Journal of Number Theory 208 (2020) 1–46

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

JNT Prime

Exponents for the equidistribution of shears and applications

Dubi Kelmer [a,1], Alex Kontorovich [b,c,*,2]

[a] *Boston College, Boston, MA, United States of America*
[b] *Rutgers University, New Brunswick, NJ, United States of America*
[c] *Institute for Advanced Study, Princeton, NJ, United States of America*

ARTICLE INFO

Article history:
Received 22 February 2019
Received in revised form 15 August 2019
Accepted 18 August 2019
Available online 27 September 2019
Communicated by S.J. Miller

Keywords:
Equidistribution
Quadratic forms
Shears

ABSTRACT

In [KK18], the authors introduced "soft" methods to prove the effective (i.e. with power savings error) equidistribution of "shears" in cusped hyperbolic surfaces. In this paper, we study the same problem but now allow full use of the spectral theory of automorphic forms to produce explicit exponents, and uniformity in parameters. We give applications to counting square values of quadratic forms.

Contents

* Corresponding author.
E-mail addresses: kelmer@bc.edu (D. Kelmer), alex.kontorovich@rutgers.edu (A. Kontorovich).
[1] Kelmer is partially supported by NSF CAREER grant DMS-1651563.
[2] Kontorovich is partially supported by an NSF CAREER grant DMS-1455705, an NSF FRG grant DMS-1463940, a BSF grant 2014099, a Simons Fellowship, a von Neumann Fellowship at IAS, and the IAS's NSF grant DMS-1638352.

https://doi.org/10.1016/j.jnt.2019.08.024
0022-314X/

1. Introduction

1.1. Equidistribution

Let $\Gamma < G := \mathrm{PSL}_2(\mathbb{R})$ be a non-uniform lattice, and let

$$\mathbf{x}_0 \in T^1(\Gamma\backslash\mathbb{H}) \cong \Gamma\backslash G$$

be a base point in the unit tangent bundle of the punctured surface $\Gamma\backslash\mathbb{H}$, so that the visual limit point

$$\mathfrak{a} = \lim_{t\to\infty} \mathbf{x}_0 \cdot a_t$$

is a cusp of Γ. Here $a_t = \mathrm{diag}(e^{t/2}, e^{-t/2})$ is the geodesic flow; let $A^+ = \{a_t, t > 0\}$. As in [KK18], we define a **shear** of the cuspidal geodesic ray $\mathbf{x}_0 \cdot A^+$ to be its left-translate by

$$\mathfrak{s}_T := \begin{pmatrix} (T^2+1)^{-1/4} & T(T^2+1)^{-1/4} \\ 0 & (T^2+1)^{1/4} \end{pmatrix}.$$

Note that $\mathfrak{s}_T$ arises naturally as $\mathfrak{s}_T = a_{-\frac{1}{2}\log(T^2+1)} \cdot n_T$, where $n_x = \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$. For example, identifying G/K with $\mathbb{H}$ under $g \mapsto g \cdot i$ (here $K = \mathrm{SO}(2)$ is a maximal compact subgroup) and taking $\mathbf{x}_0 = e$, the identity element of G, we have for a right-K-invariant test function $\Psi \in C_c(\Gamma\backslash G)^K$ that its evaluation along such a shear is given by

$$\int\limits_{a\in A^+} \Psi(\mathbf{x}_0 \cdot a \cdot \mathfrak{s}_T)da = \int\limits_{1/\sqrt{T^2+1}}^{\infty} \Psi\left(Ty + iy\right)\frac{dy}{y}.$$

In [KK18], the authors proved the effective equidistribution of shears (as $T \to \infty$) using “soft” ergodic methods (e.g. mixing) and basic properties of Eisenstein series.[3] The goal of this paper is to use more of the spectral theory of automorphic forms to produce explicit exponents in this problem. For ease of exposition, and to write the best exponents that come from our method, we restrict below to Γ conjugate to the congruence group

[3] See also [OS14], where an asymptotic formula is obtained by different means, with an error term weaker than power savings.

$$\Gamma_0(p) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{PSL}_2(\mathbb{Z}) : c \equiv 0 (\operatorname{mod} p)\},$$

with p a prime number. A special case of Theorem 4.1 below gives the following bound.

Theorem 1.1. *Assume the Ramanujan conjecture for the exponent bounding the Fourier coefficients of Maass forms on* Γ *(see §3). Then for any* $\Psi \in C_c^\infty(\Gamma\backslash G)^K$*, any* $T \geq 2$*, and any* $\epsilon > 0$*, there are constants* $C_j = C_j(\Psi)$*,* $j = 1, 2$*, so that*

$$\int\limits_{a \in A^+} \Psi(\mathbf{x}_0 \cdot a \cdot \mathfrak{s}_T) da = C_1 \log T + C_2 + O_{\Psi,\epsilon}(T^{-1/4+\epsilon}).$$

Remark 1.2. We take the opportunity here to correct an error in the analysis in [KK18], which has no effect on the qualitative power gain, but does affect the exponents as explicitly quantified here. In particular, [KK18, Remark 1.7] is incorrect as stated, and we do not know how to obtain square-root cancellation by this approach. See Remark 4.5 for the error and how to correct it.

1.2. Counting

As is standard, such equidistribution results can be applied to counting problems in discrete orbits. In particular, see [KK18, §1.3.1] where we explain that proving the effective equidistribution of shears settles the remaining lacunary cases of the Duke-Rudnick-Sarnak/Eskin-McMullen program [DRS93,EM93,Mar04] of effectively counting discrete orbits on quadrics in archimedean balls. In smooth form, one can produce from Theorem 1.1 above some rather sharp error exponents, as we now illustrate.

Let F be a real ternary indefinite quadratic form, let $G = \mathrm{SO}_F^\circ(\mathbb{R})$ be the connected component of the real special orthogonal group preserving F, and assume that $\Gamma < G$ is the image of $\Gamma_0(p)$ under a spin morphism $\mathrm{PSL}_2(\mathbb{R}) \to \mathrm{SO}_F^\circ(\mathbb{R})$ (see §6.1). Let $\psi : G \to \mathbb{R}_{\geq 0}$ be a smooth bump function about a sufficiently small bi-K-invariant neighborhood of the identity in G (that is, a region of the form (2.3)), with $\int_G \psi dg = 1$. Fix $\mathbf{v}_0 \in \mathbb{R}^3$ so that the orbit

$$\mathcal{O} := \mathbf{v}_0 \cdot \Gamma \ \subset \ \mathbb{R}^3$$

is discrete, the stabilizer of $\mathbf{v}_0$ in G is a split torus, H, say, and $\Gamma \cap H$ is finite. Fix a right-K-invariant archimedean norm $\|\cdot\|$ on $H\backslash G \cong \mathbf{v}_0 \cdot G$. This data induces a smoothed indicator function of a norm-T ball on $\mathbf{v} \in \mathbf{v}_0 \cdot G$ via convolution with ψ:

$$\widetilde{\psi}_T(\mathbf{v}) := \int\limits_{g \in G} \mathbf{1}_{\{\|\mathbf{v}g\| < T\}} \psi(g) dg. \tag{1.3}$$

In §5 (see Proposition 5.12), we prove the following

Theorem 1.4. *Again assume the Ramanujan conjecture for Maass forms on* Γ. *Then for any* $\epsilon > 0$, *there are constants* $C_j = C_j(\psi)$, $j = 1, 2$, *so that*

$$\sum_{\mathbf{v}\in\mathcal{O}} \widetilde{\psi}_T(\mathbf{v}) = C_1 T \log T + C_2 T + O_{\epsilon,\psi}\left(T^{\frac{3}{4}+\epsilon}\right).$$

Unconditionally, the error exponent $\frac{3}{4}+\epsilon$ *can be replaced by* $\frac{3+4\theta}{4+8\theta}+\epsilon$, *where* $\theta = 7/64$ *is the best currently known bound towards the Ramanujan Conjecture (which stipulates that* $\theta = 0$ *holds).*

1.3. Explicit constants

In certain settings of classical interest, one can go a step further and explicitly identify the constants C_j appearing in the main terms above. To showcase this fact, we count integer points on the inhomogeneous Pythagorean quadric:

$$W_d \ : \ x^2 + y^2 - z^2 = d,$$

when d is a perfect square (which corresponds to $\Gamma \cap H$ finite as above). After unsmoothing the count in Theorem 1.4 to make the constants independent of the smoothing function ψ, we obtain the following.

Theorem 1.5. *Let* $d = n^2$ *be a square, with* $n \in \mathbb{Z}_{>0}$. *Define the counting function*

$$\mathcal{N}_d(T) \ := \ \#W_d(\mathbb{Z}) \cap \{x^2 + y^2 + z^2 < T^2\}.$$

Again let $\theta = 7/64$ *be the bound towards the Ramanujan Conjecture (that* $\theta = 0$*). For any* $\eta < \frac{3}{40+72\theta}$, $\beta > \frac{3}{2} + 2\theta$, *and* $T \geq d^\beta$, *we have that*

$$\mathcal{N}_d(T) \ = \ \mathcal{M}_d(T) + O\left(T^{1-\eta} d^{\beta\eta}\right), \tag{1.6}$$

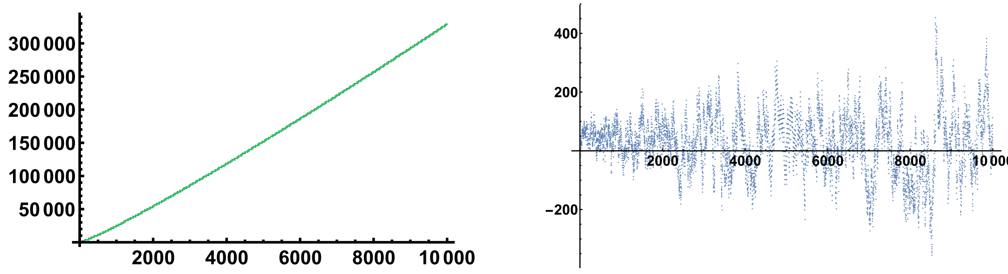
where the "main term" is given by

$$\mathcal{M}_d(T) \ := \ \frac{\sqrt{128}T}{\pi}\left(\log(T) + C - D(n) + \log(2)\left(\tfrac{1}{3} - \tfrac{1}{2^{\nu+2}}\right)\right).$$

Here $\nu = \nu_2(n)$ *is the* 2*-adic valuation of* n, *the constant* C *is given by:*

$$C = 2\gamma - 1 - 2\frac{\zeta'}{\zeta}(2) - \tfrac{\log(2)}{2} - \log\left(\tfrac{|\Gamma(1/4)|^4}{4\pi^3}\right) \ = \ 0.616174\ldots,$$

where $\gamma = 0.577\ldots$ *is Euler's constant,* $\zeta(s)$ *is the Riemann zeta function,* $\Gamma(s)$ *is the Gamma function, and* ϕ *is the Euler totient function, and the factor* $D(n)$ *is the Dirichlet coefficient of* $(\zeta^2 \cdot \zeta')(s-1)$, *that is,*



(a) ■=$\mathcal{N}_d(T)$ and ■=$\mathcal{M}_d(T)$ (indistinguishable)

(b) $\mathcal{N}_d(T) - \mathcal{M}_d(T)$

Fig. 1. Plots for $d = 12^2 = 144$ and $T < 10,000$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$D(n) = \frac{1}{n} \sum_{a|n} \phi(\tfrac{n}{a}) \log(a).$$

To illustrate the validity of this complicated formula, we verify it numerically with plots of $\mathcal{N}_d(T)$, $\mathcal{M}_d(T)$, and their difference, for $d = 144$; see Fig. 1. For T as large as $10,000$, the counting function reaches around $350,000$, while the difference $\mathcal{N}_d(T) - \mathcal{M}_d(T)$ remains of size around $400 \asymp \sqrt{T}$, suggesting perhaps that (1.6) may remain valid with any $\eta < 1/2$ and $\beta = 1/2$.

Remark 1.7. Our results are meaningful as long as $T > d^{3/2+2\theta+\epsilon}$. In particular, assuming Ramanujan we can take T to be almost as small as $d^{3/2}$. When $T < \sqrt{d}$ we trivially have $\mathcal{N}_d(T) = 0$ and it is an interesting problem to obtain meaningful asymptotics also for the range $\sqrt{d} < T < d^{3/2}$. (In the rather different setting of d being a fundamental discriminant, Friedlander-Iwaniec [FI13], using different tools showed an asymptotic formula which is effective for T almost as small as $\sqrt{d}$.)

Along the way to proving Theorem 1.5, we need to establish the following result, counting the number of binary quadratic forms of a fixed square discriminant d with coefficients in a norm ball. As it is possibly of independent interest, we record it here. Let

$$Q(a, b, c) = b^2 - 4ac,$$

and

$$\mathcal{N}_{Q,d}(T) = \{(a, b, c) \in \mathbb{Z}^3 : Q(a, b, c) = d,\ 2a^2 + b^2 + 2c^2 \leq T^2\}. \tag{1.8}$$

Theorem 1.9. *With all notation and assumptions as in Theorem 1.5, we have*

$$\mathcal{N}_{Q,d}(T) = \frac{\sqrt{72}T}{\pi} \left(\log(T) + C - D(n) + O(\frac{d^{\eta\beta}}{T^\eta}) \right). \tag{1.10}$$

As discussed in [KK18, Remark 1.10], there are many other methods for counting such expressions. For example, Hulse et al. [HKKL16] used a Multiple Dirichlet Series technique to count binary forms with a fixed discriminant. In smooth form and counting in a slightly different region, they obtain a version of (1.10) with square-root error in the T aspect, but with no visible uniformity in d. It is also not clear how easy it would be to convert their constants into the completely numerically explicit values as in (1.10).

1.4. Notation

Throughout this paper we denote by $G = \mathrm{PSL}_2(\mathbb{R})$ and Γ a non-uniform lattice in G. For ease of exposition (and for our applications), we assume that Γ is conjugate to $\Gamma_0(p)$; minor modifications are needed to handle a more general setting. We will use the notation $A(t) \ll B(t)$ to mean that there is some constant $c > 0$ such that $A(t) \leq cB(t)$, and we will use subscripts to indicate the dependance of the constant on parameters. For $B(t) \geq 0$ the notation $A(t) = O(B(t))$ means that $|A(t)| \ll B(t)$. We write $A(t) \asymp B(t)$ for $A(t) \ll B(t) \ll A(t)$.

1.5. Organization

The preliminary §2 reviews spectral decompositions, Eisenstein series, and Sobolev norms. Then in §3, we improve on our analysis in [KK18], giving stronger estimates for Fourier expansions at various cusps in terms of approximations to the Ramanujan Conjectures. We use these, together with some slight improvements on the method in [KK18], to prove in §4 the sharp equidistribution of shears claimed in Theorem 1.1. The counting Theorem 1.4 is derived from this in §5, and then used in §6 to prove the explicit counting Theorems 1.5 and 1.9. Calculations of Fourier expansions for Eisenstein series on $\Gamma_0(p)$ are reserved for the Appendix.

Acknowledgments

We thank Zeev Rudnick for comments on an earlier draft, and the referee for an extremely thorough and thoughtful report.

2. Preliminaries

2.1. Coordinates

Let $K, A, N \leq G$ denote the orthogonal group, the diagonal group, and the unipotent group respectively. Explicitly let

$$k_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \; a_y = \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix}, \text{ and } n_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},$$

parametrize elements in K, A and N respectively. The decomposition $G = NAK$ gives coordinates $g = n_x a_y k_\theta$ on G, and the Haar measure in these coordinates is

$$dg = \frac{dxdyd\theta}{2\pi y^2}. \tag{2.1}$$

The group G acts on the upper half space $\mathbb{H} = \{z = x + iy \in \mathbb{C} : y > 0\}$ by linear fractional transformations, explicitly, for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $gz = \frac{az+b}{cz+d}$ preserving the hyperbolic area $d\mu = \frac{dxdy}{y^2}$.

For any lattice Γ we identify the quotient $\Gamma\backslash\mathbb{H}$ with $\Gamma\backslash G/K$ (in particular, we will think of functions on $\Gamma\backslash\mathbb{H}$ as right K-invariant functions on $\Gamma\backslash G$). Our normalization (2.1) is such that the Haar measure of $\Gamma\backslash G$ is equal to the hyperbolic area $v_\Gamma = \mu(\Gamma\backslash\mathbb{H})$.

Remark 2.2. For $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$ we recall that $v_{\Gamma_1} = \frac{\pi}{3}$.

2.2. Sobolev norms

Fix a basis $\mathscr{B} = \{X_1, X_2, X_3\}$ for the Lie algebra $\mathfrak{g}$ of G, and given a smooth test function $\Psi \in C^\infty(\Gamma\backslash G)$, define the "$L^p$, order-$d$" Sobolev norm $\mathcal{S}_{p,d}(\Psi)$ as

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

Here $\mathscr{D}$ ranges over monomials in $\mathscr{B}$ of order at most d. Note that since the right action of $\mathscr{B}$ commutes with the left action of G, all norms are invariant in the sense that $\mathcal{S}_{p,d}(\Psi^\tau) = \mathcal{S}_{p,d}(\Psi)$ where $\Psi^\tau(x) = \Psi(\tau x)$.

We will work with various norms that are convex combinations of these Sobolev norms and it will be convenient to classify these norms with respect to how large they become on functions approximating a small bump function. For small $0 < \delta < 1$ let

$$B_\delta = KA_\delta K \tag{2.3}$$

denote a (spherical) δ-neighborhood of the identity, where

$$A_\delta = \{a_y : |\log(y)| < \delta\}.$$

Definition 2.4. Fix a family of smooth functions ψ_δ on $K\backslash G/K$, supported on B_δ, with average

$$\int_G \psi_\delta(g)dg = 1,$$

and denote the corresponding periodized function by

$$\Psi_\delta(g) = \sum_{\gamma\in\Gamma} \psi_\delta(\gamma g).$$

We say that a norm $\mathcal{S}$ is of **degree** α if these periodized functions have norms growing like $\mathcal{S}(\Psi_\delta) \asymp \delta^{-\alpha}$. We will slightly abuse notation and sometimes denote by $\mathcal{S}_\alpha$ a norm of degree α, without specifying the norm explicitly.

Note for future reference that the Sobolev norm $\mathcal{S}_{p,d}$ is of degree $d+2-\frac{2}{p}$, in particular, the L^2 norm has degree 1, while the L^∞ norm has degree 2. Note that if $\mathcal{S}_\alpha$ and $\mathcal{S}_\beta$ are of degrees α and β respectively, then the convex combination $\mathcal{S}_\alpha^q\mathcal{S}_\beta^{1-q}$ is of degree $q\alpha+(1-q)\beta$. Moreover, if $\alpha\leq\beta$ then $\tilde{\mathcal{S}}_\beta = \max\{\mathcal{S}_\alpha,\mathcal{S}_\beta\}$ is also of degree β, hence, after perhaps replacing $\mathcal{S}_\beta$ with $\tilde{\mathcal{S}}_\beta$ we may assume without loss of generality that $\mathcal{S}_\alpha\leq\mathcal{S}_\beta$ whenever $\alpha\leq\beta$.

