AVERAGING OVER HEEGNER POINTS IN THE HYPERBOLIC CIRCLE PROBLEM

YIANNIS N. PETRIDIS AND MORTEN S. RISAGER

Abstract. For $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ the hyperbolic circle problem aims to estimate the number of elements of the orbit Γz inside the hyperbolic disc centered at z with radius $\cosh^{-1}(X/2)$. We show that, by averaging over Heegner points z of discriminant D, Selberg's error term estimate can be improved, if D is large enough. The proof uses bounds on spectral exponential sums, and results towards the sup-norm conjecture of eigenfunctions, and the Lindelöf conjecture for twists of the L-functions attached to Maaß cusp forms.

1. Introduction

For $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ we consider the standard point-pair invariant on the upper half-plane $\mathbb{H}$ given by

$$u(z,w) = \frac{|z-w|^2}{4\Im(z)\Im(w)}.$$

The hyperbolic circle problem aims to find good estimates, as $X \to \infty$, on

$$N(z,w,X) - \frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}, \tag{1.1}$$

where

$$N(z,w,X) = \#\{\gamma \in \Gamma; 4u(\gamma z, w) + 2 \leq X\}.$$

The best known estimate in this direction is due to Selberg who proved that

$$N(z,w,X) - \frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} = O(X^{2/3}). \tag{1.2}$$

The corresponding estimate for the error term of the hyperbolic circle problem modified to exclude small eigenvalues is also known for any cofinite group and has not been improved for any group or any z, w. The conjectural bound for the error term $O(X^{1/2+\varepsilon})$ would be optimal except for X^ε possibly being replaced by powers of $\log X$, see [27].

We investigate averages of (1.1) with $z = w$ when z averages over the set of Heegner points: consider the Γ-orbits of binary quadratic forms

$$ax^2 + bxy + cy^2$$

Date: January 26, 2017.
2010 *Mathematics Subject Classification.* Primary 11F72; Secondary 11E45.
The second author was supported by a Sapere Aude grant from The Danish Council for Independent Research (Grant-id:0602-02161B).

of negative fundamental discriminant $b^2-4ac=D<0$, and with $a>0$. Given such an orbit one associates the corresponding Heegner point in $\mathbb{H}$ given by

$$z=\frac{-b+i\sqrt{|D|}}{2a},$$

where we choose the representative of the orbit such that z lies in the usual fundamental domain of Γ. Denote the set of Heegner points of discriminant D by Λ_D and the class number by $h(D)=\#\Lambda_D$. The order of growth of $h(D)$ is controlled by the estimates

$$|D|^{1/2-\varepsilon}\ll_\varepsilon h(D)\ll |D|^{1/2}\log|D|, \tag{1.3}$$

where the lower bound is a strong but ineffective result of Siegel, see e.g. [8, Ch. 21].

Duke [9] showed that Heegner points become equidistributed on $\Gamma\backslash\mathbb{H}$, i.e. for f a smooth compactly supported function on $\Gamma\backslash\mathbb{H}$ we have, as $D\to\infty$,

$$\frac{1}{h(D)}\sum_{z\in\Lambda_D}f(z)\to\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}}fd\mu(z). \tag{1.4}$$

In [26, Th. 1.1] we have improved on Selberg's bound (1.2) when we average the center locally in $\Gamma\backslash\mathbb{H}$. Therefore, we may also suspect an improvement in (1.2) if we make a *discrete* average over Heegner points. Our main theorem confirms that this is indeed the case.

Theorem 1.1. *Let f be a smooth compactly supported non-negative function on $\Gamma\backslash\mathbb{H}$. Then*

$$\frac{1}{h(D)}\sum_{z\in\Lambda_D}f(z)\left(N(z,z,X)-\frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\right)=O_{f,\varepsilon}(X^{7/12+\varepsilon}+X^{4/5+\varepsilon}D^{-4/165+\varepsilon}).$$

Remark 1.2. We notice that for $D\geq X^{11/2+\varepsilon}$ this is better than what we would get using Selberg's bound (1.2).

Let $\lambda_j=1/4+t_j^2$, $t_j\geq 0$, $j\geq 1$ be the cuspidal eigenvalues of the automorphic Laplacian on $L^2(\Gamma\backslash\mathbb{H})$ listed with multiplicity, and u_j the corresponding Hecke–Maaß cusp forms normalized to have L^2-norm equal to one. We remark that the cuspidal eigenvalues satisfy Weyl's law, see e.g. [18, Eq. (11.5)]

$$\sum_{|t_j|\leq T}1=\frac{\mathrm{vol}(\Gamma\backslash\mathbb{H})}{4\pi}T^2+O(T\log T). \tag{1.5}$$

We also set $\lambda_0=0$, $t_0=i/2$ to correspond to the constant eigenfunction $(\mathrm{vol}(\Gamma\backslash\mathbb{H}))^{-1/2}$.

Let $L(u_j\times\chi_D,1/2)$ be the central value of the L-function of u_j twisted by the odd primitive quadratic character χ_D with conductor $|D|$.

Let u be a Maaß cusp form with eigenvalue $1/4+t^2$ or an Eisenstein series $E(z,1/2+it)$. In our proof of Theorem 1.1 we are using approximations to the following three conjectures:

(C1) The Lindelöf conjecture for $L(u\times\chi_D,1/2)$:

$$L(u\times\chi_D,1/2)=O_\varepsilon(((1+|t|)D)^\varepsilon).$$

(C2) The sup-norm conjecture, i.e.

$$\sup_{z\in K}|u(z)| = O_{K,\varepsilon}((1+|t|)^{\varepsilon})$$

for any compact set K.

(C3) Bounds on spectral exponential sums

$$\sum_{|t_j|\le T} X^{it_j} = O_{\varepsilon}(X^{\varepsilon}(1+T)^{1+\varepsilon}),$$

see [26].

Assuming these three conjectures we may improve Theorem 1.1 and prove the following conditional result:

Theorem 1.3. *Let f be a smooth compactly supported non-negative function on $\Gamma\backslash\mathbb{H}$. Then assuming (C1), (C2), and (C3) we have*

$$\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left(N(z,z,X)-\frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\right) = O_{f,\varepsilon}(X^{1/2+\varepsilon}+X^{4/5+\varepsilon}D^{-1/10+\varepsilon}).$$

Remark 1.4. Theorem 1.3 implies that if $D\ge X^{4/3+\varepsilon}$ then this is better than what we would get using Selberg's bound (1.2), and if $D\ge X^3$ we have the bound $O(X^{1/2+\varepsilon})$.

