges of integration $\int_{-\infty}^{-q-\sqrt{q^2-1}}$, $\int_{-q-\sqrt{q^2-1}}^{-q}$ respectively. In the case $q < -1$, we write $H_- = I_3 + I_4$ with I_3, I_4 having ranges of integration $\int_{-\infty}^{-1}$, $\int_{1}^{-q}$ respectively. Writing $I_5^\pm$ for $H_\pm(q,v)$ in the case $-1 < q < 1$, we thus have 8 integrals, $I_1^\pm, I_2^\pm, I_3, I_4, I_5^\pm$, to analyze.

We start by considering the functions

$$v_1(s) = \frac{s^2+2qs+1}{1-s^2}$$

and

$$v_2(s) = \frac{1}{2}\frac{s^2+2qs+1}{s+q},$$

which are clearly related to the conditions (4.19), (4.20). We note that $v_1(s)$ is increasing for all s if $q > 1$ and decreasing for all s if $q < -1$. When $0 < q < 1$, $v_1(s)$ is decreasing for $-\frac{1}{q} - \sqrt{\frac{1}{q^2}-1} < s < -\frac{1}{q} + \sqrt{\frac{1}{q^2}-1}$ and increasing otherwise, and when $-1 < q < 0$, $v_1(s)$ is increasing for $-\frac{1}{q} - \sqrt{\frac{1}{q^2}-1} < s < -\frac{1}{q} + \sqrt{\frac{1}{q^2}-1}$ and decreasing otherwise. As for $v_2(s)$, it is increasing for all s if $|q| > 1$, and when $|q| < 1$, $v_2(s)$ is decreasing for $-q-\sqrt{1-q^2} < s < -q+\sqrt{1-q^2}$ and is increasing otherwise.

We also define the functions

$$s_1(v) = \frac{-q+\sqrt{v^2+q^2-1}}{v+1}$$

and

$$s_2(v) = v - q - \sqrt{v^2+q^2-1}$$

for all v if $|q| > 1$ and $|v| > \sqrt{1-q^2}$ if $|q| < 1$. We note that $s_2(v)$ is the smaller solution to $v = v_2(s)$ and that $s_1(v)$ is the smaller solution to $v = v_1(s)$ when $v < -1$ and the larger solution when $v > -1$.

Finally, we define, for s in the range of the integrals $I_1^{\pm}, I_2^{\pm}, I_3, I_4, I_5^{\pm}$,

$$h_q(s) = \log \frac{s+q}{1-s^2},$$

so that

$$h_q'(s) = \frac{s^2+2qs+1}{(s+q)(1-s^2)}.$$

We note that $|h_q'(s)| = h_q'(s)$ in the range of integration for $I_2^+, I_1^-, I_3, I_5^{\pm}$ and $|h_q'(s)| = -h_q'(s)$ in the range of integration for I_1^+, I_2^-, I_4.

For I_1^+ we note that the integral is 0 unless $v < 0$, and the conditions (4.19), (4.20) become $v \le v_1(s)$, $v \le v_2(s)$. We observe that in this case $v_1(s) \le v_2(s)$ for all s in the range of integration, and thus

$$I_1^+ = \begin{cases} h_q(s_1(v)) - h_q(-q+\sqrt{q^2-1}) & \text{if } v < 0 \\ 0 & \text{if } v > 0. \end{cases}$$

Similarly, for I_2^+ we note that the integral is 0 unless $v > 0$, and the conditions (4.19), (4.20) become $v \ge v_1(s)$, $v \ge v_2(s)$. Here we find $v_1(s) \ge v_2(s)$ for all s in the range of integration and hence

$$I_2^+ = \begin{cases} 0 & \text{if } v < 0 \\ h_q(s_1(v)) - h_q(-q+\sqrt{q^2-1}) & \text{if } v > 0. \end{cases}$$

For I_1^-, we note that the integral is 0 unless $v < 0$, and the conditions (4.19), (4.20) become $v \le v_1(s)$ and $v \le v_2(s)$. In this case we have $v_2(s) \le v_1(s)$ for all s in the range of integration, and we find

$$I_1^- = \begin{cases} h_q(-q-\sqrt{q^2-1}) - h_q(s_2(v)) & \text{if } v < 0 \\ 0 & \text{if } v > 0. \end{cases}$$

Similarly, for I_2^- we note that the integral is 0 unless $v > 0$, and the condition (4.19), (4.20) become $v \ge v_1(s)$ and $v \ge v_2(s)$. Here we have $v_2(s) \ge v_1(s)$ for all s in the range of integration, and

$$I_2^- = \begin{cases} 0 & \text{if } v < 0 \\ h_q(-q-\sqrt{q^2-1}) - h_q(s_2(v)) & \text{if } v > 0. \end{cases}$$

For I_3 we note that the integral is 0 unless $v < 0$, and the conditions (4.19), (4.20) become $v \le v_1(s)$, $v \le v_2(s)$. In the range of integration, $s = -1 - \sqrt{2-2q}$ is the unique point at which $v_1(s) = v_2(s)$, and their value at this point is $-\sqrt{2-2q}$. Since $v_1(s)$ is decreasing and $v_2(s)$ is increasing, we obtain

$$I_3 = \begin{cases} h_q(s_1(v)) - h_q(s_2(v)) & \text{if } v < -\sqrt{2-2q} \\ 0 & \text{if } v > -\sqrt{2-2q}. \end{cases}$$

Similarly, for I_4 we note that the integral is 0 unless $v > 0$ and the conditions (4.19), (4.20) become $v \ge v_1(s)$, $v \ge v_2(s)$. In the range of integration, $s = -1 + \sqrt{2-2q}$ is the unique point at which $v_1(s) = v_2(s)$, and their value at this point is $\sqrt{2-2q}$. Since $v_1(s)$ is decreasing and $v_2(s)$ is increasing, we obtain

$$I_4 = \begin{cases} 0 & \text{if } v < \sqrt{2-2q} \\ h_q(s_1(v)) - h_q(s_2(v)) & \text{if } v > \sqrt{2-2q}. \end{cases}$$

For I_5^+ we note that the integral is 0 unless $v > 0$ and the conditions (4.19), (4.20) become $v \geq v_1(s)$, $v \geq v_2(s)$. In the range of integration, $s = -1 + \sqrt{2-2q}$ is the unique point at which $v_1(s) = v_2(s)$, and their value at this point is $\sqrt{2-2q}$. We find that $v_1(s)$ is increasing on $-1 + \sqrt{2-2q} < s < 1$ and that $v_2(s)$ is decreasing on $-q < s < -1 + \sqrt{2-2q}$, and so

$$I_5^+ = \begin{cases} 0 & \text{if } v < \sqrt{2-2q} \\ h_q(s_1(v)) - h_q(s_2(v)) & \text{if } v > \sqrt{2-2q}. \end{cases}$$

Similarly, for I_5^- we note that the integral is 0 unless $v < 0$ and the conditions (4.19), (4.20) become $v \leq v_1(s)$, $v \leq v_2(s)$. In the range of integration, $s = -1 - \sqrt{2-2q}$ is the unique point at which $v_1(s) = v_2(s)$, and their value at this point is $-\sqrt{2-2q}$. We find that $v_1(s)$ is decreasing on $-1 - \sqrt{2-2q} < s < -1$ and that $v_2(s)$ is increasing on $-\infty < s < -1 - \sqrt{2-2q}$, and so

$$I_5^- = \begin{cases} h_q(s_1(v)) - h_q(s_2(v)) & \text{if } v < -\sqrt{2-2q} \\ 0 & \text{if } v > -\sqrt{2-2q}. \end{cases}$$

Putting these calculations together, we obtain (4.14) and (4.15). □

We remark that the cross-ratio r in (4.11) is a classical quantity and $q = \frac{r+1}{r-1}$ appears very naturally when counting geodesics, see for example the results by Good [11] proving asymptotic formula for the number of $\gamma \in \Gamma_{\boldsymbol{c}_{l_1}} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}}$ with $q(\gamma, l_1, l_2) \leq Q$ as $Q \to \infty$. In what follows we need an upper bound for this count of the correct order of magnitude, and in the interest of being self-contained, we give a proof of this bound despite [11] providing much stronger information.

Lemma 4.5. *For $1 \leq l_1, l_2 \leq h$ and $Q > 1$, the number of $\gamma \in \Gamma_{\boldsymbol{c}_{l_1}} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}}$ such that $|q(\gamma, l_1, l_2)| \leq Q$ is $O(Q)$ with implied constant depending on Γ and the geodesics $\boldsymbol{c}_{l_1}$, $\boldsymbol{c}_{l_2}$.*

Proof. We recall from the properties of the cross-ratio $r = r(\gamma, l_1, l_2)$, (4.11), that there is $g \in G$ such that $g\boldsymbol{c}_{l_1}$ is the geodesic from 0 to ∞ and $g\gamma\boldsymbol{c}_{l_2}$ is the geodesic from ± 1 to $\pm r$, the sign being the same as that of $g_{l_1}^{-1}\gamma g_{l_2}(0)$. It follows that when $g_{l_1}^{-1}\gamma g_{l_2}(0) > 0$ there is $g \in G$ such that $g\boldsymbol{c}_{l_1}$ is the geodesic from -1 to 1 and $g\gamma\boldsymbol{c}_{l_2}$ is the geodesic from ∞ to $-q$, and when $g_{l_1}^{-1}\gamma g_{l_2}(0) < 0$ there is $g \in G$ such that $g\boldsymbol{c}_{l_1}$ is the geodesic from 1 to -1 and $g\gamma\boldsymbol{c}_{l_2}$ is the geodesic from ∞ to q. In either case we see that $\boldsymbol{c}_{l_1}$ intersects $\gamma\boldsymbol{c}_{l_2}$ if and only if $-1 < q < 1$, and when this occurs $|q| = |\cos\theta|$ with θ the intersection angle between $\boldsymbol{c}_{l_1}$ and $\gamma\boldsymbol{c}_{l_2}$. We note that there are only finitely many $q = q(\gamma, l_1, l_2)$ with $|q| < 1$. Indeed, these correspond to intersections between $\boldsymbol{c}_{l_1}$ and $\gamma\boldsymbol{c}_{l_2}$ in $\mathbb{H}$, or intersections between the geodesics $\boldsymbol{c}_{l_1}$ and $\boldsymbol{c}_{l_2}$ in $\Gamma\backslash\mathbb{H}$. As these distinct geodesics are closed, there can only be finitely many such intersections.