2.3. Eisenstein series

For any cusp $\mathfrak{a}$ of Γ let $\Gamma_\mathfrak{a}$ denote the stabilizer of $\mathfrak{a}$ in Γ and $\tau_\mathfrak{a}\in G$ be a corresponding scaling matrix such that $\tau_\mathfrak{a}\infty=\mathfrak{a}$ and

$$\tau_\mathfrak{a}^{-1}\Gamma\tau_\mathfrak{a}\cap N = \{n_k : k\in\mathbb{Z}\}.$$

In particular, for the congruence groups $\Gamma_0(p)$ there are two cusps, one at ∞ of width 1 (that is, the stabilizer of ∞ is generated by $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$), and the other at 0 of width p with scaling matrix $\tau_0 = \begin{pmatrix}0 & 1/\sqrt{p}\\ -\sqrt{p} & 0\end{pmatrix} = k_{\pi/2}\cdot a_p$.

The Eisenstein series corresponding to a cusp $\mathfrak{a}$ is defined for $\mathfrak{Re}(s)>1$ by

$$E_{\Gamma,\mathfrak{a}}(z,s) = \sum_{\gamma\in\Gamma_\mathfrak{a}\backslash\Gamma} (\mathfrak{Im}(\tau_\mathfrak{a}^{-1}\gamma z))^s \tag{2.5}$$

and has a meromorphic continuation to $\mathbb{C}$ with a simple pole at $s=1$ with residue $\frac{1}{v_\Gamma}$ and (since Γ is congruence) no other poles in $\mathfrak{Re}(s)>\frac{1}{2}$.

One can regularize the Eisenstein series by subtracting the pole at $s=1$, and we define the corresponding Kronecker limit by

$$\mathcal{K}_{\Gamma,\mathfrak{a}}(z) = \lim_{s\to1}\left(E_{\Gamma,\mathfrak{a}}(s,z) - \frac{1}{v_\Gamma(s-1)}\right). \tag{2.6}$$

2.4. Spectral decomposition

The hyperbolic Laplace operator Δ acts (after unique extension) on the space $L^2(\Gamma\backslash\mathbb{H})$ of square-integrable automorphic functions, and is self-adjoint and positive semi-definite. The spectrum of Δ is composed of the constant functions, the continuous part (spanned by Eisenstein series), and discrete part (spanned, in the congruence case, by Maass cusp forms).

We denote by $\mathcal{E}(\Gamma\backslash\mathbb{H})$ the space spanned by the Eisenstein series and by $\mathcal{C}(\Gamma\backslash\mathbb{H})$ its orthogonal complement which is the space of cusp forms. For $\Gamma = \Gamma_0(p)$ the space of cusp forms further decomposes into the space of old forms $\mathcal{C}_{\mathrm{old}}(\Gamma\backslash\mathbb{H})$ spanned by the set $\{\varphi(z), \varphi(pz) : \varphi \in \mathcal{C}(\Gamma_1\backslash\mathbb{H})\}$, and its orthogonal complement $\mathcal{C}_{\mathrm{new}}(\Gamma\backslash\mathbb{H})$.

For congruence Γ, the space of cusp forms has an orthonormal basis composed of Hecke-Maass forms, that are joint eigenfunctions of the Laplacian and all Hecke operators. We have the following spectral decomposition (see [Iwa95, Theorems 4.7 and 7.3]).

Proposition 2.7. *For $\Psi \in L^2(\Gamma\backslash\mathbb{H})$,*

$$\Psi(z) = \mu_\Gamma(\Psi) + \sum_k \langle \Psi, \varphi_k \rangle \varphi_k(z) \tag{2.8}$$

$$+ \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2} + ir) \right\rangle E_{\Gamma,\mathfrak{a}}(z, \tfrac{1}{2} + ir) dr,$$

where $\mu_\Gamma(\Psi) = \frac{1}{v_\Gamma} \int_{\Gamma\backslash\mathbb{H}} \Psi(z) d\mu(z)$, and the first sum is over an orthonormal basis of Maass cusp forms with eigenvalues $\lambda_k = \frac{1}{4} + r_k^2$ with $r_k \in i(0, \frac{1}{2}) \cup [0, \infty)$.

The equality is in $L^2(\Gamma\backslash\mathbb{H})$ and pointwise for $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$. As a direct consequence, we have the following

Corollary 2.9. *For $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$,*

$$\|\Psi\|^2 = |\mu_\Gamma(\Psi)|^2 + \sum_k |\langle \Psi, \varphi_k \rangle|^2 +$$

$$\sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} |\left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2} + ir) \right\rangle|^2 dr.$$

3. Fourier coefficients

In this section, we derive general bounds for Fourier coefficients of test functions at various cusps. In principle, most of the steps are standard, but we did not find a reference in the literature which carries out each of the necessary calculations, so we give details for the reader's benefit. Another reason for restricting to Γ conjugate to $\Gamma_0(p)$ is that the general theory of Fourier coefficients at arbitrary cusps becomes extremely cumbersome (see, e.g., [GHL15, Theorem 49]).

First we specify what we mean by the "Ramanujan conjectures." Let φ be a Hecke-Maass cusp newform for $\Gamma_0(p)$, with Laplace Eigenvalue $\frac{1}{4} + r^2$. For $m \neq 0$, its mth Fourier coefficient (at the cusp $\mathfrak{a} = \infty$) satisfies

$$a_{\varphi,\infty}(m,y) := \int_0^1 \varphi(x+iy)e(-mx)dx = a_{\varphi,\infty}(m)\sqrt{y}K_{ir}(2\pi|m|y),$$

with $K_s(y)$ the Bessel function of the second kind. The coefficient further decomposes as

$$a_{\varphi,\infty}(m) = a_{\varphi,\infty}(1)\lambda(m),$$

where $\lambda(m)$ is the corresponding Hecke eigenvalue. Let $\theta \in [0,1/2]$ be a number so that

$$|\lambda(m)| \ll_\epsilon |m|^{\theta+\epsilon}. \tag{3.1}$$

In particular,

$$\theta = 7/64$$

is known [KS03], while the Ramanujan conjecture predicts that $\theta = 0$ holds.

Remark 3.2. Selberg's eigenvalue conjecture is the Ramanujan conjecture "at infinity," and asserts that for the congruence groups $\Gamma_0(p)$ there are no Maass cusp forms with eigenvalue $\lambda < 1/4$. While the conjecture is known for $\mathrm{PSL}_2(\mathbb{Z})$ and some small values of p, for general congruence groups we currently only know that a hypothetical exceptional Maass form with eigenvalue $\lambda = 1/4 + r^2$, $r \in i(0, \frac{1}{2})$, has $r \in i(0,\theta]$, $\theta = 7/64$.

Now, let $\mathfrak{a}$ be a cusp of a lattice Γ, and let $\tau_\mathfrak{a}$ denote the corresponding scaling matrix. Then for any test function $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$, the translated function $\Psi^{\tau_\mathfrak{a}}(z) := \Psi(\tau_\mathfrak{a} z)$ is periodic in x with period one and hence has a Fourier expansion

$$\Psi^{\tau_\mathfrak{a}}(z) = \sum_{m\in\mathbb{Z}} a_{\Psi,\mathfrak{a}}(m,y)e^{2\pi imx}. \tag{3.3}$$

In [KK18, Prop. 2.2], we proved that there are constants $0 < c_\Gamma < \infty$ and $0 < \eta_\Gamma < 1$ and some norm $\mathcal{S}$ (a convex combination of Sobolev norms) such that these coefficients satisfy

$$|a_{\Psi,\mathfrak{a}}(m,y)| \ll_\Gamma \mathcal{S}(\Psi)|m|^{c_\Gamma} y^{\eta_\Gamma}$$

uniformly for all $0 \neq m \in \mathbb{Z}$ and $y > 0$. The argument there was quite soft (using mixing) and applied to any lattice. Now we specialize to Γ conjugate to $\Gamma_0(p)$ to improve the exponents c_Γ and η_Γ above, as well as to have better control on the degree of the Sobolev norm $\mathcal{S}$. Our main result is the following.

Proposition 3.4. *Let Γ be conjugate to $\Gamma_0(p)$. For any $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$ and any cusp $\mathfrak{a}$ of Γ, we have that*

$$a_{\Psi,\mathfrak{a}}(0,y) = \mu_\Gamma(\Psi) + O(\|\Psi\|_2^{3/4}\|\triangle\Psi\|_2^{1/4}y^{1/2}). \tag{3.5}$$

Moreover, for any $m \neq 0$, any $\epsilon > 0$ and any

$$\alpha_0 > 5/3,$$

we have

$$a_{\Psi,\mathfrak{a}}(m,y) = \sum_{r_k \in i(0,\frac{1}{2})} \langle \Psi, \varphi_k \rangle a_{\varphi_k,\mathfrak{a}}(m,y) + O_{\alpha_0,\epsilon,p}\left(\mathcal{S}_{\alpha_0}(\Psi) y^{\frac{1}{2}-\epsilon}|m|^\theta\right), \tag{3.6}$$

where $\mathcal{S}_{\alpha_0}$ is a norm of degree α_0. Moreover, for each of the exceptional forms $a_{\varphi_k,\mathfrak{a}}(m,y) = O_\epsilon(|m|^{\theta-|r_k|+\epsilon}y^{1/2-|r_k|}e^{-2\pi|m|y})$.

In order to prove Proposition 3.4 we consider the spectral decomposition of Ψ into Maass forms and Eisenstein series and bound the Fourier coefficients of each. Explicitly, for the cusp forms we show the following.

Lemma 3.7. *Let φ_k be a Hecke-Maass cusp form on $\Gamma_0(p)$ with eigenvalue $\frac{1}{4} + r_k^2$. Then for any cusp $\mathfrak{a}$, any $m \neq 0$, and any $\epsilon > 0$, we have for $r_k \geq 0$*

$$|a_{\varphi_k,\mathfrak{a}}(m,y)| \ll_{\epsilon,p} (r_k+1)^{-1/3+\epsilon}y^{1/2-\epsilon}|m|^\theta \tag{3.8}$$

while for $r_k = i\sigma_k$ with $\sigma_k \in (0,1/2)$

$$|a_{\varphi_k,\mathfrak{a}}(m,y)| \ll_\epsilon |m|^{\theta-\sigma_k+\epsilon}y^{1/2-\sigma_k}e^{-2\pi|m|y}. \tag{3.9}$$

Proof. When φ is an eigenfunction of the Laplacian with eigenvalue $s(1-s)$ we have that $a_{\varphi,\mathfrak{a}}(0,y)$ is a linear combination of y^s and y^{1-s} and for $m \neq 0$ it takes the form

$$a_{\varphi,\mathfrak{a}}(m,y) = a_{\varphi,\mathfrak{a}}(m)\sqrt{y}K_{s-1/2}(2\pi|m|y), \tag{3.10}$$

with $K_s(y)$ the Bessel function of the second kind.

We first consider the case when φ_k is a new form. Recall that $\Gamma_0(p)$ has two inequivalent cusps, one at ∞ and one at 0. First assume that $\mathfrak{a}$ is equivalent to ∞. Combining (3.1) with Hoffstein-Lockhart's [HL94, Corollary 0.3] stating that $a_{\varphi_k,\infty}(1) \ll_{p,\epsilon} (r_k^2+1/4)^\epsilon e^{\pi r_k/2}$, we obtain the bound

$$|a_{\varphi_k,\infty}(m)| \ll_{p,\epsilon} (|m|r_k)^\epsilon |m|^\theta e^{\pi r_k/2} \tag{3.11}$$

for $r_k \geq 0$ and $|a_{\varphi_k,\infty}(m)| = O(|m|^{\theta+\epsilon})$ for $r_k \in i(0,1/2)$. Now for $r_k > 0$ we use the bound [Str04, eq. 4.15] for the Bessel function,

$$|K_{ir}(y)| \ll_\epsilon e^{-\pi r/2}(r+1)^{-1/3+\epsilon} y^{-\epsilon} \min\{1, e^{\pi r/2 - y}\}, \tag{3.12}$$

to see that (3.8) holds in this case. While for $r_k = i\sigma_k$ we use the bound $K_\sigma(y) \ll y^{-|\sigma|}e^{-y}$ (which follows directly from the formula $K_\sigma(y) = \int_0^\infty e^{-y\cosh(t)}\cosh(\sigma t)dt$) to get (3.9).

Next, for the cusp at 0 we note that the scaling matrix $\tau_0 = \left(\begin{smallmatrix} 0 & 1/\sqrt{p} \\ -\sqrt{p} & 0 \end{smallmatrix}\right)$ commutes with the Hecke operators $T(n)$ with $(n,p) = 1$ and satisfies $\tau_0^{-1}\Gamma_0(p)\tau_0 = \Gamma_0(p)$ (see [Asa76]). Hence $\varphi_k^{\tau_0}$ is also Hecke eigenfunction with the same eigenvalues, and from multiplicity one for new forms we get that $\varphi_k^{\tau_0} = c\varphi_k$ with some scalar c of modulus 1. Hence, in absolute value, $|a_{\varphi_k,\infty}(m)| = |a_{\varphi_k,0}(m)|$ so we have the same bounds also for the cusp at 0.

Finally, the bound for old forms follows directly from the bound for new forms of $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$. Explicitly, let φ be a Hecke-Maass form for Γ_1 with Fourier coefficients $a_\varphi(m,y)$. From this form we get two companion forms $\varphi_1(z) = \varphi(z)$ and $\varphi_2(z) = \varphi(pz)$ invariant under $\Gamma_0(p)$. For the cusp at infinity $\varphi_1 = \varphi$ has the same Fourier expansion at infinity as φ. For the second form

$$\varphi_2(z) = \varphi(pz) = \sum_m a_\varphi(m, py)e^{2\pi i mpx} = \sum_{m\equiv 0(p)} a_\varphi(\tfrac{m}{p}, py)e^{2\pi i mx},$$

hence $a_{\varphi_2,\infty}(m,y) = a_\varphi(m/p, py)$ if $p|m$ and is zero otherwise. The cusp at zero has scaling matrix $\tau_0 = k_{\pi/2}\cdot a_p$. Write $\sigma = k_{\pi/2}$ so that $\tau_0 = \sigma a_p$. Since $\varphi^\sigma = \varphi$ we get that $\varphi_1^{\tau_0} = \varphi^{\sigma a_p} = \varphi^{a_p} = \varphi_2$, whence $a_{\varphi_1,0} = a_{\varphi_2,\infty}$. Similarly $\varphi_2^{\tau_0} = \varphi^{a_p\sigma a_p} = \varphi^\sigma = \varphi$ and $a_{\varphi_2,0} = a_{\varphi_1,\infty}$. Thus the same bound holds also for the cusp at zero. □

Next, we need to bound the Fourier coefficients of Eisenstein series. For each pair of cusps $\mathfrak{a}, \mathfrak{b}$ of Γ the Fourier expansion of the Eisenstein series $E_{\Gamma,\mathfrak{b}}$ with respect to the cusp at $\mathfrak{a}$, is given by

$$E^{\tau_\mathfrak{a}}_{\Gamma,\mathfrak{b}}(z,s) = \delta_{\mathfrak{a},\mathfrak{b}}y^s + \phi_{\mathfrak{a},\mathfrak{b}}(s)y^{1-s} + \sum_{m\neq 0} a_{\mathfrak{a},\mathfrak{b}}(s;m,y)e(mx).$$

Lemma 3.13. *For Γ conjugate to $\Gamma_0(p)$ and any two cusps $\mathfrak{a}, \mathfrak{b}$, we have for $r \in \mathbb{R}$ that*

$$|a_{\mathfrak{a},\mathfrak{b}}(\tfrac{1}{2} + ir; m, y)| \ll_\epsilon y^{1/2-\epsilon}(1+|r|)^{-1/3+\epsilon}. \tag{3.14}$$

Proof. Since $E_{\Gamma,\mathfrak{a}}(z,s)$ is an eigenfunction with eigenvalue $s(1-s)$ we can write

$$a_{\mathfrak{a},\mathfrak{b}}(s;m,y) = \phi_{\mathfrak{a},\mathfrak{b}}(s;m)2\sqrt{y}K_{s-\frac{1}{2}}(2\pi|m|y).$$

For the full modular group $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$ there is just one cusp at ∞ and the Fourier coefficients are given explicitly by $\phi(s) = \frac{\zeta^*(2s-1)}{\zeta^*(2s)}$ and

$$\phi(s;m) = \frac{\tau_{s-1/2}(m)}{\zeta^*(2s)}, \tag{3.15}$$

where $\zeta^*(s) = \pi^{-s/2}\zeta(s)\Gamma(s/2)$ is the completed Riemann zeta function and $\tau_s(m) = \sum_{ab=|m|}(\frac{a}{b})^s$ is the divisor function [Iwa95, page 67]. In particular, using the Stirling approximation for the Γ-function

$$|\Gamma(\tfrac{1}{2}+ir)| \asymp e^{-\pi|r|/2},$$

the bound (3.12) for the Bessel function, together with the standard bounds $\frac{1}{|\zeta(1+ir)|} \ll |r|^\epsilon$ and $|\tau_{ir}(m)| \leq \tau_0(m) \ll_\epsilon |m|^\epsilon$, gives (3.14) in this case. For the congruence groups $\Gamma_0(p)$ the coefficients $\phi_{\mathfrak{a},\mathfrak{b}}(s;m)$ are given by a similar explicit formula (see Proposition A.7 below), resulting in the same bound. □

Combining the above bounds for Fourier coefficients of Maass forms and Eisenstein series we can use the spectral decomposition to bound the Fourier coefficients of any smooth function as follows.

Proof of Proposition 3.4. First, noting that $\Psi \in C_c^\infty(\Gamma_0(p)\backslash\mathbb{H})$ iff $\Psi^\tau \in C_c^\infty(\Gamma\backslash\mathbb{H})$ and the Fourier coefficients satisfy $|a_{\Psi,\mathfrak{a}}(m,y)| = |a_{\Psi^\tau,\mathfrak{b}}(m,y)|$ with $\mathfrak{b} = \tau^{-1}\mathfrak{a}$, we may assume that $\Gamma = \Gamma_0(p)$.