In order to prove the unconditional bound in Theorem 1.1 we use approximations to (C1), (C2), and (C3). For an approximation to (C1) we use a recent theorem by Young [36] on the cubic moment of the L-function over short sums, see Theorem 2.1 below. To address (C3) we use the following estimate due to Sarnak and Luo [23, Eq. (58)]:

$$\sum_{|t_j|\le T} X^{it_j} = O_{\varepsilon}(X^{1/8}T^{5/4+\varepsilon}). \tag{1.6}$$

As an approximation to (C2) we prove the following average bound, which may be of independent interest:

Theorem 1.5. *We have*

$$T^{-2}\sum_{1\le t_j\le T}\|u_j\|_{\infty}^2 = O_{\varepsilon}(T^{2(1/2-1/8)+\varepsilon}).$$

Remark 1.6. This shows that on average $\|u_j\|_\infty$ is of size $O(|t|^{3/8+\varepsilon})$. This is an improvement on average of the individual bound $\|u_j\|_\infty = O(|t|^{5/12+\varepsilon})$ due to Iwaniec and Sarnak [20, Eq. (A.15)]. The same average bound on sup norms over compact sets follows from recent results of Jung, in particular [22, Cor. 1.15]. The individual bound $\|u_j\|_\infty = O(|t|^{3/8+\varepsilon})$ would follow if we could prove expected lower bounds on the mean square for the Hecke eigenvalues, see [20, Remark 1.6]. Moreover, Young [37, Th. 1.1] proved the following bound for the sup-norm of Eisenstein series on a compact set K:

$$\|E(z,1/2+it)\|_{\infty,K} = O_{K,\varepsilon}((1+|t|)^{1/2-1/8+\varepsilon}). \tag{1.7}$$

This has more recently been improved to exponent $1/2-1/6+\varepsilon$ by Blomer [2, Th. 1.1].

Remark 1.7. The situation is much easier when we fix w and average over $z \in \Lambda_D$. We get the following estimate:

$$\frac{1}{h(D)} \sum_{z\in\Lambda_D} \left(N(z,w,X) - \frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} \right) = O_\varepsilon(X^{1/2+\varepsilon}) \tag{1.8}$$

for $D \geq X^3$. On GRH we have (1.8) for $D \geq X^1$. We omit the proof. It is simpler than the proof of Theorem 1.1, because we can separate the average over $z \in \Lambda_D$ for $u_j(z)$ and $E(z, 1/2+it)$ from $u_j(w)$ and $E(w, 1/2+it)$ in the pre-trace formula.

Acknowledgements. We thank Gergely Harcos and Péter Maga for useful comments on sup-norm bounds. Also we thank the anonymous referees for useful comments, and for informing us about the recent work of Blomer [2] and Jung [22].

2. Equidistribution of Heegner points

Duke's proof of the equidistribution of Heegner points (1.4) involves non-trivial bounds on Fourier coefficients of half-integral weight Maaß forms and uses the Kuznetsov formula. The technique is due to Iwaniec [16], who proved non-trivial bounds for the Fourier coefficients of holomorphic forms of half-integral weight by using the Petersson formula to relate them to Kloosterman sums $K(m,n,c)$ and exhibited cancellations for sums of these as c varies. Duke proved the following bounds on the average of eigenfunctions, the so-called 'Weyl sums', see [9, p. 89]:

$$\frac{1}{h(D)} W_c(D,t_j) = \frac{1}{h(D)} \sum_{z\in\Lambda_D} u_j(z) \ll_\varepsilon |t_j|^A D^{-1/28+\varepsilon},$$
$$\frac{1}{h(D)} W_E(D,t) = \frac{1}{h(D)} \sum_{z\in\Lambda_D} E(z,1/2+it) \ll_\varepsilon |t|^A D^{-1/28+\varepsilon}.$$

One can then use standard approximation techniques to prove (1.4).

To improve on these bounds we use the fact that Weyl sums are connected to L-functions. We assume that the Maaß cusp forms u_j are also eigenfunctions of all Hecke operators. We quote [36, Eq. (2.2)] for the following Waldspurger–Zhang type formula:

$$|W_c(D,t_j)|^2 = \frac{\sqrt{|D|} L(u_j \times \chi_D, 1/2) L(u_j, 1/2)}{2L(\mathrm{sym}^2 u_j, 1)}. \tag{2.1}$$

Similarly we have

$$W_E(D,t) = \left(\frac{\sqrt{|D|}}{2} \right)^{1/2+it} \frac{L(1/2+it,\chi_D)\zeta(1/2+it)}{\zeta(1+2it)}, \tag{2.2}$$

see [19, Eq. 22.45].

Young recently proved the following bound:

Theorem 2.1. [36, Thm. 1.1] *For the third moment of twists of the L-functions of Maaß forms we have*

$$\sum_{T\leq t_j\leq T+1} L(u_j\times\chi_D, 1/2)^3 + \int_T^{T+1} |L(1/2+it,\chi_D)|^6 \ll_\varepsilon (|D|(1+T))^{1+\varepsilon}.$$

In our formulation we have used the positivity of central values, see [21, Th. 1] to restrict to Maaß forms for the full modular group.

Lemma 2.2. *The following estimates on short averages of the Weyl sums hold:*

$$\frac{1}{h(D)^2}\sum_{T\leq t_j\leq T+1}|W_c(D,t_j)|^2 \ll_\varepsilon D^{-\frac{1}{6}+\varepsilon}(1+T)^{1+\varepsilon},$$

$$\frac{1}{h(D)^2}\int_T^{T+1}|W_E(D,t)|^2dt \ll_\varepsilon D^{-\frac{1}{6}+\varepsilon}(1+T)^{1+\varepsilon}.$$

Proof. It follows from the work of Ivić [15, Eq. (1.9)] that

$$\sum_{T\leq t_j\leq T+1} L(u_j,1/2)^3 = O_\varepsilon((1+T)^{1+\varepsilon}) \tag{2.3}$$

and the bound

$$\int_T^{T+1}|\zeta(1/2+it)|^6\,dt = O_\varepsilon((1+T)^{1+\varepsilon}) \tag{2.4}$$

follows from the classical Weyl estimate on the Riemann zeta function on the critical line, see e.g. [35, Th. 5.5].

We will also need lower bounds

$$L(\mathrm{sym}^2u_j,1)\gg_\varepsilon |t_j|^{-\epsilon}, \text{ and } \zeta(1+it)\gg \log\log|t|\,/\log|t|\,. \tag{2.5}$$

The first bound is due to Lockhart and Hoffstein [14, Th. 0.2], and the second is classical [35, Th. 5.17].

