Translates of rational points along expanding closed horocycles on the modular surface

Claire Burrin[1,2] · Uri Shapira[3] · Shucheng Yu[3,4]

Abstract

We study the limiting distribution of the rational points under a horizontal translation along a sequence of expanding closed horocycles on the modular surface. Using spectral methods we confirm equidistribution of these sample points for any translate when the sequence of horocycles expands within a certain polynomial range. We show that the equidistribution fails for generic translates and a slightly faster expanding rate. We also prove both equidistribution and non-equidistribution results by obtaining explicit limiting measures while allowing the sequence of horocycles to expand arbitrarily fast. Similar results are also obtained for translates of primitive rational points.

Contents

✉ Shucheng Yu
  shucheng.yu@math.uu.se; yushucheng@campus.technion.ac.il

Claire Burrin
  claire.burrin@rutgers.edu; claire.burrin@math.ethz.ch

Uri Shapira
  ushapira@tx.technion.ac.il

[1] Department of Mathematics, Rutgers University, 110 Frelinghuysen Rd, Piscataway, NJ 08854, USA

[2] Present Address: Department Mathematik, ETH Zürich, 8092 Zurich, Switzerland

[3] Department of Mathematics, Technion, Haifa, Israel

[4] Present Address: Department of Mathematics, Uppsala University, Box 480, 75106 Uppsala, Sweden

1 Introduction

Let $\{S_n\}_{n\in\mathbb{N}}$ be a sequence of "nice" subsets that become equidistributed in their ambient space. Given a sequence of discrete subsets $\{R_n\}_{n\in\mathbb{N}}$ with $R_n \subset S_n$, an interesting question is to study to what extent does the distribution behavior of $\{R_n\}_{n\in\mathbb{N}}$ mimic that of $\{S_n\}_{n\in\mathbb{N}}$. One naturally expects that when the size of R_n is relatively large, it is more likely that $\{R_n\}_{n\in\mathbb{N}}$ inherits some distribution property from $\{S_n\}_{n\in\mathbb{N}}$; on the other hand if R_n lies on S_n sparsely, then it is more likely that points in $\{R_n\}_{n\in\mathbb{N}}$ become decorrelated and distribute like random points on the ambient space.

In the setting of unipotent dynamics, the most typical example of a sequence $\{S_n\}_{n\in\mathbb{N}}$ is a sequence of *expanding closed horocycles* on a non-compact finite-area hyperbolic surface $\mathcal{M}$. More precisely, we can realize $\mathcal{M}$ as a quotient $\Gamma\backslash\mathbb{H}$ where Γ is a cofinite Fuchsian subgroup and $\mathbb{H} = \{z = x + iy \in \mathbb{C} : y > 0\}$ is the Poincaré upper half-plane, equipped with the hyperbolic metric $ds = |dz|/y$, where $dz = dx + i\,dy$ is the complex line element. Up to conjugating by an appropriate isometry, we may assume that $\mathcal{M} = \Gamma\backslash\mathbb{H}$ has a width one cusp at infinity, that is, that the isotropy group $\Gamma_\infty < \Gamma$ is generated by the translation sending $z \in \mathbb{H}$ to $z + 1$. A *closed horocycle of height $y > 0$* is a closed set of the form

$$\mathcal{H}_y := \{\Gamma(x + iy) : x \in \mathbb{R}/\mathbb{Z}\} \subset \mathcal{M},$$

and its period, i.e., its hyperbolic length, is y^{-1}. As $\mathcal{H}_y$ gets longer, that is, as $y \to 0^+$, it becomes equidistributed on $\mathcal{M}$ with respect to the hyperbolic area $d\mu(z) = y^{-2}dxdy$. The first effective version of this result is due to Sarnak [28] who, using spectral arguments, proved that for every $\Psi \in C_c^\infty(\Gamma \backslash \mathbb{H})$ and any $y > 0$,

$$\int_0^1 \Psi(x + iy)dx = \frac{\int_\mathcal{M} \Psi(z)d\mu(z)}{\mu(\mathcal{M})} + O\left(\mathcal{S}(\Psi)y^\alpha\right), \tag{1.1}$$

where $\mathcal{S}$ is some Sobolev norm, and $0 < \alpha < 1$ is a constant depending on the first non-trivial residual hyperbolic Laplacian eigenvalue of Γ. In the case of the modular surface $\mathrm{SL}_2(\mathbb{Z})\backslash \mathbb{H}$, $\alpha = \frac{1}{2}$, while Zagier [32] observed that the Riemann hypothesis is equivalent to the equidistribution rate $O_\epsilon\left(y^{3/4-\epsilon}\right)$.

In this setting, this problem was first investigated by Hejhal in [12] with a heuristic and numerical study of the value distribution of the sample points

$$\Gamma\left(\tfrac{x+j}{n} + iy\right) : 0 \le j \le n - 1 \tag{1.2}$$

for some Hecke triangle groups $\Gamma = \mathbb{G}_q$ under the assumption that ny is small. Set

$$S_{y,n,\Psi}(x) := \sum_{j=0}^{n-1} \Psi\left(\tfrac{x+j}{n} + iy\right),$$

where Ψ is some mean-zero step function on a fixed fundamental domain for $\Gamma \backslash \mathbb{H}$ (automorphically extended to $\mathbb{H}$). The numerics show that the value distribution of $n^{-1/2}S_{n,y,\Psi}(x)$ with respect to $x \in [0, 1)$ approaches a Gaussian curve for the non-arithmetic Hecke triangle groups $\mathbb{G}_5$ and $\mathbb{G}_7$, while this phenomenon breaks down for $\mathbb{G}_3 = \mathrm{PSL}_2(\mathbb{Z})$. Hejhal gave an explanation of this difference based on the existence of Hecke operators on $\mathbb{G}_3$. The convergence to a Gaussian distribution for general non-arithmetic Fuchsian groups was later confirmed by Strömbergsson [30, Corollary 6.5], under the assumption that the sequence $\{y_n\}_{n \in \mathbb{N}}$ decays sufficiently rapidly.

Other such problems have since been investigated. Marklof and Strömbergsson [27] proved the equidistribution of generic Kronecker sequences

$$\{\Gamma(j\beta + iy_n) \in \mathcal{M} : 1 \le j \le n\} \subset \mathcal{M} \tag{1.3}$$

along a sequence of closed horocycles expanded at a certain rate y_n on $T_1\mathcal{M}$, the unit tangent bundle of $\mathcal{M}$. The equidistribution of Hecke points proved by Clozel–Ullmo [4] (see also [3,10]) implies the equidistribution of the primitive rational points

$$\left\{\Gamma\left(\tfrac{j}{n} + \tfrac{i}{n}\right) : 1 \le j \le n - 1, \ \gcd(j, n) = 1\right\}$$

at prime steps on the modular surface, see [10, Remark on p. 171]. More recently, the equidistribution of the above sequence along the full sequence of positive integers was proved by Einsiedler–Luethi–Shah [8] in a slightly more general setting, namely

on the product of the unit tangent bundle of the modular surface and a torus. Various sparse equidistribution results have also been obtained for expanding horospheres in the space of lattices $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$ for $n \geq 3$ [7,9,22,23,26] and in Hilbert modular surfaces [24].

For each of these equidistribution results, assumptions on the expanding rate of the sequence $\{S_n\}_{n\in\mathbb{N}}$ are crucial; the discrete subsets $\{R_n\}_{n\in\mathbb{N}}$ lying on $\{S_n\}_{n\in\mathbb{N}}$ can not be too sparse.

This paper emerged from an attempt to prove a result which turned out to be false. We consider the *sparse equidistribution problem* for the subset of rational points (with denominator n) under a horizontal translation $x \in \mathbb{R}/\mathbb{Z}$ on a horocycle $\mathcal{H}_y$ on the modular surface; we denote this subset by $\mathcal{R}_n(x, y_n)$ (cf. (1.4)). We thought that since the closed horocycles $\mathcal{H}_y$ equidistribute as $y \to 0^+$, if we fix a sequence $\{y_n\}_{n\in\mathbb{N}}$ approaching zero, then the normalized counting measures on $\mathcal{R}_n(x, y_n)$ (and its primitive counterpart) should equidistribute for Lebesgue almost every x as $n \to \infty$. See the recent paper of Bersudsky [1, Theorem 1.5] for an analogue situation where such a result is true. Note the order of quantifiers; we first fix the sequence $\{y_n\}_{n\in\mathbb{N}}$ and only then choose the horizontal translation x. It is not hard to see that if one flips the quantifiers, for any fixed horizontal translation x, there are sequences $\{y_n\}_{n\in\mathbb{N}}$ (approaching zero rapidly) such that equidistribution fails. We were very surprised to learn though, that in stark contrast to our initial expectation, equidistribution fails. The main novel result of this paper (Theorem 1.5) says that there are sequences $\{y_n\}_{n\in\mathbb{N}}$ approaching zero arbitrarily fast such that for almost every horizontal translation x the normalized counting measures $\mathcal{R}_n(x, y_n)$ and its primitive counterpart do not equidistribute. In fact, we show the collection of limit measures contains the uniform measure $\mu_{\mathcal{M}}$, the zero measure and certain singular measures. Although these should be considered as the main contribution of this paper, we also complement our analysis with answering natural questions concerning sequences $\{y_n\}_{n\in\mathbb{N}}$ approaching zero in a polynomial rate.

The next subsections describe more precisely the setting and results obtained.

1.1 Context of the present paper

Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and let $\mathcal{M} = \Gamma\backslash\mathbb{H}$ be the modular surface. In this paper, generalizing the setting of [8], we study the equidistribution problem for the sets of rational and primitive rational points under an arbitrary horizontal translation $x \in \mathbb{R}/\mathbb{Z}$ along a given sequence of expanding closed horocycles on $\mathcal{M}$. The set of rational points is the obvious choice of a sparse set with identical spacings, while primitive rational points constitute the simplest pseudorandom sequence (via the linear congruential generator). For any $n \in \mathbb{N}$, $x \in \mathbb{R}/\mathbb{Z}$ and $y > 0$ we denote by

$$\mathcal{R}_n(x, y) := \left\{ \Gamma\left(x + \tfrac{j}{n} + iy\right) \in \mathcal{H}_y : 0 \leq j \leq n-1 \right\} \qquad (1.4)$$

and respectively

$$\mathcal{R}_n^{\mathrm{pr}}(x, y) := \left\{ \Gamma\left(x + \tfrac{j}{n} + iy\right) \in \mathcal{H}_y : j \in (\mathbb{Z}/n\mathbb{Z})^\times \right\}, \qquad (1.5)$$

the set of rational and respectively primitive rational points with denominator n on the closed horocycle $\mathcal{H}_y$ translated to the right by x. As usual, $(\mathbb{Z}/n\mathbb{Z})^\times$ denotes here the multiplicative group of integers modulo n.

Let $\{y_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers such that $y_n \to 0$ as $n \to \infty$. We investigate the *limiting distribution* of the sequences of sample points $\{\mathcal{R}_n(x, y_n)\}_{n\in\mathbb{N}}$ and $\{\mathcal{R}_n^{\mathrm{pr}}(x, y_n)\}_{n\in\mathbb{N}}$ under various assumptions on the expanding rate of the sequence of horocycles $\{\mathcal{H}_{y_n}\}_{n\in\mathbb{N}}$, or equivalently, the decay rate of $\{y_n\}_{n\in\mathbb{N}}$.

This problem is naturally easier when the sequence $\{y_n\}_{n\in\mathbb{N}}$ decays slowly since then at each step we have relatively more sample points on the underlying horocycle. For instance, if $ny_n \to \infty$ as $n \to \infty$, the hyperbolic distance between two adjacent points in $\mathcal{R}_n(x, y_n)$ decays to zero as $n \to \infty$. Since the points in $\mathcal{R}_n(x, y_n)$ distribute evenly on $\mathcal{H}_{y_n}$, the distribution behavior of $\mathcal{R}_n(x, y_n)$ then mimics that of $\mathcal{H}_{y_n}$. In particular, for any $x \in \mathbb{R}/\mathbb{Z}$ the sequence $\{\mathcal{R}_n(x, y_n)\}_{n\in\mathbb{N}}$ becomes equidistributed on $\mathcal{M}$ with respect to the hyperbolic area μ as $n \to \infty$, following from the equidistribution of the sequence $\{\mathcal{H}_{y_n}\}_{n\in\mathbb{N}}$.

Regarding $\{\mathcal{R}_n^{\mathrm{pr}}(x, y_n)\}_{n\in\mathbb{N}}$, its distribution behavior is well understood when $x = 0$. Indeed, it was shown by Luethi [24] that if $y_n = c/n^\alpha$ for some $c > 0$ and some $\alpha \in (0, 1)$, then $\mathcal{R}_n^{\mathrm{pr}}(0, y_n)$ becomes equidistributed on $\mathcal{M}$ with respect to μ as $n \to \infty$. Moreover, under the simple symmetry relation that for $\gcd(j, n) = 1$ and $y > 0$

$$\Gamma\left(\tfrac{j}{n} + iy\right) = \Gamma\left(-\tfrac{\bar{j}}{n} + \tfrac{i}{n^2 y}\right), \tag{1.6}$$

one can extend this equidistribution result to the range $\alpha \in (1, 2)$; this improves the previous work of Demirci Akarsu [5, Theorem 2] which confirms equdistribution of $\{\mathcal{R}_n^{\mathrm{pr}}(0, c/n^\alpha)\}_{n\in\mathbb{N}}$ for $\alpha \in (\tfrac{3}{2}, 2)$. Here $\bar{j} \in (\mathbb{Z}/n\mathbb{Z})^\times$ denotes the multiplicative inverse of $j \in (\mathbb{Z}/n\mathbb{Z})^\times$. The equidistribution for the case $\alpha = 1$ was later proved by Einsiedler–Luethi–Shah [8]; Jana [16, Theorem 1] recently gave an alternative spectral proof to this equidistribution result. We also mention that both [5, Theorem 2] and [16, Theorem 1] are valid in the same setting as [8], namely, on the product of the unit tangent bundle of the modular surface and a torus. When $\alpha = 2$ the equidistribution fails as the aforementioned symmetry implies that $\mathcal{R}_n^{\mathrm{pr}}(0, c/n^2) = \mathcal{R}_n^{\mathrm{pr}}(0, 1/c)$ is always trapped in the closed horocycle $\mathcal{H}_{1/c}$. For the same reason, when $\alpha > 2$ (or more generally for any sequence satisfying $n^2 y_n \to 0$), one has with $\mathcal{R}_n^{\mathrm{pr}}(0, c/n^\alpha) = \mathcal{R}_n^{\mathrm{pr}}(0, n^{\alpha-2}/c) \subset \mathcal{H}_{n^{\alpha-2}/c}$ a full escape to the cusp of $\mathcal{M}$ as $n \to \infty$. It is worth noting that while the symmetry (1.6) still holds for rational translates (cf. Lemma 3.6), it breaks down for irrational translates.

1.2 Statements of the results

We will state here the main results of this paper, and postpone the discussion of their proofs to the next subsection. Let $\mu_{\mathcal{M}} := \mu(\mathcal{M})^{-1}\mu$ be the normalized hyperbolic area on $\mathcal{M}$. For any $n \in \mathbb{N}$, $x \in \mathbb{R}/\mathbb{Z}$ and $y > 0$ let $\delta_{n,x,y}$ and $\delta_{n,x,y}^{\mathrm{pr}}$ denote the normalized probability counting measure supported on $\mathcal{R}_n(x, y)$ and $\mathcal{R}_n^{\mathrm{pr}}(x, y)$ respectively.

That is, for any $\Psi \in C_c^\infty(\mathcal{M})$,

$$\delta_{n,x,y}(\Psi) = \frac{1}{n} \sum_{j=0}^{n-1} \Psi\left(x + \frac{j}{n} + iy\right),$$

and

$$\delta_{n,x,y}^{\mathrm{pr}}(\Psi) = \frac{1}{\varphi(n)} \sum_{j \in (\mathbb{Z}/n\mathbb{Z})^\times} \Psi\left(x + \frac{j}{n} + iy\right),$$

where φ is Euler's totient function. Here and throughout, for any measure ν on $\mathcal{M}$, we set $\nu(\Psi) := \int_{\mathcal{M}} \Psi(z) d\nu(z)$.

Using spectral expansion and collecting estimates on the Fourier coefficients of Hecke–Maass forms and Eisenstein series, we obtain the following effective result, which yields equidistribution when the sequence is within a certain polynomial range.

Theorem 1.1 *Let $\mathcal{M}$ be the modular surface. For any $\Psi \in C_c^\infty(\mathcal{M})$, for any $n \in \mathbb{N}$, $x \in \mathbb{R}/\mathbb{Z}$ and $y > 0$ we have*

$$\left|\delta_{n,x,y}(\Psi) - \mu_{\mathcal{M}}(\Psi)\right| \ll_\epsilon \mathcal{S}_{2,2}(\Psi)\left(y^{1/2} + n^{-1}y^{-(1/2+\theta+\epsilon)}\right),$$

and

$$\left|\delta_{n,x,y}^{\mathrm{pr}}(\Psi) - \mu_{\mathcal{M}}(\Psi)\right| \ll_\epsilon \mathcal{S}_{2,2}(\Psi)\left(y^{1/2} + n^{-1+\epsilon}y^{-(1/2+\theta+\epsilon)}\right),$$

where $\theta = 7/64$ is the current best known bound towards the Ramanujan conjecture (which implies $\theta = 0$) and $\mathcal{S}_{2,2}$ is a "L^2, order-2" Sobolev norm on $C_c^\infty(\mathcal{M})$, see Sect. 2.1.

If $\{y_n\}_{n \in \mathbb{N}}$ is a sequence of positive numbers satisfying $\lim_{n \to \infty} y_n = 0$ and $y_n \gg 1/n^\alpha$ for some fixed $\alpha \in \left(0, \frac{2}{1+2\theta}\right) = (0, \frac{64}{39})$, then Theorem 1.1 implies that for any translate $x \in \mathbb{R}/\mathbb{Z}$, both $\{\mathcal{R}_n(x, y_n)\}_{n \in \mathbb{N}}$ and $\left\{\mathcal{R}_n^{\mathrm{pr}}(x, y_n)\right\}_{n \in \mathbb{N}}$ become equidistributed on $\mathcal{M}$ with respect to $\mu_{\mathcal{M}}$ as $n \to \infty$. In particular, it gives an alternative – spectral – proof to the aforementioned results of Luethi [24] and Einsiedler–Luethi–Shah [8]. The upper bound $\frac{2}{1+2\theta}$ is the natural barrier for our spectral methods. Nevertheless, when x is a rational translate, a generalization of the symmetry (1.6) allows to go beyond this barrier, and to prove unconditionally the remaining range $\alpha \in [\frac{2}{1+2\theta}, 2)$, as holds in the case of $\{\mathcal{R}_n^{\mathrm{pr}}(0, y_n)\}_{n \in \mathbb{N}}$.

Theorem 1.2 *Let $x = p/q$ be a primitive rational number, i.e. $\gcd(p, q) = 1$. Let $\{y_n\}_{n \in \mathbb{N}}$ be a sequence of positive numbers satisfying $y_n \asymp 1/n^\alpha$ for some fixed $\alpha \in [\frac{2}{1+2\theta}, 2)$. Then both $\left\{\delta_{n,x,y_n}\right\}_{n \in \mathbb{N}_q}$ and $\left\{\delta_{n,x,y_n}^{\mathrm{pr}}\right\}_{n \in \mathbb{N}_q^{\mathrm{pr}}}$ weakly converge to $\mu_{\mathcal{M}}$ as n goes to infinity, where*

$$\mathbb{N}_q := \{n \in \mathbb{N} : \gcd(n^2, q) \mid n\} \quad \text{and} \quad \mathbb{N}_q^{\mathrm{pr}} := \{n \in \mathbb{N} : \gcd(n, q) = 1\}.$$

Remark 1.7 If q is squarefree, then the condition $\gcd(n^2, q) \mid n$ is void. Thus for such q, Theorem 1.2 (together with Theorem 1.1) confirms the equidistribution of the sample points $\mathcal{R}_n(p/q, y_n)$ (with $y_n \asymp 1/n^\alpha$) along the full set of positive integers for any $0 < \alpha < 2$.

As a byproduct of our analysis, we also have the following non-equidistribution result for rational translates, giving infinitely many explicit limiting measures. Let us first fix some notation. For each $m \in \mathbb{N}$, let

$$\mathbb{P}_m := \{n = m\ell \in \mathbb{N} : \ell \text{ is a prime number and } \ell \nmid m\}. \tag{1.8}$$

For each $Y > 0$, we denote by μ_Y the uniform probability measure supported on the closed horocycle $\mathcal{H}_Y$. For each $m \in \mathbb{N}$ and $Y > 0$, we define the probability measure $\nu_{m,Y}$ on $\mathcal{M}$ by

$$\nu_{m,Y} := \frac{1}{m} \sum_{d \mid m} \varphi(\tfrac{m}{d}) \mu_{d^2 Y}. \tag{1.9}$$

Theorem 1.3 *Keep the notation as above. Let $x = p/q$ be a primitive rational number and let $\{y_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers.*

(1) *If $y_n = c/n^2$ for some constant $c > 0$, then for any $m \in \mathbb{N}_q$ and for any $\Psi \in C_c^\infty(\mathcal{M})$*

$$\lim_{\substack{n\to\infty \\ \gcd(n,q)=1}} \delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) = \mu_{\frac{1}{cq^2}}(\Psi) \quad and \quad \lim_{\substack{n\to\infty \\ n\in\mathbb{P}_m}} \delta_{n,x,y_n}(\Psi) = \nu_{m,\frac{\gcd(m,q)^2}{cq^2}}(\Psi).$$

(2) *If $\lim_{n\to\infty} n^2 y_n = 0$, then both sequences $\{\mathcal{R}_n(x, y_n)\}_{n\in\mathbb{N}}$ and $\{\mathcal{R}_n^{\mathrm{pr}}(x, y_n)\}_{n\in\mathbb{N}}$ fully escape to the cusp of $\mathcal{M}$.*

Our next result shows that, similar to the rational translate case, equidistribution fails for generic translates as soon as $\{y_n\}_{n\in\mathbb{N}}$ decays logarithmically faster than $1/n^2$.

Theorem 1.4 *Let $d_\mathcal{M}(\cdot, \cdot)$ be the distance function on $\mathcal{M}$ induced from the hyperbolic distance function on $\mathbb{H}$. Fix $\Gamma z_0 \in \mathcal{M}$. Let $\{y_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers satisfying $y_n \asymp 1/(n^2 \log^\beta n)$ for some fixed $0 < \beta < 2$. Then for almost every $x \in \mathbb{R}/\mathbb{Z}$*

$$\varlimsup_{n\to\infty} \frac{\inf_{\Gamma z \in \mathcal{R}_n(x, y_n)} d_\mathcal{M}(\Gamma z_0, \Gamma z)}{\log\log n} \geq \min\{\beta, 2 - \beta\}. \tag{1.10}$$

This implies that for almost every $x \in \mathbb{R}/\mathbb{Z}$, there exists an unbounded subsequence of $\mathbb{N}$ such that along this subsequence

$$\inf_{\Gamma z \in \mathcal{R}_n(x, y_n)} d_\mathcal{M}(\Gamma z_0, \Gamma z) \geq (\alpha - \epsilon) \log\log n,$$

where $\alpha = \min\{\beta, 2 - \beta\}$. That is, for almost every $x \in \mathbb{R}/\mathbb{Z}$, all the sample points $\mathcal{R}_n(x, y_n)$ (and hence also $\mathcal{R}_n^{\mathrm{pr}}(x, y_n)$) are moving towards the cusp of $\mathcal{M}$ along this subsequence, and eventually escape to the cusp as n in this subsequence goes to infinity.

Our proof of Theorem 1.4 relies on connections to Diophantine approximation theory. This viewpoint comes with inherent limitations; in the specific setting $y_n \asymp 1/(n^2 \log^\beta n)$, Khintchine's approximation theorem guarantees full escape to the cusp almost surely, but this argument does not extend to any sequence $\{y_n\}_{n\in\mathbb{N}}$ that decays polynomially faster than $1/n^2$, see Sect. 1.3 for a more detailed discussion. It is thus interesting to study the cases when $\{y_n\}_{n\in\mathbb{N}}$ is beyond the ranges in Theorems 1.1 and 1.4.

Indeed, the rest of our results deal with sequences $\{y_n\}_{n\in\mathbb{N}}$ that can decay *arbitrarily fast*, and give both positive and negative results. This is the main novelty of this paper; the handling of cases in which the sample points can be *arbitrarily sparse* on the closed horocycles they lie on. We now state the main novel aspect of this paper:

Theorem 1.5 *For any sequence of positive numbers $\{c_n\}_{n\in\mathbb{N}}$, there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ satisfying $0 < y_n < c_n$ for each $n \in \mathbb{N}$ and such that for almost every $x \in \mathbb{R}/\mathbb{Z}$ the set of limiting measures of $\{\delta_{n,x,y_n}\}_{n\in\mathbb{N}}$ and $\{\delta_{n,x,y_n}^{\mathrm{pr}}\}_{n\in\mathbb{N}}$ both contain the uniform measure $\mu_{\mathcal{M}}$, the zero measure, and singular probability measures.*

Theorem 1.5 is a sum of three more precise theorems, which each handles a specific limiting measure, and which we discuss in the next subsection.

1.3 Discussion of the results

Our proofs of Theorems 1.1 and 1.2 rely on spectral estimates collected in the recent paper of Kelmer and Kontorovich [18], with a necessary refinement of [18, (3.6)] in the form of Proposition 3.3, which comes at the cost of a higher degree Sobolev norm. This strategy is standard and is also found in [4,16,27,31], to name just a few recent papers on related problems. The analysis in [18] was carried out in a more general setting, namely for the congruence covers $\Gamma_0(p)\backslash\mathbb{H}$ with p a prime number. Theorem 1.1 can be extended to that more general setting, see Remark 3.11. With these spectral estimates in hand, we further prove an effective non-equidistribution result for rational translates from which part (1) of Theorem 1.3 follows, see Theorem 3.10. Part (2) of Theorem 1.3 is an easy application of the symmetry (1.6).

Remark 1.11 As was pointed out to us by Asaf Katz, we could also have used the estimates from [31, Proposition 3.1] in place of [18, Proposition 3.4], which in our specific setting, give the same equidistribution range (with a higher degree Sobolev norm). We also mention that the estimates in [31, Proposition 3.1] are valid in the setting of $\Gamma_0(q)\backslash \mathrm{SL}_2(\mathbb{R})$ with $q \in \mathbb{N}$, and thus imply an effective equidistribution result analogous to Theorem 1.1 in this generality.

As mentioned earlier, a generalization of the symmetry (1.6) is available for rational translates but breaks down for irrational translates. To handle irrational translates, we approximate them by rational ones to apply the symmetry relation, see Lemma 4.2. This is where Diophantine approximation kicks in. Similar ideas were also used in [27,

Section 7] to construct counterexamples in their setting. In fact, we prove Theorem 1.4 by proving a more general result that captures the cusp excursion rates of the sample points $\mathcal{R}_n(x, y_n)$ in terms of the Diophantine properties of the translate x, see Theorem 4.3. Theorem 1.4 will then follow from Theorem 4.3 by imposing a Diophantine condition which ensures cusp excursion, while also holds for almost every translate thanks to Khintchine's approximation theorem. This Diophantine condition accounts for the tight restrictions on $\{y_n\}_{n\in\mathbb{N}}$ in Theorem 1.4. On the other hand, assuming an even stronger Diophantine condition (which holds for a null set of translates), we can handle sequences decaying polynomially faster than $1/n^2$ with a much faster excursion rate towards the cusp, see Theorem 4.4. We also prove a non-equidistribution result (which, this time, holds for *every x*) when $y_n = c/n^2$ and the constant c is restricted to some range, see Theorem 4.5. The trade-off of this upgrade from Theorem 1.4 to the everywhere non-equidistribution result is that we can no longer prove the full escape to the cusp along subsequences as in Theorem 1.4.

As mentioned before, Theorem 1.5 follows from three more precise theorems which each handles a specific limiting measure. Our first result confirms equidistribution almost surely along a fixed subsequence of $\mathbb{N}$ for *any* sequence $\{y_n\}_{n\in\mathbb{N}}$ decaying at least polynomially.

Theorem 1.6 *Fix $\alpha > 0$. Then there exists a fixed unbounded subsequence $\mathcal{N} \subset \mathbb{N}$ such that for any sequence of positive numbers $\{y_n\}_{n\in\mathbb{N}}$ satisfying $y_n \ll n^{-\alpha}$ and for almost every $x \in \mathbb{R}/\mathbb{Z}$, both δ_{n,x,y_n} and $\delta_{n,x,y_n}^{\mathrm{pr}}$ weakly converge to $\mu_{\mathcal{M}}$ as $n \in \mathcal{N}$ goes to infinity.*

Remark 1.12 It will be clear from our proof that one can take $\mathcal{N} \subset \mathbb{N}$ to be any subsequence satisfying $\sum_{n\in\mathcal{N}} n^{-c} < \infty$ for some positive $c < \min\{\frac{\alpha}{2}, 1 - 2\theta\}$, e.g. we may take $\mathcal{N} = \{\lfloor n^\kappa \rfloor\}_{n\in\mathbb{N}}$ for any $\kappa > 1/\min\{\frac{\alpha}{2}, 1 - 2\theta\}$.