Let $\Gamma = \Gamma_0(p)$ and $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$. Using the spectral expansion we can write (for any cusp $\mathfrak{b}$) and $m \neq 0$

$$a_{\Psi,\mathfrak{b}}(m,y) = \sum_{r_k\in i(0,\theta)} \langle\Psi,\varphi_k\rangle a_{\varphi_k,\mathfrak{b}}(m,y) + \sum_{r_k\geq 0}\langle\Psi,\varphi_k\rangle a_{\varphi_k,\mathfrak{b}}(m,y)$$
$$+ \sum_{\mathfrak{a}} \frac{1}{2\pi}\int_{\mathbb{R}} \langle\Psi, E_{\Gamma,\mathfrak{a}}(\cdot,\tfrac{1}{2}+ir)\rangle a_{\mathfrak{a},\mathfrak{b}}(\tfrac{1}{2}+ir;m,y)dr.$$

To bound the contribution of the second sum fix a large parameter M (to be determined later). Applying the bound (3.8) to the Fourier coefficients we get

$$\left|\sum_{r_k>0}\langle\Psi,\varphi_k\rangle a_{\varphi_k,\mathfrak{b}}(m,y)\right| \ll_\epsilon y^{1/2-\epsilon}|m|^\theta\left(\sum_{0\leq r_k\leq M}\frac{|\langle\Psi,\varphi_k\rangle|}{(r_k+1)^{1/3-\epsilon}}\right.$$
$$\left. + \sum_{r_k\geq M}\frac{|\langle\triangle\Psi,\varphi_k\rangle|}{r_k^{7/3-\epsilon}}\right).$$

Using Cauchy-Schwarz, followed by summation by parts (using Weyl's law stating that $\#\{r_k \leq M\} \ll M^2$) we can bound the first sum by

$$\|\Psi\|_2\sqrt{\sum_{0\le r_k\le M}(r_k+1)^{-2/3+\epsilon}}\ll M^{2/3+\epsilon}\mathcal{S}_{2,0}(\Psi),$$

and the second by

$$\|\triangle\Psi\|_2\sqrt{\sum_{r_k>M} r_k^{-14/3+\epsilon}}\ll M^{-4/3+\epsilon}\mathcal{S}_{2,2}(\Psi).$$

Choosing $M=\mathcal{S}_{2,0}^{-1/2}\mathcal{S}_{2,2}^{1/2}$ we get that

$$|\sum_{r_k>0}\langle\Psi,\varphi_k\rangle a_{\varphi_k,\mathfrak{b}}(m,y)|\ll_\epsilon y^{1/2-\epsilon}|m|^\theta\mathcal{S}_{\frac{5}{3}+\epsilon}(\Psi),$$

where the norm $\mathcal{S}_{5/3+\epsilon}(\Psi)=\mathcal{S}_{2,0}(\Psi)^{2/3-\epsilon/2}\mathcal{S}_{2,2}(\Psi)^{1/3+\epsilon/2}$ is of degree $5/3+\epsilon$.

Next for the Eisenstein integrals for each pair of cusps $\mathfrak{a},\mathfrak{b}$ fix a large parameter M and use (3.14) to bound

$$\left|\int_{\mathbb{R}}\langle\Psi,E(\cdot,\tfrac12+ir)\rangle a_{\mathfrak{a},\mathfrak{b}}(\tfrac12+ir;m,y)dr\right|\ll_\epsilon y^{1/2-\epsilon}\left(\int_{|r|\le M}\frac{|\langle\Psi,E(\cdot,\frac12+ir)\rangle|}{(1+r)^{1/3-\epsilon}}dr\right.$$
$$\left.+\int_{|r|>M}\frac{|\langle\triangle\Psi,E(\cdot,\frac12+ir)\rangle|}{r^{7/3-\epsilon}}\right)dr.$$

As before we can use Cauchy-Schwarz to bound the first integral by $O_\epsilon(\mathcal{S}_{2,0}(\Psi)M^{1/6+\epsilon})$ and the second by $O_\epsilon(\mathcal{S}_{2,2}(\Psi)M^{-11/6+\epsilon})$ so taking $M=\mathcal{S}_{2,0}^{-1/2}\mathcal{S}_{2,2}^{1/2}$ the whole integral is bounded by

$$|\int_{\mathbb{R}}\langle\Psi,E(\cdot,\tfrac12+ir)\rangle a_r(m,y)dr|\ll y^{1/2-\epsilon}\mathcal{S}_{7/6+\epsilon}(\Psi),$$

where

$$\mathcal{S}_{7/6+\epsilon}(\Psi)=\mathcal{S}_{2,0}(\Psi)^{11/12-\epsilon/2}\mathcal{S}_{2,2}^{1/12+\epsilon/2},$$

is of degree $7/6+\epsilon$.

Collecting the contributions of cusp forms and Eisenstein series of all cusps we get that

$$a_{\Psi,\mathfrak{b}}(m,y)=\sum_{r_k\in i(0,\theta)}\langle\Psi,\varphi_k\rangle a_{\varphi_k,\mathfrak{b}}(m,y)+O\left(y^{1/2-\epsilon}\left(|m|^\theta\mathcal{S}_{\frac53+\epsilon}(\Psi)+\mathcal{S}_{\frac76+\epsilon}(\Psi)\right)\right).$$

Taking ϵ sufficiently small so that $\frac53+\epsilon=\alpha_0$, and noting that the second term is bounded by the first concludes the proof of (3.6).

For the trivial ($m = 0$) coefficient, again using the spectral expansion, the only contribution comes from the constant function (giving the main term) and the Eisenstein integrals. We thus need to bound for each pair of cusps

$$\int_{\mathbb{R}} \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle (y^{1/2+ir} + \phi_{\mathfrak{a},\mathfrak{b}}(\tfrac{1}{2}+ir) y^{1/2-ir}) dr$$

$$\ll y^{1/2} \int_{\mathbb{R}} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr.$$

Fix a large parameter M and separate the integral to

$$\int_{\mathbb{R}} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr = \int_{|r|<M} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr + \int_{|r|>M} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr.$$

Using Cauchy-Schwarz, we can bound the first integral by

$$\int_{|r|<M} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr \leq \sqrt{2M} \sqrt{\int_{\mathbb{R}} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right|^2 dr} \ll \sqrt{M} \|\Psi\|_2,$$

and

$$\int_{|r|>M} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr = \int_{|r|>M} \frac{\left| \left\langle \triangle\Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right|}{1/4 + r^2} dr \ll \|\triangle\Psi\|_2 M^{-3/2}.$$

We thus get that

$$\int_{\mathbb{R}} \left| \left\langle \Psi, E_{\Gamma,\mathfrak{a}}(\cdot, \tfrac{1}{2}+ir) \right\rangle \right| dr \ll \sqrt{M} \|\Psi\|_2 + M^{-3/2} \|\triangle\Psi\|_2,$$

and the optimal choice of $M = \sqrt{\frac{\|\triangle\Psi\|_2}{\|\Psi\|_2}}$ gives the desired bound. $\square$

3.1. Additional estimates

We end this section with several estimates which follow from our bounds on Fourier coefficients. First, as consequence of the estimate for $a_{\Psi,\mathfrak{a}}(0,y)$ we get the following useful estimate that we record here for future use.

Corollary 3.16. *Let Γ be conjugate to $\Gamma_0(p)$. For any cusp $\mathfrak{a}$ with scaling matrix $\tau_\mathfrak{a}$, for any $\psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$ we have the bound*

$$\left(\int_0^1 |\psi^{\tau_\mathfrak{a}}(x+iy)|^2 dx\right)^{1/2} \ll \mathcal{S}_1(\psi) + \mathcal{S}_2(\psi) y^{1/4}, \tag{3.17}$$

with $\mathcal{S}_1$ and $\mathcal{S}_2$ are suitable norms of degree 1 and 2 respectively.

Remark 3.18. When y is small and ψ approximates a bump function this is an improvement over the trivial bound of $\mathcal{S}_{\infty,0}(\psi)$ which is a norm of degree 2.

Proof. Using (3.5) with the test function $\Psi(z) = |\psi^{\tau_\mathfrak{a}}(z)|^2$, we get that

$$\int_0^1 |\psi^{\tau_\mathfrak{a}}(x+iy)|^2 dx = \mu(\Psi) + O(\|\Psi\|_2^{3/4} \|\triangle\Psi\|_2^{1/4} y^{1/2}).$$

For $\Psi = |\psi^{\tau_\mathfrak{a}}|^2$ we have

$$\mu(\Psi) = |\mathcal{S}_{2,0}(\psi)|^2, \ \|\Psi\|_2 = |\mathcal{S}_{4,0}(\psi)|^2, \ \text{ and } \|\triangle\Psi\|_2 \ll |\mathcal{S}_{4,2}(\psi)|^2.$$

Define the norms $\mathcal{S}_1(\psi) = \mathcal{S}_{2,0}(\psi)$ and

$$\mathcal{S}_2(\psi) := \mathcal{S}_{4,0}(\psi)^{3/4} \mathcal{S}_{4,2}(\psi)^{1/4}.$$

Clearly $\mathcal{S}_1$ is of degree 1 and $\mathcal{S}_2$ is of degree 2. Finally taking a square root gives the result. □

Combining Corollary 3.16 and Proposition 3.4 we obtain another estimate that we will need.

Proposition 3.19. *Let Γ be conjugate to $\Gamma_0(p)$ and $\Psi \in \mathcal{C}_c^\infty(\Gamma\backslash\mathbb{H})$ orthogonal to the exceptional spectrum. Then for any $\alpha_1 > \frac{5+6\theta}{3+6\theta}$ and $\eta_1 < \frac{1}{2+4\theta}$,*

$$\sum_{m\neq 0} \frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|} \ll_{\alpha_1,\eta_1} \mathcal{S}_{\alpha_1}(\Psi) y^{\eta_1} + \mathcal{S}_{\alpha_1+1}(\Psi) y^{\eta_1+1/4}, \tag{3.20}$$

with $\mathcal{S}_{\alpha_1}, \mathcal{S}_{\alpha_1+1}$ norms of degrees α_1 and α_1+1 respectively. Similarly, for any exceptional cusp form φ_k with eigenvalue $\frac{1}{4}+r_k^2<\frac{1}{4}$ for any $\epsilon>0$

$$\sum_{m\neq 0}\frac{|a_{\varphi_k,\mathfrak{a}}(m,y)|}{|m|}\ll_\epsilon y^{1/2-\theta-\epsilon}, \tag{3.21}$$

where the implied constant may depend on φ_k.

Proof. Replacing Ψ with $\Psi^{\tau_\mathfrak{a}}$ we may assume that $\mathfrak{a}=\infty$ and Γ has cusp of width one. We can find $\epsilon>0$ and $\alpha_0>5/3$, sufficiently small so that $\alpha_1=\frac{\alpha_0+2\theta}{1+2\theta}$ and $\eta_1=\frac{1-2\epsilon}{2(1+2\theta)}$.

Fix a parameter $M\geq 1$ and separate the sum to

$$\sum_{m\neq 0}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}=\sum_{0<|m|<M}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}+\sum_{|m|\geq M}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}.$$

For the first sum, using Proposition 3.4 we get that

$$\sum_{|m|<M}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}\ll \mathcal{S}_{\alpha_0}(\Psi)y^{1/2-\epsilon}M^\theta.$$

For the second sum using Cauchy-Schwarz, followed by Parseval, and then applying Corollary 3.16, we get

$$\begin{aligned}\sum_{|m|\geq M}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}&\leq\frac{1}{\sqrt{M}}\left(\sum_{|m|\geq M}|a_{\Psi,\mathfrak{a}}(m,y)|^2\right)^{1/2}\\&\leq\frac{1}{\sqrt{M}}\left(\int_0^1|\Psi(x+iy)|^2dx\right)^{1/2}\\&\ll\frac{\mathcal{S}_1(\Psi)+\mathcal{S}_2(\Psi)y^{1/4}}{\sqrt{M}}.\end{aligned}$$

Now, when $\mathcal{S}_1(\Psi)\geq\mathcal{S}_{\alpha_0}(\Psi)y^{1/2-\epsilon}$ let $M=(\frac{\mathcal{S}_1(\Psi)}{\mathcal{S}_{\alpha_0}(\Psi)}y^{\epsilon-1/2})^{\frac{2}{1+2\theta}}$ to get that

$$\sum_{m\neq 0}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}\ll\mathcal{S}_{\alpha_1}(\Psi)y^{\frac{1-2\epsilon}{2(1+2\theta)}}+\mathcal{S}_{\alpha_1+1}(\Psi)y^{\frac{1-2\epsilon}{2(1+2\theta)}+1/4},$$

with the norms $\mathcal{S}_{\alpha_1}(\Psi)=\mathcal{S}_1(\Psi)^{\frac{2\theta}{1+2\theta}}\mathcal{S}_{\alpha_0}(\Psi)^{\frac{1}{1+2\theta}}$ and $\mathcal{S}_{\alpha_1+1}(\Psi)=\frac{\mathcal{S}_{\alpha_1}(\Psi)\mathcal{S}_2(\Psi)}{\mathcal{S}_1(\Psi)}$.

Otherwise take $M=1$ so that we only have the second term, which is bounded by

$$\sum_{|m|\geq 1}\frac{|a_{\Psi,\mathfrak{a}}(m,y)|}{|m|}\ll\mathcal{S}_1(\Psi)+\mathcal{S}_2(\Psi)y^{1/4}.$$

But since $\mathcal{S}_1(\Psi) \leq \mathcal{S}_{\alpha_0}(\Psi) y^{1/2-\epsilon}$, we can bound

$$\mathcal{S}_1(\Psi) + \mathcal{S}_2(\Psi) y^{1/4} \leq \mathcal{S}_{\alpha_1}(\Psi) y^{\eta_1} + \mathcal{S}_{\alpha_1+1}(\Psi) y^{\eta_1+1/4},$$

also in this case.

Finally, for each exceptional form with $r_k = i\sigma \in i(0,\theta]$ we can use (3.9) to bound

$$\sum_{m\neq 0} \frac{|a_{\varphi_k,\mathfrak{a}}(m,y)|}{|m|} \ll_\epsilon \sum_{m=1}^{\infty} m^{\theta+\epsilon-\sigma-1} y^{1/2-\sigma} e^{-2\pi m y} \ll y^{1/2-\theta-\epsilon}. \quad \square$$

4. Equidistribution of shears

We now use our results on the Fourier coefficients from the previous section to improve the error term in the equidistribution result of [KK18, Theorem 1.1]. In addition to improving the error term we also take care to make the dependance of the error on width of the cusp explicit, as this will be needed for our application. We will show the following

Theorem 4.1. *Let Γ be a conjugate of $\Gamma_0(p)$ with a cusp at ∞ of width $\omega \geq 1$. For any $\frac{7+12\theta}{3+6\theta} < \alpha < 3$ and any $\frac{1}{4} < \eta_1 < \frac{1}{2+4\theta}$ and $\frac{1}{4} < \eta_2 < \frac{1}{2} - \theta$, there are norms $\mathcal{S}_0, \mathcal{S}_\alpha, \mathcal{S}_{\alpha_2}$ of degrees $0, \alpha$ and $\alpha_2 = 1 + \alpha + \frac{3-\alpha}{4\eta_1}$ respectively, so that for any $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$ and any $T \geq \omega^{1/\eta_1}$*

$$\int_{\frac{1}{\sqrt{1+T^2}}}^{\infty} \Psi(yT+iy)\frac{dy}{y} = \mu_\Gamma(\Psi)\log(T\omega) + \langle \mathcal{K}_{\Gamma,\infty}, \Psi\rangle$$
$$+O_{\alpha,\eta_1,\eta_2}\left(\frac{\mathcal{S}_\alpha(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{1}{4}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}} + \frac{\mathcal{S}_0(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_2}{2}}}\right).$$

As in [KK18], the proof of Theorem 4.1 splits into two parts. The first gives equidistribution in the strip (compare to [KK18, Theorem 3.2]) which in our setting is

Proposition 4.2. *Under the same conditions and notations as in Theorem 4.1 we have*

$$\int_{\frac{1}{\sqrt{1+T^2}}}^{\infty} \Psi(yT+iy)\frac{dy}{y} = \frac{1}{\omega}\int_0^{\omega}\int_{1/T}^{\infty} \Psi(x+iy)\frac{dydx}{y}$$
$$+O_{\alpha,\eta_1}\left(\frac{\mathcal{S}_\alpha(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{1}{4}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}} + \frac{\mathcal{S}_0(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_2}{2}}}\right).$$

The second step uses the theory of Eisenstein series to estimate the strip average by the Eisenstein distribution (see [KK18, Theorem 3.3]).

Proposition 4.3. *Under the same conditions and notations as in Theorem 4.1 we have,*

$$\frac{1}{\omega}\int_0^{\omega}\int_{1/T}^{\infty}\Psi(x+iy)\frac{dydx}{y}=\mu_\Gamma(\Psi)\log(T\omega)+\langle\mathcal{K}_{\Gamma,\infty},\Psi\rangle+O\left(\frac{\mathcal{S}_{2,0}(\Psi)}{\sqrt{\omega T}}\right).$$

The proof of Proposition 4.3 follows exactly as that of [KK18, Theorem 3.3] without any changes. The only difference is the absence of residual spectrum in this case, and the fact that we are considering K-invariant test functions, allowing us to use the Sobolev norm $\mathcal{S}_{2,0}(\Psi)=\|\Psi\|_2$ instead of $\mathcal{S}_{2,1}(\Psi)$. Theorem 4.1 follows from these two propositions after noting that the error term $O(\frac{\mathcal{S}_{2,0}(\Psi)}{\sqrt{\omega T}})$ is subsumed by the other terms. We thus devote the rest of this section to the proof of Proposition 4.2, taking advantage of the assumption that Γ is a conjugate of the congruence group $\Gamma_0(p)$.

4.1. Equidistribution in the strip

In order to prove Proposition 4.2 write $\Psi(z)=a_{\Psi,\infty}(0,\frac{y}{\omega})+\Psi^\perp(z)$. The main term will come from the constant term, and we will bound the remaining integrals of $\Psi^\perp$ using the Fourier expansion. To do this we prove the following lemma.

Lemma 4.4. *Let Γ be conjugate to $\Gamma_0(p)$ with a cusp at ∞ of width ω. For any $\alpha>\frac{7+12\theta}{3+6\theta}$, $\eta_1<\frac{1}{2+4\theta}$ and $\beta\in(0,\frac{1-\eta_1}{3})$, there are norms $\mathcal{S}_\alpha,\mathcal{S}_{\alpha+1}$ of orders α and $\alpha+1$ respectively, so that for any $T\geq 1$ and $\Psi\in C_c^\infty(\Gamma\backslash\mathbb{H})$ orthogonal to all exceptional forms, with*

$$K_0=K_0(\Psi,T)=\left(\tfrac{\mathcal{S}_{\infty,1}(\Psi)^2}{\mathcal{S}_\alpha(\Psi)^2}\omega^{-2\beta}T^{\eta_1}\right)^{\frac{1}{\eta_1+1}}\geq 1,$$

for any $C\geq 1$, and any $1\leq k<CK_0$ we have

$$\left|\int_k^{k+1}\Psi^\perp(y+i\frac{y}{T})\frac{dy}{y}\right|\ll C^{\frac{1+\eta_1}{2}}\left(\mathcal{S}_\alpha(\Psi)\frac{\omega^{1-\eta_1-\beta}}{k^{\frac{3-\eta_1}{2}}T^{\frac{\eta_1}{2}}}+\mathcal{S}_{\alpha+1}(\Psi)\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{k^{\frac{5}{4}-\frac{\eta_1}{2}}T^{\frac{1}{4}+\frac{\eta_1}{2}}}\right).$$

Proof. Fix a large parameter $N\in\mathbb{N}$ to be determined later and write

$$\int_k^{k+1}\Psi^\perp(y+i\frac{y}{T})\frac{dy}{y}=\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\Psi^\perp(y+i\frac{y}{T})\frac{dy}{y},$$

with $t_j=\frac{Nk+j}{N}$. For $t_j\leq y\leq t_{j+1}$ we can approximate

$$\Psi^\perp(y+i\frac{y}{T})=\Psi^\perp(y+i\frac{t_j}{T})+O\left(\frac{\mathcal{S}_{\infty,1}(\Psi)}{Nk+j}\right),$$

where we used that, for the hyperbolic distance,

$$d(y+i\frac{y}{T},y+i\frac{t_j}{T}) = |\log(y/t_j)| \le \log(t_{j+1}/t_j) \le (Nk+j)^{-1}.$$

Remark 4.5. It is here that we correct the error in the proof of [KK18, Lemma 3.6].