We use first (2.1) and (2.2) and the Hölder inequality with exponents $(1/3,1/3,1/3)$ so that we can apply Theorem 2.1. With the help of (2.3), (2.4), (2.5), and (1.5) we get

$$\sum_{T\leq t_j\leq T+1}|W_c(D,t_j)|^2 \ll_\varepsilon (D^{\frac{5}{6}}(1+T))^{1+\varepsilon},$$

$$\int_T^{T+1}|W_E(D,t)|^2dt \ll_\varepsilon (D^{\frac{5}{6}}(1+T))^{1+\varepsilon}.$$

Using the lower bound in (1.3) for $h(D)$, we establish the claim. $\square$

Remark 2.3. We notice that on GRH, or more precisely (C1), we may replace $D^{-1/6+\varepsilon}$ by $D^{-1/2+\varepsilon}$ in both bounds.

3. Sup-norm estimates

The general bound for the sup-norm is

$$\|u_j\|_\infty = O((1+|t_j|)^{1/2}), \tag{3.1}$$

see e.g. [32] or Theorem 4.5 below. This is sometimes called the convexity bound for sup-norms. In a ground-breaking work Iwaniec and Sarnak [20] showed that this bound may be improved for $\mathrm{PSL}_2(\mathbb{Z})$ to

$$\|u_j\|_\infty = O_\varepsilon((1+|t_j|)^{1/2-1/12+\varepsilon}). \tag{3.2}$$

Denote by $\|u\|_{\infty,K}$ the supremum of a function u restricted to the set K. They also conjectured that if we restrict to a compact set K the sup-norm is essentially bounded:

$$\|u_j\|_{\infty,K} = O_\varepsilon((1+|t_j|)^\varepsilon).$$

In [20] the restriction to compact sets is not explicit but is now known to be needed, see [31]. More precisely we have

$$\|u_j\|_\infty \geq C_\varepsilon(1+|t_j|)^{1/6-\varepsilon}.$$

This is a purely analytic fact that follows from the properties of K-Bessel functions.

After the work of Iwaniec and Sarnak there has been a lot of results about subconvexity of the sup-norm in the level aspect [3], weight aspect [7], for holomorphic forms, hybrid bounds [34, 29], as well as other groups [5, 24]. The bound (3.2) has not been improved.

We now discuss average bounds on sup-norms. In particular we prove Theorem 1.5, which states that on average in a window of size T we can improve (3.2) to $\|u_j\|_\infty = O((1+|t_j|)^{1/2-1/8+\varepsilon})$.

The Maaß cusp forms have Fourier expansions

$$u_j(z) = \sum_{n\neq 0} \rho_j(n)\sqrt{y}K_{it_j}(2\pi\,|n|\,y)e^{2\pi inx}.$$

If u_j is also a Hecke eigenform with Hecke eigenvalues $\lambda_j(n)$, then $\rho_j(n) = \rho_j(\mathrm{sign}(n))\lambda_j(|n|)$. It is known [10, Prop. 19.6] that

$$\sum_{n\leq x} |\lambda_j(n)|^2 = O_\varepsilon(x^{1+\varepsilon}\,|t_j|^\varepsilon). \tag{3.3}$$

Moreover, we set $\rho_j(n) = \cosh(\pi t_j/2)v_j(n)$, so that $v_j(n) = v_j(1)\lambda_j(n)$. It is known [17, 14] that

$$|t_j|^{-\varepsilon} \ll_\varepsilon |v_j(1)| \ll_\varepsilon |t_j|^\varepsilon\,. \tag{3.4}$$

Proof of Theorem 1.5. This is an adaptation of the proof for individual bounds in [20]. We quote [4, p. 678] for the following crucial inequality: assume that $T \leq t_j \leq T+1$. Then

$$|u_j(z)|^2\left|\sum_{l\leq L}\alpha_l\lambda_j(l)\right|^2 \ll_\varepsilon (LT)^\varepsilon\left(T\sum_{l\leq L}|\alpha_l|^2 + (L+y)T^{1/2}\Big(\sum_{l\leq L}|\alpha_l|\Big)^2\right)$$

for every sequence α_n. This improves on [20, Eq. (A.12)] by replacing $yL^{1/2}$ by y.

Now we choose a smooth non-negative function h supported in $[1,2]$ with integral $\int_{\mathbb{R}} h(t)dt = 1$, and consider $h_N(t) = h(t/N)$. By choosing $\alpha_n = h_N(n)\lambda_j(n)\,|v_j(1)|^2$ and using Cauchy–Schwarz on the last sum, we arrive at

$$
\begin{aligned}
&|u_j(z)|^2 \left| \sum_{n\le 2N} h_N(n)\,|v_j(n)|^2 \right|^2 \\
&\ll_\epsilon (TN)^\varepsilon \left(T+(N+y)T^{1/2}N\right) \sum_{n\le 2N} h_N(n)^2\,|v_j(n)|^2 \\
&\ll_\epsilon (NT)^\epsilon \left(T+(N+y)T^{1/2}N\right) N,
\end{aligned} \tag{3.5}
$$

where we have used (3.3) and (3.4). Luo and Sarnak [23, p. 233] proved Iwaniec' mean Lindelöf conjecture for the Rankin–Selberg convolution in the spectral aspect, which allowed them to prove that

$$
\sum_{n=1}^{2N} h_N(n)\,|v_j(n)|^2 = \frac{12}{\pi^2}N + r(t_j,N), \tag{3.6}
$$

where the reminder is of size $N^{1/2}$ on average:

$$
\sum_{t_j\le T} |r(t_j,N)| = O_\varepsilon(T^{2+\varepsilon}N^{1/2}). \tag{3.7}
$$

When we square the right-hand side of (3.6), the term $(r(t_j,N))^2$ may be dropped by positivity and we find

$$
|u_j(z)|^2 \ll_\varepsilon N^{-2}\left((NT)^\epsilon \left(T+(N+y)T^{1/2}N\right)N + |u_j(z)|^2\,N\,|r(t_j,N)|\right). \tag{3.8}
$$

We first consider the set $A_N=\{z\in\Gamma\backslash\mathbb{H}; y\le N\}$. We use the subconvexity bound (3.2) on u_j on the right-hand side of (3.8) to see that

$$
\|u_j\|^2_{\infty,A_N} \ll_\varepsilon N^{-2}\left((NT)^\epsilon \left(T+NT^{1/2}N\right)N + T^{2(1/2-1/12)}N\,|r(t_j,N)|\right). \tag{3.9}
$$

Averaging over t_j we find by (3.7) and Weyl's law (1.5) that

$$
\sum_{1\le t_j\le T} \|u_j\|^2_{\infty,A_N} \ll_\varepsilon (NT)^\varepsilon N^{-2}\left(T^3N + N^3T^{5/2} + T^{2+2(1/2-1/12)}N^{3/2}\right).
$$

We choose $N=T^{1/4}$ so that the right-hand side is $O(T^{3-1/4+\varepsilon})$.