When $|q| > 1$ we find $g \in G$ such that $g\boldsymbol{c}_{l_1}$ is the geodesic from -1 to 1 (or 1 to -1) and $g\gamma\boldsymbol{c}_{l_2}$ is a geodesic connecting $\pm \log d$, d consequently being the distance between the geodesics. Expressing d in terms of q, we find that $|q| = \cosh d$.

Let us fix points $z_1 \in \boldsymbol{c}_{l_1}$ and $z_2 \in \boldsymbol{c}_{l_2}$. Now let $z_1' \in \boldsymbol{c}_{l_1}$ and $z_2' \in \boldsymbol{c}_{l_2}$ be such that the distance between $\boldsymbol{c}_{l_1}$ and $\gamma\boldsymbol{c}_{l_2}$, namely $\cosh^{-1}(|q|)$, is the hyperbolic distance between z_1' and $\gamma z_2'$. Let $\gamma_1 \in \Gamma_{\boldsymbol{c}_{l_1}}$ and $\gamma_2 \in \Gamma_{\boldsymbol{c}_{l_2}}$ be such that the distance between $\gamma_1 z_1$ and z_1' is at most $\log \varepsilon_{l_1}$ and the distance between $\gamma_2 z_2$ and z_2' is at most $\log \varepsilon_{l_2}$. Allowing the implied constant to depend on $\boldsymbol{c}_{l_1}, \boldsymbol{c}_{l_2}$, it follows that the distance between z_1 and $\gamma_1^{-1}\gamma\gamma_2 z_2$ is $\cosh^{-1}(|q|) + O(1)$.

Since $\gamma_1^{-1}\gamma\gamma_2 \in \Gamma_{\boldsymbol{c}_{l_1}}\gamma\Gamma_{\boldsymbol{c}_{l_2}}$, the above shows that

$$\#\{\gamma \in \Gamma_{\boldsymbol{c}_1} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}} : 1 < |q(\gamma, l_1, l_2)| \leq Q\} \leq \#\{\gamma \in \Gamma : \mathrm{d}_{\mathbb{H}}(z_1, \gamma z_2) \leq \cosh^{-1}(Q) + O(1)\}$$

where $\mathrm{d}_{\mathbb{H}}$ is the hyperbolic metric. As the set of $z \in \mathbb{H}$ such that $\mathrm{d}_{\mathbb{H}}(z_1, z) \leq \cosh^{-1}(Q) + O(1)$ has area $O(Q)$, the lemma follows. □

Lemma 4.6. *For $q = q(\gamma, l_1, l_2)$ and $\frac{v}{\alpha_1}$, $\frac{v}{\alpha_2}$ sufficiently small, we have*

$$H_\pm(q, \alpha_1^{-1}v, \alpha_2^{-1}v) = \begin{cases} 0 & \text{if } q < 1 \\ O(\frac{v^2}{q^2}) & \text{if } q > 1, \end{cases} \tag{4.21}$$

where the implied constants depend only on the group Γ and the geodesics $\boldsymbol{c}_l$. In particular (4.18) converges absolutely and uniformly for v near 0.

Moreover, we have

$$H_\pm(q, v, v) = \begin{cases} 0 & \text{if } q < 1 \\ \frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2 + O(\frac{v^3}{q^2}) & \text{if } q > 1, \end{cases} \tag{4.22}$$

and thus

$$W_{1,1}(v) = \frac{1}{\ell}\sum_{l_1,l_1=1}^{h} \sum_{\substack{q=q(\gamma,l_1,l_2)>1 \\ \gamma\in\Gamma_{\boldsymbol{c}_{l_1}}\backslash\Gamma/\Gamma_{\boldsymbol{c}_{l_2}} \\ \gamma\boldsymbol{c}_{l_2}\neq\boldsymbol{c}_{l_1},\overline{\boldsymbol{c}_{l_1}}}} \frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}} + O(v).$$

Proof. For the proof of this lemma we break into the same cases as in the proof of proposition 4.4, namely the eight integrals $I_1^\pm, I_2^\pm, I_3, I_4, I_5^\pm$. We recall that

$$\begin{aligned}
H_+ &= I_1^+ + I_2^+ = \int_{-1}^{-q+\sqrt{q^2-1}} + \int_{-q+\sqrt{q^2-1}}^{1} \quad \text{when } q > 1, \\
H_- &= I_1^- + I_2^- = \int_{-\infty}^{-q-\sqrt{q^2-1}} + \int_{-q-\sqrt{q^2-1}}^{-q} \quad \text{when } q > 1, \\
H_+ &= 0 \quad \text{when } q < -1, \\
H_- &= I_3 + I_4 = \int_{-\infty}^{-1} + \int_{1}^{-q} \quad \text{when } q < -1, \\
H_+ &= I_5^+ = \int_{-q}^{1} \quad \text{when } -1 < q < 1, \\
H_- &= I_5^+ = \int_{-\infty}^{-1} \quad \text{when } -1 < q < 1.
\end{aligned}$$

We recall that the integrand of all the above integrals is $\chi_1(s,v)\chi_2(s,v)$ where $\chi_1(s,v)$ is the indicator function of the condition

$$2v\frac{s+q}{s^2+2qs+1} \geq \alpha_1 \tag{4.23}$$

and $\chi_2(s,v)$ is the indicator function of the condition

$$v\frac{1-s^2}{s^2+2qs+1} \geq \alpha_2. \tag{4.24}$$

We first show that for v sufficiently small, $I_3 = I_4 = I_5^\pm = 0$, thus verifying the first case in (4.21).

We have that $I_3 = 0$ if $v \geq 0$ and the conditions (4.23), (4.24) become $v \leq \alpha_1 v_2(s)$ and $v \leq \alpha_2 v_1(s)$. We note however that $v_1(s) \leq 1$ for all $s < -1$, and so we conclude that $I_3 = 0$ for all $\frac{v}{\alpha_2}$ sufficiently small. Similarly, we have that $I_4 = 0$ if $v \leq 0$ and the conditions (4.23), (4.24) become $v \geq \alpha_1 v_2(s)$ and $v \geq \alpha_2 v_1(s)$. We note however that $v_2(s) \geq 1$ for all $1 < s < -q$, and so $I_4 = 0$ for all $\frac{v}{\alpha_1}$ sufficiently small.

We have that $I_5^+ = 0$ if $v \leq 0$ and the conditions (4.23), (4.24) become $v \geq v_2(s)\alpha_1$ and $v \geq \alpha_2 v_1(s)$. We note that in this case we have $v_2(s) \geq \sqrt{1-q^2}$, and since by lemma 4.5 there are only finitely many $-1 < q < 1$, it follows that $I_5^+ = 0$ for all $\frac{v}{\alpha_1}$ sufficiently small. Similarly, we have that $I_5^- = 0$ if $v \geq 0$ and the conditions (4.23), (4.24) become $v \leq v_2(s)\alpha_1$ and $v \leq v_1(s)\alpha_2$. We note that $v_2(s) \leq -\sqrt{1-q^2}$, so as before we have $I_5^+ = 0$ for $\frac{v}{\alpha_1}$ sufficiently small.

It remains to consider the cases $I_1^\pm$ and $I_2^\pm$, i.e. when $q > 1$. We have that $I_1^+ = 0$ if $v \geq 0$ and the conditions (4.23), (4.24) become $v \leq \alpha_1 v_2(s)$ and $v \leq \alpha_2 v_1(s)$. To upper bound I_1^+ we only consider the condition $v \leq \alpha_2 v_1(s)$ and we obtain

$$I_1^+ \leq h_q(s_1(\alpha_2^{-1}v)) - h_q(-q + \sqrt{q^2-1}). \tag{4.25}$$

Similarly, we have that $I_2^+ = 0$ if $v \leq 0$ and the conditions (4.23), (4.24) become $v \leq \alpha_1 v_2(s)$ and $v \leq \alpha_2 v_1(s)$. By dropping the first condition we obtain the same bound as before,

$$I_2^+ \leq h_q(s_1(\alpha_2^{-1}v)) - h_q(-q + \sqrt{q^2-1}). \tag{4.26}$$

We have that $I_1^- = 0$ if $v \geq 0$ and the conditions (4.23), (4.24) become $v \leq \alpha_1 v_2(s)$ and $v \leq \alpha_2 v_1(s)$. Dropping the second condition we obtain the upper bound

$$I_1^- \leq h_q(-q - \sqrt{q^2-1}) - h_q(s_2(\alpha_1^{-1}v)). \tag{4.27}$$

Similarly, we have that $I_2^- = 0$ if $v \leq 0$ and the conditions (4.23), (4.24) become $v \geq \alpha_1 s_2(v)$ and $v \geq \alpha_2 s_1(v)$. Dropping the second condition we obtain the same bound

$$I_2^- \leq h_q(-q - \sqrt{q^2-1}) - h_q(s_2(\alpha_1^{-1}v)). \tag{4.28}$$

We remark that, by analysing when $v_1(s) \leq v_2(s)$, $v_1(s) \geq v_2(s)$ for all s in the range of integration of $I_1^\pm, I_2^\pm$, these upper bounds are in fact equalities for $q > \frac{\alpha_1^2+\alpha_2^2}{2\alpha_1\alpha_2}$. By lemma 4.5 this condition is satisfied for all but finitely many double cosets.