Plugging this in gives

$$\int_{t_j}^{t_{j+1}} \Psi^\perp(y+i\frac{y}{T})\frac{dy}{y} = \int_{t_j}^{t_{j+1}} \Psi^\perp(y+i\frac{t_j}{T})\frac{dy}{y} + O\left(\frac{\mathcal{S}_{\infty,1}(\Psi)}{(Nk+j)^2}\right).$$

Now for the first term, expanding $\Psi^\perp$ into its Fourier series, using integration by parts to bound $|\int_{t_j}^{t_{j+1}} e^{2\pi i \frac{my}{\omega}}\frac{dy}{y}| \ll \frac{\omega}{mk}$, and using Proposition 3.19 with η_1 and $\alpha_1 = 2\alpha - 3$ we can bound

$$\begin{aligned}
\left|\int_{t_j}^{t_{j+1}} \Psi^\perp(y+i\frac{t_j}{T})\frac{dy}{y}\right| &= \left|\sum_{m\neq 0} a_{\Psi,\infty}(m,\frac{t_j}{\omega T}) \int_{t_j}^{t_{j+1}} e^{2\pi i\frac{my}{\omega}}\frac{dy}{y}\right| \\
&\ll \frac{\omega}{k}\sum_{m\neq 0}\frac{|a_{\Psi,\infty}(m,\frac{t_j}{\omega T})|}{|m|} \\
&\ll \frac{\omega}{k}(\mathcal{S}_{\alpha_1}(\Psi)(\tfrac{k}{\omega T})^{\eta_1} + \mathcal{S}_{\alpha_1+1}(\Psi)(\tfrac{k}{\omega T})^{\eta_1+1/4} \\
&= \frac{\omega^{1-\eta_1}}{k^{1-\eta_1}}\frac{\mathcal{S}_{\alpha_1}(\Psi)}{T^{\eta_1}} + \frac{\omega^{3/4-\eta_1}}{k^{3/4-\eta_1}}\frac{\mathcal{S}_{\alpha_1+1}(\Psi)}{T^{\eta_1+1/4}}.
\end{aligned}$$

Summing over $0 \le j < N$ we get that

$$\left|\int_k^{k+1} \Psi^\perp(y+i\frac{y}{T})\frac{dy}{y}\right| \ll N\frac{\omega^{1-\eta_1}}{k^{1-\eta_1}}\frac{\mathcal{S}_{\alpha_1}(\Psi)}{T^{\eta_1}} + N\frac{\omega^{3/4-\eta_1}}{k^{3/4-\eta_1}}\frac{\mathcal{S}_{\alpha_1+1}(\Psi)}{T^{\eta_1+1/4}} + \frac{\mathcal{S}_{\infty,1}(\Psi)}{Nk^2}.$$

Now define the norms

$$\mathcal{S}_\alpha(\Psi) = \sqrt{\mathcal{S}_{\alpha_1}(\Psi)\mathcal{S}_{\infty,1}(\Psi)}, \quad \mathcal{S}_{\alpha+1}(\Psi) = \sqrt{\frac{\mathcal{S}_{\infty,1}(\Psi)\mathcal{S}_{\alpha_1+1}(\Psi)^2}{\mathcal{S}_{\alpha_1}(\Psi)}},$$

(that are indeed of degrees α and $\alpha+1$). First, for $k \le K_0$ we can take

$$N = N_0(k) = [\sqrt{\tfrac{\mathcal{S}_{\infty,1}(\Psi)}{\mathcal{S}_{\alpha_1}(\Psi)}}\tfrac{T^{\frac{\eta_1}{2}}}{\omega^\beta k^{\frac{1+\eta_1}{2}}}] \ge 1,$$

to get that

$$\left|\int_k^{k+1} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}\right| \ll \mathcal{S}_\alpha(\Psi)\frac{\omega^{1-\eta_1-\beta}}{k^{\frac{3}{2}-\frac{\eta_1}{2}}T^{\frac{\eta_1}{2}}} + \mathcal{S}_{\alpha+1}(\Psi)\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{k^{\frac{5}{4}-\frac{\eta_1}{2}}T^{\frac{1}{4}+\frac{\eta_1}{2}}}.$$

Next for $K_0 \leq k < CK_0$ we take instead $N = [C^{\frac{1+\eta_1}{2}}N_0] \geq 1$ to get that

$$\left|\int_k^{k+1} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}\right| \ll C^{\frac{1+\eta_1}{2}}\mathcal{S}_\alpha(\Psi)\frac{\omega^{1-\eta_1-\beta}}{k^{\frac{3}{2}-\frac{\eta_1}{2}}T^{\frac{\eta_1}{2}}} + C^{\frac{1+\eta_1}{2}}\mathcal{S}_{\alpha+1}(\Psi)\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{k^{\frac{5}{4}-\frac{\eta_1}{2}}T^{\frac{1}{4}+\frac{\eta_1}{2}}}. \qquad \square$$

Using this estimate it is possible to evaluate $\int_1^{K_0} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}$. We now show how to boot strap this to extend the range all the way up to $K_1 = K_0^{\frac{1+\eta_1}{2\eta_1}}$.

Proposition 4.6. *With the same assumptions and notations as in Lemma 4.4, let* $K_1 = K_1(\Psi,T) = K_0(\Psi,T)^{\frac{1+\eta_1}{2\eta_1}}$. *Then for* $K_0 \geq 1$, *there is a constant* $c \in (1/e^2, 1)$ *such that*

$$\left|\int_1^{cK_1} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}\right| \ll \max\{1, \log\log(K_0)\}\left(\frac{\mathcal{S}_\alpha(\Psi)\omega^{1-\eta_1-\beta}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{3}{4}-\eta_1-\beta-\frac{\beta}{4\eta_1}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}}\right),$$

where $\mathcal{S}_\alpha$ *is as above and* $\mathcal{S}_{\alpha_2}$ *is a norm of degree* $\alpha_2 = \frac{4\eta_1+3}{4\eta_1} + \alpha\frac{4\eta_1-1}{4\eta_1}$.

Proof. As a first step, a simple application of Lemma 4.4 with $C = 1$, noting that for both terms the power of k in the denominator is greater than one gives

$$\begin{aligned}\left|\int_1^{K_0} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}\right| &\leq \sum_{k=1}^{K_0-1}\left|\int_k^{k+1} \Psi(y+i\frac{y}{T})\frac{dy}{y}\right| \\ &\ll \mathcal{S}_\alpha(\Psi)\frac{\omega^{1-\eta_1-\beta}}{T^{\frac{\eta_1}{2}}} + \mathcal{S}_{\alpha+1}(\Psi)\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{T^{\frac{1}{4}+\frac{\eta_1}{2}}}.\end{aligned}$$

Next let s_ℓ denote the partial sums of geometric series $s_\ell = \sum_{j=0}^{\ell}(\frac{1-\eta_1}{1+\eta_1})^j$ converging to $s_\infty = \frac{1+\eta_1}{2\eta_1}$ and let $C_\ell = K_0^{s_\ell - 1}$. Applying Lemma 4.4 with $C = C_\ell$ we get the bound

$$\begin{aligned}\left|\int_{C_{\ell-1}K_0}^{C_\ell K_0} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}\right| &\leq \sum_{k=C_{\ell-1}K_0}^{C_\ell K_0 - 1}\left|\int_k^{k+1} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}\right| \\ &\ll \frac{C_\ell^{\frac{1+\eta_1}{2}}\mathcal{S}_\alpha(\Psi)}{(C_{\ell-1}K_0)^{\frac{1-\eta_1}{2}}}\frac{\omega^{1-\eta_1-\beta}}{T^{\frac{\eta_1}{2}}} + \frac{C_\ell^{\frac{1+\eta_1}{2}}\mathcal{S}_{\alpha+1}(\Psi)}{(C_{\ell-1}K_0)^{\frac{1-2\eta_1}{4}}}\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{T^{\frac{1}{4}+\frac{\eta_1}{2}}}.\end{aligned}$$

From our choice of the constants we have that $C_\ell = (K_0 C_{\ell-1})^{\frac{1-\eta_1}{1+\eta_1}}$ so that $\frac{C_\ell^{1+\eta_1}}{(C_{\ell-1}K_0)^{1-\eta_1}} = 1$ and $\frac{C_\ell^{1+\eta_1}}{(C_{\ell-1}K_0)^{\frac{1-2\eta_1}{2}}} = (C_{\ell-1}K_0)^{1/2}$, hence

$$\left|\int_{C_{\ell-1}K_0}^{C_\ell K_0} \Psi(y+i\frac{y}{T})\frac{dy}{y}\right| \ll \mathcal{S}_\alpha(\Psi)\frac{\omega^{1-\eta_1-\beta}}{T^{\frac{\eta_1}{2}}} + (C_{\ell-1}K_0)^{1/4}\mathcal{S}_{\alpha+1}(\Psi)\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{T^{\frac{1}{4}+\frac{\eta_1}{2}}}.$$

Now bounding $C_{\ell-1}K_0 \leq K_0^{\frac{1+\eta_1}{2\eta_1}} = K_1$ and plugging in $K_0 = \left(\frac{\mathcal{S}_{\infty,1}(\Psi)^2}{\mathcal{S}_\alpha(\Psi)^2}\omega^{-2\beta}T^{\eta_1}\right)^{\frac{1}{\eta_1+1}}$ we can bound the second term by

$$K_0^{\frac{1+\eta_1}{8\eta_1}}\mathcal{S}_{\alpha+1}(\Psi)\frac{\omega^{\frac{3}{4}-\eta_1-\beta}}{T^{\frac{1}{4}+\frac{\eta_1}{2}}} = \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{3}{4}-\eta_1-\beta-\frac{\beta}{4\eta_1}}}{T^{\frac{1}{8}+\frac{\eta_1}{2}}},$$

with

$$\mathcal{S}_{\alpha_2}(\Psi) = \mathcal{S}_{\alpha+1}(\Psi)\mathcal{S}_{\infty,1}(\Psi)^{\frac{1}{4\eta_1}}\mathcal{S}_\alpha(\Psi)^{\frac{-1}{4\eta_1}}.$$

We thus get that for each $\ell \geq 1$ we have

$$\left|\int_{C_{\ell-1}K_0}^{C_\ell K_0} \Psi(y+i\frac{y}{T})\frac{dy}{y}\right| \ll \mathcal{S}_\alpha(\Psi)\frac{\omega^{1-\eta_1-\beta}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{3}{4}-\eta_1-\beta-\frac{\beta}{4\eta_1}}}{T^{\frac{1}{8}+\frac{\eta_1}{2}}}.$$

Finally, taking $\ell = \max\{1, [\frac{\log\log(K_0)}{\log(\frac{1+\eta_1}{1-\eta_1})}]+1\}$ we get that

$$s_\ell = \tfrac{1+\eta_1}{2\eta_1} - \tfrac{1-\eta_1}{2\eta_1}(\tfrac{1-\eta_1}{1+\eta_1})^\ell \geq \tfrac{1+\eta_1}{2\eta_1} - 2(\tfrac{1-\eta_1}{1+\eta_1})^\ell \geq \tfrac{1+\eta_1}{2\eta_1} - \tfrac{2}{\log(K_0)},$$

so that $e^{-2}K_1 \leq C_\ell K_0 \leq K_1$, hence, $C_\ell K_0 = cK_1$ for some $c > 1/e^2$. Summing up these $\ell = O(\max\{1, \log\log(K_0)\})$ terms concludes the proof. □

For large values of $k > K_1$ this estimate is no longer optimal, and instead we will use the following alternative bound.

Lemma 4.7. *Let Γ be conjugate of $\Gamma_0(p)$ with a cusp at ∞ of width ω. For any $\alpha_1 > \frac{5+6\theta}{3+6\theta}$ and $\eta_1 < \frac{1}{2+4\theta}$ there are norms $\mathcal{S}_{\alpha_1}, \mathcal{S}_{\alpha_1+1}$ such that for any $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$ orthogonal to all exceptional forms, for any $k \geq 1$ we have*

$$\left|\int_{\omega k}^{\omega(k+1)} \Psi^\perp(y+i\frac{y}{T})\frac{dy}{y}\right| \ll \frac{\mathcal{S}_{\alpha_1}(\Psi)}{k^{2-\eta_1}T^{\eta_1}} + \frac{\mathcal{S}_{\alpha_1+1}(\Psi)}{k^{7/4-\eta_1}T^{\eta_1+1/4}} + \frac{\mathcal{S}_{\infty,1}(\Psi)}{k^2}.$$

Proof. For $k\omega \leq y \leq (k+1)\omega$ we can estimate

$$\Psi(y+i\frac{y}{T}) = \Psi(y+i\frac{k\omega}{T}) + O\left(\frac{\mathcal{S}_{\infty,1}(\Psi)}{k}\right),$$

so that

$$\int_{\omega k}^{\omega(k+1)} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y} = \int_{\omega k}^{\omega(k+1)} \Psi^{\perp}(y+i\frac{\omega k}{T})\frac{dy}{y} + O\left(\frac{\mathcal{S}_{\infty,1}(\Psi)}{k^2}\right).$$

Expanding the first term in Fourier series

$$\int_{\omega k}^{\omega(k+1)} \Psi^{\perp}(y+i\frac{\omega k}{T})\frac{dy}{y} = \sum_{m\neq 0} a_{\Psi,\infty}(m,\frac{k}{T}) \int_{\omega k}^{\omega(k+1)} e^{\frac{2\pi i m y}{\omega}}\frac{dy}{y}.$$

Integrating by parts we can bound

$$\left|\int_{\omega k}^{\omega(k+1)} e^{\frac{2\pi i m y}{\omega}}\frac{dy}{y}\right| = \left|\int_{k}^{k+1} e^{2\pi i m y}\frac{dy}{y}\right| \leq \frac{1}{|m|k^2},$$

and using Proposition 3.19 we bound

$$\begin{aligned}
\left|\int_{\omega k}^{\omega(k+1)} \Psi^{\perp}(y+i\frac{\omega k}{T})\frac{dy}{y}\right| &\leq \sum_{m\neq 0} |a_{\Psi,\infty}(m,\frac{k}{T})| \left|\int_{\omega k}^{\omega(k+1)} e^{\frac{2\pi i m y}{\omega}}\frac{dy}{y}\right| \\
&\leq \frac{1}{k^2}\sum_{m\neq 0}\frac{|a_{\Psi,\infty}(m,\frac{k}{T})|}{|m|} \\
&\ll \frac{\mathcal{S}_{\alpha_1}(\Psi)}{k^{2-\eta_1}T^{\eta_1}} + \frac{\mathcal{S}_{\alpha_1+1}(\Psi)}{k^{7/4-\eta_1}T^{\eta_1+1/4}},
\end{aligned}$$

concluding the proof. □

We can now give the

Proof of Proposition 4.2. First assume that Ψ is orthogonal to all exceptional forms (if any such exist). Noting that

$$\int_{\frac{T}{\sqrt{T^2+1}}}^{\infty} \Psi(y+i\tfrac{y}{T})\frac{dy}{y} = \int_{1}^{\infty} \Psi(y+i\frac{y}{T})\frac{dy}{y} + O\left(\frac{\mathcal{S}_{\infty,0}(\Psi)}{T^2}\right),$$

and that

$$\int_{1}^{\infty} a_{\Psi,\infty}(0,\frac{y}{\omega T})\frac{dy}{y} = \frac{1}{\omega}\int_{0}^{\omega}\int_{1/T}^{\infty} \Psi(x+iy)\frac{dydx}{y},$$

we just need to bound $\int_1^{\infty} \Psi^{\perp}(y+i\frac{y}{T})\frac{dy}{y}$.

Let $\beta = \frac{1}{2} - \eta_1$ and let $K_0 = K_0(\Psi, T) = \left(\frac{\mathcal{S}_{\infty,1}(\Psi)^2}{\mathcal{S}_\alpha(\Psi)^2}\omega^{-2\beta}T^{\eta_1}\right)^{\frac{1}{\eta_1+1}}$ be as in Lemma 4.4 and $K_1 = K_0^{\frac{1+\eta_1}{2\eta_1}} = \left(\frac{\mathcal{S}_{\infty,1}(\Psi)}{\mathcal{S}_\alpha(\Psi)}\right)^{\frac{1}{\eta_1}}\omega^{-\frac{\beta}{\eta_1}}T^{\frac{1}{2}}$ be as in Proposition 4.6. Note that the condition that $T \geq \omega^{1/\eta_1}$ implies that $\omega^{-2\beta}T^{\eta_1} \geq \omega^{2\eta_1} \geq 1$ and since $\mathcal{S}_{\infty,1}$ is of degree 3, and $\alpha \leq 3$ we also have that $\frac{\mathcal{S}_{\infty,1}(\Psi)}{\mathcal{S}_\alpha(\Psi)} \geq 1$ so that indeed $K_0 \geq 1$. Now by Proposition 4.6

$$\left|\int_1^{cK_1} \Psi^\perp(y + i\frac{y}{T})\frac{dy}{y}\right| \ll \max\{1, \log\log(K_0)\}\left(\frac{\mathcal{S}_\alpha(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{1}{4}-\frac{1-2\eta_1}{8\eta_1}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}}\right).$$

Next, using Lemma 4.7 we can bound

$$\begin{aligned}\left|\int_{cK_1}^{\infty} \Psi^\perp(y + i\frac{y}{T})\frac{dy}{y}\right| &\leq \sum_{k > cK_1/\omega}\left|\int_{\omega k}^{\omega(k+1)} \Psi^\perp(y + i\frac{y}{T})\frac{dy}{y}\right| \\ &\ll \frac{\mathcal{S}_{\alpha_1}(\Psi)}{T^{\eta_1}} + \frac{\mathcal{S}_{\alpha_1+1}(\Psi)}{T^{\eta_1+1/4}} + \frac{\omega\mathcal{S}_{\infty,1}(\Psi)}{K_1}.\end{aligned}$$

The first two terms are clearly bounded by the similar terms appearing above and plugging in the value of K_1, the third term is

$$\frac{\omega\mathcal{S}_{\infty,1}(\Psi)}{K_1} = \mathcal{S}_{\infty,1}(\Psi)^{1-\frac{1}{\eta_1}}\mathcal{S}_\alpha(\Psi)^{\frac{1}{\eta_1}}\left(\frac{\omega^{\eta_1+\beta}}{T^{\frac{\eta_1}{2}}}\right)^{\frac{1}{\eta_1}}.$$

Notice that $\mathcal{S}_{\infty,1}(\Psi)^{1-\frac{1}{\eta_1}}\mathcal{S}_\alpha(\Psi)^{\frac{1}{\eta_1}}$ is a norm of degree $\frac{\alpha}{\eta_1} - 3(\frac{1}{\eta_1} - 1)$ (which is smaller than α as long as $\alpha < 3$), and with our choice of $\beta = \frac{1}{2} - \eta_1$ we see that

$$\left(\frac{\omega^{\eta_1+\beta}}{T^{\frac{\eta_1}{2}}}\right)^{\frac{1}{\eta_1}} = \left(\frac{\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}}\right)^{\frac{1}{\eta_1}} \leq \frac{\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}}.$$

Hence, after perhaps replacing $\mathcal{S}_\alpha$ by a different norm of degree α we get that

$$\frac{\omega\mathcal{S}_{\infty,1}(\Psi)}{K_1} \leq \frac{\mathcal{S}_\alpha(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}},$$

whence

$$\left|\int_1^{\infty} \Psi^\perp(y + i\frac{y}{T})\frac{dy}{y}\right| \ll \log\log(K_0)\left(\frac{\mathcal{S}_\alpha(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{1}{4}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}}\right).$$

Finally, since K_0 grows polynomially with T and the norms of Ψ, replacing the exponent η_1 with a slightly smaller exponent and, if needed, slightly increasing the degrees α and α_2, we may drop the $\log\log(K_0)$ term.