Theorem 1.6 follows from a second moment estimate for the discrepancies $|\delta_{n,x,y} - \mu_{\mathcal{M}}|$ and $|\delta_{n,x,y}^{\mathrm{pr}} - \mu_{\mathcal{M}}|$ along the closed horocycle $\mathcal{H}_y$ (Theorem 5.2) together with a standard Borel–Cantelli type argument. This was also the strategy used in [27] when studying the Kronecker sequences in (1.3). Along these lines, they deduce from spectral estimates the equidistribution for almost every $\beta \in \mathbb{R}$ along a fixed subsequence $\{n^k\}_{n\in\mathbb{N}}$ when $y_n \asymp n^{-\alpha}$ with $k \in \mathbb{N}$ depending on $\alpha > 0$. Then, using a continuity argument, this result is upgraded to the equidistribution along the full sequence of positive integers, see [27, Section 4]. This continuity argument fails in our situation. Instead of applying directly spectral estimates to the second moment formulas, we express the latter in terms of certain Hecke operators (Proposition 5.1), and rely on available (spectral) bounds for their operator norm, see [10]. Contrarily to spectral estimates, the recourse to Hecke operators allows us to have a uniform subsequence $\mathcal{N}$ which is valid for all $\{y_n\}_{n\in\mathbb{N}}$ decaying at least polynomially.

Next, we show that there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ decaying arbitrarily rapidly such that for almost every x, $\mathcal{R}_n(x, y_n)$ (and thus also $\mathcal{R}_n^{\mathrm{pr}}(x, y)$) escapes to the cusp with a certain rate along subsequences.

Theorem 1.7 *Fix $\Gamma z_0 \in \mathcal{M}$. For any sequence of positive numbers $\{c_n\}_{n\in\mathbb{N}}$, there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ satisfying $0 < y_n < c_n$ for each $n \in \mathbb{N}$ and such that for*

almost every $x \in \mathbb{R}/\mathbb{Z}$

$$\varlimsup_{n \to \infty} \frac{\inf_{\Gamma z \in \mathcal{R}_n(x, y_n)} d_{\mathcal{M}}(\Gamma z_0, \Gamma z)}{\log \log n} \geq 1. \tag{1.13}$$

Finally, we show that escape to the cusp is not the only obstacle to equidistribution.

Theorem 1.8 *Let $m \in \mathbb{N}$ and $Y > 0$ satisfy $m^2 Y > 1$. Let $\mathbb{P}_m \subset \mathbb{N}$ and $\nu_{m,Y}$ be as defined in (1.8) and (1.9) respectively. For any sequence of positive numbers $\{c_n\}_{n \in \mathbb{P}_m}$, there exists a sequence $\{y_n\}_{n \in \mathbb{P}_m}$ satisfying $0 < y_n < c_n$ for all $n \in \mathbb{P}_m$ such that for almost every $x \in \mathbb{R}/\mathbb{Z}$, the set of limiting measures of $\{\delta_{n,x,y_n}\}_{n \in \mathbb{P}_m}$ contains $\nu_{m,Y}$.*

Remark 1.14 We note that $\mathbb{P}_1$ is the set of prime numbers and $\nu_{1,Y} = \mu_Y$. Since

$$\delta_{p,x,y}^{\text{pr}}(\Psi) = \frac{p}{p-1} \delta_{p,x,y}(\Psi) + O(p^{-1} \|\Psi\|_\infty)$$

whenever p is a prime number, when $m = 1$ the conclusion of Theorem 1.8 also holds for the sequence $\{\delta_{n,x,y_n}^{\text{pr}}\}_{n \in \mathbb{P}_1}$. We also note that it will be clear from our proof that Theorems 1.7 and 1.8 can be combined. In fact, our argument shows that there always exists a sequence $\{y_n\}_{n \in \mathbb{N}}$ decaying faster than any prescribed sequence such that for almost every $x \in \mathbb{R}/\mathbb{Z}$ the set of limiting measures of $\{\delta_{n,x,y_n}\}_{n \in \mathbb{N}}$ contains the trivial measure and $\nu_{m,Y}$ for any finitely many pairs $(m, Y) \in \mathbb{N} \times \mathbb{R}_{>0}$ with $m^2 Y > 1$, see Remark 7.26. Moreover, in view of Theorem 1.6 if $y_n \ll n^{-\alpha}$ for some $\alpha > 0$, then it also contains the hyperbolic area $\mu_{\mathcal{M}}$ almost surely.

For the rest of this introduction we describe the strategy of our proof to Theorem 1.7 (Theorem 1.8 follows from similar ideas). To detect cusp excursions, we study for each $n \in \mathbb{N}$ the occurrence of the events

$$\Gamma\left(x + \frac{j}{n} + i y_n\right) \in \mathcal{C} \quad \text{for all } 0 \leq j \leq n - 1, \tag{1.15}$$

where $\mathcal{C} \subset \mathcal{M}$ is some fixed cusp neighborhood of $\mathcal{M}$. More precisely, we determine when the limsup set $I_\infty = \varlimsup_{n \to \infty} I_n$ is of full measure, where for each $n \in \mathbb{N}$,

$$I_n := \{x \in \mathbb{R}/\mathbb{Z} : \mathcal{R}_n(x, y_n) \subset \mathcal{C}\}$$

consists of translates $x \in \mathbb{R}/\mathbb{Z}$ for which the events in (1.15) occur. This requires to study the left regular $u_{1/n}$-action on $\mathcal{C} \subset \mathcal{M}$ and thus calls for the underlying lattice to be normalized by $u_{1/n}$. Therefore, we construct an explicit tower of coverings $\{\Gamma_n \backslash \mathbb{H}\}_{n \in \mathbb{N}}$ in which each Γ_n is a congruence subgroup normalized by $u_{1/n}$. We note that the existence of such $\Gamma_n < \Gamma$ is the starting point of our proof and it relies on the assumption that $\Gamma = SL_2(\mathbb{Z})$; this construction would fail for Γ replaced by a non-arithmetic lattice.

The key ingredient of the proof will be a sufficient condition which states that if a point $\Gamma_n(x + i y_n) \in \Gamma_n \backslash \mathbb{H}$ visits a certain cusp neighborhood $\mathcal{C}_n$ on $\Gamma_n \backslash \mathbb{H}$, then the events in (1.15) will be realized for $x \in \mathbb{R}/\mathbb{Z}$, that is, $x \in I_n$, see Lemma 7.6.

Using this sufficient condition, we can then relate the measure of I_n to the proportion of certain closed horocycles on $\Gamma_n \backslash \mathbb{H}$ visiting the cusp neighborhood $\mathcal{C}_n \subset \Gamma_n \backslash \mathbb{H}$, which in turn, using the equidistribution of expanding closed horocycles on $\Gamma_n \backslash \mathbb{H}$, can be estimated for y_n sufficiently small. Since the sets I_n also need to satisfy certain quasi-independence conditions for I_∞ to have full measure (Lemma 2.5), we need to apply the equidistribution of certain subsegments of the expanding closed horocycles on $\Gamma_n \backslash \mathbb{H}$. More precisely, at the n-th step these subsegments will be taken to be the sets I_m for all $m < n$. These subsegment are finite disjoint unions of subintervals whose number and size depend sensitively on the height parameters $\{y_m\}_{m<n}$, see Remark 6.3. If there would exist an effective equidistribution result which would be insensitive to the geometry of these subsegments, that is, for which the error term depends only on the measure of these subsegments, then we would have an effective control on the sequence $\{y_n\}_{n\in\mathbb{N}}$ in Theorem 1.7 (and similarly also in Theorem 1.8). However, it is not clear to us whether one should expect such an effective equidistribution result.

Finally, we note that it was communicated to us by Strömbergsson that using a number theoretic interpretation of the aforementioned sufficient condition and some elementary estimates, one can alternatively prove Theorem 1.7 without going into these congruence covers, see Remark 7.17.

Structure of the paper

In Sect. 2, we collect some preliminary results that will be needed in the rest of the paper. In Sect. 3, we prove a key spectral estimate (Proposition 3.3) and proceed to prove Theorems 1.1 and 1.2. In Sect. 4, we prove Theorems 4.3 and 4.5 by examining the connections between Diophantine approximations and cusp excursions on the modular surface. In Sect. 5, we prove Theorem 1.6 by proving a second moment bound using Hecke operators. In Sect. 6, we study the left regular action of a normalizing element on the set of cusp neighborhoods of a congruence cover of the modular surface. Building on the results, we prove Theorems 1.7 and 1.8 in Sect. 7.

Notation

For two positive quantities A and B, we will use the notation $A \ll B$ or $A = O(B)$ to mean that there is a constant $c > 0$ such that $A \leq cB$, and we will use subscripts to indicate the dependence of the constant on parameters. We will write $A \asymp B$ for $A \ll B \ll A$. For any $z \in \mathbb{H}$ we denote by $e(z) := e^{2\pi i z}$. For any $n \in \mathbb{N}$, we denote by $\prod_{d|n}$ the product over all positive divisors of n, and by $\prod_{\substack{p|n \\ prime}}$ the product over all prime divisors of n. For any $x \geq 0$ and $n \in \mathbb{N}$, $\sigma_x(n) := \sum_{d|n} d^x$ is the power-x divisor function which satisfies the estimate $\sigma_x(n) \ll_\epsilon n^{x+\epsilon}$ for any small $\epsilon > 0$.

2 Preliminaries

Let $G = \mathrm{SL}_2(\mathbb{R})$. We consider the Iwasawa decomposition $G = NAK$ with

$$N = \{u_x : x \in \mathbb{R}\}, \quad A = \{a_y : y > 0\}, \quad K = \{k_\theta : 0 \le \theta < 2\pi\},$$

where $u_x = \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$, $a_y = \left(\begin{smallmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{smallmatrix}\right)$ and $k_\theta = \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$ respectively. Under the coordinates $g = u_x a_y k_\theta$ on G, the Haar measure is given (up to scalars) by

$$dg = y^{-2} dx\, dy\, d\theta.$$

The group G acts on the upper half plane $\mathbb{H} = \{z = x + iy \in \mathbb{C} : y > 0\}$ via the Möbius transformation: $gz = \frac{az+b}{cz+d}$ for any $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in G$ and $z \in \mathbb{H}$. This action preserves the hyperbolic area $d\mu(z) = y^{-2} dx\, dy$ and induces an identification between G/K and $\mathbb{H}$.

Let $\Gamma < G$ be a *lattice*, that is, Γ is a discrete subgroup of G such that the corresponding hyperbolic surface $\Gamma \backslash \mathbb{H}$ has finite area (with respect to μ). We denote by $\mu_\Gamma := \mu(\Gamma \backslash \mathbb{H})^{-1} \mu$ the normalized hyperbolic area on $\Gamma \backslash \mathbb{H}$ such that $\mu_\Gamma(\Gamma \backslash \mathbb{H}) = 1$. We note that when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ then $\mu_\Gamma = \mu_{\mathcal{M}}$ with $\mu_{\mathcal{M}}$ the normalized hyperbolic area on the modular surface $\mathcal{M}$ given as in the introduction. We note that in this case it is well known $\mu(\mathcal{M}) = \pi/3$, and hence

$$d\mu_{\mathcal{M}}(z) = \frac{3}{\pi} \frac{dx\, dy}{y^2}. \tag{2.1}$$

Using the above identification between $\mathbb{H}$ and G/K we can identify the hyperbolic surface $\Gamma \backslash \mathbb{H}$ with the locally symmetric space $\Gamma \backslash G / K$. We can thus view subsets of $\Gamma \backslash \mathbb{H}$ as right K-invariant subsets of $\Gamma \backslash G$. Similarly, we can view functions on $\Gamma \backslash \mathbb{H}$ as right K-invariant functions on $\Gamma \backslash G$. We note that using the above description of the Haar measure, the probability Haar measure on $\Gamma \backslash G$ (when restricted to the subfamily of right K-invariant subsets) coincides with the normalized hyperbolic area μ_Γ on $\Gamma \backslash \mathbb{H}$.

2.1 Sobolev norms

In this subsection we record some useful properties about Sobolev norms. Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R})$ be the Lie algebra of G. Fix a basis $\mathcal{B} = \{X_1, X_2, X_3\}$ for $\mathfrak{g}$, and given a smooth test function $\Psi \in C^\infty(\Gamma \backslash G)$ we define the "L^p, order-d" Sobolev norm $\mathcal{S}_{p,d}^\Gamma(\Psi)$ as

$$\mathcal{S}_{p,d}^\Gamma(\Psi) := \sum_{\mathrm{ord}(\mathscr{D}) \le d} \|\mathscr{D}\Psi\|_{L^p(\Gamma \backslash G)},$$

where $\mathscr{D}$ runs over all monomials in $\mathcal{B}$ of order at most d, and the L^p-norm is with respect to the normalized Haar measure on $\Gamma \backslash G$.

For any $\Psi \in C^\infty(\Gamma \backslash G)$ (which we think of a smooth left Γ-invariant function on G) and for any $h \in G$ we denote by $L_h \Psi(g) := \Psi(h^{-1}g)$ the left regular h-action on Ψ. It is easy to check that $L_h \Psi \in C^\infty(h\Gamma h^{-1} \backslash G)$, and since taking Lie derivatives commutes with the left regular action, we have

$$\mathcal{S}_{p,d}^{\Gamma}(\Psi) = \mathcal{S}_{p,d}^{h\Gamma h^{-1}}(L_h \Psi). \tag{2.2}$$

Next we note that using the product rule for Lie derivatives (see e.g. [21, p. 90]), the triangle inequality and the Cauchy–Schwarz inequality, for any monomial $\mathscr{D}$ of order $k \leq d$ we have for any smooth functions $\Psi_1, \Psi_2 \in C^\infty(\Gamma \backslash G)$

$$\|\mathscr{D}\Psi_1 \Psi_2\|_{L^p(\Gamma \backslash G)} \ll_k \mathcal{S}_{2p,k}^{\Gamma}(\Psi_1)\mathcal{S}_{2p,k}^{\Gamma}(\Psi_2) \leq \mathcal{S}_{2p,d}^{\Gamma}(\Psi_1)\mathcal{S}_{2p,d}^{\Gamma}(\Psi_2).$$

In particular this implies that

$$\mathcal{S}_{p,d}^{\Gamma}(\Psi_1 \Psi_2) \ll_d \mathcal{S}_{2p,d}^{\Gamma}(\Psi_1)\mathcal{S}_{2p,d}^{\Gamma}(\Psi_2). \tag{2.3}$$

Finally, we note that if $\Gamma' < \Gamma$ is a finite-index subgroup of Γ, then there is a natural embedding $C^\infty(\Gamma \backslash G) \hookrightarrow C^\infty(\Gamma' \backslash G)$ since each $\Psi \in C^\infty(\Gamma \backslash G)$ can be viewed as a smooth left Γ'-invariant function on G. Since the Sobolev norms are defined with respect to the normalized Haar measure on the corresponding homogeneous space, we have for $\Gamma' < \Gamma$ of finite index and $\Psi \in C^\infty(\Gamma \backslash G)$

$$\mathcal{S}_{p,d}^{\Gamma'}(\Psi) = \mathcal{S}_{p,d}^{\Gamma}(\Psi). \tag{2.4}$$

2.2 Spectral decomposition

Let $\Gamma < G$ be a *non-uniform* lattice, that is, Γ is a lattice and $\Gamma \backslash \mathbb{H}$ is not compact. Let $\Delta = -y^2(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2})$ be the hyperbolic Laplace operator. It is a second order differential operator acting on $C^\infty(\Gamma \backslash \mathbb{H})$ and extends uniquely to a self-adjoint and positive semi-definite operator on $L^2(\Gamma \backslash \mathbb{H})$. Since Γ is non-uniform, the spectrum of Δ is composed of a continuous part (spanned by Eisenstein series) and a discrete part (spanned by Maass forms) which further decomposes as the cuspidal spectrum and the residual spectrum. The residual spectrum always contains the constant functions (coming from the trivial pole of the Eisenstein series). If Γ is a *congruence subgroup*, that is, Γ contains a *principal congruence subgroup*

$$\Gamma(n) := \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv I_2 \pmod{n}\}$$

for some $n \in \mathbb{N}$, then the residual spectrum consists only of the constant functions, see e.g. [15, Theorem 11.3].

Let $\{\phi_k\}$ be an orthonormal basis of the space of cusp forms that are eigenfunctions of the Laplace operator Δ. Explicitly, for each ϕ_k there exists $\lambda_k \geq 0$ such that

$$\Delta \phi_k = \lambda_k \phi_k = s_k(1 - s_k)\phi_k = \left(\tfrac{1}{4} + r_k^2\right)\phi_k.$$

Selberg's eigenvalue conjecture states that for congruence subgroups, $\lambda_k \geq 1/4$, or equivalently, there is no $r_k \in i(0, 1/2)$. Selberg's conjecture is known to be true for the modular surface $\mathcal{M}$, and more generally, the best known bound towards this conjecture is currently $\lambda_k \geq \frac{1}{4} - \theta^2$, with $\theta = 7/64$, which follows from the bound of Kim and Sarnak towards the Ramanujan conjecture, see [19, p. 176].

Let now $\Gamma = \text{SL}_2(\mathbb{Z})$. In the notation introduced at the beginning of this section, the Eisenstein series for the modular group Γ at the cusp ∞ is defined for $\mathfrak{Re}(s) > 1$ by

$$E(z, s) = \sum_{\gamma \in (\Gamma \cap \pm N) \backslash \Gamma} \mathfrak{Im}(\gamma z)^s \tag{2.5}$$

with a meromorphic continuation to $s \in \mathbb{C}$. Moreover, for any $s \in \mathbb{C}$, $E(\cdot, s)$ is an eigenfunction of the Laplace operator with eigenvalue $s(1 - s)$.

Let $\Psi \in L^2(\mathcal{M})$ and we have the following spectral decomposition (see [15, Theorems 4.7 and 7.3])

$$\Psi(z) = \mu_{\mathcal{M}}(\Psi) + \sum_{r_k \geq 0} \langle \Psi, \phi_k \rangle \phi_k(z)$$
$$+ \frac{1}{4\pi} \int_{-\infty}^{\infty} \langle \Psi, E(\cdot, \tfrac{1}{2} + ir) \rangle E(z, \tfrac{1}{2} + ir) dr, \tag{2.6}$$

where the convergence holds in the L^2-norm topology, and is pointwise if $\Psi \in C_c^{\infty}(\mathcal{M})$. As a direct consequence we have for $\Psi \in L^2(\mathcal{M})$,

$$\|\Psi\|_2^2 = |\mu_{\mathcal{M}}(\Psi)|^2 + \sum_{r_k \geq 0} |\langle \Psi, \phi_k \rangle|^2 + \frac{1}{4\pi} \int_{-\infty}^{\infty} \left| \langle \Psi, E(\cdot, \tfrac{1}{2} + ir) \rangle \right|^2 dr. \tag{2.7}$$

2.3 Hecke operators

The spectral theory of $\mathcal{M}$ has extra structure due to the existence of *Hecke operators*. The main goal of this subsection is to prove an operator norm bound for Hecke operators and the main reference is [15, Section 8.5]. For any $n \in \mathbb{N}$ define the set

$$\mathcal{L}_n := \left\{ n^{-1/2} g : g \in M_2(\mathbb{Z}), \ \det(g) = n \right\} \subset G, \tag{2.8}$$

where $M_2(\mathbb{Z})$ is the space of two by two integral matrices. The *n-th Hecke operator* T_n is defined by that for any $\Psi \in L^2(\mathcal{M})$

$$T_n(\Psi)(z) = \frac{1}{n^{1/2}} \sum_{\gamma \in \Gamma \backslash \mathcal{L}_n} \Psi(\gamma z).$$

The Hecke operator T_n is a self-adjoint operator on $L^2(\mathcal{M})$ and since T_n commutes with the Laplace operator Δ (since Δ is defined via right multiplication and T_n is defined via left multiplication) the orthonormal basis of the space of cusp forms $\{\phi_k\}$ can be chosen consisting of joint eigenfunctions of all T_n, that is,

$$T_n \phi_k = \lambda_{\phi_k}(n) \phi_k.$$

On the other hand, for any $r \in \mathbb{R}$ the Eisenstein series $E(z, 1/2 + ir)$ is an eigenfunction of T_n with eigenvalue $\lambda_r(n) := \sum_{d|n} \left(\frac{n}{d^2}\right)^{ir}$, see [15, Equation (8.33)]. It is clear that $|\lambda_r(n)| \leq \sigma_0(n)$ with $\sigma_0(n)$ the divisor function. For the eigenvalue of cusp forms it is conjectured (Ramanujan-Petersson) that for any above ϕ_k and for any $n \in \mathbb{N}$

$$\left|\lambda_{\phi_k}(n)\right| \leq \sigma_0(n).$$

The aforementioned bound of Sarnak and Kim [19] implies that

$$\left|\lambda_{\phi_k}(n)\right| \leq \sigma_0(n) n^{7/64}.$$

Using these bounds on eigenvalues and the above spectral decomposition (2.6) and (2.7) we have the following bound on the operator norm of the Hecke operator, see also [10, pp. 172-173].

Proposition 2.1 *For any $\Psi \in L^2(\mathcal{M})$ and for any $n \in \mathbb{N}$ we have*

$$\langle \Psi_0, T_n(\Psi_0) \rangle_{L^2(\mathcal{M})} \ll_\epsilon n^{\theta+\epsilon} \|\Psi\|_2^2,$$

where $\Psi_0 := \Psi - \mu_\mathcal{M}(\Psi)$ and $\theta = 7/64$ as before.

2.3.1 Hecke operators attached to a group element

Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and let $\mathcal{M} = \Gamma \backslash \mathbb{H}$ be the modular surface as above. There is another type of Hecke operators on $L^2(\mathcal{M})$ defined via a group element in $\mathrm{SL}_2(\mathbb{Q})$. Namely, for each $h \in \mathrm{SL}_2(\mathbb{Q})$ the *Hecke operator attached to h*, denoted by $\widetilde{T}_h$, is defined by that for any $\Psi \in L^2(\mathcal{M})$

$$\widetilde{T}_h(\Psi)(z) = \frac{1}{\#(\Gamma \backslash \Gamma h \Gamma)} \sum_{g \in \Gamma \backslash \Gamma h \Gamma} \Psi(gz), \tag{2.9}$$

where $\Gamma h \Gamma = \{\gamma_1 h \gamma_2 : \gamma_1, \gamma_2 \in \Gamma\}$ is the double coset attached to h. We note that $\widetilde{T}_h$ is well-defined since Ψ is left Γ-invariant.

For our purpose, we will need another expression for $\widetilde{T}_h$. For any $h \in \mathrm{SL}_2(\mathbb{Q})$ we denote by $\Gamma^h := \Gamma \cap h^{-1} \Gamma h$. We note that the map from Γ to $\Gamma \backslash \Gamma h \Gamma$ sending $\gamma \in \Gamma$ to $\Gamma h \gamma$ induces an identification between $\Gamma^h \backslash \Gamma$ and $\Gamma \backslash \Gamma h \Gamma$. This identification induces

the following alternative expression for $\widetilde{T}_h$:

$$\widetilde{T}_h(\Psi)(g) = \frac{1}{[\Gamma : \Gamma^h]} \sum_{\gamma \in \Gamma^h \backslash \Gamma} \Psi(h\gamma g). \tag{2.10}$$

It is clear from the definition that $\widetilde{T}_h$ is defined only up to representatives for the double coset $\Gamma h \Gamma$, that is, $\widetilde{T}_h = \widetilde{T}_{h'}$ whenever $\Gamma h \Gamma = \Gamma h' \Gamma$. For a fixed $h \in \mathrm{SL}_2(\mathbb{Q})$, we call $n \in \mathbb{N}$ *the degree of h* if n is the smallest positive integer such that $nh \in M_2(\mathbb{Z})$. Using elementary column and row operations one can see that for $h \in \mathrm{SL}_2(\mathbb{Q})$ with degree n

$$\Gamma h \Gamma = \Gamma \operatorname{diag}(1/n, n) \Gamma = \left\{ n^{-1} g : g \in M_2(\mathbb{Z}), \ det(g) = n^2, \ \gcd(g) = 1 \right\} \subset G, \tag{2.11}$$

where $\gcd(g)$ is the greatest common divisor of the entries of g. Thus we can parameterize the Hecke operators by their degrees, that is, we will denote by $\widetilde{T}_n := \widetilde{T}_h$ for any $h \in \mathrm{SL}_2(\mathbb{Q})$ with degree n. We also note that by direct computation when $h = \operatorname{diag}(1/n, n)$ we have $\Gamma^h = \Gamma_0(n^2)$, implying that for any $h \in \mathrm{SL}_2(\mathbb{Q})$ with degree n (see e.g. [6, Section 1.2])

$$\nu_n := \#(\Gamma \backslash \Gamma h \Gamma) = [\Gamma : \Gamma^h] = [\Gamma : \Gamma_0(n^2)] = n^2 \prod_{\substack{p | n \\ \text{prime}}} \left(1 + p^{-1}\right). \tag{2.12}$$

Now using the description (2.11) we have the double coset decomposition

$$\mathcal{L}_{n^2} = \bigsqcup_{d | n} \Gamma \begin{pmatrix} d^{-1} & 0 \\ 0 & d \end{pmatrix} \Gamma.$$

This decomposition together with the definitions (2.8), (2.9) and (2.12) implies the relation

$$n T_{n^2} = \sum_{d | n} \nu_d \widetilde{T}_d.$$

Thus by the Möbius inversion formula we have

$$\widetilde{T}_n = \frac{n}{\nu_n} \sum_{d | n} \frac{\mu(d)}{d} T_{n^2 / d^2}. \tag{2.13}$$

Using this relation and Proposition 2.1 we can prove the following operator norm bounds for $\widetilde{T}_n$ which we will later use, see also [3, Theorem 1.1] for such bounds in a much greater generality.

Proposition 2.2 *Keep the notation as in Proposition* 2.1. *For any* $\Psi \in L^2(\mathcal{M})$ *and for any* $n \in \mathbb{N}$ *we have*

$$\langle \Psi_0, \widetilde{T}_n(\Psi_0) \rangle_{L^2(\mathcal{M})} \ll_\epsilon n^{-1+2\theta+\epsilon} \|\Psi\|_2^2.$$

Proof. By Proposition 2.1 and using the relation (2.13), the trivial estimates $|\mu(d)| \leq 1$ and $\nu_n \geq n^2$ and the triangle inequality we have

$$\langle \Psi_0, \widetilde{T}_n(\Psi_0) \rangle \leq n^{-2} \sum_{d|n} (n/d) \langle \Psi_0, T_{n^2/d^2}(\Psi_0) \rangle \ll_\epsilon n^{-2} \sum_{d|n} (n/d)^{1+2\theta+2\epsilon} \|\Psi\|_2^2$$

$$= n^{-1+2\theta+2\epsilon} \sigma_{-1+2\theta+2\epsilon}(n) \|\Psi\|_2^2 \ll_\epsilon n^{-1+2\theta+\epsilon} \|\Psi\|_2^2. \qquad \square$$

2.4 Equidistribution of subsegments of expanding closed horocycles

We record a special case of Sarnak's result [28, Theorem 1] on effective equidistribution of expanding closed horocycles, namely:

Proposition 2.3 *Let* $\Gamma < \mathrm{SL}_2(\mathbb{Z})$ *be a congruence subgroup and assume that* Γ *has a cusp at* ∞ *with width one. Then for any* $\Psi \in C^\infty(\Gamma \backslash \mathbb{H}) \cap L^2(\Gamma \backslash \mathbb{H})$ *satisfying* $\|\Delta \Psi\|_2 < \infty$ *and for any* $0 < y < 1$ *we have*

$$\left| \int_0^1 \Psi(x+iy)dx - \mu_\Gamma(\Psi) \right| \ll \|\Psi\|_2^{3/4} \|\Delta \Psi\|_2^{1/4} y^{1/2}, \qquad (2.14)$$

where the implied constant is absolute, independent of Γ, Ψ *and* y, *and the* L^2*-norm is with respect to the normalized hyperbolic area* μ_Γ.

Remark 2.15 We omit the proof here and refer the reader to [18, (3.5)]. We note that while [18] only deals with the case when $\Gamma = \Gamma_0(p)$ with p a prime number, the proof there works for general congruence subgroups, given that they have trivial residual spectrum; see [15, Theorem 11.3].