We now compute the upper bounds (4.25), (4.26) and (4.27), (4.28) asympotically as $v \to 0$. This will not only verify (4.21) but also (4.22) since these bounds are equalities for $\alpha_1 = \alpha_2 = 1$. We have

$$\begin{aligned} &h_q(s_1(v)) - h_q(-q + \sqrt{q^2-1}) \\ &\qquad = \log(v+1) + \log\frac{qv + \sqrt{v^2+q^2-1}}{\sqrt{q^2-1}} - \log\frac{v + q\sqrt{v^2+q^2-1} - q^2 + 1}{q\sqrt{q^2-1} - q^2 + 1} \end{aligned} \tag{4.29}$$

for $q > 1$ and v sufficiently small. Using that $q \geq 1 + \eta$ for some $\eta > 0$ depending only on Γ and the $\boldsymbol{c}_l$ by lemma (4.5), we expand

$$\sqrt{v^2+q^2-1} = \sqrt{q^2-1}\left(1 + \frac{v^2}{2(q^2-1)} + O(\frac{v^4}{q^4})\right). \tag{4.30}$$

We find that the first and second term in (4.29) is

$$\begin{aligned} &\log(v+1) + \log\left(1 + \frac{qv}{\sqrt{q^2-1}} + \frac{v^2}{2(q^2-1)} + O(\frac{v^4}{q^4})\right) \\ &\qquad = \log\left(1 + \left(1 + \frac{q}{\sqrt{q^2-1}}\right)v + \left(\frac{1}{2(q^2-1)} + \frac{q}{\sqrt{q^2-1}}\right)v^2\right) + O(\frac{v^3}{q^2}). \end{aligned}$$

We also find that the third term is

$$\log\left(1+\frac{v}{q\sqrt{q^2-1}-q^2+1}+\frac{qv^2}{2\sqrt{q^2-1}(q\sqrt{q^2-1}-q^2+1)}+O(\frac{v^4}{q^2})\right)$$
$$=\log\left(1+\left(1+\frac{q}{\sqrt{q^2-1}}\right)v+\frac{1}{2}\left(\frac{q}{\sqrt{q^2-1}}+\frac{q^2}{q^2-1}\right)v^2\right)+O(\frac{v^4}{q^2}).$$

We now have

$$h_q(s_1(v))-h_q(-q+\sqrt{q^2-1})=\log\frac{1+\left(1+\frac{q}{\sqrt{q^2-1}}\right)v+\left(\frac{1}{2(q^2-1)}+\frac{q}{\sqrt{q^2-1}}\right)v^2}{1+\left(1+\frac{q}{\sqrt{q^2-1}}\right)v+\frac{1}{2}\left(\frac{q}{\sqrt{q^2-1}}+\frac{q^2}{q^2-1}\right)v^2}+O(\frac{v^3}{q^2}).$$

Writing

$$\frac{1+\left(1+\frac{q}{\sqrt{q^2-1}}\right)v+\left(\frac{1}{2(q^2-1)}+\frac{q}{\sqrt{q^2-1}}\right)v^2}{1+\left(1+\frac{q}{\sqrt{q^2-1}}\right)v+\frac{1}{2}\left(\frac{q}{\sqrt{q^2-1}}+\frac{q^2}{q^2-1}\right)v^2}$$
$$=1+\frac{\frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2}{1+\left(1+\frac{q}{\sqrt{q^2-1}}\right)v+\frac{1}{2}\left(\frac{q}{\sqrt{q^2-1}}+\frac{q^2}{q^2-1}\right)v^2}=1+\frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2+O(\frac{v^3}{q^2}),$$

we obtain

$$h_q(s_1(v))-h_q(-q+\sqrt{q^2-1})=\frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2+O(\frac{v^3}{q^2}).$$

This establishes (4.21) and (4.22) in the + case.

It remains to analyze

$$h_q(-q-\sqrt{q^2-1})-h_q(s_2(v))$$
$$=-\log\frac{-v+\sqrt{v^2+q^2-1}}{\sqrt{q^2-1}}+\log\frac{q^2-1-qv+v^2+(q-v)\sqrt{v^2+q^2-1}}{q^2-1+q\sqrt{q^2-1}}.$$

Using (4.30), the first term here is

$$\log\left(1-\frac{v}{\sqrt{q^2-1}}+\frac{v^2}{2(q^2-1)}\right)+O(\frac{v^4}{q^4}),$$

and the second term is

$$\log\left(1-\frac{v}{\sqrt{q^2-1}}+(\frac{q-\sqrt{q^2-1}}{\sqrt{q^2-1}}+\frac{q^2-q\sqrt{q^2-1}}{2(q^2-1)})v^2\right)+O(\frac{v^3}{q^2}).$$

We now have

$$h_q(-q-\sqrt{q^2-1})-h_q(s_1(v))=\log\frac{1-\frac{v}{\sqrt{q^2-1}}+(\frac{q-\sqrt{q^2-1}}{\sqrt{q^2-1}}+\frac{q^2-q\sqrt{q^2-1}}{2(q^2-1)})v^2}{1-\frac{v}{\sqrt{q^2-1}}+\frac{v^2}{2(q^2-1)}}+O(\frac{v^3}{q^2}).$$

Writing

$$\frac{1-\frac{v}{\sqrt{q^2-1}}+(\frac{q-\sqrt{q^2-1}}{\sqrt{q^2-1}}+\frac{q^2-q\sqrt{q^2-1}}{2(q^2-1)})v^2}{1-\frac{v}{\sqrt{q^2-1}}+\frac{v^2}{2(q^2-1)}}$$

$$=1+\frac{\frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2}{1-\frac{v}{\sqrt{q^2-1}}+\frac{v^2}{2(q^2-1)}}=1+\frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2+O(\frac{v^3}{q^2}),$$

it follows that

$$h_q(-q-\sqrt{q^2-1})-h_q(s_1(v))=\frac{q-\sqrt{q^2-1}}{2\sqrt{q^2-1}}v^2+O(\frac{v^3}{q^2}).$$

This is enough to establish (4.21) and (4.22). □

4.3 Moments

Let us return to the moments of the counting functions $\mathcal{N}_{I,\alpha,\beta}(\xi,y)$, now averaging $\xi\in X_{\alpha_0,\beta_0}(y)$ over the sequence itself rather than the absolutely continuous measure λ. For $y>0$, $\boldsymbol{s}=(s_1,\dots,s_r)\in\mathbb{C}^r$, $\mathcal{I}=I_1\times\cdots\times I_r$, $I_j\subset\mathbb{R}$ a finite interval so that $0\notin\partial I_j$, and $\boldsymbol{\alpha}=(\alpha_0,\dots,\alpha_r)$, $\boldsymbol{\beta}=(\beta_0,\dots,\beta_r)$ satisfying $1\le\alpha_j<\beta_j\le\infty$, define the moments

$$M^{a,b}_{\mathcal{I},\boldsymbol{\beta},\boldsymbol{\alpha}}(\boldsymbol{\kappa},y)=y\sum_{\xi\in X_{\alpha_0,\beta_0}(y)\cap([a,b)+\mathbb{Z})}\prod_{1\le j\le r}\mathcal{N}_{I_j,\alpha_j,\beta_j}(\xi,y)^{\kappa_j},$$

$$M^{0}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa})=\int_{\Gamma\backslash G}\prod_{1\le j\le r}\mathcal{N}_{I_j,\alpha_j,\beta_j}(g)^{\kappa_j}\mathrm{d}\widehat{\nu}_{\alpha_0,\beta_0}(g);$$

and the moment generating functions

$$G^{a,b}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s},y)=y\sum_{\xi\in X_{\alpha_0,\beta_0}(y)\cap([a,b)+\mathbb{Z})}\exp\left(\sum_{1\le j\le r}s_j\mathcal{N}_{I_j,\alpha_j,\beta_j}(x,y)\right),$$

$$G^{0}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s})=\int_{\Gamma\backslash G}\exp\left(\sum_{1\le j\le r}s_j\mathcal{N}_{I_j,\alpha_j,\beta_j}(g)\right)\mathrm{d}\widehat{\nu}_{\alpha_0,\beta_0}(g).$$

Theorem 4.7. *Let $0\le a<b\le 1$ and $\mathcal{I}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ as above. Then there is a constant $c_0>0$ such that for $\boldsymbol{s}=(s_1,\dots,s_r)\in\mathbb{C}^r$ satisfying $\sum_j\max\{\mathrm{Re}(s_j),0\}<c_0$ the function $G^0_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s})$ is analytic and we have*

$$\lim_{y\to 0}G^{a,b}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s},y)=\kappa_\Gamma(b-a)\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)G^0_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{s}).$$

As in section 2.2, theorem 4.7 implies the convergence as $y\to 0$ of the mixed moments.

Corollary 4.8. *If $\boldsymbol{\kappa}=(\kappa_1,\dots,\kappa_r)\in\mathbb{R}^r$ with $\kappa_j\ge 0$, then $M^0_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa})$ is finite and*

$$\lim_{y\to 0}M^{a,b}_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa},y)=\kappa_\Gamma(b-a)\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)M^0_{\mathcal{I},\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{\kappa}).$$

We highlight the special case of the first moment $r=1$ and $\kappa=1$.

Corollary 4.9. *For any finite interval I such that $0 \notin \partial I$, $1 \leq \alpha < \beta \leq \infty$ and $0 \leq a < b \leq 1$, we have*

$$\lim_{y \to 0} y \sum_{\xi \in X_{\alpha_0,\beta_0}(y) \cap ([a,b)+\mathbb{Z})} \mathcal{N}_{I,\alpha,\beta}(\xi, y) = \kappa_\Gamma (b-a)\left(\frac{1}{\alpha} - \frac{1}{\beta}\right) \int_{\Gamma \backslash G} \mathcal{N}_{I,\alpha,\beta}(g) \mathrm{d}\widehat{\nu}_{1,\infty}(g).$$

The proof of theorem 4.7 follows from the same argument as in the continuous case, with lemma 2.9 replaced by lemma 4.3.