Next, to deal with exceptional forms. Repeating the same arguments for $\Psi = \varphi_k$ an exceptional cusp form, and using (3.21) instead of (3.20), we get that for any $\eta_2 < \frac{1}{2} - \theta$

$$\left| \int_1^\infty \varphi_k(y + i\frac{y}{T}) \frac{dy}{y} \right| \ll_{\eta_2} \frac{\omega^{\frac{1}{2}}}{T^{\frac{\eta_2}{2}}}.$$

We should note that, although the implied constant will depend on φ_k, or more precisely on some norms $\mathcal{S}(\varphi_k)$, we may bound any such contributions by $O_p(1)$, since for any fixed level p there are at most finitely many exceptional forms (and p is treated as fixed).

Now for the general case write $\Psi = \Psi_{\mathrm{ex}} + \Psi_{\mathrm{pr}}$ with

$$\Psi_{\mathrm{ex}} = \sum_{r_k \in i(0,1/2)} \langle \Psi, \varphi_k \rangle \varphi_k,$$

the projection of Ψ to the space spanned by the exceptional forms. We can thus bound

$$\left| \int_1^\infty \Psi(y + i\frac{y}{T}) \frac{dy}{y} \right| \le \left| \int_1^\infty \Psi_{\mathrm{ex}}(y + i\frac{y}{T}) \frac{dy}{y} \right| + \left| \int_1^\infty \Psi_{\mathrm{pr}}(y + i\frac{y}{T}) \frac{dy}{y} \right|.$$

Estimating $|\langle \Psi, \varphi_k \rangle \le \|\Psi\|_1 \|\varphi_k\|_\infty \ll \|\Psi\|_1$ and noting that $\mathcal{S}_0(\Psi) = \|\Psi\|_1$ is a norm of degree 0 (and that there are at most finitely many exceptional forms), we have

$$\left| \int_1^\infty \Psi_{\mathrm{ex}}(y + i\frac{y}{T}) \frac{dy}{y} \right| \ll_{p,\eta_2} \frac{\mathcal{S}_0(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_2}{2}}}.$$

Since Ψ_{pr} is orthogonal to all exceptional forms we can apply the first part to bound

$$\left| \int_1^\infty \Psi_{\mathrm{pr}}^{\perp}(y + i\frac{y}{T}) \frac{dy}{y} \right| \ll \left(\frac{\mathcal{S}_\alpha(\Psi_{\mathrm{pr}})\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi_{\mathrm{pr}})\omega^{\frac{1}{4}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}} \right).$$

To complete the proof, observe that for any of our norms, we have $\mathcal{S}(\Psi_{\mathrm{ex}}) \ll \mathcal{S}_0(\Psi)$, and also that $\mathcal{S}_\alpha(\Psi_{\mathrm{pr}}) \asymp \mathcal{S}_\alpha(\Psi)$ holds for any $\alpha > 0$. □

4.2. Two sided cuspidal geodesics

For the cases of interest here, the lattice Γ has cusps at ∞ and at 0 and it is natural to consider shears of the two sided cuspidal geodesic connecting them, that is,

$$\mu_T(\Psi) = \int_0^\infty \Psi(Ty + iy) \frac{dy}{y}.$$

It is easy to see that μ_T is invariant under scaling, $\mu_T(\Psi) = \mu_T(\Psi^{a_t})$ and a simple computation shows that under $\sigma = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ it transforms via $\mu_T(\Psi^\sigma) = \mu_{-T}(\Psi)$. In fact, using that $\Psi^\sigma(z) = \Psi(\frac{-1}{z})$ and making a change of variables gives the identity.

$$\mu_T(\Psi) = \int_{\frac{1}{\sqrt{T^2+1}}}^{\infty} \Psi(Ty+iy)\frac{dy}{y} + \int_{\frac{1}{\sqrt{T^2+1}}}^{\infty} \Psi^\sigma(-Ty+iy)\frac{dy}{y}. \tag{4.8}$$

Applying our results on shears of cuspidal geodesic rays we get the following.

Corollary 4.9. *Let Γ be conjugate to $\Gamma_0(p)$ with cusps at ∞ and 0. Let ω_1, ω_2 denote the widths of the cusps of Γ and Γ^σ at ∞ respectively and let $\omega = \sqrt{\omega_1\omega_2}$. For any $\alpha \in (\frac{7+12\theta}{3+6\theta}, 3)$, $\eta_1 \in (\frac{1}{4}, \frac{1-2\theta}{2})$, $\eta_2 \in (\frac{1}{4}, \frac{1-2\theta}{2})$ and $T \geq \omega^{1/\eta_1}$ we have for any $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$,*

$$\begin{aligned}\mu_T(\Psi) = 2\mu_\Gamma(\Psi)\log(T\omega) + \langle \mathcal{K}_{\Gamma,\infty}, \Psi\rangle + \langle \mathcal{K}_{\Gamma,0}, \Psi\rangle \\ + O_{\alpha,\eta_1,\eta_2}\left(\frac{\mathcal{S}_\alpha(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)\omega^{\frac{1}{4}}}{T^{\frac{\eta_1}{2}+\frac{1}{8}}} + \frac{\mathcal{S}_0(\Psi)\omega^{\frac{1}{2}}}{T^{\frac{\eta_2}{2}}}\right),\end{aligned}$$

where $\mathcal{S}_\alpha, \mathcal{S}_{\alpha_2}$ are as in Theorem 4.1.

Proof. Note that if $\Psi(z)$ has period ω_1 and Ψ^σ has period ω_2 then $\tilde{\Psi}(z) = \Psi(\sqrt{\frac{\omega_1}{\omega_2}}z)$ satisfies that both $\tilde{\Psi}$ and $\tilde{\Psi}^\sigma$ have a period of $\omega = \sqrt{\omega_1\omega_2}$. Since $\mu_T(\Psi) = \mu_T(\tilde{\Psi})$, using Theorem 4.1 with $\tilde{\Psi}$ and $\tilde{\Psi}^\sigma$ in each part of (4.8) gives the claimed result. □

5. Lattice points in cones

We now apply our results on equidistribution of shears to get effective counting estimates for counting lattice points in cones of the form

$$\mathcal{C}_T = \{z \in \mathbb{H} : |\mathfrak{Re}(z)| \leq T\mathfrak{Im}(z)\}, \qquad T \to \infty. \tag{5.1}$$

For $\Gamma \leq \mathrm{PSL}_2(\mathbb{Z})$ conjugate to $\Gamma_0(p)$ and $\tau \in \mathrm{PSL}_2(\mathbb{Q})$, we define the counting function

$$\mathcal{N}^\tau_{\mathcal{C}_T}(\Gamma) = \#\{\gamma \in \Gamma : \tau^{-1}\gamma i \in \mathcal{C}_T\}. \tag{5.2}$$

Theorem 5.3. *For any $\eta < \frac{3}{40+72\theta}$ and $\beta_1 > 1+2\theta$, we have for $T > (\omega_\tau\omega_{\tau\sigma})^{\beta_1}$ that*

$$\mathcal{N}^\tau_{\mathcal{C}_T}(\Gamma) = \frac{2T}{v_\Gamma}\left(\log(T^2\omega_\tau\omega_{\tau\sigma}) - 2 + v_\Gamma(\mathcal{K}_{\Gamma,\mathfrak{a}_\tau}(i) + \mathcal{K}_{\Gamma,\mathfrak{b}_\tau}(i)) + O\big(\frac{(\omega_\tau\omega_{\tau\sigma})^{\beta_1\eta}}{T^\eta}\big)\right).$$

Here ω_τ denotes the width of the cusp at ∞ of $\Gamma^\tau = \tau^{-1}\Gamma\tau$, $\mathfrak{a}_\tau = \tau\infty$, and $\mathfrak{b}_\tau = \tau 0$.

As a first step we show that the cones $\mathcal{C}_T$ are well rounded.

Lemma 5.4. *Let $g \in G$ such that $gi \in \mathcal{C}_T$ with $T \geq 1$. Then for any $h \in B_\delta$ with $0 < \delta \leq \frac{1}{10}$, we have $ghi \in \mathcal{C}_{T(1+5\delta)}$.*

Proof. For any $h \in B_\delta$ we have that $|\mathfrak{Re}(hi)| \leq \sinh(\delta) \leq 2\delta$ and $|\mathfrak{Im}(hi) - 1| \leq 2\delta$. Write $gi = x + iy$ so that $g = n_x a_y k_\theta$ for some θ. Since B_δ is K-invariant, $k_\theta h = \tilde{h} \in B_\delta$ as well. Hence writing $\tilde{h}i = \xi + i\eta$, we have that $ghi = x + y(\xi + i\eta) = x' + iy'$ with $|\xi| \leq 2\delta$ and $|\eta - 1| \leq 2\delta$. We can thus write

$$\frac{|x'|}{y'} = \frac{|x + \xi y|}{\eta y} \leq \frac{|x|}{y}\left(\frac{1 + |\xi y/x|}{\eta}\right).$$

We now consider two cases, first assume that $\frac{|x|}{y} \geq 1$, in which case

$$\frac{|x'|}{y'} \leq T\left(\frac{1+2\delta}{1-2\delta}\right) \leq T(1+5\delta).$$

Next, when $|x| \leq y$ we bound

$$\frac{|x'|}{y'} = \frac{|x + \xi y|}{\eta y} \leq \frac{y(1+|\xi|)}{\eta y} \leq \frac{1+2\delta}{1-2\delta} \leq 1 + 5\delta. \quad \square$$

Now let $\chi_{\mathcal{C}_T}$ denote the indicator function of $\mathcal{C}_T$ and consider the function

$$F_{T,\tau}(g) = \sum_{\gamma \in \Gamma} \chi_{\mathcal{C}_T}(\tau^{-1}\gamma g.i) \tag{5.5}$$

Note that $F_{T,\tau} \in L^2(\Gamma\backslash G/K)$ and that evaluating at the identity we have $F_{T,\tau}(1) = \mathcal{N}^\tau_{\mathcal{C}_T}(\Gamma)$. From the well roundedness of $\mathcal{C}_T$ we get the following:

Lemma 5.6. *For $\delta > 0$ small, let $\psi_\delta \in C_c^\infty(K\backslash G/K)$ be supported on B_δ with $\int_G \psi_\delta = 1$, and let*

$$\Psi_\delta(g) = \sum_{\gamma \in \Gamma} \psi_\delta(\gamma g). \tag{5.7}$$

Then $\Psi_\delta \in L^2(\Gamma\backslash G/K)$ and

$$\langle F_{T(1-5\delta),\tau}, \Psi_\delta\rangle \leq \mathcal{N}^\tau_{\mathcal{C}_T}(\Gamma) \leq \langle F_{T(1+5\delta),\tau}, \Psi_\delta\rangle, \tag{5.8}$$

where the inner product the standard inner product in $L^2(\Gamma\backslash G/K)$.

Proof. Unfolding Ψ_δ we can write

$$\langle F_{T,\tau}, \Psi_\delta\rangle = \int_G F_{T,\tau}(g)\psi_\delta(g)dg.$$

By Lemma 5.4 we have that for any $g \in B_\delta$

$$F_{T(1-5\delta),\tau}(g) \leq \mathcal{N}^{\tau}_{\mathcal{C}_T}(\Gamma) \leq F_{T(1+5\delta),\tau}(g),$$

concluding the proof. $\square$

Remark 5.9. As is well known, this inner product $\langle F_{T,\tau}, \Psi_\delta\rangle$ not only is an approximation to the sharp cutoff $\mathcal{N}^{\tau}_{\mathcal{C}_T}(\Gamma)$, but it is also itself a smooth counting function, since

$$\langle F_{T,\tau}, \Psi_\delta\rangle = \sum_{\gamma\in\Gamma} \widetilde{\chi_{\mathcal{C}_T}}(\gamma),$$

where

$$\widetilde{\chi_{\mathcal{C}_T}}(\gamma) = \int_{g\in G} \chi_{\mathcal{C}_T}(\tau^{-1}\gamma g)\psi_\delta(g)dg,$$

is a smoothed cutoff, cf. (1.3).

We thus need to evaluate the inner product $\langle F_{T,\tau}, \Psi_\delta\rangle$. The following simple lemma, relates these to the shears of two sided cuspidal geodesics.

Lemma 5.10. *For any $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$ we have*

$$\langle F_{T,\tau}, \Psi\rangle = \int_{-T}^{T}\int_{0}^{\infty} \Psi^\tau(xy+iy)\tfrac{dy}{y}dx. \tag{5.11}$$

Proof. Unfolding $F_{T,\tau}$ and making some changes of variables gives

$$\begin{aligned}\langle F_{T,\tau}, \Psi\rangle &= \int_G \chi_{\mathcal{C}_T}(\tau^{-1}gi)\Psi(gi)dg\\ &= \int_G \chi_{\mathcal{C}_T}(gi)\Psi^\tau(gi)dg\\ &= \int_0^\infty\int_{-yT}^{yT} \Psi^\tau(x+iy)\tfrac{dxdy}{y^2}\end{aligned}$$

$$= \int_0^\infty \int_{-T}^{T} \Psi^\tau(xy+iy)\tfrac{dxdy}{y}$$

$$= \int_{-T}^{T} \int_0^\infty \Psi^\tau(xy+iy)\tfrac{dy}{y}dx$$

as claimed. □

We are in position now to prove Theorem 1.4, which follows immediately (see, e.g., [Kon09, (4.8)] and Remark 5.9) from the following:

Proposition 5.12. *Let $\Psi \in C_c^\infty(\Gamma\backslash\mathbb{H})$ be positive with mean one and supported the unit neighborhood of some point. Assume that $\alpha > \frac{7+12\theta}{3+6\theta}$, $\eta_1 \in (\frac{1}{4}, \frac{1}{2+4\theta})$, and $\eta_2 \in (\frac{1}{4}, \frac{1}{2}-\theta)$, and let α_2 be as in Theorem 4.1.*

Then

$$\begin{aligned}\langle F_{T,\tau}, \Psi\rangle = \frac{2T}{v_\Gamma}\Big(&\log(T^2\omega_\tau\omega_{\tau\sigma}) - 2 + v_\Gamma(\langle K_{\Gamma^\tau,\infty}, \Psi^\tau\rangle + \langle K_{\Gamma^\tau,0}, \Psi^\tau\rangle)\Big)\\ &+ O_{\alpha,\eta_1}\Big(\mathcal{S}_\alpha(\Psi)(\omega_\tau\omega_{\tau\sigma})^{\frac14}T^{1-\frac{\eta_1}{2}} + \mathcal{S}_{\alpha_2}(\Psi)(\omega_\tau\omega_{\tau\sigma})^{\frac18}T^{7/8-\frac{\eta_1}{2}}\Big)\\ &+ O_{\eta_1}\Big((\omega_\tau\omega_{\tau\sigma})^{1/(2\eta_1)}\log(\omega_\tau\omega_{\tau\sigma})\|\Psi\|_\infty\Big) + O_{\eta_2}\Big(\mathcal{S}_0(\Psi)(\omega_\tau\omega_{\tau\sigma})^{\frac14}T^{1-\frac{\eta_2}{2}}\Big).\end{aligned} \tag{5.13}$$

Proof. By Lemma 5.10, we have

$$\langle F_{T,\tau}, \Psi\rangle = \int_{-T}^{T}\int_0^\infty \Psi^\tau(xy+iy)\tfrac{dy}{y}dx. \tag{5.14}$$

Let

$$M = (\omega_\tau\omega_{\tau\sigma})^{1/(2\eta_1)}.$$

Then for $|x| > M$, we may apply Corollary 4.9 to the inner integral to obtain

$$\begin{aligned}\int_0^\infty \Psi^\tau(xy+iy)\tfrac{dy}{y} = \frac{\log(|x|^2\omega_\tau\omega_{\tau\sigma})}{v_\Gamma} + \langle \mathcal{K}_{\Gamma^\tau,\infty}, \Psi^\tau\rangle + \langle \mathcal{K}_{\Gamma^\tau,0}, \Psi^\tau\rangle\\ + O_{\alpha,\eta_1}\left(\frac{\mathcal{S}_\alpha(\Psi)(\omega_\tau\omega_{\tau\sigma})^{\frac14}}{|x|^{\frac{\eta_1}{2}}} + \frac{\mathcal{S}_{\alpha_2}(\Psi)(\omega_\tau\omega_{\tau\sigma})^{\frac18}}{|x|^{1/8+\frac{\eta_1}{2}}}\right)\end{aligned}$$

$$+O_{\eta_2}\left(\frac{\mathcal{S}_0(\Psi)(\omega_\tau\omega_{\tau\sigma})^{\frac{1}{4}}}{|x|^{\frac{\eta_2}{2}}}\right).$$

Integrating over $M \le |x| \le T$ gives the first three terms in (5.13) plus

$$O\Big(M\log(M\omega_\tau\omega_{\tau\sigma}) + M(|\langle K_{\Gamma^\tau,\infty},\Psi^\tau\rangle| + |\langle K_{\Gamma^\tau,0},\Psi^\tau\rangle|)\Big).$$

Since the Kronecker limit $K_{\Gamma^\tau,\mathfrak{a}} \in L^2(\Gamma^\tau\backslash\mathbb{H})$, this term is bounded by

$$O\Big(M\log(M\omega_\tau\omega_{\tau\sigma}) + M\|\Psi\|_2\Big),$$

which is subsumed by the last term in (5.13).