We will also need the following (non-effective) equidistribution result replacing the whole closed horocycle by a fixed subsegment:

Proposition 2.4 *Let* $\Gamma < \mathrm{SL}_2(\mathbb{Z})$ *be as in Proposition* 2.3. *Let* $I \subset (0, 1)$ *be an open interval, then for any* $\Psi \in C_c(\Gamma \backslash \mathbb{H})$ *we have*

$$\lim_{y \to 0^+} \frac{1}{|I|} \int_I \Psi(x+iy)dx = \mu_\Gamma(\Psi). \qquad (2.16)$$

The proof of Proposition 2.4 uses Margulis' thickening trick [25] and mixing property of the geodesic flow on the unit tangent bundle of $\Gamma \backslash \mathbb{H}$; this approach is also effective, see e.g. [17, Proposition 2.3]. A proof of (2.16) using spectral methods was also sketched in [12, Theorem 1′]. We also note that both equidistribution results in Propositions 2.3 and 2.4 can be lifted to the unit tangent bundle of $\Gamma \backslash \mathbb{H}$ (with necessary

modifications to the error term in (2.14)); since we will be only working in the hyperbolic surface level, we state these two results in the current format for convenience of our discussion. We further refer the reader to [13,30] for some much stronger effective equidistribution results regarding long enough (varying) subsegments on expanding closed horocycles.

Remark 2.17 Proposition 2.4 can be equivalently stated as following: For any fixed open interval $I \subset (0, 1)$, the measures $\mu_{I,y}$ weakly converge to μ_Γ as $y \to 0^+$, where for any $y \in (0, 1)$ and $\Psi \in C_c(\Gamma\backslash\mathbb{H})$, $\mu_{I,y}(\Psi) := \frac{1}{|I|} \int_I \Psi(x + iy)dx$. Thus by the Portmanteau theorem, (2.16) extends to $\Psi = \chi_B$ with $B \subset \Gamma\backslash\mathbb{H}$ a Borel subset with boundary of measure zero. More generally, let $\rho : [0, 1) \to \mathbb{R}$ be a Riemann integrable function. Since ρ can be weakly approximated from both above and below by step functions, we have

$$\lim_{y\to 0^+} \int_0^1 \rho(x)\chi_B(x + iy)dx = \mu_\Gamma(B) \int_0^1 \rho(x)dx$$

with $B \subset \Gamma\backslash\mathbb{H}$ a Borel set with boundary of measure zero.

2.5 A quantitative Borel–Cantelli lemma

Finally we record here a quantitative Borel–Cantelli lemma which ensures for the limsup set of certain sequence of events to have full measure given certain quasi-independence conditions.

Lemma 2.5 [29, Chapter I, Lemma 10] *Let $(X, \mathcal{B}, \nu)$ be a probability space with $\mathcal{B}$ a σ-algebra of subsets of X and $\nu : X \to [0, 1]$ a probability measure on X with respect to $\mathcal{B}$. Let $\{A_i\}_{i\in\mathbb{N}}$ be a sequence of measurable subsets in $\mathcal{B}$. For any $n, m \in \mathbb{N}$ we denote by $R_{n,m} := \nu(A_n \cap A_m) - \nu(A_n)\nu(A_m)$. Suppose that*

$$\exists\, C > 0 \text{ such that for all } k_2 > k_1 \geq 1, \ \sum_{n,m=k_1}^{k_2} R_{n,m} \leq C \sum_{n=k_1}^{k_2} \nu(A_n), \quad (2.18)$$

then $\sum_{n\in\mathbb{N}} \nu(A_n) = \infty$ implies that $\nu\left(\overline{\lim}_{n\to\infty} A_n\right) = 1$.

Remark 2.19 Keep the notation as in Lemma 2.5. It was shown in [20, Proposition 5.4] that if

$$\exists C' > 0 \text{ and } \eta > 1 \text{ such that for any } n \neq m, \ R_{n,m} \leq C' \frac{\sqrt{\nu(A_n)\nu(A_m)}}{|n - m|^\eta},$$

then the sequence $\{A_i\}_{i\in\mathbb{N}}$ satisfies the condition (2.18).

We will use the following slightly modified version of quantitative Borel–Cantelli lemma which has the flexibility to consider sequence of measurable sets $\{A_n\}_{n\in\mathbb{S}}$ indexed by a general unbounded subset $\mathbb{S} \subset \mathbb{N}$.

Corollary 2.6 *Let $(X, \mathcal{B}, \nu)$ be as in Lemma 2.5. Let $\mathbb{S} \subset \mathbb{N}$ be an unbounded subset and let $\{A_n\}_{n \in \mathbb{S}}$ be a sequence of measurable subsets in $\mathcal{B}$. Suppose that*

$$\exists C' > 0 \text{ and } \eta > 1 \text{ such that } \forall n, m \in \mathbb{S} \text{ with } m < n, R_{n,m} \leq C' \frac{\nu(A_n)\nu(A_m)}{n^\eta}, \quad (2.20)$$

then $\sum_{n \in \mathbb{S}} \nu(A_n) = \infty$ implies that $\nu \left(\overline{\lim}_{\substack{n \in \mathbb{S} \\ n \to \infty}} A_n \right) = 1$.

Proof For any $i \in \mathbb{N}$ let $a_i \in \mathbb{S}$ be the i-th integer in $\mathbb{S}$ and let $B_i := A_{a_i}$. For any $i, j \in \mathbb{N}$ let $R'_{i,j} := \nu(B_i \cap B_j) - \nu(B_i)\nu(B_j)$ so that $R'_{i,j} = R_{a_i,a_j}$. Then by for any $i < j$ we have

$$R'_{i,j} = R_{a_i,a_j} \leq C' \frac{\nu(A_{a_i})\nu(A_{a_j})}{a_j^\eta} = C' \frac{\nu(B_i)\nu(B_j)}{a_j^\eta} < C' \frac{\sqrt{\nu(B_i)\nu(B_j)}}{|i - j|^\eta},$$

where for the first inequality we used the assumption (2.20) and for the second inequality we used the estimates $a_j \geq j > j - i$ and $\sqrt{\nu(B_i)\nu(B_j)} \leq 1$. Thus in view of Remark 2.19 and Lemma 2.5 we have $\sum_{i \in \mathbb{N}} \nu(B_i) = \infty$ implies that $\nu \left(\overline{\lim}_{i \to \infty} B_i \right) = 1$ which is equivalent to the conclusion of this corollary in view of the relation $B_i = A_{a_i}$. $\qquad\qquad\square$

3 Equidistribution range

Let $\mathcal{M} = \mathrm{SL}_2(\mathbb{Z}) \backslash \mathbb{H}$. Since we fix $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ throughout this section, we abbreviate the Sobolev norm $\mathcal{S}^\Gamma_{p,d}$ by $\mathcal{S}_{p,d}$. In this section, we prove Theorems 1.1 and 1.2. The main ingredient of our proof is an explicit bound of Fourier coefficients which follows from a slight modification of the estimates obtained in [18].

3.1 Bounds on Fourier coefficients

Let $\Psi \in C_c^\infty(\mathcal{M})$. Since Ψ is left Γ-invariant, it is invariant under the transformation determined by $u_1 : z \mapsto z+1$, and it thus has a Fourier expansion for Ψ in the variable $x = \mathfrak{Re}(z)$:

$$\Psi(x + iy) = \sum_{m \in \mathbb{Z}} a_\Psi(m, y)e(mx), \quad (3.1)$$

where

$$a_\Psi(m, y) = \int_0^1 \Psi(x + iy)e(-mx)dx.$$

Similarly we denote by $a_{\phi_k}(m, y)$ and $a(s; m, y)$ the mth Fourier coefficients of the Hecke-Maass form ϕ_k and the Eisenstein series $E(\cdot, s)$ respectively. Estimates on these

Fourier coefficients yield, via the spectral expansion (2.6), estimates on the Fourier coefficients of Ψ. Namely,

$$a_\Psi(m, y) = \sum_{r_k \geq 0} \langle \Psi, \phi_k \rangle a_{\phi_k}(m, y) + \frac{1}{4\pi} \int_{-\infty}^{\infty} \langle \Psi, E(\cdot, \tfrac{1}{2} + ir) \rangle a(\tfrac{1}{2} + ir; m, y) dr.$$

We record the following bounds for $a_{\phi_k}(m, y)$ and $a(s; m, y)$:

Lemma 3.1 [18, Lemmata 3.7 and 3.13] *For any $m \neq 0$ and for any $\epsilon > 0$ we have*

$$|a_{\phi_k}(m, y)| \ll_\epsilon |m|^\theta y^{1/2-\epsilon} (r_k + 1)^{-1/3+\epsilon} \min\{1, e^{\pi r_k/2 - 2\pi |m| y}\}, \qquad (3.2)$$

and

$$\left| a\left(\tfrac{1}{2} + ir; m, y\right) \right| \ll_\epsilon y^{1/2-\epsilon} (1 + |r|)^{-1/3+\epsilon} \min\{1, e^{\pi |r|/2 - 2\pi |m| y}\}, \qquad (3.3)$$

where $\theta = 7/64$ is the best known bound towards the Ramanujan conjecture as before.

Remark 3.4 Contrarily to [18] that uses the trivial bound $\min\{1, e^{\pi r/2 - 2\pi |m| y}\} \leq 1$, we keep this term.

Proposition 3.2 [18, Proposition 3.4] *For any $\Psi \in C_c^\infty(\mathcal{M})$, we have that*

$$a_\Psi(0, y) = \mu_\mathcal{M}(\Psi) + O\left(\|\Psi\|_2^{3/4} \|\Delta\Psi\|_2^{1/4} y^{1/2}\right). \qquad (3.5)$$

Moreover, for any $m \neq 0$, and any $\epsilon > 0$ and any $\alpha_0 > 5/3$, we have

$$a_\Psi(m, y) \ll_{\alpha_0, \epsilon, p} \mathcal{S}_{\alpha_0}(\Psi) y^{1/2-\epsilon} |m|^\theta, \qquad (3.6)$$

where $\mathcal{S}_{\alpha_0}$ is a Sobolev norm of degree α_0.

Remark 3.7 The Sobolev norm $\mathcal{S}_{\alpha_0}$ is explicit from the proof of [18, Proposition 3.4]: Writing $\alpha_0 = 5/3 + \epsilon$ with $\epsilon > 0$, then $\mathcal{S}_{\alpha_0}(\Psi) = \mathcal{S}_{2,0}(\Psi)^{2/3-\epsilon/2} \mathcal{S}_{2,2}(\Psi)^{1/3+\epsilon/2}$ for any $\Psi \in C_c^\infty(\mathcal{M})$. In particular, using the estimate $\mathcal{S}_{2,0}(\Psi) \leq \mathcal{S}_{2,2}(\Psi)$ we have $\mathcal{S}_{\alpha_0}(\Psi) \leq \mathcal{S}_{2,2}(\Psi)$.

The following refinement of this last estimate allows to estimate the Fourier coefficients when $|m| > y^{-1}$ is large. This refinement is crucial for our later results, and the price we pay is a Sobolev norm of higher degree.

Proposition 3.3 *Let $\Psi \in C_c^\infty(\mathcal{M})$. Whenever $|m| y > 1$ and for any $\epsilon > 0$, we have*

$$|a_\Psi(m, y)| \ll_\epsilon \mathcal{S}_{2,2}(\Psi) |m|^{-4/3+\theta+\epsilon} y^{-5/6}.$$

Proof For the contribution from the cusp forms we apply the bound (3.2) to the Fourier coefficients and the bound

$$\min\{1, e^{\pi r/2 - 2\pi|m|y}\} \leq \begin{cases} e^{-\pi|m|y} & 0 \leq r \leq 2|m|y \\ 1 & r > 2|m|y, \end{cases} \tag{3.8}$$

and the relation $\langle \Delta\Psi, \phi_k \rangle = \langle \Psi, \Delta\phi_k \rangle = (1/4 + r_k^2)\langle \Psi, \phi_k \rangle$ to get that

$$\left| \sum_{r_k \geq 0} \langle \Psi, \phi_k \rangle a_{\phi_k}(m, y) \right| \ll_\epsilon \sum_{0 \leq r_k \leq 2|m|y} |\langle \Psi, \phi_k \rangle| \, |m|^\theta y^{1/2-\epsilon} (r_k + 1)^{-1/3+\epsilon} e^{-\pi|m|y}$$
$$+ \sum_{r_k > 2|m|y} |\langle \Delta\Psi, \phi_k \rangle| \, |m|^\theta y^{1/2-\epsilon} r_k^{-7/3+\epsilon}. \tag{3.9}$$

Now using Cauchy–Schwarz followed by summation by parts (together with Weyl's law stating that $\#\{r_k : r_k \leq M\} \ll M^2$ (see e.g. [15, Corollary 11.2]) we can bound

$$\sum_{0 \leq r_k \leq 2|m|y} |\langle \Psi, \phi_k \rangle| \, (r_k + 1)^{-1/3+\epsilon} \leq \|\Psi\|_2 \left(\sum_{0 \leq r_k \leq 2|m|y} \frac{1}{(r_k + 1)^{2/3-2\epsilon}} \right)^{1/2}$$
$$\ll_\epsilon \|\Psi\|_2 \, (|m|y)^{2/3+\epsilon}.$$

Similarly, for the second sum we can bound

$$\sum_{r_k > 2|m|y} |\langle \Delta\Psi, \phi_k \rangle| \, r_k^{-7/3+\epsilon}$$
$$\leq \|\Delta\Psi\|_2 \left(\sum_{r_k > 2|m|y} r_k^{-14/3+2\epsilon} \right)^{1/2} \ll_\epsilon \|\Delta\Psi\|_2 \, (|m|y)^{-4/3+\epsilon}.$$

To summarize, the left-hand side of (3.9) is bounded by

$$\ll_\epsilon \|\Psi\|_2 |m|^{2/3+\theta+\epsilon} y^{7/6} e^{-\pi|m|y} + \|\Delta\Psi\|_2 |m|^{-4/3+\theta+\epsilon} y^{-5/6}. \tag{3.10}$$

For the contribution from the continuous spectrum using the estimates (3.3), (3.8), the relation $\langle \Delta\Psi, E(\cdot, \frac{1}{2} + ir) \rangle = (\frac{1}{4} + r^2)\langle \Psi, E(\cdot, \frac{1}{2} + ir) \rangle$ and Cauchy–Schwarz we can similarly bound $\left| \int_{-\infty}^\infty \langle \Psi, E(\cdot, \frac{1}{2} + ir) \rangle a(\frac{1}{2} + ir; m, y) dr \right|$ by

$$\ll_\epsilon e^{-\pi|m|y} y^{1/2-\epsilon} \int_{|r| \leq 2|m|y} \left| \langle \Psi, E\left(\cdot, \tfrac{1}{2} + ir\right) \rangle \right| (|r| + 1)^{-1/3+\epsilon} dr$$
$$+ y^{1/2-\epsilon} \int_{|r| > 2|m|y} \left| \langle \Delta\Psi, E\left(\cdot, \tfrac{1}{2} + ir\right) \rangle \right| |r|^{-7/3+\epsilon} dr$$
$$\ll_\epsilon y^{1/2-\epsilon} \left(\|\Psi\|_2 \, (|m|y)^{1/6+\epsilon} e^{-\pi|m|y} + \|\Delta\Psi\|_2 \, (|m|y)^{-11/6+\epsilon} \right),$$

which is subsumed by the right-hand side of (3.10) (since $|m|y > 1$). Finally, we conclude the proof by applying the bounds $\max\{\|\Psi\|_2, \|\Delta\Psi\|_2\} \leq \mathcal{S}_{2,2}(\Psi)$ and $e^{-\pi|m|y} \ll (|m|y)^{-2}$ (again since $|m|y > 1$) to the right hand side of (3.10). $\qquad\square$

The following corollary of Proposition 3.3 is the key estimate that we will use to prove Theorem 1.1.

Corollary 3.4 *Let q be a positive integer. For any $\Psi \in C_c^\infty(\mathcal{M})$, $y > 0$, and any $\epsilon > 0$, we have*

$$\sum_{m \neq 0} |a_\Psi(qm, y)| \ll_\epsilon \mathcal{S}_{2,2}(\Psi) q^{-1} y^{-(1/2 + \theta + \epsilon)}.$$

Proof If $qy \leq 1$ we can separate the above sum into two parts to get

$$\sum_{m \neq 0} |a_\Psi(qm, y)| = \sum_{1 \leq |m| \leq (qy)^{-1}} |a_\Psi(qm, y)| + \sum_{|m| > (qy)^{-1}} |a_\Psi(qm, y)|.$$

Applying (3.6) (and the estimate $\mathcal{S}_{\alpha_0}(\Psi) \leq \mathcal{S}_{2,2}(\Psi)$ by Remark 3.7) to the first sum and Proposition 3.3 to the second, we have

$$\sum_{m \neq 0} |a_\Psi(qm, y)|$$

$$\ll_\epsilon \mathcal{S}_{2,2}(\Psi) \left(\sum_{1 \leq |m| \leq (ny)^{-1}} |qm|^\theta y^{1/2 - \epsilon} + \sum_{|m| > (qy)^{-1}} |qm|^{-4/3 + \theta + \epsilon} y^{-5/6} \right)$$

$$\asymp \mathcal{S}_{2,2}(\Psi) \left(q^\theta y^{1/2 - \epsilon} (qy)^{-(1+\theta)} + q^{-4/3 + \theta + \epsilon} y^{-5/6} (qy)^{1/3 - \theta - \epsilon} \right)$$

$$= \mathcal{S}_{2,2}(\Psi) q^{-1} y^{-(1/2 + \theta + \epsilon)},$$

where for the second estimate we used that $4/3 - \theta - \epsilon > 1$. If $qy > 1$ then we have $|qm|y > 1$ for all $m \neq 0$. We can apply Proposition 3.3 to $a_\Psi(qm, y)$ for all integers $m \neq 0$ to get

$$\sum_{m \neq 0} |a_\Psi(qm, y)| \ll_\epsilon \mathcal{S}_{2,2}(\Psi) \sum_{|m| \neq 0} |qm|^{-4/3 + \theta + \epsilon} y^{-5/6}$$

$$\ll \mathcal{S}_{2,2}(\Psi) q^{-4/3 + \theta + \epsilon} y^{-5/6} \ll \mathcal{S}_{2,2}(\Psi) q^{-1} y^{-(1/2 + \theta + \epsilon)},$$

where for the last estimate we used that $\theta < 1/3 - \epsilon$. $\qquad\square$

Remark 3.11 The estimates in [18] hold more generally for any Γ conjugate to some $\Gamma_0(p)$. In this generality, there might be (finitely many) exceptional cusp forms with $r_k \in i(0, \theta]$. For such forms, it was shown in [18, Lemma 3.7] that for any $m \neq 0$

$$\left| a_{\phi_k}(m, y) \right| \ll_{\epsilon, p} \|\Psi\|_2 |m|^\theta y^{1/2 - \epsilon} (|m|y)^{-|r_k| + \epsilon} e^{-2\pi|m|y}.$$

Using the estimates $(|m|y)^{-|r_k|+\epsilon}e^{-2\pi|m|y} < (|m|y)^{-\theta}$ when $|m|y \leq 1$ and $(|m|y)^{-|r_k|+\epsilon}e^{-2\pi|m|y} \ll (|m|y)^{-2}$ when $|m|y > 1$ one can easily recover Corollary 3.4 for ϕ_k, and hence for a general $\Psi \in C_c^\infty(\Gamma_0(p)\backslash\mathbb{H})$. Then one can easily deduce analogous estimates as in Theorem 1.1 for Ψ, see the arguments in the next subsection.

3.2 Proof of Theorem 1.1

In this subsection we prove Theorem 1.1. In view of (3.5) it suffices to prove the following proposition.

Proposition 3.5 *Let $\mathcal{M}$ be the modular surface. For any $\Psi \in C_c^\infty(\mathcal{M})$, for any $x \in \mathbb{R}/\mathbb{Z}$ and $y > 0$, we have*

$$\delta_{n,x,y}(\Psi) = a_\Psi(0, y) + O_\epsilon\left(\mathcal{S}_{2,2}(\Psi)n^{-1}y^{-(1/2+\theta+\epsilon)}\right) \tag{3.12}$$

and

$$\delta_{n,x,y}^{\mathrm{pr}}(\Psi) = a_\Psi(0, y) + O_\epsilon\left(\mathcal{S}_{2,2}(\Psi)n^{-1+\epsilon}y^{-(1/2+\theta+\epsilon)}\right). \tag{3.13}$$

Proof Let $J \subset \mathbb{R}/\mathbb{Z} \cong [0, 1)$ be a finite subset and for any $m \in \mathbb{Z}$ denote by $W_J(m) := \frac{1}{|J|}\sum_{t\in J} e(mt)$. We note that $\frac{1}{|J|}\sum_{t\in J} \Psi(t + iy)$ equals $\delta_{n,x,y}(\Psi)$ when $J = \{x + j/n : 0 \leq j \leq n - 1\}$ and equals $\delta_{n,x,y}^{\mathrm{pr}}(\Psi)$ when $J = \{x + j/n : 0 \leq j \leq n - 1, \gcd(j, n) = 1\}$. Applying the Fourier expansion (3.1) to Ψ we get that

$$\frac{1}{|J|}\sum_{t\in J} \Psi(t + iy) = \frac{1}{|J|}\sum_{t\in J}\sum_{m\in\mathbb{Z}} a_\Psi(m, y)e(mt) = \sum_{m\in\mathbb{Z}} a_\Psi(m, y)\frac{1}{|J|}\sum_{t\in J} e(mt)$$

$$= a_\Psi(0, y) + \sum_{m\neq 0} a_\Psi(m, y)W_J(m).$$

Now for (3.12) we take $J = \{x + j/n : 0 \leq j \leq n - 1\}$ and note that for such J, $|W_J(m)|$ equals 1 if $n \mid m$ and equals 0 otherwise. Hence

$$\left|\sum_{m\neq 0} a_\Psi(m, y)W_J(m)\right| \leq \sum_{\substack{m\neq 0 \\ n|m}} |a_\Psi(m, y)| \ll_\epsilon n^{-1}y^{-(1/2+\theta+\epsilon)},$$

where for the last estimate we applied Corollary 3.4.

For (3.13) we take $J = \{x + j/n : 0 \leq j \leq n - 1, \gcd(j, n) = 1\}$ and note the identity

$$\sum_{j\in(\mathbb{Z}/n\mathbb{Z})^\times} e\left(\frac{mj}{n}\right) = \frac{\mu(n_m)\varphi(n)}{\varphi(n_m)}$$

for the Ramanujan's sum, where $n_m := n/\gcd(n, m)$ and $\mu : \mathbb{N} \to \{0, \pm 1\}$ is the Möbius function; see e.g. [14, Theorem 272]. Then

$$|W_J(m)| = \left| \frac{1}{\varphi(n)} \sum_{j \in (\mathbb{Z}/n\mathbb{Z})^\times} e\left(\frac{mj}{n}\right) \right| = \frac{|\mu(n_m)|}{\varphi(n_m)} \leq \frac{1}{\varphi(n_m)}.$$

Hence we have

$$
\begin{aligned}
\left| \sum_{m \neq 0} a_\Psi(m, y) W_J(m) \right| &\leq \sum_{m \neq 0} \frac{|a_\Psi(m, y)|}{\varphi(n_m)} = \sum_{d \mid n} \frac{1}{\varphi(d)} \sum_{\substack{m \neq 0 \\ \gcd(m, n) = n/d}} |a_\Psi(m, y)| \\
&\leq \sum_{d \mid n} \frac{1}{\varphi(d)} \sum_{\substack{m \neq 0 \\ (n/d) \mid m}} |a_\Psi(m, y)| \\
&\ll_\epsilon \sum_{d \mid n} \frac{1}{\varphi(d)} \left(\frac{n}{d}\right)^{-1} y^{-(1/2+\theta+\epsilon)} \\
&\ll_\epsilon n^{-1} \sigma_{\epsilon/2}(n) y^{-(1/2+\theta+\epsilon)} \ll_\epsilon n^{-1+\epsilon} y^{-(1/2+\theta+\epsilon)},
\end{aligned}
$$

where for the second inequality we used the fact that $\gcd(m, n) = n/d$ implies that $(n/d) \mid m$, for the third inequality we applied Corollary 3.4 and for the second last inequality we applied the estimate $\varphi(d) \gg_\epsilon d^{1-\epsilon/2}$. □

3.3 Full range equidistribution for rational translates

In this subsection we prove Theorem 1.2. We fix $x = p/q$ a primitive rational number and let

$$\mathbb{N}_q = \left\{ n \in \mathbb{N} : \gcd(n^2, q) \mid n \right\}$$

be as in Theorem 1.2. As mentioned in the introduction, the key ingredient is a symmetry lemma for rational translates which generalizes the symmetry (1.6). Before stating the lemma, let us briefly explain why we need to restrict to the subsequence $\mathbb{N}_q$. Let $n \in \mathbb{N}$ and let $y > 0$. We need to study the distribution of the points $\Gamma(x + \frac{j}{n} + iy) = \Gamma(\frac{p}{q} + \frac{j}{n} + iy)$ for $0 \leq j \leq n - 1$. Let $\frac{p_j}{q_j}$ be the reduced form of $\frac{p}{q} + \frac{j}{n}$ and in view of the symmetry (1.6) we have

$$\Gamma\left(x + \frac{j}{n} + iy\right) = \Gamma\left(\frac{p_j}{q_j} + iy\right) = \Gamma\left(-\frac{\overline{p_j}}{q_j} + \frac{i}{q_j^2 y}\right),$$

where $\overline{p_j}$ is the multiplicative inverse of p_j modulo q_j. To further analyze the distribution of these points, we thus need to solve the congruence equation $xp_j \equiv 1 \pmod{q_j}$

in x. Write $k = \gcd(n, q)$ and $q' = q/k$ and $n' = n/k$. Then

$$\frac{p}{q} + \frac{j}{n} = \frac{p}{kq'} + \frac{j}{kn'} = \frac{pn'+jq'}{kq'n'},$$

implying that

$$q_j = \frac{kq'n'}{\gcd(pn'+jq',kq'n')} = \frac{kn'q'}{\gcd(pn'+jq',kn')} = q'\frac{n}{\gcd(pn'+jq',n)}$$

can be written canonically as a product of two integers. Here for the second equality we used that $\gcd(pn' + jq', q') = \gcd(pn', q') = 1$. In view of the Chinese remainder theorem, the above congruence equation modulo q_j is relatively easy to solve when the two factors q' and $n/\gcd(pn' + jq', n)$ are coprime (see the proof of Lemma 3.6 for more details). This condition can be guaranteed for any j if $\gcd(q', n) = \gcd(q/\gcd(q, n), n) = 1$ which is equivalent to the condition $n \in \mathbb{N}_q$. Finally, we also note that by writing n and q in prime decomposition forms, it is not hard to check that $n \in \mathbb{N}_q$ is equivalent to $q = kl$ with $l = \gcd(n, q) \mid n$ and $\gcd(k, n) = 1$. We now state the symmetry lemma.