For the complement of the set A_N, i.e. for $B_N=\{z\in\Gamma\backslash\mathbb{H}; y>N\}$ we argue as follows: we have set $N=T^{1/4}$. We bound $\|u_j\|^2_{\infty,B_N}$ individually as $O(T^{2(1/2-1/8)+\varepsilon})$ using the following simple upper bound

$$
|u_j(z)| \ll_\varepsilon t_j^\varepsilon\left((t_j/y)^{1/2} + t_j^{1/6}\right),
$$

see [31, Lemma A.1′]. We finish the proof by noticing that

$$
\|u_j\|^2_\infty = \max(\|u_j\|^2_{\infty,A_N}, \|u_j\|^2_{\infty,B_N}).
$$

□

If we consider sup-norms of averages instead of averages of sup-norms we have much better bounds.

Proposition 3.1. *Let $\{u_j\}$ be the L^2-normalized Maaß cusp forms for $PSL_2(\mathbb{Z})$. Then*

$$\Big\| \sum_{T\leq t_j\leq 2T} |u_j(z)|^2 \Big\|_\infty = O(T^2).$$

Proof. Notice first that, if we restrict z to a compact set K the inequality in [18, Prop. 7.2] gives

$$\sum_{T\leq t_j\leq 2T} |u_j(z)|^2 = O(T^2),$$

where the implied constant depends on K and the group, but *not* on z. Even if we do not restrict to a compact set the same bound holds for $y\leq 2T$, say, by the same inequality. To bound the average for $y\geq 2T$ we use the decay properties of the K-Bessel function. We use the integral representation

$$K_{it}(y)=\int_0^\infty e^{-y\cosh(v)}\cosh(itv)dv,$$

see [18, p. 205]. Fix $y\geq 0$. For t real we have $|\cosh(itv)|\leq 1$. We use $\cosh(v)\geq 1+v^2/2$ to get

$$|K_{it}(y)|\leq \int_0^\infty e^{-y(1+v^2/2)}dv\leq \frac{\sqrt{\pi}e^{-y}}{\sqrt{2}\sqrt{y}}.$$

The bound

$$\rho_j(n)=O(e^{\pi t_j/2}\sqrt{|n|}t_j^\varepsilon)$$

follows trivially from (3.3) and (3.4). We can now prove good decay properties for $|u_j|$ when y is large compared to t_j, e.g. when $y>2T$ and $T\leq t_j\leq 2T$ we have

$$|u_j(z)|=O\left(\sum_{n=1}^\infty e^{\pi t_j/2}\sqrt{n}t_j^\varepsilon\sqrt{y}\frac{e^{-2\pi ny}}{\sqrt{ny}}\right)=O(e^{-T}).$$

The claim follows using Weyl's law (1.5). □

Remark 3.2. We note that the proof of Proposition 3.1 is much simpler than that of Theorem 1.5. The only input is the use of the local Weyl law [18, Prop. 7.2] and bounds on the Fourier coefficients that are uniform in t_j and n.

We also need a similar result for the Eisenstein series.

Proposition 3.3. *Let K be a compact set on $\Gamma\backslash\mathbb{H}$. Then*

$$\left\|\int_T^{2T} |E(z,1/2+it)|^2\,dt\right\|_{\infty,K} = O(T^2).$$

Proof. This follows directly from [18, Prop. 7.2]. □

4. Bounds for weight k eigenfunctions.

In this section we discuss bounds on weight k eigenfunctions. We need this later when we estimate certain inner products of eigenfunctions. This is important in order to bound derivatives of cusp forms and Eisenstein series in the proof of Theorem 5.1 below.

For the following discussion we adopt the notation and terminology from Fay [11]. In this section we allow Γ to be any discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$ unless explicitly stated otherwise.

Let k be an integer and let $\mathfrak{F}_k$ be the space of functions $f : \mathbb{H} \to \mathbb{C}$ satisfying

$$f(\gamma z) = \left(\frac{cz+d}{c\bar{z}+d}\right)^k f(z), \quad \gamma \in \Gamma.$$

The k-Laplacian is defined by

$$\Delta_k = y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - 2iky\frac{\partial}{\partial x}.$$

We write the eigenvalue equation as $\Delta_k u + s(1-s)u = 0$ with $s = 1/2 + it$. The raising and lowering operators are defined on $\mathfrak{F}_k$ by

$$K_k = (z-\bar{z})\frac{\partial}{\partial z} + k, \quad L_k = (\bar{z}-z)\frac{\partial}{\partial \bar{z}} - k.$$

It is well-known that if $f \in \mathfrak{F}_k$ is differentiable, then $K_k f \in \mathfrak{F}_{k+1}$, while $L_k f \in \mathfrak{F}_{k-1}$. Moreover, $\Delta_{k+1}K_k = K_k\Delta_k$, $\Delta_k L_{k+1} = L_{k+1}\Delta_{k+1}$. It is clear that if f is an eigenfunction of Δ_k with eigenvalue λ, i.e. $(\Delta_k + \lambda)f = 0$, then $K_k f$ (resp. $L_k f$) is an eigenfunction of Δ_{k+1} (resp. Δ_{k-1}) with the same eigenvalue. For $f \in \mathfrak{F}_k$ and $g \in \mathfrak{F}_l$ we have the product rules: if k, l are integers, then

$$K_{k+l}(fg) = (K_k f)g + f(K_l g), \quad L_{k+l}(fg) = (L_k f)g + f(L_l g). \tag{4.1}$$

We use polar coordinates centered at a point z_0, defined through

$$\frac{z-z_0}{z-\bar{z}_0} = \tanh(r/2)e^{i\theta},$$

with $r = r(z, z_0)$ the hyperbolic distance and $\theta = \theta(z, z_0) \in [0, 2\pi]$. Set $v = \cosh r$. We remark that in polar coordinates $d\mu(z) = \sinh r\, dr\, d\theta$. We need the radial expansion of f, given in Theorem 1.2 in [11]. For f an eigenfunction of Δ_k with eigenvalue $s(1-s)$ on a disk $r(z, z_0) < R$ we have

$$f(z)\left(\frac{z-\bar{z}_0}{z_0-\bar{z}}\right)^k = \sum_{n\in\mathbb{Z}} f_n(z_0)P^n_{s,k}(z, z_0)e^{in\theta} \tag{4.2}$$

with $P^n_{s,k}(z, z_0) = P^n_{s,k}(r)$ given by

$$P^n_{s,k}(r) = \left(\frac{v-1}{v+1}\right)^{|n|/2}\left(\frac{2}{1+v}\right)^s F(s-k_n, s+k_n+|n|, 1+|n|; (v-1)/(v+1)).$$

Here $F(a, b, c; z)$ is the Gauss hypergeometric function and $k_n = kn/|n|$ for $n \neq 0$ and $k_0 = k$. We can recover the coefficients $f_n(z_0)$ by the formula

$$n!f_n(z_0) = n!\overline{\bar{f}_{-n}(z_0)} = K_{k+n-1}K_{k+n-2}\cdots K_{k+1}K_k f(z_0), \quad n > 0, \tag{4.3}$$

see [11, Eq. 23].