For $|x| < M$, we argue as follows. First fix x and apply (4.8) to the y integral:

$$\int_0^\infty \Psi^\tau(xy+iy)\tfrac{dy}{y} = \int_{\frac{1}{\sqrt{1+x^2}}}^\infty \Psi^\tau(xy+iy)\tfrac{dy}{y} + sim,$$

where "sim" refers to a similar term with τ replaced by $\tau\sigma$ and x by $-x$. Break the y integral into

$$\int_{\frac{1}{\sqrt{1+x^2}}}^\infty = \int_{1/\sqrt{x^2+1}}^{\omega_\tau} + \int_{\omega_\tau}^\infty .$$

Now, since Γ^τ has a cusp at infinity of width ω_τ the quotient $\Gamma^\tau\backslash\mathbb{H}$ has a fundamental domain contained in the set $\{x+iy : |x| \le \omega_\tau/2,\ y > 0\}$, and since the Siegel set $\{x+iy : |x| \le \omega_\tau/2,\ y \ge \omega_\tau\}$ has hyperbolic measure one, it is contained in this fundamental domain (recall that the fundamental domain for $\mathrm{PSL}_2(\mathbb{Z})$ and hence also for Γ^τ has hyperbolic measure > 1). On this one fundamental domain, Ψ^τ is supported on the unit neighborhood of some point $x_0 + iy_0$, which is contained in some strip of the form

$$\{x+iy : x\in\mathbb{R},\ y \in (\tfrac{y_0}{2}, 2y_0)\}.$$

Hence $\Psi^\tau(xy+iy) = 0$ for all $y \ge \omega_\tau$ outside of this strip implying that

$$\left|\int_{\omega_\tau}^\infty \Psi^\tau(xy+iy)\tfrac{dy}{y}\right| \ll \|\Psi\|_\infty \int_{y_0/2}^{2y_0} \frac{dy}{y} \ll \|\Psi\|_\infty.$$

For the first interval, we trivially bound

$$\left|\int_{1/\sqrt{x^2+1}}^{\omega_\tau} \Psi^\tau(xy+iy)\frac{dy}{y}\right| \ll \|\Psi\|_\infty \log M\omega_\tau.$$

Repeating the same argument with τ replaced by $\tau\sigma$ we see that for $|x| \leq M$

$$\left|\int_0^\infty \Psi^\tau(xy+iy)\frac{dy}{y}\right| \ll \|\Psi\|_\infty \log(\omega_\tau\omega_{\tau\sigma}),$$

and hence

$$\left|\int_{-M}^{M}\int_0^\infty \Psi^\tau(xy+iy)\frac{dy}{y}\right| \ll \|\Psi\|_\infty(\omega_\tau\omega_{\tau\sigma})^{1/(2\eta_1)}\log(\omega_\tau\omega_{\tau\sigma})$$

concluding the proof. □

We are finally in position to give the following

Proof of Theorem 5.3. Let α and η_1 be as in Proposition 5.12. Let $\eta = \frac{\eta_1}{2(\alpha+1)}$ and $\beta_1 = \frac{1}{2\eta_1}$ and denote by $\omega = \sqrt{\omega_\tau\omega_{\tau\sigma}}$. We can also choose $\eta < \eta_2 < 1/2-\theta$. Our δ-bump function Ψ_δ satisfies:

$$\mathcal{S}_\alpha(\Psi_\delta) \asymp \delta^{-\alpha}, \quad \mathcal{S}_{\alpha_2}(\Psi_\delta) \asymp \delta^{-\alpha_2}, \quad \|\Psi_\delta\|_\infty \asymp \delta^{-2}$$

where $\mathcal{S}_\alpha, \mathcal{S}_{\alpha_2}$ are as in Theorem 4.1.

Now we further simplify the Eisenstein terms appearing in Proposition 5.12. Since $E_{\Gamma^\tau,\infty}(z,s) = E^\tau_{\Gamma,\mathfrak{a}_\tau}(z,s)$ removing the residue we get that $\mathcal{K}_{\Gamma^\tau,\infty}(z) = \mathcal{K}^\tau_{\Gamma,\mathfrak{a}_\tau}(z)$ and hence, $\langle \mathcal{K}_{\Gamma^\tau,\infty}, \Psi^\tau_\delta\rangle_{\Gamma^\tau} = \langle \mathcal{K}_{\Gamma,\mathfrak{a}_\tau}, \Psi_\delta\rangle_\Gamma$. Since $\mathcal{K}_{\Gamma,\mathfrak{a}_\tau}(z)$ is smooth and Ψ_δ is supported on a δ-neighborhood of i we can estimate $\langle \mathcal{K}_{\Gamma,\mathfrak{a}_\tau}, \Psi_\delta\rangle_\Gamma = \mathcal{K}_{\Gamma,\mathfrak{a}_\tau}(i) + O(\delta)$ so the Eisenstein term can be approximated by

$$\langle \mathcal{K}_{\Gamma^\tau,\infty}, \Psi^\tau_\delta\rangle_{\Gamma^\tau} = \mathcal{K}_{\Gamma,\mathfrak{a}_\tau}(i) + O(\delta).$$

Making an optimal choice of

$$\delta = \omega^{\frac{1}{2(1+\alpha)}} T^{-\frac{\eta_1}{2(1+\alpha)}}, \tag{5.15}$$

the first error term dominates and we get

$$\langle F_{T,\tau}, \Psi_\delta\rangle = \frac{2T}{v_\Gamma}\left(\log(T^2\omega^2) - 2 + v_\Gamma(\mathcal{K}_{\Gamma,\mathfrak{a}_\tau}(i) + \mathcal{K}_{\Gamma,\mathfrak{a}_{\tau\sigma}}(i))\right)$$
$$+O\left(\omega^{2\beta_1\eta}T^{1-\eta}\right).$$

Finally, using (5.8) relating the counting problem to the inner product (after replacing T with $T(1 \pm 5\delta)$) we get that

$$\mathcal{N}^{\tau}_{\mathcal{C}_T}(\Gamma) = \frac{2T}{v_\Gamma}\left(\log(T^2\omega) - 2 + v_\Gamma(\mathcal{K}_{\Gamma,\mathfrak{a}_\tau}(i) + \mathcal{K}_{\Gamma,\mathfrak{b}_\tau}(i)) + O((\tfrac{\omega^{2\beta_1}}{T})^\eta\right). \quad \square$$

Remark 5.16. As can be expected, the main term does not depend on τ. The secondary term does depend on τ, but only involves knowledge of which cusps are used for the Eisenstein term, and the widths of these cusps.

6. Counting integer solutions

In this section, we establish the results claimed in §1.3, handling Theorems 1.5 and 1.9 simultaneously.

6.1. Decomposition into orbits

Consider the variety

$$V_d \ : \ b^2 - 4ac = d,$$

where d is fixed. Identifying the triple (a,b,c) with the quadratic form

$$Q(x,y) = ax^2 + bxy + cy^2,$$

gives a natural PSL_2 action on V_d, via

$$Q^g(v) = Q(vg^t),$$

where $g \in \mathrm{PSL}_2$ acts linearly on $v = (x,y)$ from the right (here g^t is the transpose of g). More explicitly, the action of $g \in \mathrm{PSL}_2$ on triples $(a,b,c) \in V_d$ is given by the linear action

$$(a,b,c)^g := (a,b,c)\iota(g),$$

where $\iota : \mathrm{PSL}_2 \to \mathrm{SO}_{b^2-4ac}$ is the spin morphism given by

$$\iota\begin{pmatrix} a & b \\ c & d\end{pmatrix} = \begin{pmatrix} a^2 & 2ab & b^2 \\ ac & ad+bc & bd \\ c^2 & 2cd & d^2 \end{pmatrix}.$$

In particular, $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$ acts on the integer points $V_d(\mathbb{Z})$ and we can decompose

$$V_d(\mathbb{Z}) = \cup_{i=1}^{h(d)} v_i\Gamma_1,$$

into finitely many orbits. The number $h(d)$ of orbits is, in general, very mysterious; for instance, when d is a square free fundamental discriminant then $h(d)$ is the class number of the quadratic extension $\mathbb{Q}(\sqrt{d})$. However when $d = n^2$ is a perfect square, the number of orbits can be computed explicitly. In the following lemma, we compute it and give a full set of representatives for the orbits.

Lemma 6.1. *For $d = n^2$ a square we have $h(n^2) = n$. Moreover, the set $\{(0, n, 0)^{\tau_j} | 0 \le j < n\}$ with $\tau_j = \left(\begin{smallmatrix} 1 & j/n \\ 0 & 1 \end{smallmatrix}\right)$ is a full set of representatives for the classes of $V_d(\mathbb{Z})/\operatorname{PSL}_2(\mathbb{Z})$.*

Proof. We identify a point $(a, b, c) \in V_d(\mathbb{Z})$ with the corresponding quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ having discriminant $d = b^2 - 4ac$. Then $\operatorname{PSL}_2(\mathbb{Z})$ acts on the set of quadratic forms by $Q^\gamma(v) = Q(v\gamma^t)$ where $v = (x, y)$ and $\gamma \in \operatorname{PSL}_2(\mathbb{Z})$ is acting linearly on the right.

Recall that a binary quadratic form Q has a square discriminant, if and only if the form factors as a product of linear forms

$$Q(x, y) = (Ax + By)(Cx + Dy),$$

in which case the discriminant is given by $(AD - BC)^2$. We thus get a map, from the set $\mathcal{M}_n(\mathbb{Z})$ of 2×2 integral matrices with determinant n, onto the set of integral quadratic forms of discriminant n^2, sending $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ to the form $Q_M(x, y) = (Ax + By)(Cx + Dy)$ (this map is not injective since the same form can have several factorizations). A direct computation shows that $Q_M^\gamma = Q_{M\gamma}$, and since

$$\Delta_n = \{\left(\begin{smallmatrix} A & 0 \\ C & D \end{smallmatrix}\right) \in \mathcal{M}_n(\mathbb{Z}) : AD = n,\ 0 \le C < D\},$$

is a full set of representatives for $\mathcal{M}_n(\mathbb{Z})/\operatorname{PSL}_2(\mathbb{Z})$, the set $\{Q_M : M \in \Delta_n\}$ is a full set of representatives for classes of quadratic form of discriminant n^2 (some of these might be equivalent though). A form in this set of representatives can be written explicitly as

$$Q_M(x, y) = Ax(Cx + Dy) = ACx^2 + nxy,$$

and we see that the classes with $A \ne 1$ are redundant. A smaller full set of representatives is thus given by the forms $Cx^2 + nxy$ with $0 \le C < n$, and it is not hard to verify directly that these are all inequivalent. Finally, observing that $Cx^2 + nxy$ is equivalent to $nxy - Cy^2$ concludes the proof. □

Remark 6.2. Instead of looking at all integer points in $V_d(\mathbb{Z})$ one can consider only primitive points (i.e., points with gcd $(a, b, c) = 1$). It is easy to see that Γ_1 also acts on the set of primitive points, and the same proof shows that the set $\{(0, n, j) | 0 \le j < n,\ (n, j) = 1\}$ is a full set of representatives for the orbits of primitive points.

Using our orbit decomposition of V_d we can also get a corresponding decomposition of the variety

$$W_d \;:\; x^2 + y^2 - z^2 = d.$$

The map $(x, y, z) \mapsto (\frac{z+y}{2}, x, \frac{z-y}{2})$ is a bijection between W_d and V_d and the integer points $W_d(\mathbb{Z})$ map to the set

$$\widetilde{V}_d(\mathbb{Z}) := \{(a, b, c) \in V_d | b \in \mathbb{Z},\ a, c \in \tfrac{1}{2}\mathbb{Z},\ a + c \in \mathbb{Z}\}.$$

From this map we see that the congruence subgroup

$$\Gamma_2 = \{\gamma \in \Gamma_1 : \bar{\gamma} \in \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\}\},$$

(with $\bar{\gamma} \in \mathrm{PSL}_2(\mathbb{Z}/2\mathbb{Z})$ the projection of γ), acts on $\widetilde{V}_d(\mathbb{Z})$ (and hence also on $W_d(\mathbb{Z})$). Using the classification of the orbit of the Γ_1 action on $V_d(\mathbb{Z})$ we get the following classification for the Γ_2 action on $\widetilde{V}_d(\mathbb{Z})$.

Lemma 6.3. *For $d = n^2$ a complete set of representatives for the Γ_2 orbits of $\widetilde{V}_d(\mathbb{Z})$ is given by*

$$\{(0, n, 0)^{\tau_j},\ 0 \le j < 2n\} \cup \{(0, n, 0)^{\widetilde{\tau}_j} : 0 \le j < 2n\},$$

where $\tau_j = \left(\begin{smallmatrix} 1 & j/n \\ 0 & 1 \end{smallmatrix}\right)$ is as above and $\widetilde{\tau}_j = \left(\begin{smallmatrix} 1+\frac{j}{2n} & \frac{j}{2n} \\ 1 & 1 \end{smallmatrix}\right)$.

Proof. Let $(a, b, c) \in \widetilde{V}_d(\mathbb{Z})$ then $(2a, 2b, 2c) \in V_{4d}(\mathbb{Z})$. From the classification of Γ_1 orbits of $V_{4d}(\mathbb{Z})$ there is $\gamma \in \Gamma_1$ with $(a, b, c) = (0, n, j/2)^{\gamma}$ with $0 \le j < 2n$. We can write γ as $\sigma\widetilde{\gamma}$ with $\widetilde{\gamma} \in \Gamma_2$ and $\sigma \in \{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\}$ in the set of representatives for Γ_1/Γ_2. We can thus write the point (a, b, c) as $(0, n, j/2)^{\widetilde{\gamma}}$ or as $(0, n, n + j/2)^{\widetilde{\gamma}}$ or $(n + \frac{j}{2}, n + j, \frac{j}{2})^{\widetilde{\gamma}}$ for some $0 \le j < 2n$. Now, $(a, b, c) \in \widetilde{V}_d(\mathbb{Z})$ and $\widetilde{\gamma}$ preserves this space, so in the first two cases we must have that $j/2 \in \mathbb{Z}$. Hence, a full set of representatives is indeed $(0, n, j) = (0, n, 0)^{\tau_j}$ and $(n + \frac{j}{2}, n + j, \frac{j}{2}) = (0, n, 0)^{\widetilde{\tau}_j}$ for $0 \le j < 2n$. □

Remark 6.4. Consider the norm $\|\cdot\|_*$ on $\mathbb{R}^3$ defined by

$$\|(a, b, c)\|_*^2 = 2a^2 + b^2 + 2c^2,$$

and let $B_T = \{(a, b, c) \in \mathbb{R}^3 : \|(a, b, c)\|_* < T\}$. Note that $\widetilde{V}_d(\mathbb{Z}) \cap B_T$ is in bijection with the set $\{(x, y, z) \in W_d(\mathbb{Z}) : x^2 + y^2 + z^2 \le T^2\}$ so $\mathcal{N}_d(T) = \#(\widetilde{V}_d(\mathbb{Z}) \cap B_T)$. Moreover, under this bijection, the primitive points of $W_d(\mathbb{Z})$ correspond exactly to the Γ_2-orbits of the classes $(0, n, 0)^{\tau_j}$ and $(0, n, 0)^{\widetilde{\tau}_j}$ with $(n, j) = 1$.

Decomposing the integer points of $V_d(\mathbb{Z}) \cap B_T$ and $\widetilde{V}_d(\mathbb{Z}) \cap B_T$ into the finitely many orbits, it is enough to count points in each orbit separately. For this we need to estimate terms of the form

$$\#\{\gamma \in \Gamma : \|(0,n,0)^{\tau\gamma}\|_* \leq T\}$$

with $\tau = \tau_j$ or $\tau = \widetilde{\tau}_j$ as above and the lattice $\Gamma = \Gamma_1$ or $\Gamma = \Gamma_2$ respectively. We now show that these counting functions are given in terms of the cone counting function defined in (5.2).

Lemma 6.5. *For any lattice* Γ *and* $\tau \in G$ *we have*

$$\#\{\gamma \in \Gamma : \|(0,n,0)^{\tau\gamma}\|_* \leq T\} = \mathcal{N}^{\tau^{-1}}_{\mathcal{C}_{T_n}}(\Gamma)$$

with $T_n = \sqrt{\frac{T^2}{2n^2} - \frac{1}{2}}$.

Proof. Write $\tau\gamma = a_y n_x k$ so that $\tau\gamma i = yx + iy$. Since $(0,n,0)^{a_y} = (0,n,0)$ and our norm is K-invariant we can explicitly compute

$$\|(0,n,0)^{\tau\gamma}\|^2 = n^2(1+2x^2),$$

so that indeed $\|(0,n,0)^{\tau\gamma}\|_* \leq T$ if and only if $|x| \leq T_n$ which is equivalent to $\tau\gamma.i \in \mathcal{C}_{T_n}$. □

6.2. Square discriminants

Our goal here is to prove Theorem 1.9 by estimating

$$\begin{aligned}\#V_d(\mathbb{Z}) \cap B_T &= \sum_{j=0}^{n-1} \#\{\gamma \in \Gamma_1 : \|(0,n,0)^{\tau_j\gamma}\|_* \leq T\} \\ &= \sum_{j=0}^{n-1} \mathcal{N}^{\tau_j^{-1}}_{\mathcal{C}_{T_n}}(\Gamma_1).\end{aligned}$$

Note that $\Gamma_1^{\tau_j^{-1}}$ has a cusp at ∞ of width 1 and $\Gamma_1^{\tau_j^{-1}\sigma}$ has a cusp at ∞ of width $\omega_j = \frac{n^2}{(n,j)^2}$. Hence, from Theorem 5.3 we get that

$$\mathcal{N}^{\tau_j^{-1}}_{\mathcal{C}_{T_n}}(\Gamma_1) = \frac{2T_n}{v_{\Gamma_1}}\left(2\log(T_n) + \log(\frac{n^2}{(n,j)^2}) + 2v_{\Gamma_1}\mathcal{K}_{\Gamma_1}(i) + O(\frac{n^{2\beta_1\eta}}{T_n^\eta})\right),$$

with $\beta_1 = \beta - \frac{1}{2} > 1 + 2\theta$. Recalling the assumption $T \geq d^\beta = n^{2\beta}$ in the statement of Theorem 1.9, we have that $T_n \gg 1$ and we can estimate $T_n = \frac{T}{\sqrt{2}n} + O(\frac{n}{T})$, and

$$\log(T_n) = \log(T) - \tfrac{1}{2}\log(2) - \log(n) + O(\frac{n^2}{T^2}),$$

so that

$$\mathcal{N}^{\tau_j^{-1}}_{\mathcal{C}_{T_n}}(\Gamma_1) = \frac{\sqrt{2}T}{nv_{\Gamma_1}}\left(2\log(T) - \log(2) - 2\log((n,j)) + 2v_{\Gamma_1}\mathcal{K}_{\Gamma_1}(i) + O(\tfrac{n^{(2\beta_1+1)\eta}}{T^\eta})\right).$$

Summing over all orbits we get that

$$\#V_d(\mathbb{Z}) \cap B_T = \frac{\sqrt{72}T}{\pi}\left(\log(T) - \tfrac{1}{2}\log(2) + \tfrac{\pi}{3}\mathcal{K}_{\Gamma_1}(i) - \frac{1}{n}\sum_{j=1}^{n}\log(n,j) + O(\tfrac{n^{2\beta\eta}}{T^\eta})\right).$$

Plugging in the value of $\mathcal{K}_{\Gamma_1}(i)$ from (A.10) and noting that

$$\sum_{j=1}^{n-1}\log(n,j) = \sum_{a|n}\phi(\tfrac{n}{a})\log(a),$$

concludes the proof of (1.10).