Lemma 3.6 *Let $\frac{m}{kl}$ be a primitive rational number and let $n \in \mathbb{N}$ such that $l \mid n$ and $\gcd(k, n) = 1$. Then for any $0 \leq j \leq n - 1$ and for any $y > 0$ we have*

$$\Gamma\left(\frac{m}{kl} + \frac{j}{n} + iy\right) = \Gamma\left(-\frac{dl\overline{mn}a}{k} - \frac{\left((m\frac{n}{l}+jk)/d\right)^* b}{n/d} + i\frac{d^2}{k^2n^2y}\right), \qquad (3.14)$$

where $d = d_j := \gcd(m\frac{n}{l} + jk, n)$ and $a = a_d$, $b = b_d \in \mathbb{Z}$ are some fixed integers such that $a\frac{n}{d} + bk = 1$. Here, for any integer x, $\overline{x}$ denotes the multiplicative inverse of x modulo k, x^* denotes the multiplicative inverse of x modulo n/d. If we further assume $\gcd(j, n) = l = 1$, then $d_j = \gcd(mn + jk, n) = 1$ and

$$\Gamma\left(\frac{m}{k} + \frac{j}{n} + iy\right) = \Gamma\left(-\frac{\overline{mn}a}{k} - \frac{(jk)^* b}{n} + i\frac{i}{k^2n^2y}\right). \qquad (3.15)$$

Proof Since $l \mid n$, by direct computation we have $\frac{m}{kl} + \frac{j}{n} = \frac{mn/l+jk}{kn}$. Note that since $\gcd(k, mn) = 1$ we have $\gcd(m\frac{n}{l} + jk, k) = \gcd(m\frac{n}{l}, k) = 1$. This implies that $\gcd(m\frac{n}{l} + jk, kn) = \gcd(m\frac{n}{l} + jk, n) = d$. Hence let $\frac{p}{q}$ be the reduced form of $\frac{m}{kl} + \frac{j}{n}$, then we have $(p, q) = ((m\frac{n}{l} + jk)/d, kn/d)$. Now since $\gcd(p, q) = 1$, there exist some integers $v, w \in \mathbb{Z}$ such that $\gamma = \left(\begin{smallmatrix} w & v \\ -q & p \end{smallmatrix}\right) \in \Gamma$. By direct computation we have

$$\gamma\left(\frac{m}{kl} + \frac{j}{n} + iy\right) = \gamma\left(\frac{p}{q} + iy\right) = -\frac{w}{q} + \frac{i}{q^2y}.$$

implying that

$$\Gamma\left(\frac{m}{kl} + \frac{j}{n} + iy\right) = \Gamma\left(-\frac{w}{q} + \frac{i}{q^2y}\right) = \Gamma\left(-\frac{w}{kn/d} + i\frac{d^2}{k^2n^2y}\right), \qquad (3.16)$$

where for the second equality we used the relation $q = kn/d$. Moreover, since $\gamma \in \Gamma$ we have $wp + vq = 1$, implying that (again using the relation $(p, q) = ((m\frac{n}{l} + jk)/d, kn/d))$

$$w\left((m\tfrac{n}{l} + jk)/d\right) \equiv 1 \pmod{k\tfrac{n}{d}}.$$

We claim that

$$w \equiv dl\overline{mn}\tfrac{n}{d}a + \left((m\tfrac{n}{l} + jk)/d\right)^{*} kb \pmod{k\tfrac{n}{d}}. \tag{3.17}$$

In view of the Chinese Remainder Theorem, since $\gcd(k, n/d) = 1$, it suffices to check

$$\left(dl\overline{mn}\tfrac{n}{d}a + \left((m\tfrac{n}{l} + jk)/d\right)^{*} kb\right)\left((m\tfrac{n}{l} + jk)/d\right) \equiv 1 \pmod{k}$$

and

$$\left(dl\overline{mn}\tfrac{n}{d}a + \left((m\tfrac{n}{l} + jk)/d\right)^{*} kb\right)\left((m\tfrac{n}{l} + jk)/d\right) \equiv 1 \pmod{\tfrac{n}{d}}.$$

For the first equation we have

$$\left(dl\overline{mn}\tfrac{n}{d}a + \left((m\tfrac{n}{l} + jk)/d\right)^{*} kb\right)\left((m\tfrac{n}{l} + jk)/d\right) \equiv dl\overline{mn}\tfrac{n}{d}amn\overline{ld} \equiv a\tfrac{n}{d}$$
$$= 1 - bk \equiv 1 \pmod{k},$$

where for the first equality we used the fact that $\gcd(dl, k) = 1$ (since $d \mid n$, $l \mid n$ and $\gcd(k, n) = 1$). The second equation follows similarly. Now plugging relation (3.17) into (3.16) we get (3.14).

For the second half we note that $d_j = \gcd(mn + jk, n) = \gcd(jk, n) = 1$. The first equality is true since $l = 1$, and the second equality is true since by assumption $\gcd(k, n) = \gcd(j, n) = 1$. Thus in view of (3.14), to prove (3.15) it suffices to note that $(mn + jk)^{*} \equiv (jk)^{*} \pmod{n}$, or equivalently, $mn + jk \equiv jk \pmod{n}$. $\qquad\square$

Remark 3.18 When $k = 1$ we can take $(a, b) = (0, 1)$, then (3.15) recovers the symmetry (1.6). We also note that for the point $\Gamma(x + j/n + iy)$ with x irrational, the above symmetry clearly breaks.

Proposition 3.7 *Let p/q be a primitive rational number and let $n \in \mathbb{N}_q$. Then for any $y > 0$ we have*

$$\mathcal{R}_n\left(\tfrac{p}{q}, y\right) = \bigcup_{d \mid n} \mathcal{R}^{\mathrm{pr}}_{n/d}\left(x_d, \tfrac{d^2}{k^2 n^2 y}\right), \tag{3.19}$$

where $x_d \in \mathbb{R}/\mathbb{Z}$ is some number depending on d (and also on p, q, n) and $k := q/\gcd(n, q)$. If we further assume $\gcd(n, q) = 1$, then

$$\mathcal{R}^{\mathrm{pr}}_n\left(\tfrac{p}{q}, y\right) = \mathcal{R}^{\mathrm{pr}}_n\left(-\tfrac{\overline{pn}a}{q}, \tfrac{1}{q^2 n^2 y}\right), \tag{3.20}$$

where $\overline{x}$ denotes the multiplicative inverse of x modulo q and $a \in \mathbb{Z}$ is as in Lemma 3.6.

Proof Relation (3.20) follows immediately from (3.15) by taking $(m, k) = (p, q)$ and noting that

$$\{(-[(qj)^*b] \in (\mathbb{Z}/n\mathbb{Z})^\times : j \in (\mathbb{Z}/n\mathbb{Z})^\times\} = (\mathbb{Z}/n\mathbb{Z})^\times,$$

which follows from the fact that $\gcd(bq, n) = 1$ (since $\gcd(bq, n) = \gcd(1 - an, n) = 1$). Here $(qj)^*$ denotes the multiplicative inverse of qj modulo n and $b \in \mathbb{Z}$ is as in Lemma 3.6.

For (3.19), we set $m = p, l = \gcd(n, q)$ (so that $k = q/l$). As mentioned above, the condition $\gcd(n^2, q) \mid n$ implies that $\gcd(k, n) = 1$. Thus the pair $(\frac{m}{kl}, n)$ satisfies the assumptions in Lemma 3.6 and we can apply (3.14) for the points

$$\Gamma\left(\frac{p}{q} + \frac{j}{n} + iy\right) = \Gamma\left(\frac{m}{kl} + \frac{j}{n} + iy\right), 0 \le j \le n - 1.$$

Now for any $d \mid n$ define

$$D_d := \left\{0 \le j \le n - 1 : d_j = \gcd(m\tfrac{n}{l} + jk, n) = d\right\}$$

so that

$$\mathcal{R}_n\left(\frac{p}{q}, y\right) = \bigcup_{d|n}\left\{\Gamma\left(\frac{p}{q} + \frac{j}{n} + iy\right) \in \mathcal{M} : j \in D_d\right\}. \tag{3.21}$$

Moreover, we note that since $\gcd(k, n) = 1$, we have

$$\left\{[m\tfrac{n}{l} + jk] \in \mathbb{Z}/n\mathbb{Z} : 0 \le j \le n - 1\right\} = \mathbb{Z}/n\mathbb{Z}$$

and hence

$$\left\{[m\tfrac{n}{l} + jk] \in \mathbb{Z}/n\mathbb{Z} : j \in D_d\right\} = \{[j] \in \mathbb{Z}/n\mathbb{Z} : \gcd(j, n) = d\}. \tag{3.22}$$

On the other hand, by (3.14) we have

$$\left\{\Gamma\left(\frac{p}{q} + \frac{j}{n} + iy\right) \in \mathcal{M} : j \in D_d\right\}$$
$$= \left\{\Gamma\left(-\frac{dl\overline{mn}a_d}{k} - \frac{\left((m\frac{n}{l}+jk)/d\right)^* b_d}{n/d} + i\frac{d^2}{k^2n^2y}\right) \in \mathcal{M} : j \in D_d\right\},$$

where for any integer $x, \overline{x}$ denotes the multiplicative inverse of x modulo k, x^* denotes the multiplicative inverse of x modulo n/d, and $a_d, b_d \in \mathbb{Z}$ are some fixed integers such that $a_d\frac{n}{d} + b_d k = 1$. Now for each $d \mid n$ we let $x_d \in [0, 1), x_d \equiv -\frac{dl\overline{mn}a_d}{k} \pmod 1$ so that it remains to show

$$\left\{-[((m\tfrac{n}{l} + jk)/d)^* b_d] \in (\mathbb{Z}/(n/d)\mathbb{Z})^\times : j \in D_d\right\} = (\mathbb{Z}/(n/d)\mathbb{Z})^\times.$$

We can thus conclude the proof by noting that the above relation follows immediately from (3.22) together with the fact $\gcd(b_d, \frac{n}{d}) = 1$ (since $\gcd(b_d, \frac{n}{d}) = \gcd(b_d k, \frac{n}{d}) = \gcd(1 - a_d \frac{n}{d}, \frac{n}{d}) = 1$).   □

Using these two relations and the estimate (3.13) one gets the following effective estimates.

Proposition 3.8 *Let $x = p/q$ be a primitive rational number and let $n \in \mathbb{N}_q$. Then for any $\Psi \in C_c^\infty(\mathcal{M})$ and $y > 0$ we have*

$$\delta_{n,x,y}(\Psi) = \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y}\right) + O_{\epsilon,q}\left(\mathcal{S}_{2,2}(\Psi) n^{2\theta+4\epsilon} y^{1/2+\theta+\epsilon}\right),$$

where $k := q/\gcd(n, q)$. If we further assume that $\gcd(n, q) = 1$, then

$$\delta_{n,x,y}^{\mathrm{pr}}(\Psi) = a_\Psi\left(0, \frac{1}{q^2 n^2 y}\right) + O_{\epsilon,q}\left(\mathcal{S}_{2,2}(\Psi) n^{2\theta+3\epsilon} y^{1/2+\theta+\epsilon}\right).$$

Proof For any positive divisor $d \mid n$, let $y_d = d^2/(k^2 n^2 y)$ with $k := q/\gcd(n, q)$ as above and let $x_d \in \mathbb{R}/\mathbb{Z}$ be as in (3.19). Then by (3.19) for $x = p/q$ we have

$$\delta_{n,x,y}(\Psi) = \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) \delta_{n/d, x_d, y_d}^{\mathrm{pr}}(\Psi)$$

$$= \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) \left(a_\Psi(0, y_d) + O_\epsilon\left(\mathcal{S}_{2,2}(\Psi) \left(\frac{n}{d}\right)^{-1+\epsilon} y_d^{-(1/2+\theta+\epsilon)}\right)\right)$$

$$= \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi(0, y_d) + O_\epsilon\left(\mathcal{S}_{2,2}(\Psi) n^{-1} \sum_{d|n} \left(\frac{n}{d}\right)^\epsilon y_d^{-(1/2+\theta+\epsilon)}\right),$$

where for the second estimate we applied (3.13) and for the third estimate we used the trivial estimate $\varphi(n/d) < n/d$. Now plugging $y_d = d^2/(k^2 n^2 y)$ into the above equation we get

$$\delta_{n,x,y}(\Psi) = \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y}\right) + O_{\epsilon,q}\left(\mathcal{S}_{2,2}(\Psi) n^{-1} \sigma_{1+2\theta+3\epsilon}(n) y^{1/2+\theta+\epsilon}\right)$$

$$= \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y}\right) + O_{\epsilon,q}\left(\mathcal{S}_{2,2}(\Psi) n^{2\theta+4\epsilon} y^{1/2+\theta+\epsilon}\right),$$

where the dependence on k in the first estimate is absorbed into the dependence on q (since $k := q/\gcd(n, q) \leq q$). The second estimate follows from similar (but easier) analysis with the relation (3.20) in place of (3.19).   □

We are now in the position to prove Theorem 1.2. We will prove the following proposition from which Theorem 1.2 follows, see also Remark 3.23.

Theorem 3.9 *Let* $x = p/q$ *be a primitive rational number and let* $n \in \mathbb{N}_q$. *Let* $y_n = c/n^\alpha$ *for some* $1 < \alpha < 2$ *and* $c > 0$. *Then for any* $\Psi \in C_c^\infty(\mathcal{M})$ *we have*

$$\left| \delta_{n,x,y_n}(\Psi) - \mu_{\mathcal{M}}(\Psi) \right| \ll_{\epsilon,q,c,\Psi} n^{\alpha/2-1+\epsilon} + n^{2\theta+4\epsilon-\alpha(1/2+\theta+\epsilon)}.$$

If we further assume $\gcd(n, q) = 1$, *then we have*

$$\left| \delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) - \mu_{\mathcal{M}}(\Psi) \right| \ll_{\epsilon,q,c} \mathcal{S}_{2,2}(\Psi) \left(n^{\alpha/2-1} + n^{2\theta+3\epsilon-\alpha(1/2+\theta+\epsilon)} \right).$$

Remark 3.23 The dependence on Ψ in the first estimate can also be made explicit. In fact, we can remove this dependence by adding a factor of $\mathcal{S}_{2,2}(\Psi) + \|\Psi\|_\infty$ to the right hand side of this estimate. We also note that since we may take $\theta = 7/64$, the right hand side of these two estimates decays to zero as $n \to \infty$ for any $1 < \alpha < 2$.

Proof of Theorem 3.9. In view of Proposition 3.8 and the assumption $y_n = c/n^\alpha$, it suffices to show that

$$\frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y_n}\right) = \mu_{\mathcal{M}}(\Psi) + O_{\epsilon,c,\Psi}\left(n^{\alpha/2-1+\epsilon}\right)$$

with $k := q/\gcd(n, q)$, and that (under the extra assumption $\gcd(n, q) = 1$)

$$a_\Psi\left(0, \frac{1}{q^2 n^2 y_n}\right) = \mu_{\mathcal{M}}(\Psi) + O_c\left(\mathcal{S}_{2,2}(\Psi) n^{\alpha/2-1}\right).$$

The second estimate follows immediately from (3.5) and the trivial estimate $|q| \geq 1$. For the first estimate we separate the sum into two parts to get

$$\frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y_n}\right) = \frac{1}{n} \left(\sum_{\substack{d|n \\ d < n^{1-\alpha/2}}} + \sum_{\substack{d|n \\ d \geq n^{1-\alpha/2}}} \right) \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y_n}\right).$$

Applying (3.5) (and the trivial estimate $|k| \geq 1$) for the first sum and applying the estimate

$$\left| a_\Psi\left(0, \frac{d^2}{k^2 n^2 y_n}\right) \right| = \left| \int_0^1 \Psi\left(t + i\frac{d^2}{k^2 n^2 y_n}\right) dt \right| \leq \|\Psi\|_\infty$$

for the second sum we get $\frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) a_\Psi\left(0, \frac{d^2}{k^2 n^2 y_n}\right)$ equals

$$\frac{1}{n}\left(\sum_{\substack{d|n \\ d<n^{1-\alpha/2}}} \varphi\left(\tfrac{n}{d}\right)\left(\mu_{\mathcal{M}}(\Psi) + O_{c,\Psi}\left(\left(\tfrac{n}{d}\right)^{-1} n^{\alpha/2}\right)\right) + O_\Psi\left(\sum_{\substack{d|n \\ d\geq n^{1-\alpha/2}}} \varphi\left(\tfrac{n}{d}\right)\right)\right)$$

$$= \mu_{\mathcal{M}}(\Psi) + \frac{1}{n} O_{c,\Psi}\left(n^{\alpha/2} \sum_{\substack{d|n \\ d<n^{1-\alpha/2}}} 1 + \sum_{\substack{d|n \\ d\geq n^{1-\alpha/2}}} \tfrac{n}{d}\right)$$

$$= \mu_{\mathcal{M}}(\Psi) + O_{c,\Psi}\left(n^{\alpha/2-1}\sigma_0(n)\right) = \mu_{\mathcal{M}}(\Psi) + O_{\epsilon,c,\Psi}\left(n^{\alpha/2-1+\epsilon}\right),$$

finishing the proof, where for the first estimate we used the identity that $\sum_{d|n} \varphi(n/d) = n$ and the estimate that $\varphi(n/d) < n/d$, and for the second estimate we used the estimates $\sum_{\substack{d|n \\ d<n^{1-\alpha/2}}} 1 \leq \sigma_0(n)$ and

$$\sum_{\substack{d|n \\ d\geq n^{1-\alpha/2}}} \frac{n}{d} = \sum_{\substack{d|n \\ d\leq n^{\alpha/2}}} d \leq n^{\alpha/2} \sum_{\substack{d|n \\ d\leq n^{\alpha/2}}} 1 \leq n^{\alpha/2}\sigma_0(n).$$

$\square$

3.4 Quantitative non-equidistribution for rational translates

As a direct consequence of the analysis in the previous subsection we also have the following quantitative non-equidistribution result for rational translates when $\{y_n\}_{n\in\mathbb{N}}$ is beyond the above range, generalizing the situation for $\{\mathcal{R}_n^{\mathrm{pr}}(0, y_n)\}_{n\in\mathbb{N}}$. As before, for any $Y > 0$ we denote by μ_Y the probability uniform distribution measure supported on $\mathcal{H}_Y$.

Theorem 3.10 *Let $x = p/q$ be a primitive rational number and let $y_n = c/n^2$ for some constant $c > 0$. Let $\Psi \in C_c^\infty(\Psi)$. Then for any $n \in \mathbb{N}_q$ we have*

$$\delta_{n,x,y_n}(\Psi) = \frac{1}{n}\sum_{d|n} \varphi\left(\tfrac{n}{d}\right) \mu_{\frac{d^2}{ck_n^2}}(\Psi) + O_{\epsilon,q,c}\left(\mathcal{S}_{2,2}(\Psi)n^{-1+2\epsilon}\right)$$

with $k_n = q/\gcd(n^2, q)$. If we further assume that $\gcd(n, q) = 1$, then

$$\delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) = \mu_{\frac{1}{cq^2}}(\Psi) + O_{\epsilon,q,c}\left(\mathcal{S}_{2,2}(\Psi)n^{-1+\epsilon}\right).$$

Proof These two effective estimates follow immediately from Proposition 3.8 by plugging in $y_n = c/n^2$ and noting that $a_\Psi(0, Y) = \int_0^1 \Psi(x + iY)dx = \mu_Y(\Psi)$. $\square$

We can now give the

Proof of Theorem 1.3. For part (1), in view of Theorem 3.10 only the second equation needs a proof. Since we are taking $n \in \mathbb{P}_m$ going to infinity, it is sufficient to consider $n = m\ell \in \mathbb{P}_m$ with the prime number $\ell > q$ (so that $\ell \nmid q$). For such n, we have $\gcd(n^2, q) = \gcd(m^2\ell^2, q) = \gcd(m^2, q)$. Since by assumption $\gcd(m^2, q) \mid m$ and $m \mid n$, we can apply the first effective estimate in Theorem 3.10 for such $n = m\ell \in \mathbb{P}_m$. Moreover, for any such n we have

$$k_n = \frac{q}{\gcd(n^2, q)} = \frac{q}{\gcd(m^2, q)} = \frac{q}{\gcd(m, q)}$$

is a fixed number only depending on m and q. Here for the last equality we used the assumption that $\gcd(m^2, q) \mid m$. Now let $n = m\ell \in \mathbb{P}_m$ with $\ell \gg q$ sufficiently large such that $\mu_Y(\Psi) = 0$ whenever $Y > \ell^2/(ck_n)^2$ (this can be guaranteed since k_n is a fixed number and Ψ is compactly supported). In particular, for any $d \mid n$, $\mu_{d^2/(ck_n^2)}(\Psi) = 0$ whenever $\ell \mid d$. This, together with the first estimate in Theorem 3.10 implies that for all such sufficiently large $n = m\ell \in \mathbb{P}_m$

$$\delta_{n,x,y_n}(\Psi) = \frac{1}{m\ell} \sum_{d|m} \varphi\left(\frac{m\ell}{d}\right) \mu_{\frac{d^2}{ck_n^2}}(\Psi) + O_{\epsilon,q,c,\Psi,m}\left(\ell^{-1+2\epsilon}\right)$$

$$= \frac{\ell - 1}{\ell} \nu_{m, \frac{1}{ck_n^2}}(\Psi) + O_{\epsilon,q,c,\Psi,m}\left(\ell^{-1+2\epsilon}\right),$$

where for the second estimate we used that $\gcd(m, \ell) = 1$ and ℓ is a prime number. We can now finish the proof by taking $n = m\ell \to \infty$ along the subsequence $\mathbb{P}_m$ (equivalently, taking $\ell \to \infty$) and plugging in the relation $k_n = q/\gcd(m, q)$.

For part (2), since $\mathcal{R}_n^{\mathrm{pr}}(x, y_n) \subset \mathcal{R}_n(x, y_n)$, we only need to prove the full escape to the cusp for the sequence $\{\mathcal{R}_n(x, y_n)\}_{n \in \mathbb{N}}$. Identify (up to a null set) $\mathcal{M}$ with the standard fundamental domain $\mathcal{F}_\Gamma := \{z \in \mathbb{H} : \mathfrak{Re}(z) < \frac{1}{2}, |z| > 1\}$. For any $n \in \mathbb{N}$ and $0 \le j \le n - 1$ let $\frac{p_j}{q_j}$ be the reduced form of $x + \frac{j}{n} = \frac{p}{q} + \frac{j}{n} = \frac{pn+qj}{qn}$ so that by (1.6)

$$\Gamma\left(x + \frac{j}{n} + iy_n\right) = \Gamma\left(-\frac{\overline{p_j}}{q_j} + \frac{i}{q_j^2 y_n}\right).$$

Thus using the trivial inequality $|q_j| \le |q|n$ for all $0 \le j \le n - 1$ and the assumption $\lim_{n \to \infty} n^2 y_n = 0$, we have

$$\mathcal{R}(x, y_n) \subset \left\{z \in \mathcal{F}_\Gamma : \mathfrak{Im}(z) \ge \frac{1}{q^2 n^2 y_n}\right\} \xrightarrow{n \to \infty} \text{cusp of } \mathcal{M}.$$

$\square$

4 Negative results: in connection with Diophantine approximations

Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and $\mathcal{M} = \Gamma \backslash \mathbb{H}$ be the modular surface. Let $\mu_{\mathcal{M}}$ be the normalized hyperbolic area on $\mathcal{M}$ as before. In this section we prove a general result which captures the cusp excursion rate for the sample points $\mathcal{R}_n(x, y_n)$ in terms of the Diophantine property of the translate $x \in \mathbb{R}/\mathbb{Z} \cong [0, 1)$, see Theorem 4.3. Theorem 1.4 will then be an easy consequence of this result.

4.1 Notation and a preliminary result on cusp excursions

In this subsection we prove a preliminary lemma relating cusp excursions on the modular surface to Diophantine approximations. Let us first fix some notation. For any $Y > 0$, we denote by $\mathcal{C}_Y \subset \mathcal{M}$ the image of the region

$$\{z \in \mathbb{H} : \mathfrak{Im}(z) > Y\}$$

under the natural projection from $\mathbb{H}$ to $\mathcal{M} = \Gamma \backslash \mathbb{H}$. As Y goes to infinity, the sets $\mathcal{C}_Y$ diverge to the cusp of $\mathcal{M}$, and we call $\mathcal{C}_Y$ a *cusp neighborhood of* $\mathcal{M}$. Similarly, for any $Y' > Y > 0$, we denote by $\mathcal{C}_{Y,Y'}$ the projection onto $\mathcal{M}$ of the open set

$$\left\{z \in \mathbb{H} : Y < \mathfrak{Im}(z) < Y'\right\}.$$

For any primitive rational number m/n, and for any $r > 0$ we denote by

$$H_{m/n,r} := \left\{z = x + iy \in \mathbb{H} : (x - m/n)^2 + (y - r)^2 = r^2\right\}$$

the horocycle tangent to $\partial \mathbb{H}$ at m/n with Euclidean radius r. We denote by

$$H^\circ_{m/n,r} := \left\{z = x + iy \in \mathbb{H} : (x - m/n)^2 + (y - r)^2 < r^2\right\}$$

the open horodisc enclosed by $H_{m/n,r}$. We have the following geometric description of Lemma 3.6: Let $\gamma = \left(\begin{smallmatrix} m & * \\ n & * \end{smallmatrix}\right)$ be an element in Γ. Then γ sends the horizontal horocycle $\{z \in \mathbb{H} : \mathfrak{Im}(z) = Y\}$ to the horocycle $H_{m/n,r}$ with $r = 1/(2Yn^2)$, while the open region $\{z \in \mathbb{H} : \mathfrak{Im}(z) > Y\}$ is mapped to the horodisc $H^\circ_{m/n,r}$. On the other hand, for any primitive rational number m/n, there is $\gamma \in \Gamma$ of the form $\gamma = \left(\begin{smallmatrix} m & * \\ n & * \end{smallmatrix}\right)$. Thus for any $Y > 0$ and for any $z \in \mathbb{H}$, $\Gamma z \in \mathcal{C}_Y$ if and only if $z \in H^\circ_{m/n,r}$ for some primitive rational number m/n with $r = 1/(2Yn^2)$.

Finally, we record a distance formula that we will later use. Let $d_{\mathcal{M}}(\cdot, \cdot)$ be the distance function on $\mathcal{M}$ induced from the hyperbolic distance function $d_{\mathbb{H}}$ on $\mathbb{H}$, i.e.,

$$d_{\mathcal{M}}(\Gamma z_1, \Gamma z_2) = \inf_{\gamma \in \Gamma} d_{\mathbb{H}}(\gamma z_1, z_2).$$

Lemma 4.1 *Let $\Gamma z_0 \in \mathcal{M}$ be a fixed base point. Then there exists a constant $c > 0$ (which may depend on Γz_0) such that for any $Y > 1$ and for any $\Gamma z \in \mathcal{C}_Y$*

$$d_{\mathcal{M}}(\Gamma z_0, \Gamma z) \geq \log Y - c. \tag{4.1}$$

The estimate (4.1) holds for a general non-compact finite-volume hyperbolic manifold using reduction theory after Garland and Raghunathan [11, Theorem 0.6] combined with a distance estimate by Borel [2, Theorem C]. We give here a self-contained elementary proof for the special case of the modular surface.

Proof of Lemma 4.1. In view of the triangle inequality, we may assume $\Gamma z_0 = \Gamma i$. Note that $d_{\mathbb{H}}(i, z) \geq \log Y$ for any $z \in \mathbb{H}$ with $\mathfrak{Im}(z) \in (0, 1/Y) \cup (Y, \infty)$. Thus it suffices to show that if $\Gamma z \in \mathcal{C}_Y$, then $\mathfrak{Im}(\gamma z) \in (0, 1/Y) \cup (Y, \infty)$ for any $\gamma \in \Gamma$. By the definition of $\mathcal{C}_Y$, we may assume $z = x + iy \in \mathbb{H}$ with $y > Y$. Now let $\gamma = \left(\begin{smallmatrix} * & * \\ a & b \end{smallmatrix}\right) \in \Gamma$. If $a = 0$, then $\mathfrak{Im}(\gamma z) = \mathfrak{Im}(z) > Y$. If $a \neq 0$, then

$$\mathfrak{Im}(\gamma z) = \frac{\mathfrak{Im}(z)}{|az + b|^2} = \frac{y}{(ax + b)^2 + a^2 y^2} \leq \frac{1}{y} < \frac{1}{Y}.$$

$\square$

The following simple lemma is the key observation relating cusp excursions with Diophantine approximation.

Lemma 4.2 *Let $x \in [0, 1)$ be a real number. Suppose there exist a primitive rational number m/n and $n > 0$, and a real number $Y > 0$ satisfying*

$$\left| x - \frac{m}{n} \right| < \frac{1}{2Yn^2}.$$

Then for any $0 \leq j \leq n - 1$ we have

$$\Gamma\left(x + \tfrac{j}{n} + \tfrac{i}{2Yn^2}\right) \in \mathcal{C}_{Y_j, 2Y_j}, \quad \text{where } Y_j = \gcd(n, m + j)^2 Y. \tag{4.2}$$

In particular, we have

$$\left\{ \Gamma\left(x + \tfrac{j}{n} + \tfrac{i}{2Yn^2}\right) : 0 \leq j \leq n - 1 \right\} \subset \mathcal{C}_Y. \tag{4.3}$$

Proof The in particular part follows immediately from the inclusion $\mathcal{C}_{Y_j, 2Y_j} \subset \mathcal{C}_Y$, which in turn follows from the trivial bound $Y_j \geq Y$. Hence it suffices to prove the first half of the lemma. For simplicity of notation, we set $r = 1/(2Yn^2)$. Then by assumption $|x - \frac{m}{n}| < r$. Fix $0 \leq j \leq n - 1$, and let $\frac{p}{q}$ be the reduced form of $\frac{m+j}{n}$ (so that $q = \frac{n}{\gcd(n, m+j)}$). Then $x + \frac{j}{n} + ir \in H^\circ_{p/q, r}$ and $x + \frac{j}{n} + ir' \in H_{p/q, r}$ for some $r < r' < 2r$. Take $\gamma \in \Gamma$ sending $H^\circ_{p/q, r}$ to the region $\{ z \in \mathbb{H} : \mathfrak{Im}(z) > 1/(2rq^2) = Y_j \}$.