In this section we will prove the following lemma, which is crucial in the proof of Theorem 5.1.

Lemma 4.1. *Let $f \in \mathfrak{F}_k$ be an eigenfunction of Δ_k with eigenvalue $s(1-s)$ and let C be a compact subset of $\Gamma\backslash\mathbb{H}$. Then there exists a compact subset C' of $\Gamma\backslash\mathbb{H}$ containing C such that*

$$\left\|K_k f\right\|_{\infty,C} \ll_k |s| \left\|f\right\|_{\infty,C'}, \quad \left\|L_k f\right\|_{\infty,C} \ll_k |s| \left\|f\right\|_{\infty,C'}.$$

Proof. We set $h(z) = f(z)(z-\bar{z}_0)^k/(z_0-\bar{z})^k$. Since $P^n_{s,k}(z_0,z_0) = 0$ for $n \neq 0$ and $P^0_{s,k}(z_0,z_0) = 1$, we have $f(z_0) = f_0(z_0)$. By (4.2) we get

$$\int_0^{2\pi} h(z)e^{-in\theta}\,d\theta = 2\pi f_n(z_0)P^n_{s,k}(z,z_0).$$

Let $A = A(r_1,r_2) = \{z; r_1 \leq r(z,z_0) \leq r_2\}$ be a disc or annulus centered at z_0. We multiply by $\overline{P^n_{s,k}(r)}$, and integrate the radial variable to get

$$\int_A h(z)e^{-in\theta}\overline{P^n_{s,k}(z,z_0)}\,d\mu(z) = \int_A f_n(z_0)|P^n_{s,k}(z,z_0)|^2\,d\mu(z).$$

This implies the crucial identity

$$f_n(z_0) = \frac{\int_A h(z)e^{-in\theta}\overline{P^n_{s,k}(z,z_0)}\,d\mu(z)}{\int_A |P^n_{s,k}(z,z_0)|^2\,d\mu(z)}. \tag{4.4}$$

We apply (4.4) for $n=1$ and choose A to be the disc of radius $c|s|^{-1}$, for a sufficiently small constant c to be chosen. We get

$$|f_1(z_0)| \leq \frac{\|h\|_{\infty,A}\int_A |P^1_{s,k}(r)|d\mu(z)}{\int_A |P^1_{s,k}(r)|^2 d\mu(z)}.$$

If we can prove that

$$\int_A |P^1_{s,k}(r)|d\mu(z) \ll_k |s|^{-3}, \quad \int_A |P^1_{s,k}(r)|^2 d\mu(z) \gg_k |s|^{-4}, \tag{4.5}$$

then, by (4.3), we have $|K_k f(z_0)| \ll_k |s| \left\|f\right\|_{\infty,A}$. By compactness of C, we can find a compact set $C' \subset \mathbb{H}$ containing C such that

$$\left\|K_k f\right\|_{\infty,C} \ll_k |s| \left\|f\right\|_{\infty,C'}.$$

To prove (4.5) we need to study the asymptotics of $P^n_{s,k}(r)$ jointly for r small and $t = \Im(s) \to \infty$. For simplicity let n be nonnegative. We can approximate the hypergeometric function $F(a,b,c;z)$ by its Taylor polynomials, see [12, Eq. (4.13), (4.14)]:

$$F(a,b,c;z) = \sum_{j=0}^{J-1}\frac{(a)_j(b)_j}{(c)_j j!}z^j + O\left(\left|\frac{(a)_J(b)_J z^J}{(c)_J J!}\right|\right)$$

uniformly in a,b,c as long as

$$|z|\max_{j\geq 0}\left|\frac{(a+j)(b+j)}{(c+j)(j+1)}\right| \leq \frac{1}{2}. \tag{4.6}$$

It is easily verified that if $|r| < c|s|^{-1}$, then the condition of (4.6) is satisfied and we can apply the Taylor series for $F(a,b,c,z)$ with $J = 2$ to get

$$P^n_{s,k}(r) = \left(\frac{r}{2}\right)^n \left(1 + \left(\frac{(s-k)(s+k+n)}{4(1+n)} - \frac{s}{4} - \frac{n}{12}\right) r^2 + O_k(s^4 r^4)\right),$$

cf. [11, Eq. 17]. We specialize to $n = 1$ in the simpler form

$$P^1_{s,k}(r) = (r/2)(1 + O_k(s^2 r^2)).$$

We have

$$\int_0^{c/|s|} r^a \sinh r dr \sim c^{a+2} \frac{1}{(a+2)|s|^{a+2}}, \quad |s| \to \infty. \tag{4.7}$$

To investigate (4.5) we integrate in the θ variables. Then we apply (4.7) for $a = 1$ and $a = 3$ to get the first inequality in (4.5). For the second inequality we apply it for $a = 2, 4, 6$. The terms with $a = 4, 6$ get multiplied by $|s|^2$ and $|s|^4$ respectively. All three terms are of the same order of decay, i.e. $|s|^{-4}$. However, when we take into account the constant in the O_k and the powers of c, we can choose c sufficiently small to make the first term the dominant term.