4.4 Palm distribution and pair correlation

We now return to the setting of a one-dimensional point process discussed in section 2.3. Instead of

$$\Xi_{N,\lambda} = \sum_{j=1}^{N} \sum_{k \in \mathbb{Z}} \delta_{N(\xi_j - \xi + k)}$$

with λ distributed at random with respect to λ, we consider the point process

$$\Xi_N^0 = \sum_{j=1}^{N} \sum_{k \in \mathbb{Z}} \delta_{N(\xi_j - \xi_p + k)}$$

where p is a random integer drawn from $\{1, 2, \ldots, N\}$ with uniform probability. Note that the expectation measure is given by

$$\mathbb{E}\Xi_N^0 = \frac{1}{N} \sum_{j,p=1}^{N} \sum_{k \in \mathbb{Z}} \delta_{N(\xi_j - \xi_p + k)} = \delta_0 + R_{2,N},$$

with the pair correlation measure

$$R_{2,N} = \frac{1}{N} \sum_{\substack{j,p=1 \\ j \neq p}}^{N} \sum_{k \in \mathbb{Z}} \delta_{N(\xi_j - \xi_p + k)}.$$

We furthermore define the point process process Ξ^0 by

$$\Xi^0(I) = \Theta_{1,\infty}^0(B_{\kappa_\Gamma^{-1} I, 1, \infty})$$

with κ_Γ as in (2.6). The process Ξ^0 is in fact distributed according to the Palm measure of the process Ξ; see section 31 of [16] for a general discussion of Palm measures, or [21] for a context more closely analogous to ours. The intensity measure of the Palm measure is in view of proposition 4.4 given by

$$\mathbb{E}\Xi^0(I) = \mathbb{E}\Theta_{1,\infty}^0(B_{\kappa_\Gamma^{-1} I, 1, \infty}) = \delta_0(I) + \int_I w(v) \mathrm{d}v, \tag{4.31}$$

where $w(v) = \kappa_\Gamma^{-1} W(\kappa_\Gamma^{-1} v)$, and so

$$w(v) = \frac{1}{v^2 \mathrm{vol}_{\mathbb{H}}(\Gamma \backslash \mathbb{H})} \sum_{l_1, l_2 = 1}^{h} \sum_{\substack{\gamma \in \Gamma_{\boldsymbol{c}_{l_1}} \backslash \Gamma / \Gamma_{\boldsymbol{c}_{l_2}} \\ \gamma \boldsymbol{c}_{l_2} \neq \boldsymbol{c}_{l_1}, \overline{\boldsymbol{c}_{l_1}}}} H_{\mathrm{sign}(g_{l_1}^{-1} \gamma g_{l_2}(0))}(q(\gamma, l_1, l_2), \kappa_\Gamma^{-1} v). \tag{4.32}$$

This shows that the following statements are special cases of theorem 4.1 (convergence in distribution) and corollary 4.9 (pair correlation). As in the proof of theorem 2.10, the key observation is that we can move from the scaling by y^{-1} to the scaling by N by the regularity of the limiting distributions.

Theorem 4.10. *The point processes $\Xi_N^0 \to \Xi^0$ converge in distribution, together with all moments, as $N \to \infty$.*

Specifically, for all $k_1, \ldots, k_r \in \mathbb{Z}_{\geq 0}$ and finite intervals I_i, we have that

$$\lim_{N\to\infty} \frac{1}{N} \#\big(\{p \leq N : \mathcal{N}_{I_i}(\xi_p, N) = k_i \ \forall i\}\big) = \mathbb{P}\big(\Xi^0(I_i) = k_i \ \forall i\big),$$

and for every finite interval I

$$\lim_{N\to\infty} R_{2,N}(I) = \int_I w(v)\mathrm{d}v.$$

The translation-stationarity of Ξ implies cycle-stationarity of the Palm distributed Ξ^0. Furthermore the fact that $\Xi(\{0\}) = 1$ almost surely (this follows from (4.31)) implies that Ξ^0 and hence Ξ are simple. For further details on this connection see [21] and references therein. We will explore a dynamical interpretation of this in the next section.

4.5 Entry and return times

Athreya and Cheung [1] constructed a Poincaré section (different from $\widetilde{S}_{\alpha,\beta}$) for the horocycle flow on $\mathrm{SL}(2,\mathbb{Z})\backslash G$, so that the return times yield the gaps in the Farey sequence. In fact, the return times for the horocycle flow for our section $\widetilde{S}_{\alpha,\beta}$ generate the gaps between the roots of quadratic congruences when $D > 0$. To see this, define for initial condition $g \in G$ the *first entry time* for the section $\widetilde{S}_{\alpha,\beta}$ by the Γ-invariant function

$$\tau_{\alpha,\beta}(g) = \inf\{u > 0 : \Gamma g n(u) \in \widetilde{S}_{\alpha,\beta}\}.$$

For $\Gamma g \in \widetilde{S}_{\alpha,\beta}$, this is also called the *first return time.* In view of the discussion in section 3.1, $\tau_{\alpha,\beta}$ is a measurable function. It is a general fact, see [21], that the ergodicity of the horocycle flow $\Gamma g \mapsto \Gamma g n(u)$ implies that

$$\int_{\Gamma\backslash G} f(g)\mathrm{d}\mu_\Gamma(g) = \frac{1}{\overline{\tau}_{\alpha,\beta}} \int_{\Gamma\backslash G} \bigg(\int_0^{\tau_{\alpha,\beta}(g)} f(gn(u))\,\mathrm{d}u \bigg) \mathrm{d}\nu_{\alpha,\beta}(g),$$

with the average return time (mean free path length)

$$\overline{\tau}_{\alpha,\beta} = \int_{\Gamma\backslash G} \tau_{\alpha,\beta}(g)\mathrm{d}\nu_{\alpha,\beta}(g).$$

Let us define the jth entry time recursively by $\tau_{\alpha,\beta}^{(j)}(g) = \tau_{\alpha,\beta}(gn(\tau_{\alpha,\beta}^{(j-1)}(g)))$ with $\tau_{\alpha,\beta}^{(0)}(g) = 0$, and $j \in \mathbb{Z}$ (i.e., backward and forward in time). For g random, the point process

$$\sum_{j\in\mathbb{Z}} \delta_{\tau_j}, \qquad \tau_j := \frac{\tau_{\alpha,\beta}^{(j)}(g)}{\overline{\tau}_{\alpha,\beta}},$$

then has, by construction, the same distribution as the process Ξ or Ξ^0, if g is a random element in $\Gamma\backslash G$ distributed according to μ_Γ or $\widehat{\nu}_{\alpha,\beta}$, respectively. In summary, the random processes generated by the real parts of the tops of our geodesic line processes are the same as the entry times of the horocycle flow to the Poincaré section $\widetilde{S}_{\alpha,\beta}$.

5 Geodesics and roots of quadratic congruences

The aim of this section is to express the roots $\frac{\mu}{m}$ in terms of the tops of certain geodesics in the Poincaré upper half-plane $\mathbb{H}$. The key result is the following theorem 5.1. It shows that the results of sections 2 and 4 yield the convergence of all fine-scale statistics for the roots of quadratic congruences, including moments, gap distributions and pair correlation measures. In particular, the limit theorems 2.10 and 4.10 for the one-dimensional point processes, turn into our main theorems 1.1 and 1.2 in the introduction for the special choice $n = 1$.

Theorem 5.1. *Let $D > 0$ be square-free and $D \not\equiv 1 \bmod 4$. Fix $n > 0$ and $\nu \bmod n$ such that $\nu^2 \equiv D \bmod n$. Then there exists a finite set of geodesics $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_h\}$, all of which are closed and have the same length in $\Gamma_0(n)\backslash\mathbb{H}$, with the following properties:*

(i) *For any integer $m > 0$ and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$, $m \equiv 0 \bmod n$, and $\mu \equiv \nu \bmod n$, there is a unique l and double coset $\Gamma_\infty\gamma\Gamma_{\boldsymbol{c}_l} \in \Gamma_\infty\backslash\Gamma_0(n)/\Gamma_{\boldsymbol{c}_l}$ such that*

$$z_{\gamma\boldsymbol{c}_l} \equiv \frac{\mu}{m} + \mathrm{i}\frac{\sqrt{D}}{m} \bmod \Gamma_\infty. \tag{5.1}$$

(ii) *Conversely, given l and double coset $\Gamma_\infty\gamma\Gamma_{\boldsymbol{c}_l} \in \Gamma_\infty\backslash\Gamma_0(n)/\Gamma_{\boldsymbol{c}_l}$ with $\gamma\boldsymbol{c}_l$ positively oriented, there exist unique $m > 0$ and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$, $m \equiv 0 \bmod n$, and $\mu \equiv \nu \bmod n$ such that (5.1) holds.*

We will see below that the number h of the closed geodesics is equal to the narrow class number $h^+(D)$ of the quadratic number field $\mathbb{Q}(\sqrt{D})$. The narrow class group is defined as the quotient of the group of fractional ideals in $\mathbb{Q}(\sqrt{D})$ modulo the group of fractional ideals of the form $\xi\mathbb{Z}[\sqrt{D}]$ with $\xi \in \mathbb{Q}(\sqrt{D})$ totally positive; here $\mathbb{Z}[\sqrt{D}]$ is a maximal order by our assumption on D. Moreover, the stabiliser subgroups $\Gamma_{\boldsymbol{c}_l}$ are all conjugate in G to the (projectively) cyclic subgroup generated by $\pm\begin{pmatrix}\varepsilon_0 & 0\\ 0 & \varepsilon_0^{-1}\end{pmatrix}$, where ε_0 is the generator of the totally positive units of $\mathbb{Z}[\sqrt{D}]$, chosen so that $\varepsilon_0 > 1$.