Then we have $\mathfrak{Im}\left(\gamma(x + \frac{j}{n} + ir)\right) > Y_j$ and $\mathfrak{Im}\left(\gamma(x + \frac{j}{n} + ir')\right) = Y_j$. Since $r < r' < 2r$ we can bound the hyperbolic distance

$$d_{\mathbb{H}}\left(\gamma(x + \tfrac{j}{n} + ir), \gamma(x + \tfrac{j}{n} + ir')\right) = \log\left(\tfrac{r'}{r}\right) < \log 2,$$

implying that

$$\gamma\left(x + \tfrac{j}{n} + ir\right) \in \{z \in \mathbb{H} : Y_j < \mathfrak{Im}(z) < 2Y_j\},$$

which implies (4.2).   □

4.2 Full escape to the cusp along subsequences for almost every translate

In this subsection we prove Theorem 4.3. Before stating this theorem, we first recall a definition from Diophantine approximation. Let $\psi : \mathbb{N} \to (0, 1/2)$ be a non-increasing function. We say that $x \in \mathbb{R}$ is *primitive ψ-approximable* if there exist infinitely many $n \in \mathbb{N}$ such that the inequality

$$\left|x - \frac{m}{n}\right| < \frac{\psi(n)}{n} \tag{4.4}$$

is satisfied by some $m \in \mathbb{Z}$ coprime to n. Since we assume $\psi(\mathbb{N}) \subset (0, 1/2)$, the existence of such an m implies its uniqueness. We prove the following:

Theorem 4.3 *Let $\psi : \mathbb{N} \to (0, 1/2)$ be a non-increasing function such that $\lim_{n \to \infty} n\psi(n) = 0$. Let $\{y_n\}_{n \in \mathbb{N}}$ be a sequence of positive numbers satisfying*

$$r_n := \frac{1}{2} \min\left\{\psi(n)^{-2} y_n, n^{-2} y_n^{-1}\right\} \xrightarrow{n \to \infty} \infty. \tag{4.5}$$

If $x \in [0, 1)$ is primitive ψ-approximable, then $\mathcal{R}_n(x, y_n) \subset \mathcal{C}_{r_n}$ infinitely often.

Remark 4.6 Since $\mathcal{R}_n^{\mathrm{pr}}(x, y) \subset \mathcal{R}_n(x, y)$ for any $n \in \mathbb{N}$, $x \in \mathbb{R}$ and $y > 0$, Theorem 4.3 also holds for translates of the primitive rational points.

Proof of Theorem 4.3 Let $x \in [0, 1)$ be primitive ψ-approximable. Then for $Y_n = 1/(2n\psi(n))$, we have by (4.3) that

$$\left\{\Gamma\left(x + \tfrac{j}{n} + i\tfrac{\psi(n)}{n}\right) \in \mathcal{M} : 0 \le j \le n - 1\right\} \subset \mathcal{C}_{Y_n} \tag{4.7}$$

for infinitely many n's.

For every $n \in \mathbb{N}$, set $d_n := Y_n/r_n = \max\{\psi(n)/(ny_n), ny_n/\psi(n)\}$. Then

$$d_{\mathbb{H}}(t + i\psi(n)/n, t + iy_n) = \log(d_n) \tag{4.8}$$

for any $t \in \mathbb{R}$. As in the proof of Lemma 4.2, by (4.7) and (4.8) we have $\mathcal{R}_n(x, y_n) \subset \mathcal{C}_{Y_n/d_n}$ for any n in (4.7).   □

We now give a short

Proof of Theorem 1.4 Let $\alpha = \min\{\beta, 2 - \beta\}$. For each $n \geq 2$, let $\psi(n) = 1/(n \log n)$ and let $\{y_n\}_{n \in \mathbb{N}}$ be a sequence of positive numbers satisfying $y_n \asymp 1/(n^2 \log^\beta n)$. Then r_n as in (4.5) is given by $r_n = \frac{1}{2} \min\{\psi(n)^{-2} y_n, n^{-2} y_n^{-1}\} \asymp \log^\alpha n$. By Theorem 4.3, for any $x \in [0, 1)$ primitive ψ-approximable, we have that $\mathcal{R}_n(x, y_n) \subset \mathcal{C}_{r_n}$ infinitely often. Hence by (4.1), for each such $x \in \mathbb{R}/\mathbb{Z}$, we have

$$\inf_{\Gamma z \in \mathcal{R}_n(x, y_n)} d_\mathcal{M}(\Gamma z_0, \Gamma z) \geq \log(r_n) + O(1) = \alpha \log \log n + O(1)$$

infinitely often, implying the inequality (1.10). Finally, since $\sum_{n \in \mathbb{N}} \psi(n) = \infty$ and ψ is decreasing, the set of primitive ψ-approximable numbers in $[0, 1)$ is of full measure by Khintchine's approximation theorem. □

For every irrational $x \in \mathbb{R}$, the *Diophantine exponent* $\kappa_x > 0$ is the supremum of $\kappa' > 0$ for which x is primitive $n^{-\kappa'}$-approximable. Dirichlet's approximation theorem implies that $\kappa_x \geq 1$ for any irrational x and by Khintchine's theorem, $\kappa_x = 1$ for almost every $x \in \mathbb{R}$. When $\kappa_x > 1$, we have the following result that yields much faster cusp excursion rates for our sample points while handling sequences $\{y_n\}_{n \in \mathbb{N}}$ decaying polynomially faster than $1/n^2$.

Theorem 4.4 *Let* $\Gamma z_0 \in \mathcal{M}$ *be a fixed base point. Let* $x \in [0, 1)$ *with Diophantine exponent* $\kappa_x > 1$ *and let* $\{y_n\}_{n \in \mathbb{N}}$ *be a sequence of positive numbers satisfying* $y_n \asymp n^{-\beta}$ *for some fixed* $2 < \beta < 2\kappa_x$. *Then*

$$\varlimsup_{n \to \infty} \frac{\inf_{\Gamma z \in \mathcal{R}_n(x, y_n)} d_\mathcal{M}(\Gamma z_0, \Gamma z)}{\log n} \geq \min\{2\kappa_x - \beta, \beta - 2\}.$$

Proof Take $\kappa \in (1, \kappa_x)$ and set $\alpha = \min\{2\kappa - \beta, \beta - 2\}$. Let $\psi(n) = 1/n^\kappa$. Then x is primitive ψ-approximable since $\kappa < \kappa_x$. By Theorem 4.3, we have $\mathcal{R}_n(x, y_n) \subset \mathcal{C}_{r_n}$ infinitely often with $r_n = \frac{1}{2} \min\{\psi(n)^{-2} y_n, n^{-2} y_n^{-1}\} \asymp n^\alpha$. This implies that

$$\varlimsup_{n \to \infty} \frac{\inf_{\Gamma z \in \mathcal{R}_n(x, y_n)} d_\mathcal{M}(\Gamma z_0, \Gamma z)}{\log n} \geq \alpha = \min\{2\kappa - \beta, \beta - 2\}.$$

Taking $\kappa \to \kappa_x$ finishes the proof. □

4.3 A non-equidistribution result for all translates

In this subsection we prove the following result which, together with part (1) of Theorem 1.3 implies non-equidistribution for all translates:

Theorem 4.5 *Let* $1/\sqrt{5} \leq c < 3/2$ *and let* $y_n = c/n^2$. *Then there exists a closed measurable subset* $\mathcal{E}_c \subset \mathcal{M}$, *depending only on* c, *with* $\mu_\mathcal{M}(\mathcal{E}_c) < 1$, *and such that for each irrational* $x \in [0, 1)$, $\mathcal{R}_n(x, y_n) \subset \mathcal{E}_c$ *infinitely often.*

The set $\mathcal{E}_c$ in Theorem 4.5 is explicit: For any $c > 0$, $\mathcal{E}_c \subset \mathcal{M}$ is defined to be the image of the closed set

$$\{z \in \mathbb{H} : \mathfrak{Im}(z) \in [1/(2c), 1/c] \cup [2/c, 4/c] \cup [9/(2c), \infty)\}$$

under the natural projection from $\mathbb{H}$ to $\mathcal{M}$. It is clear from the definition that $\mathcal{E}_c \subset \mathcal{M}$ is closed. Theorem 4.5 is a direct consequence of the following two lemmas.

Lemma 4.6 *For any $c > 0$ let $y_n = c/n^2$ and let $\psi_c(n) = c/n$. Then if $x \in [0, 1)$ is primitive ψ_c-approximable, we have $\mathcal{R}_n(x, y_n) \subset \mathcal{E}_c$ infinitely often.*

Proof Let $x \in [0, 1)$ be primitive ψ_c-approximable, that is, there exist infinitely many $n \in \mathbb{N}$ satisfying $|x - m/n| < c/n^2 = y_n$ with some uniquely determined $m \in \mathbb{Z}$ satisfying $\gcd(m, n) = 1$. For each such n, and for any $0 \le j \le n - 1$, let $k = \gcd(n, m + j)^2$. Then by (4.2), $\Gamma(x + j/n + iy_n) \in \mathcal{C}_{k^2/(2c),k^2/c}$. Moreover, since $(k^2/(2c), k^2/c) \subset [1/(2c), 1/c] \cup [2/c, 4/c] \cup [9/(2c), \infty)$ for any $k \in \mathbb{N}$, we have $\mathcal{C}_{k^2/(2c),k^2/c} \subset \mathcal{E}_c$ for any $k \in \mathbb{N}$, implying that $\mathcal{R}_n(x, y_n) \subset \mathcal{E}_c$ for these infinitely many $n \in \mathbb{N}$. □

Lemma 4.7 *For any $0 < c < 3/2$, we have $\mu_{\mathcal{M}}(\mathcal{E}_c) \le 1 - \frac{3}{\pi} \left(\frac{1}{\max\{2c,4/c\}} - \frac{2c}{9} \right) < 1$.*

Proof Let $\mathcal{U} \subset \mathcal{M}$ be the projection of the open set

$$\{z \in \mathbb{H} : \max\{2c, 4/c\} < \mathfrak{Im}(z) < 9/(2c)\}.$$

Since $0 < c < 3/2$ we have $\max\{2c, 4/c\} < 9/(2c)$ implying that $\mathcal{U}$ is nonempty. We will show that $\mathcal{E}_c$ is disjoint from $\mathcal{U}$. Let $I_1 = [1/(2c), 1/c]$, $I_2 = [2/c, 4/c]$ and $I_3 = [9/(2c), \infty)$, and for $1 \le j \le 3$, define $\mathcal{E}_c^j$ to be the projection onto $\mathcal{M}$ of $\{z \in \mathbb{H} : \mathfrak{Im}(z) \in I_j\}$ such that $\mathcal{E}_c = \bigcup_{j=1}^3 \mathcal{E}_c^j$. It thus suffices to show that $\mathcal{E}_c^j \cap \mathcal{U} = \emptyset$ for each $1 \le j \le 3$. For this, we identify (up to a null set) $\mathcal{M}$ with the standard fundamental domain $\mathcal{F}_\Gamma := \{z \in \mathbb{H} : \mathfrak{Re}(z) < \frac{1}{2}, |z| > 1\}$. Since $0 < c < 3/2$, we have $\max\{2c, 4/c\} > 2/c > 2/(3/2) > 1$. Thus we have

$$\mathcal{U} = \{z \in \mathcal{F}_\Gamma : \max\{2c, 4/c\} < \mathfrak{Im}(z) < 9/(2c)\}, \qquad \mathcal{E}_c^j = \{z \in \mathcal{F}_\Gamma : \mathfrak{Im}(z) \in I_j\}$$

for $j = 2, 3$. Moreover, since the interval $(\max\{2c, 4/c\}, 9/2c)$ intersects I_2 and I_3 trivially, we have $\mathcal{E}_c^j \cap \mathcal{U} = \emptyset$ for $j = 2, 3$. It thus remains to show that $\mathcal{E}_c^1 \cap \mathcal{U} = \emptyset$. For this we note that $z \in \mathcal{F}_\Gamma$ satisfies the property that

$$\mathfrak{Im}(z) = \max_{\gamma \in \Gamma} \mathfrak{Im}(\gamma z).$$

Hence to show $\mathcal{E}_c^1 \cap \mathcal{U} = \emptyset$, it suffices to show that $\max_{\gamma \in \Gamma} \mathfrak{Im}(\gamma z) \le \max\{2c, 4/c\}$ for any $z = s + it \in \mathbb{H}$ with $\mathfrak{Im}(z) = t \in I_1 = [1/(2c), 1/c]$. For this, using the

same discussion as in the proof of Lemma 4.1 we have for any $z = s + it \in \mathbb{H}$ with $t \in [1/(2c), 1/c]$

$$\max_{\gamma \in \Gamma} \Im(\gamma z) \leq \max \left\{ t, t^{-1} \right\} \leq \max \{1/c, 2c\} \leq \max \{2c, 4/c\}.$$

Finally, using the above description of $\mathcal{U}$ and (2.1) we have by direct computation

$$\mu_{\mathcal{M}}(\mathcal{U}) = \frac{3}{\pi} \left(\frac{1}{\max\{2c, 4/c\}} - \frac{2c}{9} \right)$$

implying that $\mu_{\mathcal{M}}(\mathcal{E}_c) \leq 1 - \frac{3}{\pi} \left(\frac{1}{\max\{2c, 4/c\}} - \frac{2c}{9} \right) < 1$ (again since $0 < c < 3/2$). $\square$

Proof of Theorem 4.5 Let $\psi_c(n) = c/n$. Since $c \geq 1/\sqrt{5}$, any irrational number is primitive ψ_c-approximable by the Hurwitz's approximation theorem; see, e.g., [14, Theorem 193]. Hence by Lemma 4.6, for each irrational $x \in [0, 1)$, we have $\mathcal{R}_n(x, y_n) \subset \mathcal{E}_c$ infinitely often. Moreover, since $c < 3/2$ by Lemma 4.7 we have $\mu_{\mathcal{M}}(\mathcal{E}_c) < 1$, finishing the proof. $\square$

Remark 4.9 The condition on the sequence $\{y_n\}_{n \in \mathbb{N}}$ in Theorem 4.5 is quite restrictive and the proof of Theorem 4.5 is much more involved than that of Theorem 4.3. We note that this is because we need to take care of the *badly approximable numbers*, that is, the set of irrational numbers that are not primitive ψ_c-approximable for some $c > 0$. If $x \in [0, 1)$ is not badly approximable, then a similar argument as in the proof of Theorem 4.3 using only the crude estimate (4.3) would already be sufficient to prove non-equidistribution of the sample points $\mathcal{R}_n(x, y_n)$ for any sequence $\{y_n\}_{n \in \mathbb{N}}$ satisfying $y_n \asymp 1/n^2$.

5 Second moments of the discrepancy

Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and let $\mathcal{M} = \Gamma \backslash \mathbb{H}$ be the modular surface as before. In this section we prove Theorem 1.6. Our proof relies on a second moment computation of the discrepancies $|\delta_{n,x,y} - \mu_{\mathcal{M}}|$ and $|\delta^{\mathrm{pr}}_{n,x,y} - \mu_{\mathcal{M}}|$ along the closed horocycle $\mathcal{H}_y$. Throughout this section, we abbreviate the second moments $\int_0^1 |\delta_{n,x,y}(\Psi) - \mu_{\mathcal{M}}(\Psi)|^2 \, dx$ and $\int_0^1 |\delta^{\mathrm{pr}}_{n,x,y}(\Psi) - \mu_{\mathcal{M}}(\Psi)|^2 \, dx$ by $D_{n,y}(\Psi)$ and $D^{\mathrm{pr}}_{n,y}(\Psi)$ respectively. Since we assume $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ we will also use the notation μ_{Γ} for $\mu_{\mathcal{M}}$.

5.1 Relation to Hecke operators

In this subsection we prove two preliminary estimates relating these second moments to the Hecke operators defined in Sect. 2.3.

Proposition 5.1 *For any $n \in \mathbb{N}$, $y > 0$ and $\Psi \in C_c^\infty(\mathcal{M})$, we have*

$$D_{n,y}(\Psi) = \frac{1}{n} \sum_{j=0}^{n-1} \langle \Psi_0, \widetilde{T}_{u_{j/n}}(\Psi_0) \rangle + O\left(\mathcal{S}(\Psi) y^{1/2} \right), \tag{5.1}$$

and

$$D_{n,y}^{\mathrm{pr}}(\Psi) \leq \frac{1}{\varphi(n)} \sum_{j=0}^{n-1} \left| \langle \Psi_0, \widetilde{T}_{u_{j/n}}(\Psi_0) \rangle \right| + O\left(\mathcal{S}(\Psi) y^{1/2} \right). \tag{5.2}$$

where $\Psi_0 = \Psi - \mu_\Gamma(\Psi)$, $\widetilde{T}_{u_{j/n}}$ is the Hecke operator associated to $u_{j/n} \in \mathrm{SL}_2(\mathbb{Q})$ defined as in (2.9), the Sobolev norm $\mathcal{S}(\Psi)$ is defined by

$$\mathcal{S}(\Psi) := \mathcal{S}_{4,2}^\Gamma(\Psi)^2 + \mathcal{S}_{2,2}^\Gamma(\Psi) \mathcal{S}_{1,0}^\Gamma(\Psi), \tag{5.3}$$

and the implied constants are absolute.

Proof Without loss of generality we may assume that Ψ is real-valued. Expanding the square in the left hand side of (5.1), doing a change of variables, and using the left u_1-invariance of Ψ, we have that $D_{n,y}(\Psi)$ equals

$$\frac{1}{n^2} \sum_{j_1,j_2=0}^{n-1} \int_0^1 \Psi(x + \tfrac{j_1}{n} + iy) \Psi(x + \tfrac{j_2}{n} + iy) dx$$

$$- 2\mu_\Gamma(\Psi) \frac{1}{n} \sum_{j=0}^{n-1} \int_0^1 \Psi(x + \tfrac{j}{n} + iy) dx + \mu_\Gamma(\Psi)^2$$

$$= \frac{1}{n} \sum_{j=0}^{n-1} \int_0^1 \Psi(x + iy) \Psi(x + \tfrac{j}{n} + iy) dx - 2\mu_\Gamma(\Psi) \int_0^1 \Psi(x + iy) dx + \mu_\Gamma(\Psi)^2.$$

Applying (2.14) to the term $\int_0^1 \Psi(x + iy) dx$ and using the trivial estimate

$$\|\Psi\|_2^{3/4} \|\Delta\Psi\|^{1/4} |\mu_\Gamma(\Psi)| \leq \mathcal{S}_{2,2}^\Gamma(\Psi) \mathcal{S}_{1,0}^\Gamma(\Psi) \leq \mathcal{S}(\Psi), \tag{5.4}$$

we get

$$D_{n,y}(\Psi) = \frac{1}{n} \sum_{j=0}^{n-1} \int_0^1 \Psi(x + iy) \Psi(x + \tfrac{j}{n} + iy) dx - \mu_\Gamma(\Psi)^2 + O(\mathcal{S}(\Psi) y^{1/2}).$$

$$\tag{5.5}$$

For each $0 \leq j \leq n - 1$, let $\Gamma^j := \Gamma^{u_{j/n}} = \Gamma \cap u_{j/n}^{-1} \Gamma u_{j/n}$ and define $F_j(\Psi) := \Psi L_{u_{j/n}^{-1}} \Psi \in C^\infty(\mathbb{H})$. Since Ψ is left Γ-invariant, and $L_{u_{j/n}^{-1}} \Psi$ is left

$u_{j/n}^{-1}\Gamma u_{j/n}$-invariant, we have $F_j(\Psi) \in C^\infty(\Gamma^j\backslash\mathbb{H})$. Moreover,

$$F_j(\Psi)(x+iy) = \Psi(x+iy)\Psi(x+\tfrac{j}{n}+iy).$$

For each $0 \le j \le n-1$, it is easy to check that $u_1 \in \Gamma^j$ and Γ^j contains the principal congruence subgroup $\Gamma(n^2)$, hence Γ^j satisfies the assumptions in Proposition 2.3. Then by (2.14),

$$\int_0^1 F_j(\Psi)(x+iy)dx$$
$$= \int_{\Gamma^j\backslash\mathbb{H}} F_j(\Psi)(z)d\mu_{\Gamma^j}(z) + O\left(\|F_j(\Psi)\|_2^{3/4}\|\Delta F_j(\Psi)\|_2^{1/4}y^{1/2}\right).$$

Next we note that by (2.3),

$$\|F_j(\Psi)\|_2^{3/4}\|\Delta F_j(\Psi)\|_2^{1/4} \le \mathcal{S}_{2,2}^{\Gamma^j}\left(F_j(\Psi)\right)$$
$$= \mathcal{S}_{2,2}^{\Gamma^j}\left(\Psi L_{u_{j/n}^{-1}}\Psi\right) \le \mathcal{S}_{4,2}^{\Gamma^j}(\Psi)\,\mathcal{S}_{4,2}^{\Gamma^j}\left(L_{u_{j/n}^{-1}}\Psi\right).$$

Using the fact that Ψ is left Γ-invariant and Γ^j is a finite-index subgroup of Γ, by (2.4), $\mathcal{S}_{4,2}^{\Gamma^j}(\Psi) = \mathcal{S}_{4,2}^{\Gamma}(\Psi)$. Similarly, we have

$$\mathcal{S}_{4,2}^{\Gamma^j}\left(L_{u_{j/n}^{-1}}\Psi\right) = \mathcal{S}_{4,2}^{u_{j/n}^{-1}\Gamma u_{j/n}}\left(L_{u_{j/n}^{-1}}\Psi\right) = \mathcal{S}_{4,2}^{\Gamma}(\Psi),$$

where for the second equality we used (2.2). Hence we have

$$\|F_j(\Psi)\|_2^{3/4}\|\Delta F_j(\Psi)\|_2^{1/4} \le \mathcal{S}_{2,2}^{\Gamma^j}\left(F_j(\Psi)\right) \le \mathcal{S}_{4,2}^{\Gamma}(\Psi)^2 \le \mathcal{S}(\Psi) < \infty. \quad (5.6)$$

Thus applying (2.14) to $F_j \in C^\infty(\Gamma^j\backslash\mathbb{H})$ and using (5.6) we get

$$\int_0^1 \Psi(x+iy)\Psi\left(x+\tfrac{j}{n}+iy\right)dx = \left\langle\Psi, L_{u_{j/n}^{-1}}\Psi\right\rangle_{L^2(\Gamma^j\backslash\mathbb{H})} + O\left(\mathcal{S}(\Psi)y^{1/2}\right). (5.7)$$

Plugging (5.7) into (5.5) and using the identities $\mu_\Gamma(\Psi) = \mu_{\Gamma^j}(\Psi) = \mu_{\Gamma^j}(L_{u_{j/n}^{-1}}\Psi)$ (the second equality follows from the left G-invariance of the hyperbolic area μ_{Γ^j}) we get that

$$D_{n,y}(\Psi) = \frac{1}{n}\sum_{j=0}^{n-1}\left\langle\Psi_0, L_{u_{j/n}^{-1}}\Psi_0\right\rangle_{L^2(\Gamma^j\backslash\mathbb{H})} + O(\mathcal{S}(\Psi)y^{1/2}).$$

Let $\mathcal{F}_\Gamma \subset \mathbb{H}$ be a fundamental domain for $\Gamma\backslash\mathbb{H}$. The disjoint union $\bigsqcup_{\gamma\in\Gamma^j\backslash\Gamma}\gamma\mathcal{F}_\Gamma$ forms a fundamental domain for $\Gamma^j\backslash\mathbb{H}$. Thus we can conclude the proof of (5.1) by

noting that

$$\int_{\bigsqcup_{\gamma\in\Gamma^j\backslash\Gamma}\gamma\mathcal{F}_\Gamma}\Psi_0(z)\Psi_0(u_{j/n}z)d\mu_{\Gamma^j}(z) = \sum_{\gamma\in\Gamma^j\backslash\Gamma}\int_{\gamma\mathcal{F}_\Gamma}\Psi_0(z)\Psi_0(u_{j/n}z)d\mu_{\Gamma^j}(z)$$

$$= \int_{\mathcal{F}_\Gamma}\Psi_0(z)\left(\frac{1}{[\Gamma:\Gamma^j]}\sum_{\gamma\in\Gamma^j\backslash\Gamma}\Psi_0(u_{j/n}\gamma z)\right)d\mu_\Gamma(z)$$

$$= \int_{\mathcal{F}_\Gamma}\Psi_0(z)\widetilde{T}_{u_{j/n}}(\Psi_0)(z)d\mu_\Gamma(z),$$

where for the second equation we did a change of variable $z \mapsto \gamma z$, used the left Γ-invariance of Ψ and the relation $[\Gamma:\Gamma^j]\mu_{\Gamma^j} = \mu_\Gamma$, and for the last equality we used the expression (2.10). Similarly, applying the estimates (2.14) and (5.4) and making change of variables we see that $D_{n,y}^{\mathrm{pr}}(\Psi)$ equals

$$\frac{1}{\varphi(n)^2}\sum_{j=0}^{n-1}c(j)\int_0^1\Psi(x+iy)\Psi(x+\tfrac{j}{n}+iy)dx - \mu_\Gamma(\Psi)^2 + O\left(\mathcal{S}(\Psi)y^{1/2}\right),$$

where

$$c(j) := \#\left\{([j_1],[j_2]) \in (\mathbb{Z}/n\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times : [j_2] - [j_1] = [j]\right\}.$$

Now similar as before we can apply the estimate (5.7), the identities $\mu_\Gamma(\Psi) = \mu_{\Gamma^j}(\Psi) = \mu_{\Gamma^j}(L_{u_{j/n}^{-1}}\Psi)$ and $\sum_{j=0}^{n-1}c(j) = \varphi(n)^2$ to get

$$D_{n,y}^{\mathrm{pr}}(\Psi) = \frac{1}{\varphi(n)^2}\sum_{j=0}^{n-1}c(j)\left\langle\Psi_0, L_{u_{j/n}^{-1}}\Psi_0\right\rangle_{L^2(\Gamma^j\backslash\mathbb{H})} + O(\mathcal{S}(\Psi)y^{1/2})$$

$$= \frac{1}{\varphi(n)^2}\sum_{j=0}^{n-1}c(j)\left\langle\Psi_0, \widetilde{T}_{u_{j/n}}(\Psi_0)\right\rangle_{L^2(\Gamma\backslash\mathbb{H})} + O(\mathcal{S}(\Psi)y^{1/2}).$$

Finally we can finish the proof by noting that for each $0 \le j \le n-1$, $c(j) \le \varphi(n)$ (since for each $[j_1] \in (\mathbb{Z}/n\mathbb{Z})^\times$, there is at most one $[j_2] \in (\mathbb{Z}/n\mathbb{Z})^\times$ such that $[j_2] - [j_1] = [j]$). □

5.2 Second moment estimates

Combining Proposition 5.1 and the operator norm bound in Proposition 2.2 we have the following second moment estimates:

Theorem 5.2 *For any $n \in \mathbb{N}$, $y > 0$ and $\Psi \in C_c^\infty(\mathcal{M})$ we have*

$$\max\left\{D_{n,y}(\Psi), D_{n,y}^{\mathrm{pr}}(\Psi)\right\} \ll_\epsilon n^{-1+2\theta+\epsilon}\|\Psi\|_2^2 + \mathcal{S}(\Psi)y^{1/2}, \tag{5.8}$$

where $\theta = 7/64$ is the best bound towards the Ramanujan conjecture as before and the Sobolev norm $\mathcal{S}(\Psi)$ is as defined in (5.3).