Finally we prove $\|L_k f\|_{\infty,C} \ll_k |s| \, \|f\|_{\infty,C'}$: if $f \in \mathfrak{F}_k$, then $\bar{f} \in \mathfrak{F}_{-k}$. We only need to observe that $L_k = \overline{K_{-k}}$, see [11, Eq. (3)]. □

Remark 4.2. Let $f \in \mathfrak{F}_k$. For $j = 1, 2, \dots, m$ let A_j be either a lowering or a raising operator such that $A_m A_{m-1} \cdots A_1 f$ makes sense. Repeated use of Lemma 4.1 shows that

$$\|A_m A_{m-1} \cdots A_1 f\|_{\infty,C} \ll_{m,k} |s|^m \, \|f\|_{\infty,C'} .$$

We now consider L^2-norms of eigenfunctions. We denote by $\|f\|_{2,C}$ the L^2-norm of f restricted to C. Assume now that $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$. Moreover, for the corresponding Eisenstein series $E_k(z,s)$ of weight k we denote $E^Y_k(z,s)$ the function $E_k(z,s) - \chi_{[Y,\infty)}(y)(y^s + \phi_k(s) y^{1-s})$. It is well-known that $\phi_0(s) = \xi(2s-1)/\xi(2s)$, where $\xi(s)$ is the completed Riemann zeta function.

Lemma 4.3. *Let C be a compactly supported set. Assume that $C \subset \{z \in \mathbb{H}; \Im(z) \leq Y\}$. Then*

(i) for $f \in \mathfrak{F}_k$ an L^2-eigenfunction of Δ_k with eigenvalue $1/4 + t^2$ we have

$$\|K_k f\|_2 , \ll_k |t| \, \|f\|_2 , \quad \|L_k f\|_2 \ll_k |t| \, \|f\|_2 ,$$

(ii) for a weight k Eisenstein series $E_k(z,s)$ we have

$$\|K_k E_k(\cdot, 1/2 + it)\|_{2,C} \ll_k |t| \left\| E^Y_k(\cdot, 1/2 + it) \right\|_2 ,$$
$$\|L_k E_k(\cdot, 1/2 + it)\|_{2,C} \ll_k |t| \left\| E^Y_k(\cdot, 1/2 + it) \right\|_2 ,$$

for $|t| \geq 1$.

Proof. The claim in (i) follows from [28, Satz 3.1].

For (ii) we note that

$$K_k E_k(z,s) = (s+k) E_{k+1}(z,s), \quad L_k E_k(z,s) = (s-k) E_{k-1}(z,s),$$

cf. [28, Eq. (10.8), (10.9)]. It suffices to consider the L^2-norms of $E_{k\pm 1}^Y(z, 1/2+it)$. The Maaß–Selberg relations for Eisenstein series of weight k give

$$\left\|E_k^Y(\cdot, 1/2+it)\right\|^2 = -\frac{\phi_k'(1/2+it)}{\phi_k(1/2+it)} + O_Y(1),$$

see [28, Lemma 11.2, p. 301]. Moreover, see [28, Eq. (10.26)], we have

$$\phi_{k+1}(s) = \frac{k+1-s}{k+s}\phi_k(s). \tag{4.8}$$

This gives

$$\left\|E_{k+1}^Y(\cdot, 1/2+it)\right\|^2 = -\frac{\phi_k'(1/2+it)}{\phi_k(1/2+it)} + O_Y(1) = \left\|E_k^Y(\cdot, 1/2+it)\right\|^2 + O_Y(1).$$

It follows recursively from (4.8) that $-\frac{\phi_k'(1/2+it)}{\phi_k(1/2+it)} = -\frac{\phi_0'(1/2+it)}{\phi_0(1/2+it)} + o(1)$, as $|t| \to \infty$. Since

$$-\frac{\phi_0'(1/2+it)}{\phi_0(1/2+it)} = 2\Re\frac{\Gamma'}{\Gamma}(1/2+it) + O\left(\frac{\zeta'(1+it)}{\zeta(1+it)}\right) + O(1), \tag{4.9}$$

the result follows using Stirling's formula, which gives $\Gamma'(s)/\Gamma(s) \sim \log s$, and Weyl's bound $\zeta'(1+it)/\zeta(1+it) \ll \log t/\log\log t$. □

Remark 4.4. Let f be an L^2-eigenfunction of Δ_k or a weight k Eisenstein series $E_k(z,s)$. For $j = 1, 2, \ldots, m$ let A_j be either a lowering or a raising operator such that $A_m A_{m-1} \cdots A_1 f$ makes sense. Then repeated use of Lemma 4.3(i) shows that

$$\|A_m A_{m-1} \cdots A_1 f\|_2 \ll_{m,k} |s|^m \|f\|_2,$$

if f is an L^2-eigenfunction. Moreover, if f is an Eisenstein series, then $A_m A_{m-1} \cdots A_1 f$ is another Eisenstein series of appropriate weight, times a polynomial of degree m is s. A similar argument to Lemma 4.3(ii) gives

$$\|A_m A_{m-1} \cdots A_1 f\|_{2,C} \ll_{m,k} |s|^m \left\|f^Y\right\|_2,$$

Here C is any compactly supported set.

For completeness we also state and prove the convexity bound for weight k eigenfunctions:

Theorem 4.5. *Let $f \in \mathfrak{F}_k$ be an eigenfunction of Δ_k with eigenvalue $s(1-s)$ and let C be a compact subset of $\Gamma\backslash\mathbb{H}$. Then there exists a compact subset C' of $\Gamma\backslash\mathbb{H}$ containing C such that*

$$\|f(z)\|_{\infty,C} \ll_{C,k} |t|^{1/2} \left(\int_{C'} |f(z)|^2 d\mu(z)\right)^{1/2}.$$

Proof. We imitate the argument in [31]. We use (4.4) for $n = 0$ and apply the Cauchy–Schwarz inequality to get:

$$\begin{aligned}|f(z_0)| &= \left| \frac{\int_A h(z)\overline{P^0_{s,k}(z,z_0)}\,d\mu(z)}{\int_A |P^0_{s,k}(z,z_0)|^2\,d\mu(z)} \right| \\ &\leq \frac{(\int_A |h(z)|^2\,d\mu(z))^{1/2}(\int_A |P^0_{s,k}(z,z_0)|^2\,d\mu(z))^{1/2}}{\int_A |P^0_{s,k}(z,z_0)|^2\,d\mu(z)} \\ &= \frac{(\int_A |f(z)|^2\,d\mu(z))^{1/2}}{(\int_A |P^0_{s,k}(z,z_0)|^2\,d\mu(z))^{1/2}} \ll |t|^{1/2}\left(\int_A |f(z)|^2\,d\mu(z)\right)^{1/2},\end{aligned}$$

if we can show for some annulus A that

$$\int_A |P^0_{s,k}(z,z_0)|^2\,d\mu(z) \gg |t|^{-1}. \tag{4.10}$$

To prove this we need the asymptotic behavior of $P^0_{s,k}(r)$ as $\Im(s)\to\infty$, $\Re(s)$ fixed and $r > r_0$, see [11, Eq. 27]:

$$P^0_{1/2+it,k}(r) = \frac{2}{|t|^{1/2}\sqrt{2\pi\sinh r}}\cos(rt-\pi/4) + O(|t|^{-1}). \tag{4.11}$$

Since the asymptotics in (4.11) hold for r away from 0 it is convenient to work in an annulus $A = A(r_1,r_2) = \{z; r_1 < r(z,z_0) < r_2\}$ centered at z_0. We have

$$\begin{aligned}\int_{r_1}^{r_2} |P^0_{s,k}(r)|^2 \sinh r dr &= \int_{r_1}^{r_2}\left(\frac{4}{|t|2\pi\sinh r}\cos^2(rt-\pi/4) + O(|t|^{-3/2})\right)\sinh r dr \\ &= \frac{r_2-r_1}{\pi|t|} + O(|t|^{-3/2}).\end{aligned}$$

We integrate in polar coordinates to get (4.10).