6.3. Sum of squares

Next we prove Theorem 1.5 by estimating $\mathcal{N}_d(T) = \widetilde{V}_d(\mathbb{Z}) \cap B_T$. Again, split the integral points into the finitely many Γ_2 orbit and count in each orbit. We thus need to estimate $\mathcal{N}^{\tau_j^{-1}}_{\mathcal{C}_{T_n}}(\Gamma_2)$ and $\mathcal{N}^{\widetilde{\tau}_j^{-1}}_{\mathcal{C}_{T_n}}(\Gamma_2)$.

Let $\mathfrak{a}_j = \tau_j^{-1}\infty$, $\mathfrak{b}_j = \tau_j^{-1}0$ and let ω_j, ω_j' denote the width of the cusps at ∞ of $\Gamma_2^{\tau_j^{-1}}$ and $\Gamma_2^{\tau_j^{-1}\sigma}$, appearing in the formula for $\mathcal{N}^{\tau_j^{-1}}_{\mathcal{C}_{T_n}}(\Gamma_2)$. Similarly let $\widetilde{\mathfrak{a}}_j = \widetilde{\tau}_j^{-1}\infty$, $\widetilde{\mathfrak{b}}_j = \widetilde{\tau}_j^{-1}0$ and let $\widetilde{\omega}_j, \widetilde{\omega}_j'$ the corresponding cusp widths.

Recall that Γ_2 has only two inequivalent cusps, one at ∞ and another at 1. After verifying which pair of cusps we get for each orbit, we show that the contribution of the Kronecker terms to the counting function is as follows.

Lemma 6.6. *With the above notation we have*

$$\sum_{j=0}^{2n-1}(\mathcal{K}_{\Gamma_2,\mathfrak{a}_j}(i) + \mathcal{K}_{\Gamma_2,\mathfrak{b}_j}(i) + \mathcal{K}_{\Gamma_2,\widetilde{\mathfrak{a}}_j}(i) + \mathcal{K}_{\Gamma_2,\widetilde{\mathfrak{b}}_j}(i)) = 4n(\mathcal{K}_{\Gamma_2,\infty}(i) + \mathcal{K}_{\Gamma_2,1}(i)).$$

Proof. Since $\tau_j^{-1}\infty = \infty$ and $\widetilde{\tau}_j^{-1}\infty = -1$ (which is Γ_2 equivalent to the cusp at 1) we see that $\mathfrak{a}_j = \infty$ and $\widetilde{\mathfrak{a}}_j = 1$. Next $\tau_j^{-1}0 = \frac{j}{n}$ is Γ_1-equivalent to the cusp at infinity by the action of some $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1$ such that $\frac{j}{n} = \gamma\infty = \frac{a}{b}$, that is, with $a = \frac{j}{(n,j)}$ and $b = \frac{n}{(n,j)}$. Such an element γ lies in Γ_2 iff $\frac{nj}{(n,j)^2}$ is even, so that

$\mathfrak{b}_j = \begin{cases} \infty & \frac{nj}{(n,j)^2} = 0 \pmod 2 \\ 1 & \frac{nj}{(n,j)^2} = 1 \pmod 2 \end{cases}$. Similarly, $\widetilde{\tau}_j^{-1} 0 = \frac{-j}{2n+j}$ is Γ_1-equivalent to the cusp at ∞ by $\gamma \in \Gamma_1$ with $\frac{a}{b} = \frac{-j}{2n+j}$ so that $a = \frac{-j}{(j,2n)}$ and $b = \frac{2n+j}{(j,2n)}$. Hence γ can be taken from Γ_2 iff $\frac{j(j+2n)}{(j,2n)^2}$ is even implying that $\widetilde{\mathfrak{b}}_j = \begin{cases} \infty & \frac{j(j+2n)}{(2n,j)^2} = 0 \pmod 2 \\ 1 & \frac{(j+2n)j}{(2n,j)^2} = 1 \pmod 2 \end{cases}$. We thus get that the term $\mathcal{K}_j = \mathcal{K}_{\Gamma_2,\mathfrak{a}_j}(i) + \mathcal{K}_{\Gamma_2,\mathfrak{b}_j}(i) + \mathcal{K}_{\Gamma_2,\widetilde{\mathfrak{a}}_j}(i) + \mathcal{K}_{\Gamma_2,\widetilde{\mathfrak{b}}_j}(i)$ is given by

$$\mathcal{K}_j = \begin{cases} 3\mathcal{K}_{\Gamma_2,\infty}(i) + \mathcal{K}_{\Gamma_2,1}(i) & \frac{nj}{(n,j)^2} = \frac{j(j+2n)}{(2n,j)^2} = 0 \pmod 2 \\ \mathcal{K}_{\Gamma_2,\infty}(i) + 3\mathcal{K}_{\Gamma_2,1}(i) & \frac{nj}{(n,j)^2} = \frac{j(j+2n)}{(2n,j)^2} = 1 \pmod 2 \\ 2\mathcal{K}_{\Gamma_2,\infty}(i) + 2\mathcal{K}_{\Gamma_2,1}(i) & \frac{nj}{(n,j)^2} \neq \frac{j(j+2n)}{(2n,j)^2} \pmod 2. \end{cases}$$

Writing $n = 2^a m$ with m odd we see that the first case happens when $j = 0 \pmod{2^{a+1}}$ (hence for $\frac{n}{2^a}$ values of $0 \le j < 2n$), the second case when $j = 0 \pmod{2^a}$ but $j \neq 0 \pmod{2^{a+1}}$ (for another $\frac{n}{2^a}$ values of j) and the last case when $j \neq 0 \pmod{2^a}$ (for $2n(1 - \frac{1}{2^a})$ values of j). Now summing over all $0 \le j < 2n$ we get our result. $\square$

Summing up the contributions from the widths of the cusps, we get

Lemma 6.7. *With notation as above, let $n = 2^\nu m$ with m odd. Then*

$$\sum_{j=0}^{2n-1} \log(\omega_j \omega_j' \widetilde{\omega_j} \widetilde{\omega}_j') = 8n \log(2n) - \tfrac{2n\log(2)}{2^\nu} - 8 \sum_{a|n} \phi(\tfrac{n}{a}) \log(a).$$

Proof. Fix $0 \le j < 2n$. Since τ_j commutes with N the width of the cusp at ∞ of $\Gamma_2^{\tau_j^{-1}}$ is the same as for Γ_2 so $\omega_j = 2$. Similarly, $\widetilde{\tau}_j^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \widetilde{\tau}_j = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \in \Gamma_2$ so $\widetilde{\omega}_j = 1$.

Next ω_j' is the smallest integer k such that

$$\tau_j^{-1} \sigma \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \sigma^{-1} \tau_j = \begin{pmatrix} 1 + \frac{kj}{n} & \frac{j^2 k}{n^2} \\ -k & 1 - \frac{kj}{n} \end{pmatrix} \in \Gamma_2,$$

hence

$$\omega_j' = \begin{cases} \frac{n^2}{(n,j)^2}, & \frac{jn}{(n,j)^2} = 1 \pmod 2 \\ \frac{2n^2}{(n,j)^2}, & \frac{jn}{(n,j)^2} = 0 \pmod 2. \end{cases}$$

Similarly, $\widetilde{\omega}_j'$ is the smallest integer k such that

$$\widetilde{\tau}_j^{-1} \sigma \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \sigma \widetilde{\tau}_j = \begin{pmatrix} 1 - \frac{kj}{2n}(1 + \frac{j}{2n}) & -\frac{j^2 k}{4n^2} \\ -k(1 + \frac{j}{2n})^2 & 1 - \frac{kj}{2n}(1 + \frac{j}{2n}). \end{pmatrix} \in \Gamma_2, \tag{6.8}$$

and a similar computation gives that

$$\widetilde{\omega}'_j = \begin{cases} \frac{4n^2}{(2n,j)^2}, & \frac{j}{(j,2n)} = 1 \pmod 2 \text{ and } \frac{2n}{(j,2n)} = 0 \pmod 2 \\ \frac{8n^2}{(2n,j)^2}, & \text{otherwise.} \end{cases}$$

Writing $n = 2^\nu m$ with m odd, it is not hard to see that

$$\omega_j \widetilde{\omega}_j \omega'_j \widetilde{\omega}'_j = \begin{cases} \frac{8n^4}{(n,j)^4}, & j = 0 \pmod{2^\nu} \\ \frac{16n^4}{(n,j)^4}, & \text{otherwise,} \end{cases}$$

so that

$$\begin{aligned} \sum_{j=0}^{2n-1} \log(\omega_j \omega'_j \widetilde{\omega_j} \widetilde{\omega}'_j) &= \sum_{j=0}^{2n-1} \log(\frac{8n^4}{(n,j)^4}) + 2n\log(2)(1 - \tfrac{1}{2^\nu}) \\ &= 2n\log(8n^4) - 8\sum_{j=0}^{n-1} \log((n,j)) + 2n\log(2) - \tfrac{2n\log(2)}{2^\nu} \\ &= 8n\log(2n) - \tfrac{2n\log(2)}{2^\nu} - 8\sum_{a|n} \phi(\tfrac{n}{a})\log(a), \end{aligned}$$

as claimed. □

Proof of Theorem 1.5. Partitioning $\widetilde{V}_d(\mathbb{Z}) \cap B_T$ into Γ_2 orbits and summing in each orbit gives

$$\mathcal{N}_d(T) = \sum_{j=0}^{2n-1} \mathcal{N}_{\mathcal{C}_{T_n}}^{\tau_j^{-1}}(\Gamma_2) + \mathcal{N}_{\mathcal{C}_{T_n}}^{\widetilde{\tau}_j^{-1}}(\Gamma_2).$$

Using Theorem 5.3 and the estimates $T_n = \frac{T}{\sqrt{2}n} + O(\frac{n}{T})$ and $|\omega_j \omega'_j| \le 4n^2$ we estimate each of the cone counting functions

$$\mathcal{N}_{\mathcal{C}_{T_n}}^{\tau_j^{-1}}(\Gamma_2) = \frac{\sqrt{2}T}{\pi n}\left(\log(\tfrac{T^2\omega_j\omega'_j}{2n^2}) - 2 + \pi(\mathcal{K}_{\Gamma_2,\mathfrak{a}_j}(i) + \mathcal{K}_{\Gamma_2,\mathfrak{b}_j}(i)) + O(\tfrac{n^{(2\beta_1+1)\eta}}{T^\eta})\right),$$

with $\beta_1 = \beta - \frac{1}{2}$ as before, and similarly for $\mathcal{N}_{\mathcal{C}_{T_n}}^{\widetilde{\tau}_j^{-1}}(\Gamma_2)$. Summing over $0 \le j < 2n$, by Lemma 6.6 the contribution of the Kronecker terms is $\sqrt{32}T(\mathcal{K}_{\Gamma_2,\infty}(i) + \mathcal{K}_{\Gamma_2,1}(i))$ and by Lemma 6.7 the contribution of the terms $\log(\omega_j\omega'_j\widetilde{\omega}_j\widetilde{\omega}'_j)$ is

$$\sum_{j=0}^{2n-1} \log(\omega_j\omega'_j\widetilde{\omega}_j\widetilde{\omega}'_j) = 8n\log(2n) - \tfrac{2n\log(2)}{2^\nu} - 8\sum_{a|n}\phi(\tfrac{n}{a})\log(a),$$

so that

$$\mathcal{N}_d(T) = \frac{\sqrt{32}T}{\pi}\Big(2\log(T) - 2 + \log(2) + \pi(\mathcal{K}_{\Gamma_2,\infty}(i) + \mathcal{K}_{\Gamma_2,1}(i)) \\ - \frac{\log(2)}{2^{\nu+1}} - \frac{2}{n}\sum_{a|n}\phi(\tfrac{n}{a})\log(a) + O_\eta(\tfrac{n^{2\beta}}{T^\eta})\Big).$$

We now want to express the Kronecker terms in terms of special values of Dedekind eta function. To do this, note that $\Gamma_2 = \Gamma_0(2)^\tau$ with $\tau = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and hence $E_{\Gamma_2,\mathfrak{a}}(z,s) = E_{\Gamma_0(2),\tau\mathfrak{a}}(\tau z, s)$ and also $\mathcal{K}_{\Gamma_2,\mathfrak{a}}(z) = \mathcal{K}_{\Gamma_0(2),\tau\mathfrak{a}}(z,s)$. In particular

$$\mathcal{K}_{\Gamma_2,\infty}(i) + \mathcal{K}_{\Gamma_2,1}(i) = \mathcal{K}_{\Gamma_0(2),0}(\tfrac{i+1}{2}) + \mathcal{K}_{\Gamma_0(2),\infty}(\tfrac{i+1}{2}).$$

Using Proposition A.11 for $\Gamma = \Gamma_0(2)$ we have

$$\mathcal{K}_{\Gamma,\infty}(z) = \frac{1}{\pi}\left(2\gamma - 2\frac{\zeta'}{\zeta}(2) - \log\left(\tfrac{4y|\eta(2z)|^8}{|\eta(z)|^4}\right) - \frac{8\log(2)}{3}\right),$$

and

$$\mathcal{K}_{\Gamma,0}(z) = \frac{1}{\pi}\left(2\gamma - 2\frac{\zeta'}{\zeta}(2) - \log\left(\tfrac{4y|\eta(z)|^8}{|\eta(2z)|^4}\right) + \frac{\log(2)}{3}\right).$$

Hence

$$\pi(\mathcal{K}_{\Gamma,0}(\tfrac{i+1}{2}) + \mathcal{K}_{\Gamma,\infty}(\tfrac{i+1}{2})) = 4\gamma - 4\frac{\zeta'}{\zeta}(2) - \tfrac{13\log(2)}{3} - 2\log(|\eta(i+1)\eta(\tfrac{i+1}{2})|^2).$$

Using the transformation law for the Dedekind eta function

$$|\eta(z+1)|^2 = |\eta(z)|^2, \quad |\eta(\tfrac{-1}{z})|^2 = |z||\eta(z)|^2,$$

we have that $|\eta(i+1)\eta(\frac{i+1}{2})|^2 = \sqrt{2}|\eta(i)|^4 = \sqrt{2}\frac{\Gamma(1/4)^4}{16\pi^3}$ so that

$$\pi(\mathcal{K}_{\Gamma,0}(\tfrac{i+1}{2}) + \mathcal{K}_{\Gamma,\infty}(\tfrac{i+1}{2})) = 4\gamma - 4\frac{\zeta'}{\zeta}(2) - \tfrac{13\log(2)}{3} - 2\log(2) - 2\log(\frac{\Gamma(1/4)^4}{16\pi^3}),$$

and plugging this back in we get that

$$\mathcal{N}_d(T) = \frac{\sqrt{128}T}{\pi}\Big(\log(T) + C - \frac{1}{n}\sum_{a|n}\phi(\tfrac{n}{a})\log(a) + \log(2)(\tfrac{1}{3} - \tfrac{1}{2^{\nu+2}}) + O(\tfrac{n^{2\beta\eta}}{T^\eta})\Big),$$

with the constant

$$C = 2\gamma - 1 - 2\frac{\zeta'}{\zeta}(2) - \tfrac{5\log(2)}{2} - \log(\frac{\Gamma(1/4)^4}{16\pi^3}),$$

as before. This completes the proof. □

Appendix A. Eisenstein series for $\Gamma_0(p)$

As for the full modular group, $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$, the theory of Eisenstein series for the congruence groups $\Gamma_0(p)$ is also well understood; in particular, the Fourier coefficients can be expressed explicitly and there is an explicit formula for the Kronecker limit. Since we could not find a suitable reference for these formulas, we will include short proofs here, but we claim no originality.[4]

We first note that when Γ is a finite index subgroup of Γ_1, one can express the Eisenstein series for Γ_1 in terms of the Eisenstein series corresponding to the different cusps of Γ. Explicitly, we have

Lemma A.1. *Let $\sigma_1, \dots, \sigma_k$ denote a complete set of representatives of Γ_1/Γ, and let $\mathfrak{a}_i = \sigma_j^{-1}\infty$ (these are not necessarily inequivalent cusps for Γ). Then*

$$E_{\Gamma_1}(z,s) = \sum_{j=1}^{k} \omega_j^{s-1} E_{\Gamma,\mathfrak{a}_j}(z,s), \tag{A.2}$$

where ω_j denotes the width of the cusp $\mathfrak{a}_i$.

Proof. Since $N \cap \sigma_j \Gamma \sigma_j^{-1} \subseteq N \cap \sigma_j \Gamma_1 \sigma_j^{-1} = N \cap \Gamma_1$ it is generated by $n_{\omega_j} = \left(\begin{smallmatrix} 1 & \omega_j \\ 0 & 1 \end{smallmatrix}\right)$ where the width $\omega_j \in \mathbb{N}$ of $\mathfrak{a}_j$ is the index of $\Gamma_{\mathfrak{a}_j}$ in $\Gamma_1 \cap N$. We can thus write for $\mathfrak{Re}(s) > 1$ and each coset σ_j

$$\sum_{\gamma \in \Gamma_{\mathfrak{a}_j} \backslash \Gamma} \mathfrak{Im}(\sigma_j \gamma z)^s = \sum_{\substack{\gamma \in \langle n_{\omega_j} \rangle \backslash \Gamma_1 \\ \sigma_j^{-1}\gamma \in \Gamma}} \mathfrak{Im}(\gamma z)^s = \omega_j \sum_{\substack{\gamma \in (N \cap \Gamma_1) \backslash \Gamma_1 \\ \sigma_j^{-1}\gamma \in \Gamma}} \mathfrak{Im}(\gamma z)^s,$$

and dividing by ω_j and summing over all cosets gives

$$\begin{aligned} E_{\Gamma_1}(z,s) &= \sum_{(\Gamma_1 \cap N) \backslash \Gamma_1} \mathfrak{Im}(\gamma z)^s \\ &= \sum_{j=1}^{k} \omega_j^{-1} \sum_{\gamma \in \Gamma_{\mathfrak{a}_j} \backslash \Gamma} \mathfrak{Im}(\sigma_j \gamma z)^s. \end{aligned}$$

On the other hand, since $\tau_{\mathfrak{a}_j} = \sigma_j^{-1} a_{\omega_j}$ is a scaling matrix for $\mathfrak{a}_j$ we have

$$\sum_{\gamma \in \Gamma_{\mathfrak{a}_j} \backslash \Gamma} \mathfrak{Im}(\sigma_j \gamma z)^s = \sum_{\gamma \in \Gamma_{\mathfrak{a}_j} \backslash \Gamma} \mathfrak{Im}(a_{\omega_j} \tau_{\mathfrak{a}_j}^{-1} \gamma z)^s = \omega_j^s E_{\Gamma,\mathfrak{a}_j}(z,s),$$

and the result follows. □

[4] Added in print: see [Vas96] where similar calculations are carried out.