Remark 5.9 It is also possible to approach the second moment computation using the spectral bounds on the Fourier coefficients of Ψ from Sect. 3.1 rather than Hecke operators. The spectral approach however yields a weaker estimate when $y > 0$ is small. For comparison, following the spectral approach, one obtains

$$\int_0^1 |\delta_{n,x,y}(\Psi) - \mu_\Gamma(\Psi)|^2 dx \ll_\epsilon \left(n^{-1} y^{-2(\theta+\epsilon)} + y^{1/2} \right) \mathcal{S}_{2,2}(\Psi).$$

Proof of Theorem 5.2 First we prove (5.8). For each $0 \leq j \leq n-1$, it is clear that $u_{j/n}$ is of degree $n_j := n/\gcd(n, j)$, and thus $\widetilde{T}_{u_{j/n}} = \widetilde{T}_{n_j}$. Applying (5.1), (5.2), the estimate $\varphi(n) \gg_\epsilon n^{-1+\epsilon/2}$ and the operator norm bound in Proposition 2.2 to the terms $\langle \Psi_0, \widetilde{T}_{n_j} \Psi_0 \rangle$, we get

$$\max \left\{ D_{n,y}(\Psi), D_{n,y}^{\mathrm{pr}}(\Psi) \right\} \ll_\epsilon n^{-1+\epsilon/2} \sum_{j=0}^{n-1} n_j^{-1+2\theta+\epsilon/4} \|\Psi_0\|_2^2 + \mathcal{S}(\Psi) y^{1/2}.$$

For any $d \mid n$, $\#\{0 \leq j \leq n-1 : n_j = d\} = \varphi(d)$, thus

$$\sum_{j=1}^n n_j^{-1+2\theta+\epsilon/4} = \sum_{d|n} \varphi(d) d^{-1+2\theta+\epsilon/4} < \sum_{d|n} d^{2\theta+\epsilon/4} = \sigma_{2\theta+\epsilon/4}(n) \ll_\epsilon n^{2\theta+\epsilon/2},$$

where for the first inequality we used the trivial bound $\varphi(d) < d$. Finally, we observe that $\|\Psi_0\|_2 \leq \|\Psi\|_2$. $\qquad\square$

We now give a quick

Proof of Theorem 1.6 Let $\alpha > 0$ be the fixed number as in this theorem. Let $\beta := \min\{\frac{\alpha}{2}, 1 - 2\theta\}$. Fix $0 < c < \beta$ and let $\mathcal{N} \subset \mathbb{N}$ be an unbounded subsequence such that $\sum_{n \in \mathcal{N}} n^{-c} < \infty$. We want to show that for any $\{y_n\}_{n \in \mathbb{N}}$ satisfying $y_n \ll n^{-\alpha}$ there exists a full measure subset $I \subset \mathbb{R}/\mathbb{Z}$ such that for any $x \in I$, $\delta_{n,x,y_n}(\Psi) \to \mu_{\mathcal{M}}(\Psi)$ and $\delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) \to \mu_{\mathcal{M}}(\Psi)$ for any $\Psi \in C_c^\infty(\mathcal{M})$ as $n \in \mathcal{N}$ goes to infinity. Since the function space $C_c^\infty(\mathcal{M})$ has a dense countable subset, it suffices to prove the above assertion for a fixed Ψ. Now we fix $\Psi \in C_c^\infty(\mathcal{M})$ and take $\epsilon > 0$ sufficiently small such that $\beta - 2\epsilon > c$. For any $n \in \mathbb{N}$ define $I_n = I_n^1 \cup I_n^2 \subset \mathbb{R}/\mathbb{Z}$ such that

$$I_n^1 := \left\{ x \in \mathbb{R}/\mathbb{Z} : \left| \delta_{n,x,y_n}(\Psi) - \mu_{\mathcal{M}}(\Psi) \right| > n^{-\epsilon/2} \right\},$$

and

$$I_n^2 := \left\{ x \in \mathbb{R}/\mathbb{Z} : \left| \delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) - \mu_{\mathcal{M}}(\Psi) \right| > n^{-\epsilon/2} \right\}.$$

Thus by the second moment estimate (5.8), the assumption that $y_n \ll n^{-\alpha}$ and Chebyshev's inequality we get

$$|I_n| \leq \left| I_n^1 \right| + \left| I_n^2 \right| \leq 2n^\epsilon \max \left\{ D_{n,y}(\Psi), D_{n,y}^{\mathrm{pr}}(\Psi) \right\} \ll_{\epsilon, \Psi} n^{-\beta + 2\epsilon} < n^{-c},$$

implying that $\sum_{n \in \mathcal{N}} |I_n| < \infty$. Hence taking $I \subset \mathbb{R}/\mathbb{Z}$ to be the complement of this limsup set $\overline{\lim}_{\substack{n \in \mathcal{N} \\ n \to \infty}} I_n \subset \mathbb{R}/\mathbb{Z}$ and by the Borel–Cantelli lemma we have I is of full measure. Moreover, for any $x \in I$, $x \in I_n^c$ for all $n \in \mathcal{N}$ sufficiently large, that is,

$$\max \left\{ \left| \delta_{n,x,y_n}(\Psi) - \mu_{\mathcal{M}}(\Psi) \right|, \left| \delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) - \mu_{\mathcal{M}}(\Psi) \right| \right\}$$
$$\leq n^{-\epsilon/2}, \quad \forall n \in \mathcal{N} \text{ sufficiently large.}$$

In particular for such x, $\delta_{n,x,y_n}(\Psi) \to \mu_{\mathcal{M}}(\Psi)$ and $\delta_{n,x,y_n}^{\mathrm{pr}}(\Psi) \to \mu_{\mathcal{M}}(\Psi)$ as $n \in \mathcal{N}$ goes to infinity. □

Remark 5.10 The second moment $D_{n,y}(\Psi)$ is closely related to the sample points (1.2) considered in [12]: Using the extra invariance $\delta_{n,x+1/n,y}(\Psi) = \delta_{n,x,y}(\Psi)$ and applying a change of variable, one can easily check that

$$D_{n,y}(\Psi) = \int_0^1 \left| \frac{1}{n} \sum_{j=0}^{n-1} \Psi \left(\frac{x+j}{n} + iy \right) - \mu_\Gamma(\Psi) \right|^2 dx.$$

Thus let $\mathcal{N} \subset \mathbb{N}$ be the fixed sequence as in the above proof, by Theorem 5.2 and the same Borel–Cantelli type argument we have that for almost every $x \in \mathbb{R}/\mathbb{Z}$ the sequence of sample points $\{\Gamma(\frac{x+j}{n} + iy_n : 0 \leq j \leq n-1\}$ equidistributes on $\mathcal{M}$ with respect to $\mu_{\mathcal{M}}$ as $n \in \mathcal{N}$ goes to infinity, as long as $\{y_n\}_{n \in \mathbb{N}}$ decays at least polynomially.

6 Left regular action of normalizing elements

In this section, Γ denotes a congruence subgroup, and we set by $\Gamma_1 = \mathrm{SL}_2(\mathbb{Z})$. We moreover assume that there exists some $h \in \mathrm{SL}_2(\mathbb{Q})$ normalizing Γ, that is, $h^{-1}\Gamma h = \Gamma$. It induces the left regular h-action on $\Gamma \backslash \mathbb{H}$ given by $\Gamma z \in \Gamma \backslash \mathbb{H} \mapsto \Gamma hz \in \Gamma \backslash \mathbb{H}$. Since h normalizes Γ, this map is well defined: Suppose $\Gamma z = \Gamma z'$, that is there exists some $\gamma \in \Gamma$ such that $z' = \gamma z$. Then $\Gamma hz' = \Gamma h\gamma z = \Gamma h\gamma h^{-1}hz = \Gamma hz$. The goal of this section is to describe this action on cylindrical cuspidal neighborhoods of $\Gamma \backslash \mathbb{H}$.

6.1 Cusp neighborhoods of congruence surfaces

Since Γ is a congruence subgroup, the set of cusps of Γ can be parameterized by the coset $\Gamma \backslash (\mathbb{Q} \cup \{\infty\})$ (see e.g. [21, p. 222]), where the action of Γ on $\mathbb{Q} \cup \{\infty\}$ is defined via the Möbius transformation. We denote by Ω_Γ a complete list of coset

representatives for $\Gamma \backslash (\mathbb{Q} \cup \{\infty\})$. For each cusp representative $\mathfrak{c} \in \Omega_\Gamma$, its stabilizer subgroup is given by

$$\Gamma_\mathfrak{c} := \tau_\mathfrak{c} N \tau_\mathfrak{c}^{-1} \cap \Gamma,$$

where $\tau_\mathfrak{c} \in \Gamma_1$ is such that $\tau_\mathfrak{c} \infty = \mathfrak{c}$. (More precisely, $\Gamma_\mathfrak{c}$ is an index two subgroup of the stabilizer subgroup of $\mathfrak{c}$ if $-I_2 \in \Gamma$.) The existence of such $\tau_\mathfrak{c}$ is guaranteed by the transitivity of the action of Γ_1 on $\mathbb{Q} \cup \{\infty\}$. On the other hand, $\tau_\mathfrak{c}$ is only unique up to right multiplication by any element of $\pm N$. We note that $\Gamma_\mathfrak{c}$ is independent of the choice of $\tau_\mathfrak{c}$, and since $\mathfrak{c} \in \Omega_\Gamma$ is a cusp, $\Gamma_\mathfrak{c}$ is nontrivial. Moreover, $\tau_\mathfrak{c}^{-1} \Gamma_\mathfrak{c} \tau_\mathfrak{c}$ is a subgroup of $N \cap \Gamma_1 = \langle u_1 \rangle$. Hence $\tau_\mathfrak{c}^{-1} \Gamma_\mathfrak{c} \tau_\mathfrak{c}$ is a cyclic group generated by a unipotent matrix $u_{\omega_\mathfrak{c}}$ for some positive integer $\omega_\mathfrak{c}$, which is called the *width* of the cusp $\mathfrak{c}$.

We can now define cusp neighborhoods on the hyperbolic surface $\Gamma \backslash \mathbb{H}$ around a cusp $\mathfrak{c} \in \Omega_\Gamma$. For any $Y > 0$, $C_Y^{\Gamma, \mathfrak{c}} \subset \Gamma \backslash \mathbb{H}$ denote the projection of the horodisc $\{\tau_\mathfrak{c} z \in \mathbb{H} : \Im(z) > Y\}$ onto $\Gamma \backslash \mathbb{H}$. Similarly, for any $Y' > Y > 0$, let $C_{Y,Y'}^{\Gamma, \mathfrak{c}}$ denote the projection of the cylindrical region $\{\tau_\mathfrak{c} z \in \mathbb{H} : Y < \Im(z) < Y'\}$ onto $\Gamma \backslash \mathbb{H}$. We record the following two lemmas for the later purpose of computing the measure of certain unions of cusp neighborhoods.

Lemma 6.1 *If $Y' > Y > 1$, the set $C_{Y,Y'}^{\Gamma, \mathfrak{c}}$ is in one-to-one correspondence with the set*

$$\{\tau_\mathfrak{c} z \in \mathbb{H} : \Re(z) \in \mathbb{R}/\omega_\mathfrak{c} \mathbb{Z}, \ \Im(z) \in (Y, Y')\}. \tag{6.1}$$

In particular, if $-I_2 \in \Gamma$ then for any $Y' > Y > 1$

$$\mu_\Gamma \left(C_{Y,Y'}^{\Gamma, \mathfrak{c}} \right) = \frac{3\omega_\mathfrak{c}}{\pi [\Gamma_1 : \Gamma]} \left(\frac{1}{Y} - \frac{1}{Y'} \right). \tag{6.2}$$

Proof The one-to-one correspondence is given by the projection of the above rectangular set onto $\Gamma \backslash \mathbb{H}$. Indeed, since $\Gamma_\mathfrak{c} \subset \Gamma$, this map projects the rectangular set in (6.1) onto $C_{Y,Y'}^{\Gamma, \mathfrak{c}}$. To show that it is also injective, suppose $\Gamma \tau_\mathfrak{c} z = \Gamma \tau_\mathfrak{c} z'$ for some z, z' from this rectangular set. Then there exists some $\gamma \in \Gamma$ such that $\tau_\mathfrak{c}^{-1} \gamma \tau_\mathfrak{c} z = z'$. If $\gamma \in \pm \Gamma_\mathfrak{c}$ then $\tau_\mathfrak{c}^{-1} \gamma \tau_\mathfrak{c} \in \pm \langle u_{\omega_\mathfrak{c}} \rangle$, and this implies that $z = z'$. Otherwise, let $\tau_\mathfrak{c}^{-1} \gamma \tau_\mathfrak{c} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in \Gamma_1$. Since $\gamma \notin \pm \Gamma_\mathfrak{c}, c \neq 0$. We easily see this cannot happen since it would imply

$$\Im(z') = \frac{\Im(z)}{|cz + d|^2} = \frac{\Im(z)}{(cx + d)^2 + c^2 y^2} \leq \frac{1}{y} \leq 1,$$

contradicting that $\Im(z') > Y > 1$. For the area computation, we use the definition (2.1) of μ_{Γ_1} together with $\mu_{\Gamma_1} = [\Gamma_1 : \Gamma] \mu_\Gamma$ (since $-I_2 \in \Gamma$). $\qquad \square$

Lemma 6.2 *Given two distinct cusps $\mathfrak{c}_1, \mathfrak{c}_2 \in \Omega_\Gamma$, and any $Y_1, Y_2 \geq 1$, $C_{Y_1}^{\Gamma, \mathfrak{c}_1} \cap C_{Y_2}^{\Gamma, \mathfrak{c}_2} = \emptyset$.*

Proof Since $Y_1, Y_2 \geq 1$, the sets $\{\tau_{\mathfrak{c}_1} z \in \mathbb{H} : \Im(z) > Y_1\}$ and $\{\tau_{\mathfrak{c}_2} z \in \mathbb{H} : \Im(z) > Y_2\}$ are subsets of the interior of the Ford circles based at $\mathfrak{c}_1$ and $\mathfrak{c}_2$ respectively. Two

Ford circles are either disjoint or identical. Suppose $\Gamma z \in \mathcal{C}_{Y_1}^{\Gamma, c_1} \cap \mathcal{C}_{Y_2}^{\Gamma, c_2}$. Then there exists an isometry $\gamma \in \Gamma$ that maps the Ford circle at c_1 to the Ford circle at c_2. Consequently, we must have $\gamma c_1 = c_2$, which is a contradiction.   □

Remark 6.3 We will later consider sets $I_{y, Y, c} := \left\{ x \in (0, 1) : \Gamma(x + iy) \in \mathcal{C}_Y^{\Gamma, c} \right\}$ for some $y > 0, Y > 1$ and $c \in \Omega_\Gamma$. This set is the intersection of the line segment $\{x + iy \in \mathbb{H} : 0 < x < 1\}$ with the preimage of $\mathcal{C}_Y^{\Gamma, c}$ in $\mathbb{H}$ (under the natural projection from $\mathbb{H}$ to $\Gamma \backslash \mathbb{H}$). By definition the preimage of $\mathcal{C}_Y^{\Gamma, c}$ is the disjoint (since $Y > 1$) union of the infinitely many horodiscs $\{\tau_{c'} z \in \mathbb{H} : \Im(z) > Y\} = H^\circ_{p/q, 1/(2q^2 Y)}$ for all cusps $c' = p/q \in \Gamma c$. Moreover, note that a necessary condition for such a horodisc intersecting the line segment $\{x + iy \in \mathbb{H} : 0 < x < 1\}$ is that $p/q \in \Gamma c \cap (-\frac{1}{2Y}, 1 + \frac{1}{2Y})$ and $1/(q^2 Y) > y$, i.e. $q^2 < 1/(yY)$. Thus there are only finitely many such horodiscs intersecting $\{x + iy \in \mathbb{H} : 0 < x < 1\}$. Moreover, each such intersection is an open interval and the set $I_{y, Y, c} \subset (0, 1)$ is thus the disjoint union of these finitely many open intervals. Similarly, for any $Y' > Y > 1$ the set $\left\{ x \in (0, 1) : \Gamma(x + iy) \in \mathcal{C}_{Y, Y'}^{\Gamma, c} \right\} = I_{y, Y, c} \backslash \overline{I}_{y, Y', c}$ is also a disjoint union of finitely many open intervals.

6.2 Left regular action of normalizing elements

Let $h \in \mathrm{SL}_2(\mathbb{Q})$ be a group element normalizing Γ. The action of h on $\mathbb{Q} \cup \{\infty\}$ (by Möbius transformation) induces a well-defined action on $\Gamma \backslash (\mathbb{Q} \cup \{\infty\})$, the set of cusps of Γ.

Lemma 6.3 *For each $c \in \Omega_\Gamma$, we have*

$$h \Gamma_c h^{-1} = \Gamma_{hc} \tag{6.4}$$

and

$$\tau_{hc}^{-1} h \tau_c = \begin{pmatrix} \sqrt{\omega_{hc}/\omega_c} & * \\ 0 & \sqrt{\omega_c/\omega_{hc}} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Q}). \tag{6.5}$$

Proof Since h normalizes Γ we have $h \Gamma_c h^{-1} = h \tau_c N \tau_c^{-1} h^{-1} \cap \Gamma$. Thus to prove (6.4) it suffices to show $h \tau_c N \tau_c^{-1} h^{-1} = \tau_{hc} N \tau_{hc}^{-1}$. We show that $\tau_{hc}^{-1} h \tau_c$ is an upper triangular matrix. Indeed, $\tau_{hc}^{-1} h \tau_c \infty = \tau_{hc}^{-1}(hc) = \infty$. This proves (6.4). We moreover conclude that

$$\tau_{hc}^{-1} h \tau_c = \begin{pmatrix} \lambda & * \\ 0 & \lambda^{-1} \end{pmatrix} \tag{6.6}$$

for some $\lambda \neq 0$, and it remains to show that $\lambda^2 = \omega_{hc}/\omega_c$. For this we conjugate the subgroup $\tau_{hc}^{-1} \Gamma_{hc} \tau_{h \cdot c}$ by the matrix $\tau_{hc}^{-1} h \tau_c$. We obtain with (6.4) that

$$\tau_c^{-1} h^{-1} \tau_{hc} \left(\tau_{hc}^{-1} \Gamma_{hc} \tau_{hc} \right) \tau_{hc}^{-1} h \tau_c = \tau_c^{-1} \Gamma_c \tau_c = \langle u_{\omega_c} \rangle.$$

On the other hand, using (6.6) and $\tau_{h\mathfrak{c}}^{-1}\Gamma_{h\mathfrak{c}}\tau_{h\mathfrak{c}} = \langle u_{\omega_{h\mathfrak{c}}} \rangle$, we have

$$\tau_{\mathfrak{c}}^{-1}h^{-1}\tau_{h\mathfrak{c}}\left(\tau_{h\mathfrak{c}}^{-1}\Gamma_{h\mathfrak{c}}\tau_{h\mathfrak{c}}\right)\tau_{h\mathfrak{c}}^{-1}h\tau_{\mathfrak{c}} = \left(\begin{smallmatrix} \lambda^{-1} & * \\ 0 & \lambda \end{smallmatrix}\right)\left\langle\left(\begin{smallmatrix} 1 & \omega_{h\mathfrak{c}} \\ 0 & 1 \end{smallmatrix}\right)\right\rangle\left(\begin{smallmatrix} \lambda & * \\ 0 & \lambda^{-1} \end{smallmatrix}\right) = \left\langle\left(\begin{smallmatrix} 1 & \omega_{h\mathfrak{c}}/\lambda^2 \\ 0 & 1 \end{smallmatrix}\right)\right\rangle.$$

Comparing both equations we conclude that $\lambda^2 = \omega_{h\mathfrak{c}}/\omega_{\mathfrak{c}}$. Finally replacing $\tau_{h\mathfrak{c}}$ with $-\tau_{h\mathfrak{c}}$ if necessary, we can ensure λ is positive. □

Proposition 6.4 *Let $Y' > Y > 0$ and $\mathfrak{c} \in \Omega_\Gamma$. If $\Gamma z \in \mathcal{C}_{\omega_\mathfrak{c}Y,\omega_\mathfrak{c}Y'}^{\Gamma,\mathfrak{c}}$, then $\Gamma hz \in \mathcal{C}_{\omega_{h\mathfrak{c}}Y,\omega_{h\mathfrak{c}}Y'}^{\Gamma,h\mathfrak{c}}$. Similarly, if $\Gamma z \in \mathcal{C}_{\omega_\mathfrak{c}Y}^{\Gamma,\mathfrak{c}}$, then $\Gamma hz \in \mathcal{C}_{\omega_{h\mathfrak{c}}Y}^{\Gamma,h\mathfrak{c}}$.*

Proof The second statement follows from the first one by taking $Y' \to \infty$. Since $\Gamma z \in \mathcal{C}_{\omega_\mathfrak{c}Y,\omega_\mathfrak{c}Y'}^{\Gamma,\mathfrak{c}}$, by definition there exists $z' = x' + iy' \in \mathbb{H}$ with $0 \le x' < \omega_\mathfrak{c}$ and $\omega_\mathfrak{c}Y < y' < \omega_\mathfrak{c}Y'$ and $\Gamma z = \Gamma\tau_\mathfrak{c}z'$. Consider $h\tau_\mathfrak{c}z' = \tau_{h\mathfrak{c}}z''$ with $z'' = \tau_{h\mathfrak{c}}^{-1}h\tau_\mathfrak{c}z'$. By (6.5), we have $\mathfrak{Im}(z'') = (\omega_{h\mathfrak{c}}/\omega_\mathfrak{c})\mathfrak{Im}(z') \in (\omega_{h\mathfrak{c}}Y, \omega_{h\mathfrak{c}}Y')$, implying that $\Gamma hz = \Gamma h\tau_\mathfrak{c}z' \in \mathcal{C}_{\omega_{h\mathfrak{c}}Y,\omega_{h\mathfrak{c}}Y'}^{\Gamma,h\mathfrak{c}}$. □

7 Negative results: horocycles expanding arbitrarily fast

In this section using the results from the previous section, we prove Theorems 1.7 and 1.8 which provide new limiting measures for the sequences $\{\delta_{n,x,y_n}\}_{n\in\mathbb{N}}$ and $\{\delta_{n,x,y_n}^{\mathrm{pr}}\}_{n\in\mathbb{N}}$, allowing $\{y_n\}_{n\in\mathbb{N}}$ to decay arbitrarily fast. For any $n \in \mathbb{N}$ we consider the congruence subgroup $\Gamma_n < \mathrm{SL}_2(\mathbb{Z})$ given by

$$\Gamma_n := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : n^2 \mid c, \; a \equiv d \equiv \pm 1 \pmod{n} \right\}. \tag{7.1}$$

It is clear that $\Gamma_1 = \mathrm{SL}_2(\mathbb{Z})$ and that Γ_n contains the congruence subgroup

$$\Gamma_1(n^2) := \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{n^2} \right\}.$$

7.1 Basic properties of the congruence subgroups Γ_n

First we show that Γ_n is normalized by $u_{j/n}$ for any $j \in \mathbb{Z}$. As mentioned in the introduction this simple fact is the starting point of our proofs to Theorems 1.7 and 1.8.

Lemma 7.1 *For any $n \in \mathbb{N}$ and for any $j \in \mathbb{Z}$, the unipotent matrix $u_{j/n}$ normalizes Γ_n.*

Proof By direct computation, for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1$ and for any $j \in \mathbb{Z}$ we have

$$u_{j/n}^{-1}\gamma u_{j/n} = \begin{pmatrix} a - \frac{jc}{n} & b + \frac{(a-d)j}{n} - \frac{j^2c}{n^2} \\ c & d + \frac{jc}{n} \end{pmatrix}.$$

Hence if $\gamma \in \Gamma_n$, that is, $n^2 \mid c$ and $a \equiv d \equiv \pm 1 \pmod{n}$, all the entries are integers with the bottom left entry divisible by n^2, and

$$a - \frac{jc}{n} \equiv a \equiv d \equiv d + \frac{jc}{n} \equiv \pm 1 \pmod{n}.$$

This implies that $u_{j/n}^{-1} \Gamma_n u_{j/n} \subset \Gamma_n$. □

Next we prove the following index formula for Γ_n.

Lemma 7.2 *For any integer $n \geq 3$, we have*

$$[\Gamma_1 : \Gamma_n] = \frac{n^3}{2} \prod_{\substack{p \mid n \\ prime}} \left(1 - p^{-2}\right). \tag{7.2}$$

Proof Let $J_n < \left(\mathbb{Z}/n^2\mathbb{Z}\right)^\times$ be the subgroup

$$J_n := \left\{ [a] \in \left(\mathbb{Z}/n^2\mathbb{Z}\right)^\times : a \equiv \pm 1 \pmod{n} \right\}. \tag{7.3}$$

It is easy to check that $\#(J_n) = 2n$. Consider the map $h : \Gamma_n \to J_n$ sending $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_n$ to $[a] \in \left(\mathbb{Z}/n^2\mathbb{Z}\right)^\times$. Using the definition of Γ_n, one can check that h is a group homomorphism with the kernel $\ker(h) = \Gamma_1(n^2)$. For each $0 \leq k \leq n - 1$, set $\gamma_k^\pm = \pm \left(\begin{smallmatrix} 1+kn & 1 \\ -k^2n^2 & 1-kn \end{smallmatrix}\right) \in \Gamma_n$. Then h surjects the set $\left\{\gamma_k^\pm \in \Gamma_n : 0 \leq k \leq n - 1\right\}$ onto J_n. Finally we use the index formula for $\Gamma_1(n^2)$ (see e.g. [6, Section 1.2]) to get

$$[\Gamma_1 : \Gamma_n] = \frac{[\Gamma_1 : \Gamma_1(n^2)]}{[\Gamma_n : \Gamma_1(n^2)]} = \frac{[\Gamma_1 : \Gamma_1(n^2)]}{\#J_n} = \frac{n^3}{2} \prod_{\substack{p \mid n \\ prime}} \left(1 - p^{-2}\right).$$

Next, we study the properties of Γ_n relative to its cusps. As in Sect. 6 we denote by Ω_{Γ_n} the set of cusps of Γ_n. The following lemma computes the width of each cusp of Γ_n.

Lemma 7.3 *Let $n \in \mathbb{N}$ and let $\mathfrak{c} = m/l \in \Omega_{\Gamma_n}$ with $\gcd(m, l) = 1$ (if $\mathfrak{c} = \infty$, m/l is understood as $1/0$). Then we have*

$$\omega_{\mathfrak{c}} = \frac{n^2}{\gcd(n, l)^2}.$$

Proof Let $\tau_{\mathfrak{c}} \in \Gamma_1$ be as before such that $\tau_{\mathfrak{c}}\infty = \mathfrak{c}$. Thus the left column of $\tau_{\mathfrak{c}}$ is $\left(\begin{smallmatrix} m \\ l \end{smallmatrix}\right)$. By direct computation we have

$$\tau_{\mathfrak{c}} N \tau_{\mathfrak{c}}^{-1} = \left\{ \left(\begin{smallmatrix} 1 - mlt & m^2t \\ -l^2t & 1 + mlt \end{smallmatrix}\right) \in G : t \in \mathbb{R} \right\}.$$

Thus by (7.1) an element in $(\Gamma_n)_{\mathfrak{c}} = \tau_{\mathfrak{c}} N \tau_{\mathfrak{c}}^{-1} \cap \Gamma_n$ is of the form $\gamma = \begin{pmatrix} 1-mlt & m^2t \\ -l^2t & 1+mlt \end{pmatrix} \in$ Γ_1 satisfying that $n^2 \mid l^2t$ and $1 - mlt \equiv 1 + mlt \equiv \pm 1 \pmod{n}$. Looking at the top right and bottom left entries of γ, we have that $m^2t, l^2t \in \mathbb{Z}$. Since $\gcd(m, l) = 1$, we have $t \in \mathbb{Z}$. Then the condition $n^2 \mid l^2t$ is equivalent to $\frac{n^2}{\gcd(n,l)^2} \mid t$, and the condition $n \mid mlt$ is equivalent to that $\frac{n}{\gcd(n,ml)} \mid t$. Moreover, since $\frac{n}{\gcd(n,ml)} \mid \frac{n^2}{\gcd(n,l)^2}$, the condition $\frac{n}{\gcd(n,ml)} \mid t$ is implied by the condition $\frac{n^2}{\gcd(n,l)^2} \mid t$. We conclude that $n^2 \mid l^2t$ implies $1 - mlt \equiv 1 + mlt \equiv \pm 1 \pmod{n}$. Thus

$$(\Gamma_n)_{\mathfrak{c}} = \left\{ \begin{pmatrix} 1-mlt & m^2t \\ -l^2t & 1+mlt \end{pmatrix} \in \Gamma_1 : n^2 \mid l^2t \right\}.$$

Conjugating $(\Gamma_n)_{\mathfrak{c}}$ back via $\tau_{\mathfrak{c}}$ and using the equivalence of the two conditions $n^2 \mid l^2t$ and $\frac{n^2}{\gcd(n,l)^2} \mid t$ we get

$$\tau_{\mathfrak{c}}^{-1}(\Gamma_n)_{\mathfrak{c}}\tau_{\mathfrak{c}} = \left\{ u_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in \Gamma_1 : \frac{n^2}{\gcd(n,l)^2} \mid t \right\},$$

implying that $\omega_{\mathfrak{c}} = n^2/\gcd(n, l)^2$. □

Next we compute the number of cusps of Γ_n.

Proposition 7.4 *For any integer $n \geq 3$ we have*

$$\#\Omega_{\Gamma_n} = \frac{n^2}{2} \prod_{\substack{p \mid n \\ prime}} \left(1 - p^{-2} \right).$$

Remark 7.4 It is easy to check that $\Gamma_2 = \Gamma_0(4)$. Thus $[\Gamma_1 : \Gamma_2] = 6$ and Γ_2 has three cusps which can be represented by ∞, $1/2$ and 1 respectively.

To prove Proposition 7.4 we first prove a preliminary formula for $\#\Omega_{\Gamma_n}$.