We use the same r_1 and r_2 for all $z_0 \in C$ to get a compact set $K' \subset \mathbb{H}$ such that

$$\|f(z)\|_{\infty,C} \ll_C |t|^{1/2}\left(\int_{K'} |f(z)|^2 d\mu(z)\right)^{1/2}.$$

Finally by compactness we can cover K' by a finite set of Γ-translates of a compact set $C' \subset \Gamma\backslash\mathbb{H}$. $\square$

5. Squares of eigenfunctions and Heegner points

Let f be a smooth compactly supported function on $\Gamma\backslash\mathbb{H}$. In order to use the results on equidistribution of Heegner points from section 2 we need bounds on the coefficients in the spectral expansion of $f\,|u_j|^2$ and the similar coefficients coming from Eisenstein series.

Very strong bounds (with precise exponential decay in t_k) are known on $\left\langle |u_j|^2, u_k\right\rangle$ (see [30, 25, 1]) but unfortunately they do not seem uniform enough for our purposes. In particular for a given t_j the bounds only hold for t_k large enough (depending on t_j). In this section we obtain much weaker bounds for similar expressions that hold uniformly in both t_j and t_k.

Theorem 5.1. *Let f be a smooth compactly supported function on $\Gamma\backslash\mathbb{H}$ with support K. For any $b>0$ we have the bound*

$$\left\langle f\left|\phi_1\right|^2,\phi_2\right\rangle \ll_{b,f} \left(\frac{1+|t_1|}{1+|t_2|}\right)^b \left\|\phi_1\right\|_{\infty,K}\left\|\phi_1\right\|_2\left\|\phi_2\right\|_2,$$

where, for $j=1,2$, the functions ϕ_j, equal u_j with eigenvalue $1/4+t_j^2$ or $E^Y(z,1/2+it_j)$, where Y is chosen such that the set $\{z\in\Gamma\backslash\mathbb{H}; y\leq Y\}$ contains K in its interior.

Proof. By interpolation, using $\min(x,y)\leq x^ty^{1-t}$ for $0\leq t\leq 1$, it suffices to prove the claim for $b=2n$, where n is a positive integer. We have

$$(1/4+t_2^2)^n\left\langle f\left|\phi_1\right|^2,\phi_2\right\rangle=\left\langle f\left|\phi_1\right|^2,(-\Delta)^n\phi_2\right\rangle=\left\langle(-\Delta)^n f\left|\phi_1\right|^2,\phi_2\right\rangle.$$

In case $\phi_2=E^Y(\cdot,1/2+it_2)$ we have used the fact that on the support of f we have $E^Y(\cdot,1/2+it_2)=E(\cdot,1/2+it_2)$. We see that the statement follows if we can prove that

$$\left\|\Delta^n f\left|\phi_1\right|^2\right\|_2\ll_f (1/4+t_1^2)^n\left\|\phi_1\right\|_{\infty,K}\left\|\phi_1\right\|_2.$$

But since Δ^n consists of compositions of n copies of L_1K_0 this follows from the Leibniz' rule (4.1), Remark 4.2, and Remark 4.4, and the compactness of the support of f. ☐

The following theorem makes explicit the rate of equidistribution of Heegner points with the test function $f\left|\psi\right|^2$ for ψ an eigenfunction.

Theorem 5.2. *Let $\Gamma=PSL_2(\mathbb{Z})$ and let f be a function on $\Gamma\backslash\mathbb{H}$ with compact support K. Then*

$$\begin{aligned}\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|\psi_t(z)\right|^2=&\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f\left|\psi_t\right|^2 d\mu(z)\\ &+O_{f,\varepsilon}(\left\|\psi_t\right\|_{\infty,K}D^{-1/12+\varepsilon}(1+|t|)^{1+\varepsilon}),\end{aligned}$$

where either $\psi_t(z)=E(z,1/2+it)$ or $\psi_{t_j}(z)=u_j(z)$.

Before proving Theorem 5.2 we state a bound on the L^2-norm of $E^Y(\cdot,1/2+it)$. This is one of the several places in the proof of Theorem 5.2 where arithmeticity of the group enters. Here it enters through the growth of the logarithmic derivative of the scattering determinant.

Lemma 5.3. *Let $\Gamma=PSL_2(\mathbb{Z})$. Then*

$$\left\|E^Y(\cdot,1/2+it)\right\|_2\ll_Y\sqrt{\log(2+|t|)}.$$

Proof. By the Maaß–Selberg relations we have

$$\left\|E^Y(z,1/2+it)\right\|^2=O_Y\left(1+\left|\frac{-\phi'}{\phi}(1/2+it)\right|\right),$$

see [33, Eq. (7.42′)]. For congruence groups the scattering matrix can be computed, and this leads to

$$\frac{-\phi'}{\phi}\left(\frac{1}{2}+it\right)=O(\log(2+|t|)), \tag{5.1}$$

see e.g. [13, Eq. (2.5), p. 508] for $\mathrm{PSL}_2(\mathbb{Z})$. ☐

Proof of Theorem 5.2. We use the spectral expansion of $f\psi_t^2$ to see that

$$f\psi_t^2 - \frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f(z)\psi_t(z)^2 d\mu(z)$$
$$= \sum_{t_k\neq i/2} \left\langle f\psi_t^2, u_k\right\rangle u_k + \frac{1}{4\pi}\int_{\mathbb{R}} \left\langle f\psi_t^2, E(\cdot,1/2+ir)\right\rangle E(\cdot,1/2+ir)dr.$$