The remainder of this section is dedicated to the proof of theorem 5.1. It is based on the amalgamation of two well-known correspondences. The first is that given an integral binary quadratic form $F(X, Y)$ with positive discriminant, the geodesic in $\mathbb{H}$ connecting the two roots in $\mathbb{R}$ of $F(X, 1)$ projects to a closed geodesic in $\mathrm{SL}(2, \mathbb{Z})\backslash\mathbb{H}$. The second is that if $F(X, Y) = mX^2 - 2\mu XY + cY^2$ has discriminant $4D$, then $\mu^2 \equiv D \bmod m$, so then we see that the sum of the roots of $F(X, 1)$ is $2\frac{\mu}{m}$, the difference of the roots is $2\frac{\sqrt{D}}{m}$, and thus the top of the geodesic corresponding to $F(X, Y)$ is $\frac{\mu}{m} + \mathrm{i}\frac{\sqrt{D}}{m}$. In the remainder of this section we re-establish these connections in our own notation, but using ideals in $\mathbb{Z}[\sqrt{D}]$ instead of binary quadratic forms. We then examine them further in the case when $\mathrm{SL}(2, \mathbb{Z})$ is replaced by the congruence subgroup $\Gamma_0(n)$.

5.1 Roots, ideals and binary quadratic forms

We start with the following proposition, connecting the roots of the congruence $\mu^2 \equiv D \bmod m$ with ideals in $\mathbb{Z}[\sqrt{D}]$.

Proposition 5.2. *Let $D \in \mathbb{Z}$, $D \neq 1$, be square-free and $I \subset \mathbb{Z}[\sqrt{D}]$ be an ideal not divisible by any rational integers greater than 1. Then there exists a $\mathbb{Z}$-basis of I, $\{\beta_1, \beta_2\}$, unique modulo the action of Γ_∞, with the form*

$$\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix} = \begin{pmatrix}1 & \mu\\ 0 & m\end{pmatrix}\begin{pmatrix}\sqrt{D}\\ 1\end{pmatrix}, \tag{5.2}$$

where $m > 0$ and $\mu^2 \equiv D \bmod m$. Conversely, given m and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$, the sublattice of $\mathbb{Z}[\sqrt{D}]$ with $\mathbb{Z}$-basis $\{\beta_1, \beta_2\}$ given by (5.2) is an ideal of $\mathbb{Z}[\sqrt{D}]$.

Proof. Reducing to Hermite normal form, any (rank 2) lattice $I \subset \mathbb{Z}[\sqrt{D}]$ has a unique $\mathbb{Z}$-basis $\{\beta_1, \beta_2\}$ with the form

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} \sqrt{D} \\ 1 \end{pmatrix},$$

where $0 < b_1, b_3$ and $0 \le b_2 < b_3$. We remark that the action of Γ_∞ on the left of this basis is to simply add multiples of b_3 to b_2, and so one can consider only $b_2 \bmod b_3$ if one works modulo Γ_∞.

Since $\sqrt{D}$ generates the ring $\mathbb{Z}[\sqrt{D}]$, the sublattice I being an ideal is equivalent to $\sqrt{D}I \subset I$, and since $\sqrt{D}I$ has basis

$$\begin{pmatrix} \sqrt{D}\beta_1 \\ \sqrt{D}\beta_2 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} 0 & D \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ 0 & b_3 \end{pmatrix}^{-1} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$

I being an ideal is equivalent to

$$\frac{1}{b_1 b_3} \begin{pmatrix} b_1 & b_2 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} 0 & D \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_3 & -b_2 \\ 0 & b_1 \end{pmatrix} = \frac{1}{b_1 b_3} \begin{pmatrix} b_2 b_3 & -b_2^2 + D b_1^2 \\ b_3^2 & -b_2 b_3 \end{pmatrix} \tag{5.3}$$

having integer entries. From the first column and last row, we see that for I to be an ideal, it is necessary that $b_1 \mid b_2$ and $b_1 \mid b_3$. Writing $b_2 = \mu b_1$ and $b_3 = m b_1$, we see from the top right entry of (5.3) that $\mu^2 \equiv D \bmod m$ is necessary for I to be an ideal. We note that, if I is an ideal not divisible by rational integers greater than 1, then $b_1 = 1$, and we obtain the first part of proposition 5.2. For the converse part, we simply note that if $\mu^2 \equiv D \bmod m$, then the above shows that the sublattice I with basis given by (5.2) satisfies $\sqrt{D}I \subset I$ as required. □

We now produce different bases for ideals $I \subset \mathbb{Z}[\sqrt{D}]$ than those in (5.2). Any new basis is related to (5.2) by an element of $\mathrm{SL}(2, \mathbb{Z})$. We fix representatives I_l, $1 \le l \le h^+(D)$, of the narrow class group of $\mathbb{Q}(\sqrt{D})$. This implies that for any ideal $I \subset \mathbb{Z}[\sqrt{D}]$ there is a totally positive $\xi \in I_l^{-1}$ such that $I = \xi I_l$. To obtain a basis for I in this way, we first embed $\mathbb{Q}(\sqrt{D})$ in $\mathbb{R}^2$ via $a + b\sqrt{D} \mapsto (a + b\sqrt{D}, a - b\sqrt{D})$, so that $\mathbb{Z}[\sqrt{D}]$ is a lattice in $\mathbb{R}^2$. (Here we have fixed $\sqrt{D}$ to be the positive square root of D.) We adopt the notation that for $\xi \in \mathbb{Z}[\sqrt{D}]$ we use $\xi^{(1)}$ and $\xi^{(2)}$ to denote these coordinates. With this notation, in (5.2) we replace

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \leftarrow \begin{pmatrix} \beta_1^{(1)} & \beta_1^{(2)} \\ \beta_2^{(1)} & \beta_2^{(2)} \end{pmatrix}, \quad \begin{pmatrix} \sqrt{D} \\ 1 \end{pmatrix} \leftarrow \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}.$$

We fix $\mathbb{Z}$-bases $\{\beta_{l1}, \beta_{l2}\}$ for each of the I_l so that

$$\mathfrak{B}_l = \begin{pmatrix} \beta_{l1}^{(1)} & \beta_{l1}^{(2)} \\ \beta_{l2}^{(1)} & \beta_{l2}^{(2)} \end{pmatrix}$$

has positive determinant. Then $\{\xi\beta_{l1}, \xi\beta_{l2}\}$ is a basis for ξI_l, and the corresponding basis matrix is

$$\begin{pmatrix} \xi^{(1)}\beta_{l1}^{(1)} & \xi^{(2)}\beta_{l1}^{(2)} \\ \xi^{(1)}\beta_{l2}^{(1)} & \xi^{(2)}\beta_{l2}^{(2)} \end{pmatrix} = \mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix}.$$

For $m > 0$ and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$, we let $I \subset \mathbb{Z}[\sqrt{D}]$ be the corresponding ideal via proposition 5.2. There is some I_l and totally positive ξ such that $I = \xi I_l$, and hence there is a $\gamma \in \mathrm{SL}(2, \mathbb{Z})$ such that

$$\gamma \mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}. \tag{5.4}$$

We observe that considering μ as a residue class modulo m is equivalent to considering the coset $\Gamma_\infty\gamma$. More significantly, since ξ is only determined up to a multiple of a totally positive unit, we also need only consider the coset $\gamma\Gamma_l$, where

$$\Gamma_l = \left\{ \pm\mathfrak{B}_l \begin{pmatrix} \varepsilon_0^k & 0 \\ 0 & \varepsilon_0^{-k} \end{pmatrix} \mathfrak{B}_l^{-1} : k \in \mathbb{Z} \right\}$$

with ε_0, chosen to be greater than 1, generating the group of totally positive units in $\mathbb{Z}[\sqrt{D}]$. Thus equation (5.4) defines a unique double coset $\Gamma_\infty\gamma\Gamma_l$. Furthermore Γ_l satisfies

$$\Gamma_l = \mathrm{SL}(2,\mathbb{Z}) \cap \left\{ \pm\mathfrak{B}_l \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \mathfrak{B}_l^{-1} : t > 0 \right\}$$

exactly because $\varepsilon I_l = I_l$ if and only if ε is a unit in $\mathbb{Z}[\sqrt{D}]$.

Conversely, suppose that there is a totally positive ξ (necessarily in $\mathbb{Q}(\sqrt{D})$), and $\gamma \in \mathrm{SL}(2,\mathbb{Z})$ such that

$$\gamma\mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}. \tag{5.5}$$

We claim that then we in fact have

$$\gamma\mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}, \tag{5.6}$$

where $m > 0$ and $\mu^2 \equiv D \bmod m$.

To verify this, we first note that if $\xi \in I_l^{-1}$, then the left side of (5.5) is the basis for an ideal $I \subset \mathbb{Z}[\sqrt{D}]$, so then (5.6) follows by an application of proposition 5.2. Now to show $\xi \in I_l^{-1}$ follows from (5.5), we fix a basis $\{\overline{\beta}_{l1}, \overline{\beta}_{l2}\}$ of I_l^{-1} and write $\xi = c_1\overline{\beta}_{l1} + c_2\overline{\beta}_{l2}$. Our goal then is to show that c_1 and c_2 are in fact integers. Since $I_l^{-1}I_l = \mathbb{Z}[\sqrt{D}]$, we can define integers b_{lijk} by

$$\overline{\beta}_{li}\beta_{lj} = b_{lij1}\sqrt{D} + b_{lij2},$$

and we let

$$B_{li} = \begin{pmatrix} b_{li11} & b_{li12} \\ b_{li21} & b_{li22} \end{pmatrix}.$$

It follows that

$$\mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = (c_1B_{l1} + c_2B_{l2}) \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}, \tag{5.7}$$

and so, under the assumption (5.5), if $\gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix}$, then

$$B_l \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} d \\ -c \end{pmatrix},$$

where

$$B_l = \begin{pmatrix} b_{l111} & b_{l211} \\ b_{l121} & b_{l221} \end{pmatrix}.$$

The integrality of c_1 and c_2 is now an immediate consequence of the following lemma.