Lemma 7.5 *For any integer $n \geq 3$ we have*

$$\#\Omega_{\Gamma_n} = \sum_{d \mid n^2} \frac{\varphi(n^2/d)\varphi(d)\gcd(n^2/d, d)}{2n}.$$

Proof Since $-I_2 \in \Gamma_n$ and $\Gamma_1(n^2) < \Gamma_n$, we have $\Omega_{\Gamma_n} = \Gamma_n \backslash \Omega_{\Gamma_1(n^2)}$. On the other hand, by the analysis in [6, p. 102], the set $\Omega_{\Gamma_1(n^2)}$ is in bijection with the union of cosets $\bigsqcup_{d \mid n^2} \langle \pm I_2 \rangle \backslash Z_d$, where for each $d \mid n^2$,

$$Z_d := \left\{ ([m], [l])^t : [m] \in (\mathbb{Z}/d\mathbb{Z})^\times, [l] \in \mathbb{Z}/n^2\mathbb{Z}, \gcd(n^2, l) = d \right\}$$

with $([m], [l])^t$ is the transpose of the row vector $([m], [l])$ and the bijection is induced by the map sending $m/l \in \mathbb{Q} \cup \{\infty\}$ with $\gcd(m, l) = 1$ to $([m], [l]))^t \in Z_d$ with $d = \gcd(n^2, l)$. Note that $\#Z_d = \varphi(n^2/d)\varphi(d)$.

For each $d \mid n^2$, using the definition of Γ_n, it is easy to check that the linear action of Γ_n on $\mathbb{Z}^2$ (by matrix multiplication) induces a well-defined action of Γ_n on Z_d and that the corresponding action of the subgroup $\Gamma_1(n^2)$ is trivial. From the proof of Lemma 7.2, we have $\Gamma_n / \Gamma_1(n^2) \cong J_n$, where

$$J_n = \left\{ \pm[1 + kn] \in (\mathbb{Z}/n^2\mathbb{Z})^\times : 0 \le k \le n - 1 \right\}, \tag{7.5}$$

which is of size $2n$. Hence the action of Γ_n on Z_d induces the action of J_n on Z_d given by

$$[a] \cdot ([m], [l])^t = ([am], [\bar{a}l])^t,$$

with $([m], [l])^t \in Z_d$ and $\bar{a}$ the multiplicative inverse of a modulo n^2. We note that $[am] \in (\mathbb{Z}/d\mathbb{Z})^\times$ is well-defined since $d \mid n^2$.

We conclude that $\Omega_{\Gamma_n} = \Gamma_n \backslash \Omega_{\Gamma_1(n^2)}$ is in bijection with the union of cosets

$$\bigsqcup_{d \mid n^2} \Gamma_n \backslash Z_d = \bigsqcup_{d \mid n^2} J_n \backslash Z_d,$$

implying that

$$\#\Omega_{\Gamma_n} = \sum_{d \mid n^2} \#J_n \backslash Z_d.$$

Hence we want to compute the size of the coset $J_n \backslash Z_d$ for each $d \mid n^2$. For this we claim that for any for any $([m], [l])^t \in Z_d$, the orbit $J_n \cdot ([m], [l])^t$ is of size $2n/\gcd(n^2/d, d)$, implying that

$$\#J_n \backslash Z_d = \frac{\#Z_d}{2n/\gcd(n^2/d, d)} = \frac{\varphi(n^2/d)\varphi(d)\gcd(n^2/d, d)}{2n}.$$

We note that Lemma 7.5 then follows immediately from this claim. To prove this claim, it suffices to compute the size of the stabilizer

$$(J_n)_{([m], [l])} := \left\{ [a] \in J_n : [a] \cdot ([m], [l])^t = ([m], [l])^t \in Z_d \right\}.$$

Since by definition $[a] \cdot ([m], [l])^t = ([am], [\bar{a}l])^t$, $[a] \in (J_n)_{([m], [l])}$ if and only if $am \equiv m \pmod{d}$ and $\bar{a}l \equiv l \pmod{n^2}$. Since $d = \gcd(n^2, l)$ and $[m] \in (\mathbb{Z}/d\mathbb{Z})^\times$, these two conditions are equivalent to $a \equiv 1 \pmod{d}$ and $\bar{a} \equiv 1 \pmod{n^2/d}$, which are equivalent to $a \equiv 1 \pmod{\text{lcm}(n^2/d, d)}$. Hence using the description (7.5) of J_n

and the facts that $n \mid \mathrm{lcm}(n^2/d, d)$ and $\mathrm{lcm}(n^2/d, d) \gcd(n^2/d, d) = n^2$ we have

$$(J_n)_{([m],[l])} = \left\{ [1 + \mathrm{lcm}(n^2/d, d)j] \in J_n : 0 \le j \le \gcd(n^2/d, d) - 1 \right\}$$

is of size $\gcd(n^2/d, d)$. This implies that

$$\# \left(J_n \cdot ([m], [l])^t \right) = \frac{\# J_n}{\#(J_n)_{([m],[l])}} = \frac{2n}{\gcd(n^2/d, d)},$$

proving the claim, and hence also this lemma. □

We can now give the proof of Proposition 7.4 by simplifying the formula in Lemma 7.5.

Proof of Proposition 7.4 Write $n = \prod_{i=1}^{k} p_i^{\alpha_i}$ in the prime decomposition form and apply Lemma 7.5 to get

$$\# \left(\Omega_{\Gamma_n} \right) = \frac{1}{2n} \sum_{\boldsymbol{\beta} \in \mathbb{Z}^k : 0 \le \beta_i \le 2\alpha_i} \varphi \left(\prod_{i=1}^{k} p_i^{\beta_i} \right) \varphi \left(\prod_{i=1}^{k} p_i^{2\alpha_i - \beta_i} \right) \prod_{i=1}^{k} p_i^{\min\{\beta_i, 2\alpha_i - \beta_i\}},$$

where the summation is over all vectors $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k) \in \mathbb{Z}^k$ satisfying $0 \le \beta_i \le 2\alpha_i$ for all $1 \le i \le k$, and we used that $\gcd(n^2/d, d) = \prod_{i=1}^{k} p_i^{\min\{\beta_i, 2\alpha_i - \beta_i\}}$ for $d = \prod_{i=1}^{k} p_i^{\beta_i}$. Using the fact that φ is multiplicative and interchanging the summation and product signs we get

$$\# \left(\Omega_{\Gamma_n} \right) = \frac{1}{2n} \prod_{i=1}^{k} \left(\sum_{0 \le \beta_i \le 2\alpha_i} \varphi \left(p_i^{\beta_i} \right) \varphi \left(p_i^{2\alpha_i - \beta_i} \right) p_i^{\min\{\beta_i, 2\alpha_i - \beta_i\}} \right)$$

$$= \frac{1}{2n} \prod_{i=1}^{k} \left(\sum_{1 \le \beta_i \le 2\alpha_i - 1} p_i^{2\alpha_i} (1 - p_i^{-1})^2 p_i^{\min\{\beta_i, 2\alpha_i - \beta_i\}} + 2 p_i^{2\alpha_i} (1 - p_i^{-1}) \right)$$

$$= \frac{1}{2n} \prod_{i=1}^{k} p_i^{2\alpha_i} (1 - p_i^{-1}) \left((1 - p_i^{-1}) \sum_{1 \le \beta_i \le 2\alpha_i - 1} p_i^{\min\{\beta_i, 2\alpha_i - \beta_i\}} + 2 \right),$$

where for the second equality we used that for $1 \le \beta_i \le 2\alpha_i - 1$, $\varphi \left(p_i^{\beta_i} \right) \varphi \left(p_i^{2\alpha_i - \beta_i} \right) = p^{2\alpha_i} (1 - p_i^{-1})^2$, and for $\beta_i = 0$ or $\beta_i = 2\alpha_i$, $\varphi \left(p_i^{\beta_i} \right) \varphi \left(p_i^{2\alpha_i - \beta_i} \right) = p^{2\alpha_i} (1 - p_i^{-1})$

and $\min\{\beta_i, 2\alpha_i - \beta_i\} = 0$. We note that the term $\sum_{1 \le \beta_i \le 2\alpha_i - 1} p_i^{\min\{\beta_i, 2\alpha_i - \beta_i\}}$ equals

$$\sum_{1 \le \beta_i \le \alpha_i} p_i^{\beta_i} + \sum_{\alpha_i < \beta_i \le 2\alpha_i - 1} p_i^{2\alpha_i - \beta_i} = \sum_{1 \le \beta_i \le \alpha_i} p_i^{\beta_i} + \sum_{1 \le \beta_i < \alpha_i} p_i^{\beta_i}$$

$$= 2 \sum_{1 \le \beta_i \le \alpha_i} p_i^{\beta_i} - p_i^{\alpha_i} = \frac{2p_i(p_i^{\alpha_i} - 1)}{p_i - 1} - p_i^{\alpha_i}.$$

Hence we have

$$\#\left(\Omega_{\Gamma_n}\right) = \frac{1}{2n} \prod_{i=1}^{k} p_i^{2\alpha_i}(1 - p_i^{-1})\left((1 - p_i^{-1})\left(\frac{2p_i(p_i^{\alpha_i} - 1)}{p_i - 1} - p_i^{\alpha_i}\right) + 2\right)$$

$$= \frac{1}{2n} \prod_{i=1}^{k} p_i^{2\alpha_i}(1 - p_i^{-1})p_i^{\alpha_i}(1 + p_i^{-1}) = \frac{n^2}{2} \prod_{i=1}^{k}(1 - p_i^{-2}),$$

finishing the proof.   □

7.2 Proof of Theorem 1.7

For simplicity of notation, we abbreviate the cusp neighborhoods $C_Y^{\Gamma_n, \mathfrak{c}}$ and $C_{Y,Y'}^{\Gamma_n, \mathfrak{c}}$ by $C_Y^{n, \mathfrak{c}}$ and $C_{Y,Y'}^{n, \mathfrak{c}}$ respectively and the set of cusps Ω_{Γ_n} by Ω_n. We first prove the following key lemma which says that if $\Gamma_n z$ visits a cusp neighborhood on $\Gamma_n \backslash \mathbb{H}$, then all companion points $\Gamma_1 u_{j/n} z$, $0 \le j \le n - 1$ make excursions to some cusp neighborhood on $\mathcal{M} = \Gamma_1 \backslash \mathbb{H}$, the modular surface. We recall that C_Y is the projection onto $\mathcal{M}$ of the region $\{z \in \mathbb{H} : \mathfrak{Im}(z) > Y\}$.

Lemma 7.6 *Let $Y > 0$ and $n \in \mathbb{N}$. If $\Gamma_n z \in C_{\omega_{\mathfrak{c}} Y}^{n, \mathfrak{c}}$ for some $\mathfrak{c} \in \Omega_n$ then $\Gamma_1 u_{j/n} z \in C_Y$ for all $0 \le j \le n - 1$.*

Proof Fix $0 \le j \le n - 1$. By Lemma 7.1, $u_{j/n}$ normalizes Γ_n. Assuming that $\Gamma_n z \in C_{\omega_{\mathfrak{c}} Y}^{n, \mathfrak{c}}$ and applying Proposition 6.4 to $h = u_{j/n}$, we get $\Gamma_n u_{j/n} z \in C_{\omega_{h\mathfrak{c}} Y}^{n, h\mathfrak{c}}$. By definition, there exists $z' \in \mathbb{H}$ with $\mathfrak{Im}(z') > \omega_{h\mathfrak{c}} Y \ge Y$ such that $\Gamma_n \tau_{h\mathfrak{c}} z' = \Gamma_n u_{j/n} z$. Since $\tau_{h\mathfrak{c}} \in \Gamma_1$, this implies $\Gamma_1 u_{j/n} z = \Gamma_1 z' \in C_Y$.   □

We can now give the

Proof of Theorem 1.7 For any $n \in \mathbb{N}$ let $Y_n = \max\{\log n, 1\}$, and let Ψ_n be the indicator function of the union

$$\bigcup_{\mathfrak{c} \in \Omega_n} C_{\omega_{\mathfrak{c}} Y_n, 2\omega_{\mathfrak{c}} Y_n}^{n, \mathfrak{c}} \subset \Gamma_n \backslash \mathbb{H}.$$

Since for any cusp $\mathfrak{c} \in \Omega_n$, $\omega_{\mathfrak{c}} Y_n \ge Y_n \ge 1$, by Lemma 6.1, each $C_{\omega_{\mathfrak{c}} Y_n, 2\omega_{\mathfrak{c}} Y_n}^{n, \mathfrak{c}}$ is a Borel set with boundary of measure zero; and by Lemma 6.2 the above union is disjoint. Thus Ψ_n is the indicator function of a Borel set with boundary of measure

zero. Moreover, applying the volume formula (6.2), the index formula in Lemma 7.2 and the cusp number formula in Proposition 7.4 (see also Remark 7.4 for the case when $n = 2$) we have for any $n \in \mathbb{N}$,

$$
\begin{aligned}
\mu_{\Gamma_n}(\Psi_n) &= \sum_{\mathfrak{c} \in \Omega_n} \mu_{\Gamma_n} \left(\mathcal{C}^{n, \mathfrak{c}}_{\omega_\mathfrak{c} Y_n, 2\omega_\mathfrak{c} Y_n} \right) \\
&= \sum_{\mathfrak{c} \in \Omega_n} \frac{3\omega_\mathfrak{c}}{\pi [\Gamma_1 : \Gamma_n]} \times \frac{1}{2\omega_\mathfrak{c} Y_n} = \frac{3}{2\pi Y_n} \frac{\#\Omega_n}{[\Gamma_1 : \Gamma_n]} \asymp \frac{1}{n Y_n}.
\end{aligned} \tag{7.6}
$$

For any $n \in \mathbb{N}$ and $0 < y < 1$ we define

$$
I_n(y) := \{ x \in \mathbb{R}/\mathbb{Z} : \Psi_n(x + iy) = 1 \}.
$$

By definition, $x \in I_n(y)$ if and only if $\Gamma_n(x + iy) \in \mathcal{C}^{n, \mathfrak{c}}_{\omega_\mathfrak{c} Y_n, 2\omega_\mathfrak{c} Y_n} \subset \mathcal{C}^{n, \mathfrak{c}}_{\omega_\mathfrak{c} Y_n}$ for some $\mathfrak{c} \in \Omega_n$. Thus Lemma 7.6 implies that

$$
I_n(y) \subset \{ x \in \mathbb{R}/\mathbb{Z} : \mathcal{R}_n(x, y) \subset \mathcal{C}_{Y_n} \}.
$$

This, together with our choice that $Y_n = \max\{\log n, 1\}$ and the distance formula (4.1), implies that for any $n \geq 3$ and for any $x \in I_n(y)$

$$
\inf_{\Gamma_1 z \in \mathcal{R}_n(x, y)} d_{\mathcal{M}}(\Gamma_1 z_0, \Gamma_1 z) \geq \log(Y_n) + O(1) = \log \log n + O(1).
$$

It thus suffices to show that there exists a sequence $\{y_n\}_{n \in \mathbb{N}}$ satisfying that $0 < y_n < c_n$ for all $n \in \mathbb{N}$ and that the limsup set $\overline{\lim}_{n \to \infty} I_n(y_n)$ is of full Lebesgue measure in $\mathbb{R}/\mathbb{Z}$.

For this, we will construct a sequence $\{y_n\}_{n \in \mathbb{N}}$ decaying sufficiently fast and then apply the quantitative Borel–Cantelli lemma Corollary 2.6 to the sequence $\{I_n(y_n)\}_{n \in \mathbb{N}} \subset \mathbb{R}/\mathbb{Z}$. To ensure the quasi-independence condition (2.20) in Corollary 2.6, we need, for every pair $1 \leq m < n \in \mathbb{N}$, the two quantities $|I_m(y_m) \cap I_n(y_n)|$ and $|I_m(y_m)| \, |I_n(y_n)|$ to be sufficiently close to each other. The key observations for this are the following two relations that

$$
|I_n(y_n)| = \int_0^1 \Psi_n(x + iy_n) dx \tag{7.7}
$$

and

$$
|I_m(y_m) \cap I_n(y_n)| = \int_0^1 \Psi_m(x + iy_m) \Psi_n(x + iy_n) dx = \int_{I_m(y_m)} \Psi_n(x + iy_n) dx. \tag{7.8}
$$

Assuming the limit equation (2.16) holds for the pairs $((0, 1), \Psi_n)$ and $(I_m(y_m), \Psi_n)$ (we will verify this later), then by relation (7.8) the quantity $|I_m(y_m) \cap I_n(y_n)|$ is close

to the quantity $|I_m(y_m)|\mu_{\Gamma_n}(\Psi_n)$ which in turn is close to $|I_m(y_m)||I_n(y_n)|$ by relation (7.7), provided that $y_n > 0$ is sufficiently small.

We now implement the above ideas rigorously. We first claim that there exists a sequence $\{y_n\}_{n \in \mathbb{N}}$ satisfying, for all $n \in \mathbb{N}$, $0 < y_n < c_n$ and

$$\left| \frac{1}{|I|} \int_I \Psi_n(x + iy_n)dx - \mu_{\Gamma_n}(\Psi_n) \right| \le \frac{\mu_{\Gamma_n}(\Psi_n)}{2n^2}, \tag{7.9}$$

for any subset $I \subset \mathbb{R}/\mathbb{Z}$ taken from the finite set

$$\{(0, 1)\} \bigcup \{I_m(y_m) : 1 \le m < n\}. \tag{7.10}$$

For this, first note that by Remark 2.17 for any $I \subset \mathbb{R}/\mathbb{Z} \cong [0, 1)$ a disjoint union of finitely many open intervals, we have

$$\lim_{y \to 0^+} \frac{1}{|I|} \int_I \Psi_n(x + iy)dx = \mu_{\Gamma_n}(\Psi_n). \tag{7.11}$$

We now construct such a sequence successively. For the base case $n = 1$ since (7.11) holds for the pair $((0, 1), \Psi_1)$ on $\mathcal{M} = \Gamma_1 \backslash \mathbb{H}$, there exists $0 < y_1 < c_1$ sufficiently small such that

$$\left| \int_0^1 \Psi_1(x + iy_1)dx - \mu_{\Gamma_1}(\Psi_1) \right| < \frac{1}{2}\mu_{\Gamma_1}(\Psi_1).$$

For a general integer $n \ge 2$, suppose that we already have chosen $0 < y_m < c_m$ satisfying (7.9) for all the positive integers $m < n$. By Remark 6.3 the set $I_m(y_m) \subset \mathbb{R}/\mathbb{Z}$ is a disjoint union of finitely many open intervals for any $m < n$. Thus (7.11) is satisfied for all the pairs

$$((0, 1), \Psi_n), (I_m(y_m), \Psi_n), 1 \le m < n$$

on $\Gamma_n \backslash \mathbb{H}$. Since there are only finitely many such pairs, we can take $0 < y_n < c_n$ sufficiently small such that (7.9) is satisfied for all $I \in \{(0, 1)\} \bigcup \{I_m(y_m) : 1 \le m < n\}$, which is the set in (7.10). This finishes the proof of the claim.

Now let $\{y_n\}_{n \in \mathbb{N}}$ be as in the claim. For any $n \in \mathbb{N}$ apply (7.9) to the pair $((0, 1), \Psi_n)$ we get

$$\left| |I_n(y_n)| - \mu_{\Gamma_n}(\Psi_n) \right| \le \frac{\mu_{\Gamma_n}(\Psi_n)}{2n^2}. \tag{7.12}$$

By the triangle inequality, this implies

$$\mu_{\Gamma_n}(\Psi_n) \le 2|I_n(y_n)|. \tag{7.13}$$

More generally, for each $1 \leq m < n$ apply (7.9) to the pair $(I_m(y_m), \Psi_n)$ we get

$$\left||I_m(y_m) \cap I_n(y_n)| - |I_m(y_m)| \, \mu_{\Gamma_n}(\Psi_n)\right| \leq \frac{|I_m(y_m)| \, \mu_{\Gamma_n}(\Psi_n)}{2n^2}. \qquad (7.14)$$

Using the inequalities (7.12), (7.13), (7.14) together with the triangle inequality we get

$$\left||I_m(y_m) \cap I_n(y_n)| - |I_m(y_m)| \, |I_n(y_n)|\right| \leq \frac{|I_m(y_m)| \, \mu_{\Gamma_n}(\Psi_n)}{n^2}$$
$$\leq \frac{2 \, |I_m(y_m)| \, |I_n(y_n)|}{n^2}. \qquad (7.15)$$

Hence the sequence $\{I_n(y_n)\}_{n \in \mathbb{N}} \subset \mathbb{R}/\mathbb{Z}$ satisfies the quasi-independence condition (2.20) (with the subset $\mathbb{S} = \mathbb{N}$ and the exponent $\eta = 2$). Moreover, using the inequality (7.12), the volume computation (7.6) and the estimate that $Y_n \asymp \log n$ we have that

$$\sum_{n \in \mathbb{N}} |I_n(y_n)| \geq \sum_{n \in \mathbb{N}} \frac{1}{2} \mu_{\Gamma_n}(\Psi_n) \asymp \sum_{n \in \mathbb{N}} \frac{1}{n \log n} = \infty.$$

Thus by Corollary 2.6, $\overline{\lim}_{n \to \infty} I_n(y_n) \subset \mathbb{R}/\mathbb{Z}$ is of full Lebesgue measure, finishing the proof. $\qquad \square$

Remark 7.16 It is not clear to us whether the rate $\log \log n$ is the fastest excursion rate for generic translates. We note that in principle it can be proved (or disproved) if one can compute the volume of the set

$$\mathcal{E}_Y^n := \left\{ \Gamma_n z \in \Gamma_n \backslash \mathbb{H} : \Gamma_1 u_{j/n} z \in \mathcal{C}_Y \text{ for all } 0 \leq j \leq n-1 \right\}.$$

For instance, if one can show $\mu_{\Gamma_n}(\mathcal{E}_Y^n) \asymp 1/(nY)$ for all $n \in \mathbb{N}$ and for all $Y \geq 1$, then Theorem 1.7 together with a standard application of the Borel–Cantelli lemma would imply that the inequality in (1.13) is indeed an equality for almost every $x \in \mathbb{R}/\mathbb{Z}$. We also note that our analysis (Lemmas 6.2, 7.6) shows that for any $n \in \mathbb{N}$ and for any $Y \geq 1$

$$\bigsqcup_{\mathfrak{c} \in \Omega_n} \mathcal{C}_{\omega_{\mathfrak{c}} Y}^{n, \mathfrak{c}} \subset \mathcal{E}_Y^n \subset \bigsqcup_{\mathfrak{c} \in \Omega_n} \mathcal{C}_Y^{n, \mathfrak{c}},$$

implying that $1/(nY) \ll \mu_{\Gamma_n}\left(\mathcal{E}_Y^n\right) \ll 1/Y$. On the other hand using some elementary arguments (which relies on the width computation Lemma 7.3) one can show that any $\langle u_{1/n} \rangle$-orbit contains at least one cusp of width one. This fact together with the fact that $1 \leq \omega_{\mathfrak{c}} \leq n^2$ implies that $\mathcal{E}_Y^n = \bigsqcup_{\mathfrak{c} \in \Omega_n} \mathcal{C}_Y^{n, \mathfrak{c}}$ when $Y \geq n^2$. However, both estimates are not sufficient for the purpose of obtaining an upper bound.

Remark 7.17 Here we give a very brief sketch of the argument communicated to us by Strömbergsson: For each $n \in \mathbb{N}$ and $y > 0$, it is not difficult to see that $\Gamma_n(x + iy) \in$

$C_{\omega_{\mathfrak{c}} Y_n}^{n,\mathfrak{c}}$ for some $\mathfrak{c} = \frac{p}{q} \in \Omega_n$ with $\gcd(p,q) = 1$ if and only if

$$\left| x - \frac{p}{q} \right|^2 < \frac{y}{\omega_{\mathfrak{c}} Y_n q^2} - y^2 = \frac{y \gcd(n,q)^2}{n^2 Y_n q^2} - y^2. \tag{7.18}$$

Here $Y_n = \max\{\log n, 1\}$ is as in the above proof. Define

$$\tilde{I}_n(y) := \left\{ x \in \mathbb{R}/\mathbb{Z} : \exists \text{ primitive } \frac{p}{q} \text{ s.t. } n \mid q, q < \frac{1}{2\sqrt{yY_n}}, \left| x - \frac{p}{q} \right| < \frac{\sqrt{y}}{2\sqrt{Y_n}q} \right\}.$$

One can easily check that elements in $\tilde{I}_n(y)$ satisfy the inequality (7.18). Hence by Lemma 7.6 we have

$$\tilde{I}_n(y) \subset \{ x \in \mathbb{R}/\mathbb{Z} : \mathcal{R}_n(x,y) \subset \mathcal{C}_{Y_n} \}. \tag{7.19}$$

Moreover, using some standard techniques from analytic number theory one can show that for any subinterval $I \subset \mathbb{R}/\mathbb{Z}$ (or more generally, any finite disjoint union of subintervals),

$$\lim_{y \to 0^+} |I|^{-1} \left| \tilde{I}_n(y) \cap I \right| = \frac{c_n}{Y_n}$$

with $c_n = \frac{3}{\pi^2} \frac{\varphi(n)}{n^2} \prod_{p \nmid n} (1 - p^{-2})^{-1} \gg \frac{\varphi(n)}{n^2}$. This limit equation is the analog of (7.11). Another input is the divergence of the series $\sum_{n \in \mathbb{N}} \frac{c_n}{Y_n} \gg \sum_{n \in \mathbb{N}} \frac{\varphi(n)}{n^2 \log n}$, which follows from the estimate $\varphi(n) \gg n / \log \log n$. With these two inputs one can then mimic the arguments in the above proof to construct a sequence $\{y_n\}_{n \in \mathbb{N}}$ decaying sufficiently fast and then apply Corollary 2.6 to get a full measure limsup set $\overline{\lim}_{n \to \infty} \tilde{I}_n(y_n) \subset \mathbb{R}/\mathbb{Z}$. Finally, we note that the relation (7.19) can be checked directly using the definition of the set $\tilde{I}_n(y)$. Hence this argument can be carried over without going into the congruence covers $\Gamma_n \backslash \mathbb{H}$.

7.3 Proof of Theorem 1.8

We prove Theorem 1.8 in this subsection. The strategy is similar to that of Theorem 1.7 with the sequence of cuspidal sets approaching the cusps replaced by a sequence of compact cylinders approaching certain closed horocycles. Let $n \in \mathbb{N}$ be an integer and let $\Gamma_n z \in \Gamma_n \backslash \mathbb{H}$ be a point close to a cusp $\mathfrak{c} \in \Omega_n$. For any $0 \le j \le n-1$, the analysis in Sect. 6 gives exact information about the height of the companion point $\Gamma_n u_{j/n} z$ with respect to the cusp $u_{j/n} \mathfrak{c}$. While this is sufficient for Theorem 1.7 (cusp excursions), to realize the limiting measure $\nu_{m,Y}$ in Theorem 1.8 one needs more refined information about the spacing of these companion points along the closed horocycles they lie on. For this, we further analyze the left regular $u_{1/n}$-action on points near certain type cusps which we now define.

We say $\mathfrak{c} \in \Omega_n$ is of *simple type* if $\mathfrak{c}$ can be represented by a primitive rational number m/q satisfying that $\gcd(n^2, q) \mid n$, and we denote by $\Omega_n^{\text{sim}} \subset \Omega_n$ the set of

simple type cusps. (This notion of simple type cusps is closely related to the condition $n \in \mathbb{N}_q$ in Theorem 1.2. In fact, let p/q be a primitive rational number then the condition $n \in \mathbb{N}_q$ is equivalent to that the cusp $\mathfrak{c} \in \Omega_n$ represented by p/q is of simple type.) If m'/q' is another representative for $\mathfrak{c}$, that is, m'/q' is primitive and $m'/q' = \gamma(m/q)$ for some $\gamma \in \Gamma_n$, then using the definition of Γ_n, it is easy to check that $\gcd(n^2, q) = \gcd(n^2, q')$. Hence the simple type cusps are well-defined.

As mentioned in Sect. 3.3 the condition $\gcd(n^2, q) \mid q$ implies the further decomposition $q = kl$ with $l = \gcd(n, q) \mid n$ and $k = q/l$ satisfying $\gcd(k, n) = 1$. We can thus reparameterize a simple type $\mathfrak{c}$ by $m/(kl)$ with $\gcd(m, kl) = \gcd(k, n) = 1$ and $l \mid n$. The main new ingredient of our proof to Theorem 1.8 is the following decomposition of the sample points which generalizes (3.19).

Proposition 7.7 *Fix $n \in \mathbb{N}$, $z = x + iy \in \mathbb{H}$ and $\mathfrak{c} \in \Omega_n^{\mathrm{sim}}$. Then*

$$\mathcal{R}_n(x, y) = \bigcup_{d \mid n} \mathcal{R}_{n/d}^{\mathrm{pr}}(x'_{\mathfrak{c},d}, d^2 y'/\omega_{\mathfrak{c}}),$$

where $z' = x' + iy' \in \mathbb{H}$ is such that $\Gamma_n z = \Gamma_n \tau_{\mathfrak{c}} z'$, and $x'_{d,\mathfrak{c}} \in \mathbb{R}/\mathbb{Z}$ depends only on x', $\mathfrak{c}$ and d.

We first prove a simple lemma computing the width of elements in the orbits $\langle u_{1/n} \rangle \mathfrak{c}$ when $\mathfrak{c} \in \Omega_n^{\mathrm{sim}}$ is of simple type.