Subtracting the residue and taking the limit as $s \to 1$ we get the following

Corollary A.3. *For Γ a finite index subgroup of Γ_1 the Kronecker limit satisfies*

$$\mathcal{K}_{\Gamma_1}(z) = \sum_{j=1}^{k} \mathcal{K}_{\Gamma,\mathfrak{a}_j}(z) + \frac{3}{\pi k}\sum_{j=1}^{k}\log(\omega_j).$$

In particular, applying this to the subgroup $\Gamma = \Gamma_0(p)$ of $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$ we get the following identities

$$E_{\Gamma_1,\infty}(z,s) = E_{\Gamma_0(p),\infty}(z,s) + p^s E_{\Gamma_0(p),0}(z,s), \tag{A.4}$$

$$\mathcal{K}_{\Gamma_1}(z) = \mathcal{K}_{\Gamma_0(p),\infty}(z) + p\mathcal{K}_{\Gamma_0(p),0}(z) + \frac{3p\log(p)}{(p+1)\pi}. \tag{A.5}$$

A.1. Fourier coefficients

For each pair of cusps $\mathfrak{a},\mathfrak{b}$ the Fourier expansion of the Eisenstein series $E_{\Gamma,\mathfrak{b}}$ with respect to the cusp at $\mathfrak{a}$, is given by

$$E^{\tau_\mathfrak{a}}_{\Gamma,\mathfrak{b}}(z,s) = \delta_{\mathfrak{a},\mathfrak{b}}y^s + \phi_{\mathfrak{a},\mathfrak{b}}(s)y^{1-s} + \sum_{m\neq 0} a_{\mathfrak{a},\mathfrak{b}}(s;m,y)e(mx),$$

and since $E_{\Gamma,\mathfrak{b}}(z,s)$ is an eigenfunction with eigenvalue $s(1-s)$ we can write

$$a_{\mathfrak{a},\mathfrak{b}}(s;m,y) = \phi_{\mathfrak{a},\mathfrak{b}}(s;m)2\sqrt{y}K_{s-\frac{1}{2}}(2\pi m y).$$

For the full modular group $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$ there is just one cusp at ∞ and the Fourier coefficients are given explicitly by $\phi(s) = \frac{\zeta^*(2s-1)}{\zeta^*(2s)}$ and

$$\phi(s,m) = \frac{\tau_{s-1/2}(m)}{\zeta^*(2s)}, \tag{A.6}$$

where $\zeta^*(s) = \pi^{-s/2}\zeta(s)\Gamma(s/2)$ is the completed Riemann zeta function and $\tau_s(m) = \sum_{ab=|m|}(\frac{a}{b})^s$ is the divisor function [Iwa95, page 67].

For the congruence groups $\Gamma_0(p)$ the Fourier coefficients can also be given by a similar formula and satisfy a similar bound.

Proposition A.7. *For $\Gamma = \Gamma_0(p)$ we have*

$$\phi_{\infty,\infty}(s;m) = \phi_{0,0}(s;m) = \frac{1}{p^{2s}-1}\begin{cases} -\phi(s;m) & (p,m)=1 \\ p^{s+1/2}\phi(s;\frac{m}{p}) - \phi(m;s) & p|m \end{cases},$$

$$\phi_{\infty,0}(s;m) = \phi_{0,\infty}(s;m) = \frac{1}{p^{2s}-1}\begin{cases} -p^s\phi(s;m) & (p,m)=1 \\ p^s\phi(s;m) - \sqrt{p}\phi(s;\frac{m}{p}) & p|m \end{cases}.$$

Proof. Since the scaling matrix τ_0 normalizes $\Gamma = \Gamma_0(p)$ we have that $E^{\tau_0}_{\Gamma,\infty}(z,s) = E_{\Gamma,0}(z,s)$ implying that $\phi_{\infty,0}(s;m) = \phi_{0,\infty}(s;m)$, and since τ_0^2 is the identity in $\mathrm{PSL}_2(\mathbb{R})$ then $E^{\tau_0}_{\Gamma,0}(z,s) = E_{\Gamma,\infty}(z,s)$ implying that $\phi_{\infty,\infty}(s;m) = \phi_{0,0}(s;m)$.

Now looking at the expansion at infinity of (A.4) we get that

$$\phi(s;m) = \phi_{\infty,\infty}(s;m) + p^s\phi_{0,\infty}(s;m),$$

and the expansion at 0 gives

$$\sqrt{p}\phi(s;\tfrac{m}{p}) = \phi_{\infty,0}(s;m) + p^s\phi_{0,0}(s;m),$$

where it is understood that $\phi(s;\frac{m}{p}) = 0$ when $(p,m) = 1$. We thus get that

$$\begin{pmatrix} 1 & p^s \\ p^s & 1 \end{pmatrix}\begin{pmatrix} \phi_{0,0}(s;m) \\ \phi_{0,\infty}(s;m) \end{pmatrix} = \begin{pmatrix} \phi(s;m) \\ \sqrt{p}\phi(s;\frac{m}{p}) \end{pmatrix},$$

and inverting the matrix concludes the proof. □

A.2. Kronecker limits

The Kronecker limit corresponding to a cusp $\mathfrak{a}$ is defined as the limit

$$\mathcal{K}_{\Gamma,\mathfrak{a}}(z) = \lim_{s\to 1}\left(E_{\Gamma,\mathfrak{a}}(s,z) - \frac{1}{v_\Gamma(s-1)}\right). \tag{A.8}$$

When $\Gamma_1 = \mathrm{PSL}_2(\mathbb{Z})$ the Kronecker limit formula expresses $\mathcal{K}_{\Gamma_1}(z)$ explicitly in terms of the Dedekind η-function (see, e.g., [IK04, (22.42), (22.63)–(22.69)])

$$\mathcal{K}_{\Gamma_1}(z) \;=\; \frac{3}{\pi}\left(2\gamma - 2\frac{\zeta'}{\zeta}(2) - \log(4y|\eta(z)|^4)\right), \tag{A.9}$$

where $\gamma = 0.577\cdots$ is Euler's constant, $\zeta(s)$ is the Riemann zeta function, and $\eta(z)$ is the Dedekind eta function. In particular, using the special value $\eta(i)^4 = \frac{\Gamma(1/4)^4}{16\pi^3}$ we see that at $z = i$ we have

$$\mathcal{K}_{\Gamma_1}(i) \;=\; \frac{3}{\pi}\left(2\gamma - 2\frac{\zeta'}{\zeta}(2) - \log(\frac{\Gamma(1/4)^4}{4\pi^3})\right). \tag{A.10}$$

One can derive a similar formula for the congruence groups $\Gamma = \Gamma_0(p)$.

Proposition A.11. *For $\Gamma = \Gamma_0(p)$ we have*

$$\mathcal{K}_{\Gamma,\infty}(z) = \frac{3}{(p+1)\pi}\left(2\gamma - 2\frac{\zeta'}{\zeta}(2) - \log\left(\frac{4y|\eta(pz)|^{\frac{4p}{p-1}}}{|\eta(z)|^{\frac{4}{p-1}}}\right) - \frac{2\log(p)p^2}{(p^2-1)}\right),$$

and

$$\mathcal{K}_{\Gamma,0}(z)=\frac{3}{(p+1)\pi}\left(2\gamma-2\frac{\zeta'}{\zeta}(2)-\log\left(\frac{4y|\eta(z)|^{\frac{4p}{p-1}}}{|\eta(pz)|^{\frac{4}{p-1}}}\right)+\frac{\log(p)(p-1)^2}{(p^2-1)}\right).$$

Proof. Since representatives for $\Gamma_\infty\backslash\Gamma$ are given by matrices $\left(\begin{smallmatrix}* & * \\ c & d\end{smallmatrix}\right)$ with $c\geq 0$ integer with $c=0 \bmod p$ and $d\in\mathbb{Z}$ with $(c,d)=1$, after multiplying the Eisenstein series by

$$\zeta_p(2s)=\sum_{(n,p)=1} n^{-2s}=\zeta(2s)(1-p^{-2s}),$$

and expanding we get for $\mathfrak{Re}(s)>1$

$$\begin{aligned}
\zeta_p(2s)E_{\Gamma,\infty}(z,s) &= \zeta_p(2s)y^s+\sum_{(n,p)=1}\frac{1}{n^{2s}}\sum_{\substack{(c,d)=1\\ c=0 \bmod p}}\frac{y^s}{|cz+d|^{2s}}\\
&= \zeta_p(2s)y^s+\sum_{\substack{c=1\\ c=0 \bmod p}}^{\infty}\sum_{(d,p)=1}\frac{y^s}{|cz+d|^{2s}}\\
&= y^s\zeta_p(2s)+\sum_{\substack{c=1\\ c=0 \bmod p}}^{\infty}\sum_{d\in\mathbb{Z}}\frac{y^s}{|cz+d|^{2s}}-\frac{1}{p^{2s}}\sum_{c=1}^{\infty}\sum_{d\in\mathbb{Z}}\frac{y^s}{|cz+d|^{2s}}.
\end{aligned}$$

Using Poisson summation on the inner sum gives

$$\sum_{d\in\mathbb{Z}}\frac{1}{|cz+d|^{2s}}=\frac{\sqrt{\pi}\Gamma(s-\frac{1}{2})}{\Gamma(s)}(cy)^{1-2s}+(cy)^{1-2s}\sum_{m\neq 0}e(-mcx)\int_{\mathbb{R}}\frac{e(mcyt)}{(1+t^2)^s}dt$$

and dividing by $\zeta_p(2s)$ we see that

$$\begin{aligned}
E_{\Gamma,\infty}(z,s) &= y^s+\tfrac{(p-1)\sqrt{\pi}\Gamma(s-\frac{1}{2})\zeta(2s-1)}{(p^{2s}-1)\Gamma(s)\zeta(2s)}y^{1-s}\\
&\quad+\frac{py^{1-s}}{(p^{2s}-1)\zeta(2s)}\sum_{c=1}^{\infty}c^{1-2s}\sum_{m\neq 0}e(-mpcx)\int_{\mathbb{R}}\frac{e(mpcyt)}{(1+t^2)^s}dt\\
&\quad-\frac{y^{1-s}}{(p^{2s}-1)\zeta(2s)}\sum_{c=1}^{\infty}c^{1-2s}\sum_{m\neq 0}e(-mcx)\int_{\mathbb{R}}\frac{e(mcyt)}{(1+t^2)^s}dt.
\end{aligned}$$

The only pole comes from the term containing $\zeta(2s-1)=\frac{1}{2(s-1)}+\gamma+O(s-1)$. Subtracting the residue $\mathrm{Res}_{s=1}E_{\Gamma,\infty}(z,s)=\frac{3}{\pi(p+1)}$ and taking the limit as $s\to 1$ we get that

$$\mathcal{K}_{\Gamma,\infty}(z)=y+\lim_{s\to 1}\left(\frac{(p-1)\sqrt{\pi}\Gamma(s-\frac{1}{2})\zeta(2s-1)}{(p^{2s}-1)\Gamma(s)\zeta(2s)}y^{1-s}-\frac{3}{\pi(p+1)(s-1)}\right)$$

$$+\frac{p}{(p^2-1)\zeta(2)}\sum_{c=1}^{\infty}\frac{1}{c}\sum_{m\neq 0}e(-mpcx)\int_{\mathbb{R}}\frac{e(mpcyt)}{(1+t^2)}dt$$
$$-\frac{1}{(p^2-1)\zeta(2)}\sum_{c=1}^{\infty}\frac{1}{c}\sum_{m\neq 0}e(-mcx)\int_{\mathbb{R}}\frac{e(mcyt)}{(1+t^2)}dt.$$

We can evaluate the integral

$$\int_{\mathbb{R}}\frac{e(at)}{(1+t^2)}dt=\pi e^{-2\pi|a|}$$

to get that

$$\begin{aligned}\sum_{c=1}^{\infty}\frac{1}{c}\sum_{m\neq 0}e(-mpcx)\int_{\mathbb{R}}\frac{e(mpcyt)}{(1+t^2)}dt&=\pi\sum_{c=1}^{\infty}\frac{1}{c}(\sum_{m=1}^{\infty}e^{2\pi impcz}+e^{-2\pi impc\bar{z}})\\&=\pi\sum_{m=1}^{\infty}\sum_{c=1}^{\infty}\frac{1}{c}(e^{2\pi impcz}+e^{-2\pi impc\bar{z}})\\&=-\pi\sum_{m=1}^{\infty}(\log(1-e^{2\pi impz})+\log(1-e^{-2\pi imp\bar{z}})).\end{aligned}$$

Recalling that the Dedekind η-function is given by

$$\eta(z)=e^{\pi iz/12}\prod_{m=1}^{\infty}(1-e^{2\pi imz}),$$

we have that $\sum_{m=1}^{\infty}\log(1-e^{2\pi imz})=\log(\eta(z))-\frac{\pi iz}{12}$ hence

$$\sum_{c=1}^{\infty}\frac{1}{c}\sum_{m\neq 0}e(-mpcx)\int_{\mathbb{R}}\frac{e(mpcyt)}{(1+t^2)}dt=-\pi\log(|\eta(pz)|^2)-\frac{\pi^2py}{6}.$$

Similarly, we also have

$$\sum_{c=1}^{\infty}\frac{1}{c}\sum_{m\neq 0}e(-mcx)\int_{\mathbb{R}}\frac{e(mcyt)}{(1+t^2)}dt=-\pi\log(|\eta(z)|^2)-\frac{\pi^2y}{6}$$

so that

$$\mathcal{K}_{\Gamma,\infty}(z)=\lim_{s\to 1}\left(\frac{(p-1)\sqrt{\pi}\Gamma(s-\frac{1}{2})\zeta(2s-1)}{(p^{2s}-1)\Gamma(s)\zeta(2s)}y^{1-s}-\frac{3}{\pi(p+1)(s-1)}\right)$$
$$+\frac{6\log(|\eta(z)|^2|\eta(pz)|^{-2p})}{\pi(p^2-1)}.$$

Next, to compute the limit, write $\zeta(2s-1) = \frac{1}{2(s-1)} + \gamma + O(s-1)$ to get that the limit above is given by

$$\frac{6\gamma}{(p+1)\pi} + \lim_{s\to 1} \frac{\frac{(p-1)y^{1-s}\sqrt{\pi}\Gamma(s-\frac{1}{2})}{2(p^{2s}-1)\Gamma(s)\zeta(2s)} - \frac{3}{\pi(p+1)}}{s-1} = \frac{6\gamma}{(p+1)\pi} + \frac{d}{ds}\bigg|_{s=1} \left(\frac{(p-1)y^{1-s}\sqrt{\pi}\Gamma(s-\frac{1}{2})}{2(p^{2s}-1)\Gamma(s)\zeta(2s)}\right).$$

Finally, evaluate the derivative at $s=1$

$$\frac{d}{ds}\bigg|_{s=1} \left(\frac{(p-1)y^{1-s}\sqrt{\pi}\Gamma(s-\frac{1}{2})}{2(p^{2s}-1)\Gamma(s)\zeta(2s)}\right) = \frac{3}{(p+1)\pi}\left(-\log(y) - \frac{2\log(p)p^2}{(p^2-1)} - 2\log(2) - 2\frac{\zeta'}{\zeta}(2)\right)$$

to get that

$$\mathcal{K}_{\Gamma,\infty}(z) = \frac{3}{(p+1)\pi}\left(2\gamma - 2\frac{\zeta'}{\zeta}(2) - \log\left(\frac{4y|\eta(pz)|^{\frac{4p}{p-1}}}{|\eta(z)|^{\frac{4}{p-1}}}\right) - \frac{2\log(p)p^2}{(p^2-1)}\right).$$

The formula for the cusp at 0 now follows from (A.5) and (A.9). □

References

[Asa76] Tetsuya Asai, On the Fourier coefficients of automorphic forms at various cusps and some applications to Rankin's convolution, J. Math. Soc. Japan 28 (1) (1976) 48–61.

[DRS93] W. Duke, Z. Rudnick, P. Sarnak, Density of integer points on affine homogeneous varieties, Duke Math. J. 71 (1) (1993) 143–179.

[EM93] A. Eskin, C. McMullen, Mixing, counting and equidistribution in lie groups, Duke Math. J. 71 (1993) 143–180.

[FI13] J.B. Friedlander, H. Iwaniec, Small representations by indefinite ternary quadratic forms, in: Number Theory and Related Fields, in: Springer Proc. Math. Stat., vol. 43, Springer, New York, 2013, pp. 157–164.

[GHL15] Dorian Goldfeld, Joseph Hundley, Min Lee, Fourier expansions of $GL(2)$ newforms at various cusps, Ramanujan J. 36 (1–2) (2015) 3–42.

[HL94] Jeffrey Hoffstein, Paul Lockhart, Coefficients of Maass forms and the Siegel zero, Ann. of Math. (2) 140 (1) (1994) 161–181, with an appendix by Dorian Goldfeld, Hoffstein and Daniel Lieman.

[HKKL16] Thomas A. Hulse, Chan Ieong Kuan, Eren Mehmet Kıral, Li-Mei Lim, Counting square discriminants, J. Number Theory 162 (2016) 255–274.

[Iwa95] Henryk Iwaniec, Introduction to the Spectral Theory of Automorphic Forms, Madrid, 1995.

[IK04] Henryk Iwaniec, Emmanuel Kowalski, Analytic Number Theory, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[KK18] D. Kelmer, A. Kontorovich, Equidistribution of shears and applications, Math. Ann. 340 (1–2) (2018) 381–421.

[KS03] H. Kim, P. Sarnak, Refined estimates towards the Ramanujan and Selberg conjectures, J. Amer. Math. Soc. 16 (1) (2003) 175–181.

[Kon09] A. Kontorovich, The hyperbolic lattice point count in infinite volume with applications to sieves, Duke Math. J. 149 (1) (2009) 1–36, arXiv:0712.1391.

[Mar04] Grigoriy A. Margulis, On some aspects of the theory of Anosov systems, in: Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2004, with a survey by Richard Sharp: Periodic orbits of hyperbolic flows, translated from the Russian by Valentina Vladimirovna Szulikowska.

[OS14] Hee Oh, Nimish A. Shah, Limits of translates of divergent geodesics and integral points on one-sheeted hyperboloids, Israel J. Math. 199 (2) (2014) 915–931.

[Str04] A. Strombergsson, On the uniform equidistribution of long closed horocycles, Duke Math. J. 123 (2004) 507–547.

[Vas96] I. Vassileva, Dedekind Eta Function, Kronecker Limit Formula and Dedekind Sum for the Hecke Group, U Mass Amherst PhD Thesis, 1996.