Lemma 5.3. *For all l, we have $B_l \in \mathrm{GL}(2,\mathbb{Z})$.*

Proof. We first observe that if c_1', c_2' were such that

$$B_l \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

then, setting $\xi' = c_1'\overline{\beta}_{l1} + c_2'\overline{\beta}_{l2}$, we would have from (5.7) that $\xi I_l \subset \mathbb{Z}$. This is a contradiction if $\xi' \neq 0$, and so we must have $\det B_l \neq 0$.

Now suppose that $\xi'' = c_1''\overline{\beta}_{l1} + c_2''\overline{\beta}_{l2}$ is a primitive vector in I_l^{-1}, meaning $a^{-1}\xi'' \notin I_l^{-1}$ for all rational integers $a > 1$, or equivalently $\gcd(c_1'', c_2'') = 1$. Then the ideal $\xi'' I_l \subset \mathbb{Z}[\sqrt{D}]$ is not divisible by any rational integer greater than 1. Now by the construction of the B_{li}, see (5.7), the rows of

$$(c_1''B_{l1} + c_2''B_{l2}) \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}$$

form a basis of $\xi'' I_l$. Putting $c_1''B_{l1} + c_2''B_{l2}$ into Hermite normal form, proposition 5.2 implies that the entries in the first column of $c_1''B_{l1} + c_2''B_{l2}$, i.e. the entries of $B_l \begin{pmatrix} c_1'' \\ c_2'' \end{pmatrix}$, are coprime. Hence B_l maps pairs of coprime integers to coprime integers, and thus $B_l \in \mathrm{GL}(2, \mathbb{Z})$. □

We record the above observations in the following proposition.

Proposition 5.4. *Let m be a positive integer and $\mu \bmod m$ satisfy $\mu^2 \equiv D \bmod m$. Then there exists a unique l, $1 \leq l \leq h^+(D)$, and a unique double coset $\Gamma_\infty \gamma \Gamma_l \in \Gamma_\infty \backslash \mathrm{SL}(2, \mathbb{Z}) / \Gamma_l$ such that*

$$\begin{pmatrix} 1 & \frac{\mu}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} = \gamma \mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} \tag{5.8}$$

for some totally positive ξ.

Conversely, if for $\Gamma_\infty \gamma \Gamma_l \in \Gamma_\infty \backslash \mathrm{SL}(2, \mathbb{Z}) / \Gamma_l$ there exists a totally positive ξ such that

$$\gamma \mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix},$$

then in fact

$$\gamma \mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & \frac{\mu}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \tag{5.9}$$

with m a positive integer and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$. Moreover such m and $\mu \bmod m$, if they exist, are unique for a given $\Gamma_\infty \gamma \Gamma_l \in \Gamma_\infty \backslash \mathrm{SL}(2, \mathbb{Z}) / \Gamma_l$.

Using the identification of $G/\{\pm I\}$ with the unit tangent bundle of $\mathbb{H}$, we interpret proposition 5.4 geometrically in proposition 5.5 below. We define oriented geodesics $\boldsymbol{c}_l$, $1 \leq l \leq h^+(D)$ by

$$\boldsymbol{c}_l = \left\{ (\det \mathfrak{B}_l)^{-\frac{1}{2}} \mathfrak{B}_l \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} : t > 0 \right\}.$$

In this case $\boldsymbol{c}_l$ defines a closed geodesic in $\mathrm{SL}(2, \mathbb{Z}) \backslash \mathbb{H}$ and Γ_l is the stabilizer in $\mathrm{SL}(2, \mathbb{Z})$ of the geodesic $\boldsymbol{c}_l$, i.e. $\Gamma_l = \Gamma_{\boldsymbol{c}_l}$ in the general notation introduced in section 2.

Proposition 5.5. *Let $m > 0$ and $\mu \bmod m$ satisfy $\mu^2 \equiv D \bmod m$. Then there is a unique l and double coset $\Gamma_\infty \gamma \Gamma_l \in \Gamma_\infty \backslash \mathrm{SL}(2, \mathbb{Z}) / \Gamma_l$ such that*

$$z_{\gamma \boldsymbol{c}_l} \equiv \frac{\mu}{m} + \mathrm{i} \frac{\sqrt{D}}{m}$$

modulo Γ_∞.

Conversely, given l and a double coset $\Gamma_\infty \gamma \Gamma_l$ such that $\gamma \boldsymbol{c}_l$ is positively oriented, there exist unique m and $\mu \bmod m$ satisfying the above.

Proof. We observe that in view of (2.1), the matrices

$$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix}$$

are identified in the unit tangent bundle of $\mathbb{H}$ with points having tangent vectors pointing directly to the right, and so the intersection of the set of these matrices with the set of matrices having the form

$$\gamma(\det \mathfrak{B}_l)^{-\frac{1}{2}}\mathfrak{B}_l \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$$

is exactly the point on the geodesic $\gamma \boldsymbol{c}_l$ having horizontal tangent vector pointing to the right, if such a point exists, and otherwise this intersection is empty. We note that such a point on $\gamma \boldsymbol{c}_l$ exists exactly when $\gamma \boldsymbol{c}_l$ is positively oriented and in this case gives the point on $\gamma \boldsymbol{c}_l$ with largest imaginary part. Scaling and rewriting the right side of (5.8) as

$$\begin{aligned} &(2\sqrt{D})^{-\frac{1}{2}} \begin{pmatrix} 1 & \frac{\mu}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} m^{-\frac{1}{2}} & 0 \\ 0 & m^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \frac{\mu}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \left(\frac{\sqrt{D}}{m}\right)^{\frac{1}{2}} & 0 \\ 0 & \left(\frac{m}{\sqrt{D}}\right)^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix}, \end{aligned}$$

we finish the proof by appealing to proposition 5.4. □

5.2 Congruence subgroups

So far we have only considered the orbit of geodesics under the action of $\mathrm{SL}(2,\mathbb{Z})$, however, as reflected in theorem 5.1, the orbits under $\Gamma_0(n)$ also give interesting arithmetic information. Indeed, by picking cosets $\Gamma_0(n)\gamma_l \in \Gamma_0(n)\backslash\mathrm{SL}(2,\mathbb{Z})$ correctly, one can obtain all m and roots $\mu \bmod m$ satisfying $m \equiv 0 \bmod n$ and $\mu \equiv \nu \bmod n$, for a fixed root $\nu \bmod n$, by considering the $\Gamma_0(n)$-orbits of the geodesics $\gamma_l \boldsymbol{c}_l$.

To establish this, we first show that if a root $\mu_1 \bmod m_1$ satisfies $m_1 \equiv 0 \bmod n$ and $\mu_1 \equiv \nu \bmod n$, and if $\mu_2 \bmod m_2$ is another root in the same $\Gamma_0(n)$-orbit as $\mu_1 \bmod m_1$, then $\mu_2 \bmod m_2$ also satisfies $m_2 \equiv 0 \bmod n$, $\mu_2 \equiv \nu \bmod n$. We then show that if roots $\mu_1 \bmod m_1$ and $\mu_2 \bmod m_2$ in the same $\mathrm{SL}(2,\mathbb{Z})$-orbit satisfy $m_1 \equiv m_2 \equiv 0 \bmod n$ and $\mu_1 \equiv \mu_2 \equiv \nu \bmod n$, then in fact the roots are in the same $\Gamma_0(n)$-orbit. Finally, for sake of completeness, we show that for every $1 \le l \le h^+(D)$ there is $\gamma_l = \gamma_l(n,\nu)$ such that the top of $\gamma_l \boldsymbol{c}_l$ satisfies the congruence conditions.

Although it may not be the quickest route to demonstrating the first point, we proceed by recording the following corollary of proposition 5.4.

Corollary 5.6. *As* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *runs through representatives of the double cosets in* $\Gamma_\infty\backslash\mathrm{SL}(2,\mathbb{Z})/\Gamma_l$ *such that* $\gamma \boldsymbol{c}_l$ *is positively oriented,*

$$\begin{pmatrix} \mu \\ m \end{pmatrix} = (\det B_l) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} b_{l112}b_{l211} - b_{l212}b_{l111} & b_{l112}b_{l221} - b_{l212}b_{l121} \\ b_{l122}b_{l211} - b_{l111}b_{l222} & b_{l122}b_{l221} - b_{l121}b_{l222} \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix}, \tag{5.10}$$

parametrise all $\mu \bmod m$ *satisfying* $m > 0$ *and* $\mu^2 \equiv D \bmod m$.

This corollary recovers the classic parametrisation of roots of $\mu^2 \equiv D \bmod m$ in terms of binary quadratic forms with discriminant $4D$. Indeed, (5.10) gives m as a binary quadratic form evaluated at c, d, and one can check that this form has discriminant $4D$. This binary quadratic form automatically comes with a Bhargava cube, see [2], given by the matrices B_{l1}, B_{l2}, which is associated with the ideals $\mathbb{Z}[\sqrt{D}], I_l, I_l^{-1}$. We also note that (5.10) gives a criterion on γ for the geodesic $\gamma \boldsymbol{c}_l$ to be positively oriented, namely that the m given by (5.10) is in fact positive.

Proof. We recall that for $\xi = c_1\overline{\beta}_{l1} + c_2\overline{\beta}_{l2}$,

$$\gamma\mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \gamma(c_1 B_{l1} + c_2 B_{l2}) \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix},$$

and so (5.8), (5.9) give

$$\begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} = \gamma(c_1 B_{l1} + c_2 B_{l2}). \tag{5.11}$$

Equating the first columns, we see that

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = B_l^{-1} \begin{pmatrix} d \\ -c \end{pmatrix},$$

from which (5.10) follows from the second column of (5.11). □

Having (5.10), we immediately see that if γ is replaced by $\gamma_1\gamma$ for some $\gamma_1 \in \Gamma_0(n)$, $m \bmod n$ is replaced by a multiple, and, if $m \equiv 0 \bmod n$, $\mu \bmod n$ is unchanged. Hence the conditions $m \equiv 0 \bmod n$ and $\mu \equiv \nu \bmod n$ are invariant under $\Gamma_0(n)$, thus finishing the first part of our examination of the $\Gamma_0(n)$ orbits. The second part is completed by the following lemma.