Lemma 7.8 *Fix $n \in \mathbb{N}$ and $\mathfrak{c} \in \Omega_n^{\mathrm{sim}}$ a simple type cusp. Then for any $0 \leq j \leq n - 1$ we have*

$$\omega_{u_{j/n}\mathfrak{c}} = \gcd\left(m\tfrac{n}{l} + jk, n\right)^2,$$

where $m/(kl)$ is a representative for $\mathfrak{c}$ with $\gcd(m, kl) = \gcd(k, n) = 1$ and $l \mid n$.

Proof For any $0 \leq j \leq n - 1$,

$$u_{j/n}\mathfrak{c} = \frac{m}{kl} + \frac{j}{n} = \frac{m\tfrac{n}{l} + jk}{kn} =: \frac{p_j}{q_j}$$

with $\gcd(p_j, q_j) = 1$. Let $d_j := \gcd(m\tfrac{n}{l} + jk, kn)$ such that $q_j = kn/d_j$. Since $\gcd(mn, k) = 1$, we have $\gcd(m\tfrac{n}{l} + jk, k) = \gcd(m\tfrac{n}{l}, k) = 1$. Hence $d_j = \gcd(m\tfrac{n}{l} + jk, n) \mid n$. Now by Lemma 7.3 and the assumption that $\gcd(k, n) = 1$ we have

$$\omega_{u_{j/n}\mathfrak{c}} = \frac{n^2}{\gcd(n, kn/d_j)^2} = d_j^2 = \gcd(m\tfrac{n}{l} + jk, n)^2.$$

$\square$

We can now combine ideas from Sects. 3.3 and 6 to give the

Proof of Proposition 7.7 Assume $\mathfrak{c} = m/(kl)$ with $\gcd(m, kl) = \gcd(k, n) = 1$ and $l \mid n$. Up to changing the representatives for $\mathfrak{c}$, we may assume $mkl \neq 0$. Let $\tau_{\mathfrak{c}} =$

$\left(\begin{smallmatrix} m & a \\ kl & b \end{smallmatrix} \right) \in \Gamma_1$, and for each $1 \leq j \leq n - 1$ let $\tau_{u_{j/n}\mathfrak{c}} = \left(\begin{smallmatrix} p_j & v_j \\ q_j & w_j \end{smallmatrix} \right) \in \Gamma_1$, where p_j, q_j are as in the proof of Lemma 7.8, a, b, v_j, w_j are some integers such that $\tau_{\mathfrak{c}}, \tau_{u_{j/n}\mathfrak{c}} \in \Gamma_1$, that is,

$$ mb - kla = 1 \quad \text{and} \quad \left(m\tfrac{n}{l} + jk \right) w_j - knv_j = d_j \tag{7.20} $$

with $d_j = \gcd(m\tfrac{n}{l} + jk, n)$ as in the proof of Lemma 7.8. By direct computation and using Lemmas 6.3 and 7.8 (and the relation $\omega_{\mathfrak{c}} = d_0^2 = n^2/l^2$) we have

$$ \tau_{u_{j/n}\mathfrak{c}}^{-1} u_{j/n} \tau_{\mathfrak{c}} = \begin{pmatrix} d_j l/n & w_j a + b\left(\tfrac{jw_j}{n} - v_j \right) \\ 0 & n/(d_j l) \end{pmatrix}. $$

Using the relations in (7.20) the top right entry becomes

$$ \begin{aligned} w_j a + b\left(\frac{jw_j}{n} - v_j \right) &= w_j a + \frac{1 + kla}{m} \left(\frac{jw_j}{n} - v_j \right) \\ &= \frac{a(w_j mn + jw_j kl - klv_j n) + jw_j - v_j n}{mn} \\ &= \frac{ad_j l}{mn} + \frac{1}{mn}\left(\frac{d_j l - w_j mn}{kl} \right) = \frac{bd_j}{nk} - \frac{w_j}{kl}. \end{aligned} $$

(Here we used the assumption that $mkl \neq 0$.) Hence we have for any $0 \leq j \leq n - 1$

$$ \Gamma_n u_{j/n} z = \Gamma_n u_{j/n} \tau_{\mathfrak{c}}(x' + iy') = \Gamma_n \tau_{u_{j/n}\mathfrak{c}} \tau_{u_{j/n}\mathfrak{c}}^{-1} u_{j/n} \tau_{\mathfrak{c}}(x' + iy') \tag{7.21} $$
$$ = \Gamma_n \tau_{u_{j/n}\mathfrak{c}} \left(\frac{d_j^2 l^2}{n^2} x' + \frac{d_j^2 lb}{n^2 k} - \frac{d_j w_j}{kn} + i \frac{d_j^2 l^2}{n^2} y' \right). $$

Here for the first equality we used the assumption that $\Gamma_n z = \Gamma_n \tau_{\mathfrak{c}} z'$ and the fact that $u_{j/n}$ normalizes Γ_n. Now as in the proof of Proposition 3.7 for any $d \mid n$, we define

$$ D_d := \{ 0 \leq j \leq n - 1 : d_j = d \} $$

so that

$$ \mathcal{R}_n(x, y) = \bigcup_{d \mid n} \left\{ \Gamma_1 u_{j/n} z \in \mathcal{M} : j \in D_d \right\}, \tag{7.22} $$

and

$$ \left\{ [(m\tfrac{n}{l} + jk)/d] \in (\mathbb{Z}/(n/d)\mathbb{Z})^\times : j \in D_d \right\} = (\mathbb{Z}/(n/d)\mathbb{Z})^\times. \tag{7.23} $$

Use the second relation in (7.20) to get for $j \in D_d$,

$$ w_j \left((m\tfrac{n}{l} + jk)/d \right) \equiv 1 \pmod{k\tfrac{n}{d}}. $$

Solving the above congruence equation as in the proof of Lemma 3.6 we get

$$w_j \equiv d l \overline{mn} \tfrac{n}{d} e + \left(\left(m \tfrac{n}{l} + jk \right) / d \right)^* kf \ (\mathrm{mod} \ k \tfrac{n}{d}),$$

where for any integer t, $\bar{t}$ denotes the multiplicative inverse modulo k, t^* denotes the multiplicative inverse modulo n/d, and $e = e_d$, $f = f_d \in \mathbb{Z}$ are two fixed integers such that $e \tfrac{n}{d} + fk = 1$. Plugging this relation into (7.21) and using the relation $\omega_{\mathfrak{c}} = n^2/l^2$ we get for any $d \mid n$ and for any $j \in D_d$,

$$\Gamma_n u_{j/n} z = \Gamma_n \tau_{u_{j/n}\mathfrak{c}} \left(x'_{\mathfrak{c},d} - \frac{\left(\left(m \tfrac{n}{l} + jk \right)/d \right)^* f}{n/d} + i \frac{d^2 y'}{\omega_{\mathfrak{c}}} \right),$$

where $x'_{\mathfrak{c},d} := \frac{d^2 l^2}{n^2} x' + \frac{d^2 lb}{n^2 k} - \frac{d l \overline{mn} e}{k} \ (\mathrm{mod} \ \mathbb{Z}) \in \mathbb{R}/\mathbb{Z}$. Since $\tau_{u_{j/n}\mathfrak{c}} \in \Gamma_1$ we have

$$\left\{ \Gamma_1 u_{j/n} z \in \mathcal{M} : j \in D_d \right\} = \left\{ \Gamma_1 \left(x'_{\mathfrak{c},d} - \frac{\left(\left(m \tfrac{n}{l} + jk \right)/d \right)^* f}{n/d} + i \frac{d^2 y'}{\omega_{\mathfrak{c}}} \right) \in \mathcal{M} : j \in D_d \right\}.$$

Thus in view of (7.22) and the above relation it suffices to show

$$\left\{ -\left[\left(\left(m \tfrac{n}{l} + jk \right)/d \right)^* f \right] \in (\mathbb{Z}/(n/d)\mathbb{Z})^{\times} : j \in D_d \right\} = (\mathbb{Z}/(n/d)\mathbb{Z})^{\times}.$$

But this follows from (7.23) and the fact that $\gcd(f, \tfrac{n}{d}) = 1$ (since $\gcd(f, \tfrac{n}{d}) = \gcd(fk, \tfrac{n}{d}) = \gcd(1 - e \tfrac{n}{d}, \tfrac{n}{d}) = 1$), and we have thus finished the proof. $\square$

We will also need the following lemma estimating the number of cusps in Ω_n^{sim} satisfying certain restrictions on the width.

Lemma 7.9 *Let $m \in \mathbb{N}$ be a fixed integer and let $n = m\ell \geq 3$ for some prime number ℓ not dividing m. Then we have*

$$\# \left\{ \mathfrak{c} \in \Omega_n^{\mathrm{sim}} : \omega_{\mathfrak{c}} \geq m^2 \right\} \geq \frac{\varphi(m)(\ell - 1)^2}{2}.$$

Proof Recall from the proof of Lemma 7.5 that Ω_n is in bijection with the disjoint union $\bigsqcup_{d \mid n^2} J_n \backslash Z_d$. On the other hand, by definition of the simple type cusps, Ω_n^{sim} corresponds to the subset $\sqcup_{d \mid n} J_n \backslash Z_d$. Moreover, let $\mathfrak{c} = m/l \in \Omega_n^{\mathrm{sim}}$ with $\gcd(m, l) = 1$ be a simple type cusp corresponding to an element in $J_n \backslash Z_d$ for some $d \mid n$, that is, $d = \gcd(n^2, l)$. Since $d \mid n$, this implies that $d = \gcd(n^2, l) = \gcd(n, l)$. Hence by Lemma 7.3, $\omega_{\mathfrak{c}} = n^2/d^2$. Therefore for each $d \mid n$

$$\#\{\mathfrak{c} \in \Omega_n^{\mathrm{sim}} : \omega_{\mathfrak{c}} = n^2/d^2\} = |J_n \backslash Z_d| = \frac{\varphi(n^2/d)\varphi(d)\gcd(n^2/d, d)}{2n} = \frac{\varphi(n)\varphi(d)}{2},$$

□ where for the last equality we used the identities $\gcd(n^2/d, d) = d$ (since $d \mid n$) and

$$\varphi\left(\tfrac{n^2}{d}\right) = \frac{n^2}{d} \prod_{\substack{p \mid (n^2/d) \\ prime}} (1 - p^{-1}) = \frac{n}{d} \times n \prod_{\substack{p \mid n \\ prime}} (1 - p^{-1}) = \frac{n\varphi(n)}{d},$$

where for the second equality we used the fact that n^2/d and n share the same set of prime divisors. Hence for $n = m\ell$ we have

$$\#\left\{\mathfrak{c} \in \Omega_n^{\text{sim}} : \omega_{\mathfrak{c}} \geq m^2\right\} = \frac{\varphi(n)}{2} \sum_{\substack{d \mid n \\ n^2/d^2 \geq m^2}} \varphi(d) \geq \frac{\varphi(n)\varphi(\ell)}{2} = \frac{\varphi(m)(\ell-1)^2}{2}.$$

□

Lemma 7.10 *Let* $m \in \mathbb{N}$ *and* $Y > 0$ *satisfy that* $m^2 Y > 1$. *Let*

$$\mathbb{P}_m = \{n = m\ell \in \mathbb{N} : \ell \text{ is a prime number and } \ell \nmid m\}$$

be as in (1.8). Then there exist sequences of positive numbers $\{Y_n\}_{n \in \mathbb{P}_m}$ *and* $\{Y'_n\}_{n \in \mathbb{P}_m}$ *satisfying that*

(1) $Y'_n > Y > Y_n > m^{-2}$ *for any* $n \in \mathbb{P}_m$ *and* $\lim_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} Y_n = \lim_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} Y'_n = Y$;

(2) $\sum_{n \in \mathbb{P}_m} \frac{1}{n}\left(\frac{1}{Y_n} - \frac{1}{Y'_n}\right) = \infty.$

Proof For each $n = m\ell \in \mathbb{P}_m$, take $Y'_n := (1 - (2t_n)^{-1})^{-1}Y$ and $Y_n := (1 + (2t_n)^{-1})^{-1}Y$ with

$$t_n = \max\{(m^2 Y - 1)^{-1}, \log\log\ell\}.$$

We note that the first condition is guaranteed by the facts that $t_n \geq (m^2 Y - 1)^{-1}$ and that $\lim_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} t_n = \infty$. For the second condition, we note that by the definitions of Y_n and Y'_n, $\frac{1}{Y_n} - \frac{1}{Y'_n} = \frac{1}{Yt_n}$. Moreover, using the fact that there are only finitely many prime numbers dividing m we get

$$\sum_{n \in \mathbb{P}_m} \frac{1}{n}\left(\frac{1}{Y_n} - \frac{1}{Y'_n}\right) \asymp_{m,Y} \sum_{\substack{\ell \in \mathbb{P}_1 \\ \ell \nmid m}} \frac{1}{\ell \log\log\ell} = \sum_{\ell \in \mathbb{P}_1} \frac{1}{\ell \log\log\ell} + O_m(1) = \infty,$$

where the divergence of the rightmost series follows from the estimate $\ell_j \asymp j \log j$ which is an easy consequence of the prime number theorem. Here $\ell_j \in \mathbb{P}_1$ denotes the j-th prime number. □

We now give the

Proof of Theorem 1.8 Fix throughout the proof $m \in \mathbb{N}$ and $Y > 0$ with $m^2 Y > 1$ and let $\mathbb{P}_m$ be as above. Let $\{Y_n\}_{n \in \mathbb{P}_m}$ and $\{Y_n'\}_{n \in \mathbb{P}_m}$ be two sequences satisfy the conditions in Lemma 7.10. For any $n \in \mathbb{P}_m$, let $\Psi_n \in L^2(\Gamma_n \backslash \mathbb{H})$ such that Ψ_n is the indicator function of the union

$$\bigcup_{\substack{\mathfrak{c} \in \Omega_n^{\mathrm{sim}} \\ \omega_{\mathfrak{c}} \geq m^2}} \mathcal{C}^{n,\mathfrak{c}}_{\omega_{\mathfrak{c}} Y_n, \omega_{\mathfrak{c}} Y_n'} \subset \Gamma_n \backslash \mathbb{H}.$$

Since $Y_n > m^{-2}$ for any $n \in \mathbb{P}_m$, $\omega_{\mathfrak{c}} Y_n > 1$ for any $\mathfrak{c} \in \Omega_n^{\mathrm{sim}}$ with $\omega_{\mathfrak{c}} \geq m^2$. Hence similar as in the proof of Theorem 1.7, by Lemmas 6.1 and 6.2 the above union is disjoint and Ψ_n is the indicator function of a Borel set with boundary of measure zero. By the disjointness and the volume formula (6.2) we have for any $n \in \mathbb{P}_m$

$$\mu_{\Gamma_n}(\Psi_n) = \frac{3}{\pi} \frac{\# \left\{ \mathfrak{c} \in \Omega_n^{\mathrm{sim}} : \omega_{\mathfrak{c}} \geq m^2 \right\}}{[\Gamma_1 : \Gamma_n]} \left(\frac{1}{Y_n} - \frac{1}{Y_n'} \right).$$

Note that for $n = m\ell \in \mathbb{P}_m$, by Lemma 7.2, $[\Gamma_1 : \Gamma_n] \asymp_m \ell^3$. Hence by Lemma 7.9 and the above relation we get for any $n = m\ell \in \mathbb{P}_m$

$$\mu_{\Gamma_n}(\Psi_n) \gg_{m,Y} \frac{1}{\ell} \left(\frac{1}{Y_n} - \frac{1}{Y_n'} \right) \asymp_m \frac{1}{n} \left(\frac{1}{Y_n} - \frac{1}{Y_n'} \right). \tag{7.24}$$

Similar as in the proof of Theorem 1.7 for any $n \in \mathbb{P}_m$ and $0 < y < 1$ we define

$$I_n(y) := \{ x \in \mathbb{R}/\mathbb{Z} : \Psi_n(x + iy) = 1 \}.$$

We first show that there exists a sequence $\{y_n\}_{n \in \mathbb{P}_m}$ satisfying that $0 < y_n < c_n$ for all $n \in \mathbb{P}_m$ and that the limsup set $\overline{\lim}_{n \in \mathbb{P}_m} I_n(y_n) \subset \mathbb{R}/\mathbb{Z}$ is of full measure. As in the proof of Theorem 1.7, we can use Remark 2.17, together with Remark 6.3 and Lemma 6.1, to construct a sequence $\{y_n\}_{n \in \mathbb{P}_m}$ successively satisfying for any $n \in \mathbb{P}_m$, $0 < y_n < c_n$ and that

$$\left| \frac{1}{|I|} \int_I \Psi_n(x + iy_n) dx - \mu_{\Gamma_n}(\Psi_n) \right| \leq \frac{\mu_{\Gamma_n}(\Psi_n)}{2n^2} \tag{7.25}$$

for all subsets $I \subset \mathbb{R}/\mathbb{Z}$ taken from the finite set $\{(0, 1)\} \bigcup \{I_l(y_l) : l \in \mathbb{P}_m, \ l < n\}$. Again as before one can show that condition (7.25) implies that the sequence $\{I_n(y_n)\}_{n \in \mathbb{P}_m} \subset \mathbb{R}/\mathbb{Z}$ satisfies the quasi-independence condition (2.20) (with the subset $\mathbb{S} = \mathbb{P}_m$ and exponent $\eta = 2$). Moreover, using the estimate (7.24) and our assumptions on $\{Y_n\}_{n \in \mathbb{P}_m}$ and $\{Y_n'\}_{n \in \mathbb{P}_m}$) we have

$$\sum_{n \in \mathbb{P}_m} |I_n(y_n)| \asymp \sum_{n \in \mathbb{P}_m} \mu_{\Gamma_n}(\Psi_n) \gg_{m,Y} \sum_{n \in \mathbb{P}_m} \frac{1}{n} \left(\frac{1}{Y_n} - \frac{1}{Y_n'} \right) = \infty.$$

Hence by Corollary 2.6, $\overline{\lim}_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} I_n(y_n) \subset \mathbb{R}/\mathbb{Z}$ is of full Lebesgue measure.

Now take $x \in \overline{\lim}_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} I_n(y_n)$, then there exists an unbounded subsequence $\mathcal{N}_x \subset \mathbb{P}_m$ such that $x \in I_n(y_n)$ for all $n \in \mathcal{N}_x$. It thus suffices to show that for any $\Psi \in C_c^\infty(\mathcal{M})$,

$$\lim_{\substack{n \in \mathcal{N}_x \\ n \to \infty}} \delta_{n,x,y_n}(\Psi) = \nu_{m,Y}(\Psi)$$

with $\nu_{m,Y}$ defined as in (1.9). For any $n \in \mathcal{N}_x \subset \mathbb{P}_m$, since $x \in I_n(y_n)$ by definition we have $\Gamma_n(x + iy_n) \in C_{\omega_{\mathfrak{c}} Y_n, \omega_{\mathfrak{c}} Y_n'}^{n,\mathfrak{c}}$ for some $\mathfrak{c} \in \Omega_n^{\text{sim}}$ of simple type, that is, there exist some $\mathfrak{c} \in \Omega_n^{\text{sim}}$ and $z_n' = x_n' + i\omega_{\mathfrak{c}} y_n' \in \mathbb{H}$ satisfying that $\Gamma_n(x + iy_n) = \Gamma_n \tau_{\mathfrak{c}} z_n'$ with $Y_n < y_n' < Y_n'$. Then by Proposition 7.7, we have

$$\mathcal{R}_n(x, y_n) = \bigcup_{d|n} \mathcal{R}_{n/d}^{\text{pr}}(x_{n,\mathfrak{c},d}', d^2 y_n')$$

for some $x_{n,\mathfrak{c},d}' \in \mathbb{R}/\mathbb{Z}$. This implies that for any $n \in \mathcal{N}_x$

$$\delta_{n,x,y_n}(\Psi) = \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) \delta_{n/d, x_{n,\mathfrak{c},d}', d^2 y_n'}^{\text{pr}}(\Psi).$$

Since $y_n', Y \in (Y_n, Y_n')$, $\max\{y_n'/Y, Y/y_n'\} \leq Y_n'/Y_n$. Thus by the intermediate value theorem we can estimate for $n \in \mathcal{N}_x$

$$\delta_{n,x,y_n}(\Psi) = \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) \left(\delta_{n/d, x_{n,\mathfrak{c},d}', d^2 Y}^{\text{pr}}(\Psi) + O\left(\mathcal{S}_{\infty,1}^{\Gamma_1}(\Psi) \log\left(Y_n'/Y_n\right)\right)\right)$$

$$= \frac{1}{n} \sum_{d|n} \varphi\left(\frac{n}{d}\right) \delta_{n/d, x_{n,\mathfrak{c},d}', d^2 Y}^{\text{pr}}(\Psi) + O_\Psi\left(\log\left(Y_n'/Y_n\right)\right),$$

where for the second estimate we used the identity $\sum_{d|n} \varphi(n/d) = n$. Thus for $n = m\ell \in \mathcal{N}_x$ sufficiently large such that Ψ vanishes on the cusp neighborhood $\mathcal{C}_{\ell^2 Y}$ we have

$$\delta_{n,x,y_n}(\Psi) = \frac{1}{m\ell} \sum_{d|m} \varphi\left(\frac{m\ell}{d}\right) \delta_{m\ell/d, x_{n,\mathfrak{c},d}', d^2 Y}^{\text{pr}}(\Psi) + O_\Psi\left(\log\left(Y_n'/Y_n\right)\right)$$

$$= \frac{\ell - 1}{m\ell} \sum_{d|m} \varphi\left(\frac{m}{d}\right) \left(\mu_{d^2 Y}(\Psi) + O_{\Psi,m,Y,\epsilon}\left(\ell^{-1+\epsilon}\right)\right) + O_\Psi\left(\log\left(Y_n'/Y_n\right)\right)$$

$$= \frac{\ell - 1}{\ell} \nu_{m,Y}(\Psi) + O_{\Psi,m,Y,\epsilon}\left(\ell^{-1+\epsilon} + \log\left(Y_n'/Y_n\right)\right),$$

where for the second equality we used the facts that ℓ is a prime number and $\gcd(m, \ell) = 1$ and applied the effective estimate (3.13) to each of the term

$\delta^{\mathrm{pr}}_{m\ell/d, x'_{n,c,d}, d^2 Y}(\Psi)$. We now conclude by taking $n \to \infty$ along the subsequence $\mathcal{N}_x$ and noting that $\lim_{\substack{n \in \mathcal{N}_x \\ n \to \infty}} \log \left(Y'_n / Y_n \right) = 0$ (since $\lim_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} Y'_n / Y_n = 1$ which follows from the assumption $\lim_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} Y_n = \lim_{\substack{n \in \mathbb{P}_m \\ n \to \infty}} Y'_n = Y$). $\qquad\square$

Remark 7.26 It is clear that we can take a sequence $\{y_n\}_{n \in \mathbb{N}}$ decaying sufficiently fast such that the conditions (7.9) and (7.25) (for any finitely many pairs (m, Y) with $m^2 Y > 1$) are all satisfied and hence (noting that the intersection of finitely many full measure sets is still of full measure) for such a sequence the conclusions of Theorems 1.7 and 1.8 (for any finitely many pairs (m, Y) with $m^2 Y > 1$) hold simultaneously.

Declarations

Conflict of interest On behalf of all authors, Shucheng Yu states that there is no conflict of interest.

Data availability Not applicable.

Code availability Not applicable.

References

1. Bersudsky, M.: On the image in the torus of sparse points on dilating analytic curves. arXiv:2003.04112 (arXiv preprint) (2020)
2. Borel, A.: Some metric properties of arithmetic quotients of symmetric spaces and an extension theorem. J. Differ. Geometry **6**, 543–560 (1972)
3. Clozel, L., Oh, H., Ullmo, E.: Hecke operators and equidistribution of Hecke points. Invent. Math. **144**(2), 327–351 (2001)

4. Clozel, L., Ullmo, E.: Equidistribution des pointes de hecke. In: Contribution to Automorphic Forms, Geometry and Number Theory, pp. 193–254. Johns Hopkins University Press, Baltimore (2004)

5. Demirci Akarsu, E.: Short incomplete Gauss sums and rational points on metaplectic horocycles. Int. J. Number Theory **10**(6), 1553–1576 (2014)

6. Diamond, F., Shurman, J.: A First Course in Modular Forms, Volume 228 of Graduate Texts in Mathematics. Springer, New York (2005)

7. El-Baz, D., Huang, B., Lee, M.: Effective joint equidistribution of primitive rational points on expanding horospheres. arXiv:1811.04019 (arXiv preprint) (2018)

8. Einsiedler, M., Luethi, M., Shah, N.: Primitive rational points on expanding horocycles in products of the modular surface with the torus. Ergodic Theory Dyn. Syst. **20**, 1–45 (2020)

9. Einsiedler, M., Mozes, S., Shah, N., Shapira, U.: Equidistribution of primitive rational points on expanding horospheres. Compos. Math. **152**(4), 667–692 (2016)

10. Goldstein, D., Mayer, A.: On the equidistribution of Hecke points. Forum Math. **15**(2), 165–189 (2003)

11. Garland, H., Raghunathan, M.S.: Fundamental domains for lattices in (R-)rank 1 semisimple Lie groups. Ann. Math. **2**(92), 279–326 (1970)

12. Hejhal, D.A.: On value distribution properties of automorphic functions along closed horocycles. In: XVIth Rolf Nevanlinna Colloquium (Joensuu, 1995), pp. 39–52. de Gruyter, Berlin (1996)

13. Hejhal, D.A.: On the uniform equidistribution of long closed horocycles. Asian J. Math. **4**(4), 839–853 (2000). (**Loo-Keng Hua: a great mathematician of the twentieth century**)

14. Hardy, G.H., Wright, E.M.: An Introduction to the Theory of Numbers, 6th edn. Oxford University Press, Oxford (2008).. ((**Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles**))

15. Iwaniec, H.: Spectral Methods of Automorphic Forms, Volume 53 of Graduate Studies in Mathematics, Revista Matemática Iberoamericana, Madrid, 2nd edn. American Mathematical Society, Providence (2002)

16. Jana, S.: Joint equidistribution on the product of the circle and the unit cotangent bundle of the modular surface. J. Number Theory **226**, 271–283 (2021)

17. Kelmer, D., Kontorovich, A.: Effective equidistribution of shears and applications. Math. Ann. **370**(1–2), 381–421 (2018)

18. Kelmer, D., Kontorovich, A.: Exponents for the equidistribution of shears and applications. J. Number Theory **208**, 1–46 (2020)

19. Kim, H.H., Sarnak, P.: Refined estimates towards the Ramanujan and Selberg conjectures, appendix to H. H. Kim, functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2, with appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak. J. Am. Math. Soc. **16**, 1 (2003)

20. Kelmer, D., Yu, S.: Shrinking target problems for flows on homogeneous spaces. Trans. Am. Math. Soc. **372**(9), 6283–6314 (2019)

21. Lang, S.: $SL_2(R)$. Addison-Wesley Publishing Co, Reading (1975)

22. Li, H.: Effective limit distribution of the Frobenius numbers. Compos. Math. **151**(5), 898–916 (2015)

23. Lee, M., Marklof, J.: Effective equidistribution of rational points on expanding horospheres. Int. Math. Res. Not. **21**, 6581–6610 (2018)

24. Luethi, M.: Primitive rational points on expanding horospheres in Hilbert modular surfaces. J. Number Theory **225**, 327–359 (2021)

25. Margulis, G.A.: On Some Aspects of the Theory of Anosov Systems. Springer Monographs in Mathematics. Springer, Berlin (2004).. ((**With a survey by Richard Sharp: Periodic orbits of hyperbolic flows, Translated from the Russian by Valentina Vladimirovna Szulikowska**))

26. Marklof, J.: The asymptotic distribution of Frobenius numbers. Invent. Math. **181**(1), 179–207 (2010)

27. Marklof, J., Strömbergsson, A.: Equidistribution of Kronecker sequences along closed horocycles. Geom. Funct. Anal. **13**(6), 1239–1280 (2003)

28. Sarnak, P.: Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series. Comm. Pure Appl. Math. **34**(6), 719–739 (1981)

29. Sprindžuk, V.G.: Metric theory of Diophantine approximations. V. H. Winston & Sons, Washington, D.C (1979).. ((**Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics**))

30. Strömbergsson, A.: On the uniform equidistribution of long closed horocycles. Duke Math. J. **123**(3), 507–547 (2004)

31. Sarnak, P., Ubis, A.: The horocycle flow at prime times. J. Math. Pures Appl. (9) **103**(2), 575–618 (2015)
32. Zagier, D.: Eisenstein series and the Riemann zeta function. In: Automorphic Forms, Representation Theory and Arithmetic (Bombay, 1979), Volume 10 of Tata Institute of Fundamental Research Studies in Mathematics, pp. 275–301. Tata Institute of Fundamental Research, Bombay (1981)