Lemma 5.7. *Let m_1, m_2 be positive integers and $\mu_1 \bmod m_1$, $\mu_2 \bmod m_2$ satisfy $\mu_1^2 \equiv D \bmod m_1$, $\mu_2^2 \equiv D \bmod m_2$. Suppose that $m_1 \equiv m_2 \equiv 0 \bmod n$ and that $\mu_1 \equiv \mu_2 \equiv \nu \bmod n$ where $\nu \bmod n$ is a fixed root of $\nu^2 \equiv D \bmod n$. Then, if there exists $\gamma \in \mathrm{SL}(2,\mathbb{Z})$ such that*

$$\begin{pmatrix} 1 & \mu_1 \\ 0 & m_1 \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \gamma \begin{pmatrix} 1 & \mu_2 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \tag{5.12}$$

for some totally positive ξ, then in fact $\gamma \in \Gamma_0(n)$.

Proof. Let I_1, I_2 be the ideals corresponding to $\mu_1 \bmod m_1$, $\mu_2 \bmod m_2$. Then since

$$\begin{pmatrix} 1 & \mu_j \\ 0 & m_j \end{pmatrix} = \begin{pmatrix} 1 & k_j \\ 0 & \frac{m_j}{n} \end{pmatrix} \begin{pmatrix} 1 & \nu \\ 0 & n \end{pmatrix}$$

for some integers k_j, we have $I_1, I_2 \subset I$, where I is the ideal corresponding to $\nu \bmod n$. Similarly, we have

$$\begin{pmatrix} 1 & \mu_1 \\ 0 & m_1 \end{pmatrix} = \begin{pmatrix} 1 & k \\ 0 & n \end{pmatrix} \begin{pmatrix} 1 & \nu_1 \\ 0 & n_1 \end{pmatrix},$$

where $n_1 = \frac{m_1}{n}$ and $\nu_1 \bmod n_1$ is the root corresponding to the ideal $I_1 I^{-1}$. Now (5.12) implies that $\xi I_1 = I_2$, so $\xi I_1 I^{-1} = I_2 I^{-1} \subset \mathbb{Z}[\sqrt{D}]$, whence

$$\begin{pmatrix} 1 & \nu_1 \\ 0 & n_1 \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = A \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \tag{5.13}$$

for some integer matrix A. Rewriting the left side of (5.13) as

$$\begin{aligned} &\begin{pmatrix} 1 & k \\ 0 & n \end{pmatrix}^{-1} \begin{pmatrix} 1 & \mu_1 \\ 0 & m_1 \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} \\ &= \begin{pmatrix} 1 & k \\ 0 & n \end{pmatrix}^{-1} \gamma \begin{pmatrix} 1 & \mu_2 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix}, \end{aligned}$$

we obtain

$$\begin{pmatrix} 1 & k \\ 0 & n \end{pmatrix}^{-1} \gamma \begin{pmatrix} 1 & \mu_2 \\ 0 & m_2 \end{pmatrix} = A. \tag{5.14}$$

Since A has integer entries, an inspection of the $(2,1)$ entry of (5.14) shows that $\gamma \in \Gamma_0(n)$. □

Lemma 5.8. *For each $1 \le l \le h^+(D)$, there exists $\gamma_l = \gamma_l(n,\nu)$ such that*

$$z_{\gamma_l \boldsymbol{c}_l} = \frac{\mu}{m} + \mathrm{i}\frac{\sqrt{D}}{m} \tag{5.15}$$

with $m \equiv 0 \bmod n$ and $\mu \equiv \nu \bmod n$.

Proof. By proposition 5.2, there exists an ideal $I_0 \subset \mathbb{Z}[\sqrt{D}]$ corresponding to the root $\nu \bmod n$. As $I_1 \subset \mathbb{Z}[\sqrt{D}]$ ranges over the narrow ideal class of $I_l I_0^{-1}$, the integers $N(I_1)$ are the values of an integer binary quadratic form, see corollary 5.6, and so we may choose I_1 with $\gcd(n, N(I_1)) = 1$ and without rational integer divisors. We let $\mu_1 \mod m_1$ denote the root corresponding to I_1 via proposition 5.2.

Letting $m_1\overline{m_1} + n\overline{n_1} = 1$ and setting $\mu = \nu m_1\overline{m_1} + \mu_1 n\overline{n}$, we compute

$$\begin{aligned} I_0 I_1 = (n, \nu + \sqrt{D})(m_1, \mu_1 + \sqrt{D}) &= (m_1 n, m_1\nu + m_1\sqrt{D}, n\mu_1 + n\sqrt{D}, \mu_1\nu + D + (\mu_1+\nu)\sqrt{D}) \\ &= (m_1 n, \mu + \sqrt{D}, \mu_1\nu + D + (\mu_1+\nu)\sqrt{D}) = (m_1 n, \mu + \sqrt{D}), \end{aligned}$$

where the last equality follows from $\mu_1\nu + D - (\mu_1+\nu)\mu \equiv D - \mu^2 \equiv 0$ modulo both n and m_1. The ideal $I = I_0 I_1$ therefore corresponds to the root $\mu \bmod m_1 n$ and is in the narrow ideal class of I_l. Equation (5.15) then follows by the machinery of section 5.1. □

Corollary 5.6, lemma 5.7, and lemma 5.8 show that for a fixed $\nu \bmod n$ and each l, $1 \le l \le h^+(D)$, there exists a unique coset $\Gamma_0(n)\gamma_l \in \Gamma_0(n)\backslash \mathrm{SL}(2,\mathbb{Z})$ such that roots $\mu \bmod m$ with $m \equiv 0 \bmod n$ and $\mu \equiv \nu \bmod n$ are found in the positively oriented tops of the geodesics $\gamma\gamma_l\boldsymbol{c}_l$ as γ ranges over $\Gamma_0(n)$ as in proposition 5.5. This is the bulk of theorem 5.1; it only remains to show that the closed geodesics obtained from projecting $\gamma_l\boldsymbol{c}_l$ to $\Gamma_0(n)\backslash\mathbb{H}$ all have the same length. In fact we show that the closed geodesic in $\Gamma_0(n)\backslash\mathbb{H}$ coming from $\gamma_l\boldsymbol{c}_l$ has the same length as the closed geodesic in $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$ coming from $\boldsymbol{c}_l$. This is not a trivial observation; in general the splitting behaviour of geodesics in finite covering surfaces is quite complicated, being analogous to the splitting of primes in number fields, see [29].

Understanding the length of $\gamma_l\boldsymbol{c}_l$ in $\Gamma_0(n)\backslash\mathbb{H}$ is equivalent to understanding the stabilizer of $\gamma_l\boldsymbol{c}_l$ in $\Gamma_0(n)$. It is clear that there is some positive integer k such that

$$\gamma_l \mathfrak{B}_l \begin{pmatrix} \varepsilon_0^k & 0 \\ 0 & \varepsilon_0^{-k} \end{pmatrix} \mathfrak{B}_l^{-1}\gamma_l^{-1} \in \Gamma_0(n),$$

and the smallest such k gives the generator of the stabilizer. Moreover, the length of $\gamma\boldsymbol{c}_l$ in $\Gamma_0(n)\backslash\mathbb{H}$ has length this k times the length of $\boldsymbol{c}_l$ in $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}$. The following lemma, which is essentially a corollary of lemma 5.7, shows that in our setting we have in fact $k = 1$.

Lemma 5.9. *For all l, we have*

$$\gamma_l \mathfrak{B}_l \begin{pmatrix} \varepsilon_0 & 0 \\ 0 & \varepsilon_0^{-1} \end{pmatrix} \mathfrak{B}_l^{-1}\gamma_l^{-1} \in \Gamma_0(n). \tag{5.16}$$

Proof. We let γ_0 be the matrix on the left side of (5.16), so γ_0 satisfies

$$\gamma_0\gamma_l\mathfrak{B}_l = \gamma_l\mathfrak{B}_l \begin{pmatrix} \varepsilon_0 & 0 \\ 0 & \varepsilon_0^{-1} \end{pmatrix}. \tag{5.17}$$

From the way γ_l was chosen, for some totally positive ξ we have

$$\gamma_l\mathfrak{B}_l \begin{pmatrix} \xi^{(1)} & 0 \\ 0 & \xi^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \tag{5.18}$$

with $m \equiv 0 \bmod n$. Since the diagonal matrices commute, we put (5.18) into (5.17) to obtain

$$\gamma_0 \begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{D} & -\sqrt{D} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_0 & 0 \\ 0 & \varepsilon_0^{-1} \end{pmatrix},$$

so lemma 5.7 in the special case $m_1 = m_2 = m$ and $\mu_1 = \mu_2 = \mu$ gives $\gamma_0 \in \Gamma_0(n)$. □

This completes the proof of theorem 5.1.

5.3 Negative D

For completeness, let us now briefly discuss the case of negative D. Here the situation is similar as for positive D; in fact simpler, since the tops of geodesics are replaced by Γ-orbits of points in $\mathbb{H}$, whose statistical distribution is well understood. The results mirroring those in section 2 are proved in [23], and the pair correlation density is calculated in [17, 23]. For further related studies see also [3, 4, 19, 26, 27] and references therein.

Following [5, 6], the key observation is the following. Denote by Γ_z the stabiliser in $\Gamma_0(n)$ of a point $z \in \mathbb{H}$.

Theorem 5.10. *Let $D < 0$ square-free and $D \not\equiv 1 \bmod 4$. Fix $n > 0$ and $\nu \bmod n$ such that $\nu^2 \equiv D \bmod n$. Then there exist $z_1, \ldots, z_h \in \mathbb{H}$ with the following properties:*

(i) *For any integer $m > 0$ and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$, $m \equiv 0 \bmod n$, and $\mu \equiv \nu \bmod n$, there is a unique l and double coset $\Gamma_\infty \gamma \Gamma_{z_l} \in \Gamma_\infty \backslash \Gamma_0(n) / \Gamma_{z_l}$ such that*

$$\gamma z_l \equiv \frac{\mu}{m} + \mathrm{i}\frac{\sqrt{-D}}{m} \bmod \Gamma_\infty.$$

(ii) *Conversely, given l and double coset $\Gamma_\infty \gamma \Gamma_{z_l} \in \Gamma_\infty \backslash \Gamma_0(n) / \Gamma_{z_l}$, there exist unique $m > 0$ and $\mu \bmod m$ satisfying $\mu^2 \equiv D \bmod m$, $m \equiv 0 \bmod n$, and $\mu \equiv \nu \bmod n$ such that* (5.1) *holds.*

Here h is the class number $h(D)$ of the imaginary number field $\mathbb{Q}(\sqrt{D})$, and $\mathbb{Z}[\sqrt{D}]$ is its maximal order. We also note that $\Gamma_{z_l} = \{\pm I\}$ unless $D = -1$, in which case $h(-1) = 1$ and when $n = 1$ we may choose $z_1 = \mathrm{i}$, so

$$\Gamma_{\mathrm{i}} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}.$$

The proof of the above statement is analogous to that of theorem 5.1. The algebraic manipulations leading to (5.4) are identical, although now over the complex numbers. In order to turn (5.4) into a real equation, we multiply it from the right by the matrix $\begin{pmatrix} -\frac{\mathrm{i}}{2} & \frac{1}{2} \\ \frac{\mathrm{i}}{2} & \frac{1}{2} \end{pmatrix}$ to obtain, writing $\sqrt{D} = \mathrm{i}\sqrt{-D}$ and $\xi^{(1)} = a + \mathrm{i}b$ and $\xi^{(2)} = a - \mathrm{i}b$,

$$\gamma \widetilde{\mathfrak{B}}_l \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ 0 & m \end{pmatrix} \begin{pmatrix} \sqrt{-D} & 0 \\ 0 & 1 \end{pmatrix} \tag{5.19}$$

and

$$\widetilde{\mathfrak{B}}_l = \mathfrak{B}_l \begin{pmatrix} -\frac{\mathrm{i}}{2} & \frac{1}{2} \\ \frac{\mathrm{i}}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \beta_{l1}^{(1)} & \beta_{l1}^{(2)} \\ \beta_{l2}^{(1)} & \beta_{l2}^{(2)} \end{pmatrix} \begin{pmatrix} -\frac{\mathrm{i}}{2} & \frac{1}{2} \\ \frac{\mathrm{i}}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \operatorname{Im} \beta_{l1}^{(1)} & \operatorname{Re} \beta_{l1}^{(1)} \\ \operatorname{Im} \beta_{l2}^{(1)} & \operatorname{Re} \beta_{l2}^{(1)} \end{pmatrix}.$$

Applying both sides of (5.19) to i (as fractional linear transformations) yields

$$\gamma z_l = \frac{\mu}{m} + \mathrm{i}\frac{\sqrt{-D}}{m}, \qquad \text{with } z_l = \frac{\beta_{l1}^{(1)}}{\beta_{l2}^{(1)}},$$

as required.

Theorem 5.10 shows that when $D < 0$, the normalised roots $\frac{\mu}{m}$, with the congruence conditions $m \equiv 0 \bmod m$ and $\mu \equiv \nu \bmod n$ appear as the real parts of points in a finite number of $\Gamma_0(n)$-orbits. The results of [23] in the 'cuspidal observer' setting then immediately apply to give the analogues of theorems 2.1 and 2.5. As pointed out in [23], these results are closely related to previous work on the distribution of angles in hyperbolic lattices, see [3, 4, 19, 26, 27] and the references therein. The calculation of the pair correlation density in this setting is carried out in [17] and [23], and thus give the analogue of our theorem 1.1 when $D < 0$ with the expression for w_D having a similar form to our (4.32), (4.14) and (4.15).

References

[1] Jayadev S. Athreya and Yitwah Cheung. A Poincaré section for the horocycle flow on the space of lattices. *Int. Math. Res. Not. IMRN*, (10):2643–2690, 2014.

[2] Manjul Bhargava. Higher composition laws. I. A new view on Gauss composition, and quadratic generalizations. *Ann. of Math. (2)*, 159(1):217–250, 2004.

[3] Florin P. Boca, Vicenţiu Paşol, Alexandru A. Popa, and Alexandru Zaharescu. Pair correlation of angles between reciprocal geodesics on the modular surface. *Algebra Number Theory*, 8(4):999–1035, 2014.

[4] Florin P. Boca, Alexandru A. Popa, and Alexandru Zaharescu. Pair correlation of hyperbolic lattice angles. *Int. J. Number Theory*, 10(8):1955–1989, 2014.

[5] V. A. Bykovskiĭ. Spectral expansions of certain automorphic functions and their number-theoretic applications. volume 134, pages 15–33. 1984. Automorphic functions and number theory, II.

[6] W. Duke, J. B. Friedlander, and H. Iwaniec. Equidistribution of roots of a quadratic congruence to prime moduli. *Ann. of Math. (2)*, 141(2):423–441, 1995.

[7] W. Duke, J. B. Friedlander, and H. Iwaniec. Weyl sums for quadratic roots. *Int. Math. Res. Not. IMRN*, (11):2493–2549, 2012.

[8] Alex Eskin and Curt McMullen. Mixing, counting, and equidistribution in Lie groups. *Duke Math. J.*, 71(1):181–209, 1993.

[9] Etienne Fouvry and Henryk Iwaniec. Gaussian primes. *Acta Arith.*, 79(3):249–287, 1997.

[10] John Friedlander and Henryk Iwaniec. Asymptotic sieve for primes. *Ann. of Math. (2)*, 148(3):1041–1065, 1998.

[11] Anton Good. *Local analysis of Selberg's trace formula*, volume 1040 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.

[12] Dennis A. Hejhal. Roots of quadratic congruences and eigenvalues of the non-Euclidean Laplacian. In *The Selberg trace formula and related topics (Brunswick, Maine, 1984)*, volume 53 of *Contemp. Math.*, pages 277–339. Amer. Math. Soc., Providence, RI, 1986.

[13] C. Hooley. On the distribution of the roots of polynomial congruences. *Mathematika*, 11:39–49, 1964.

[14] Christopher Hooley. On the number of divisors of a quadratic polynomial. *Acta Math.*, 110:97–114, 1963.

[15] Henryk Iwaniec. Almost-primes represented by quadratic polynomials. *Invent. Math.*, 47(2):171–188, 1978.

[16] Olav Kallenberg. *Foundations of modern probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham, ©2021. Third edition.

[17] Dubi Kelmer and Alex Kontorovich. On the pair correlation density for hyperbolic angles. *Duke Math. J.*, 164(3):473–509, 2015.

[18] E. Kowalski and K. Soundararajan. Equidistribution from the Chinese remainder theorem. *Adv. Math.*, 385:Paper No. 107776, 36, 2021.

[19] Christopher Lutsko. Directions in orbits of geometrically finite hyperbolic subgroups. *Mathematical Proc. Camb. Phil. Soc.*, pages 1–40, 2020.

[20] Grigoriy A. Margulis. *On some aspects of the theory of Anosov systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2004. With a survey by Richard Sharp: Periodic orbits of hyperbolic flows, Translated from the Russian by Valentina Vladimirovna Szulikowska.

[21] Jens Marklof. Entry and return times for semi-flows. *Nonlinearity*, 30(2):810–824, 2017.

[22] Jens Marklof and Andreas Strömbergsson. The distribution of free path lengths in the periodic Lorentz gas and related lattice point problems. *Ann. of Math. (2)*, 172(3):1949–2033, 2010.

[23] Jens Marklof and Ilya Vinogradov. Directions in hyperbolic lattices. *J. Reine Angew. Math.*, 740:161–186, 2018.

[24] Jouni Parkkonen and Frédéric Paulin. Équidistribution, comptage et approximation par irrationnels quadratiques. *J. Mod. Dyn.*, 6(1):1–40, 2012.

[25] Jouni Parkkonen and Frédéric Paulin. Counting arcs in negative curvature. In *Geometry, topology, and dynamics in negative curvature*, volume 425 of *London Math. Soc. Lecture Note Ser.*, pages 289–344. Cambridge Univ. Press, Cambridge, 2016.

[26] Morten S. Risager and Zeév Rudnick. On the statistics of the minimal solution of a linear Diophantine equation and uniform distribution of the real part of orbits in hyperbolic spaces. In *Spectral analysis in geometry and number theory*, volume 484 of *Contemp. Math.*, pages 187–194. Amer. Math. Soc., Providence, RI, 2009.

[27] Morten S. Risager and Anders Södergren. Angles in hyperbolic lattices: the pair correlation density. *Trans. Amer. Math. Soc.*, 369(4):2807–2841, 2017.

[28] Peter Sarnak. *Some applications of modular forms*, volume 99 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.

[29] Peter Clive Sarnak. *Prime Geodesic Theorems*. ProQuest LLC, Ann Arbor, MI, 1980. Thesis (Ph.D.)–Stanford University.

[30] Arpad Toth. *Equidistribution of roots of quadratic congruences*. ProQuest LLC, Ann Arbor, MI, 1997. Thesis (Ph.D.)–Rutgers The State University of New Jersey - New Brunswick.

[31] Sa'ar Zehavi. On the joint distribution of the roots of pairs of polynomial congruences, 2020.