After we average over Heegner points, it suffices to show that

$$\sum_{t_k\neq i/2} \left\langle f\psi_t^2, u_k\right\rangle \frac{W_c(D,t_k)}{h(D)} + \frac{1}{4\pi}\int_{\mathbb{R}} \left\langle f\psi_t^2, E(\cdot,1/2+ir)\right\rangle \frac{W_E(D,r)}{h(D)}dr \tag{5.2}$$

is $O_{f,\varepsilon}(\|\psi_t\|_{\infty,K}\, D^{-1/12+\varepsilon}\, |t|^{1+\varepsilon})$. We bound the discrete contribution, i.e. the sum, and notice that the continuous contribution can be bounded the same way. By Cauchy–Schwarz we have

$$\left(\sum_{L\le t_k\le 2L} \left\langle f\psi_t^2, u_k\right\rangle \frac{W_c(D,t_k)}{h(D)}\right)^2 \le \sum_{L\le t_k\le 2L} \left|\left\langle f\psi_t^2, u_k\right\rangle\right|^2 \sum_{L\le t_k\le 2L} \frac{|W_c(D,t_k)|^2}{h(D)^2}$$
$$\ll \sum_{L\le t_j\le 2L} \left|\left\langle f\psi_t^2, u_k\right\rangle\right|^2 D^{-1/6+\varepsilon}(1+L)^{2+\varepsilon},$$

where we have used Lemma 2.2. We bound the sum in two different ways: either by using Theorem 5.1 or Bessel's inequality. Using Theorem 5.1 we find

$$\frac{1}{h(D)}\sum_{L\le t_k\le 2L} \left\langle f\psi_t^2, u_k\right\rangle W_c(D,t_k) \ll \frac{(1+|t|)^{b+\varepsilon}}{(1+L)^b}\,\|\psi_t\|_{\infty,K}\, D^{-1/12+\varepsilon}(1+L)^{2+\varepsilon}. \tag{5.3}$$

We will use this for the tail of the discrete contribution in (5.2).

We also need another estimate. It follows from Lemma 5.3 for the case of Eisenstein series and elementary considerations that

$$\left\|f\psi_t^2\right\|_2 \ll_f \|\psi_t\|_{\infty,K}\,(1+|t|)^{\varepsilon}. \tag{5.4}$$

Using Cauchy–Schwarz, Bessel's inequality, and (5.4) we get the estimate

$$\frac{1}{h(D)}\sum_{L\le t_k\le 2L} \left\langle f\psi_t^2, u_k\right\rangle W_c(D,t_k) \ll \|\psi_t\|_{\infty,K}\,(1+|t|)^{\varepsilon} D^{-1/12+\varepsilon}(1+L)^{1+\varepsilon}. \tag{5.5}$$

We have similar bounds for the continuous contribution. Using (5.5) for the bulk ($|t_k|\le V$) and (5.3) for the tail ($|t_k|>V$) we find that for V bounded away from zero we have

$$f\psi_t^2 - \frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f(z)\psi_t(z)^2 d\mu(z)$$
$$\ll_{f,b} \|\psi_t\|_{\infty,K}\, D^{-1/12+\varepsilon}\Big((1+|t|)^{\varepsilon}V^{1+\varepsilon} + \frac{(1+|t|)^b}{V^b}V^{2+\varepsilon}\Big),$$

when $b>2+\varepsilon$. Choosing $V=(1+|t|)^{b/(b-1)}$ and b sufficiently large we arrive at the result. □

6. Proof of the main theorems

We now have the necessary tools to prove Theorem 1.1 and Theorem 1.3. We start by constructing appropriate test functions for the Selberg pre-trace formula.

6.1. The pre-trace formula.

This follows our previous investigations. We refer to [26, Section 5] for additional details. Let $\delta>0$ be a small parameter, which will eventually be chosen to depend on X and D. Let

$$k_\delta(u)=\frac{1}{4\pi\sinh^2(\delta/2)}1_{[0,\cosh\delta-1)/2]}(u),$$

where $1_A(u)$ denotes the indicator function of any set A. Let $Y>0$ be defined by $\cosh(Y)=X/2$. This implies that $4u(z,w)+2\leq X$ if and only if $d(z,w)\leq Y$, where $d(z,w)$ denotes the hyperbolic distance between z and w. Consider now

$$k^{\pm}(u)=\left(1_{[0,(\cosh(Y\pm\delta)-1)/2]}*k_\delta\right)(u),$$

where $*$ denotes the hyperbolic convolution

$$(k_1*k_2)(u(z,w))=\int_{\mathbb{H}}k_1(u(z,v))k_2(u(v,w))d\mu(v).$$

With this choice of kernels it follows from the triangle inequality for the hyperbolic distance that

$$k^-(u)\leq 1_{[0,(X-2)/4]}(u)\leq k^+(u),$$

see [26, Eq. (5.4)]. By summing over $\gamma\in\Gamma$ we find that

$$K^-(z,w,X)\leq N(z,w,X)\leq K^+(z,w,X), \tag{6.1}$$

where

$$K^{\pm}(z,w,X):=\sum_{\gamma\in\Gamma}k^{\pm}(\gamma z,w,X).$$

Subtracting $\pi X/\mathrm{vol}(\Gamma\backslash\mathbb{H})$ and averaging over Heegner points of discriminant D, we see that bounds for the absolute values of the two expressions

$$\frac{1}{h(D)}\sum_{z\in\Lambda_D}f(z)\left(K^{\pm}(z,z,X)-\frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\right)$$

imply the same bound for the absolute value of

$$\frac{1}{h(D)}\sum_{z\in\Lambda_D}f(z)\left(N(z,z,X)-\frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\right).$$

The advantage of approximating $N(z,w,X)$ by the automorphic kernels $K^{\pm}(z,w,X)$ is that, contrary to $1_{[0,(X-2)/4)]}$, the kernels $k^{\pm}$ are admissible in the Selberg pre-trace formula. Moreover, the corresponding Selberg–Harish-Chandra transforms $h^{\pm}$ can be computed explicitly in terms of special functions, as the Selberg–Harish-Chandra transform maps the hyperbolic convolution k_1*k_2 into the product of the transforms h_1h_2, see [6, p. 323].

Let h_R denote the Selberg–Harish-Chandra transform of $k_R = 1_{[0,(\cosh R-1)/2]}$, with $R > 0$. The function h_R can be computed:

$$h_R(t) = \sqrt{2\pi \sinh R}\,\Re\left(e^{itR}\frac{\Gamma(it)}{\Gamma(3/2+it)}F\left(\frac{-1}{2},\frac{3}{2},1-it,\frac{1}{(1-e^{2R})}\right)\right),$$

where F is the Gauss hypergeometric function, see [6, p. 321]. It follows from the series representation of F that, for large enough R, say $R > \log(2)/2$, we have

$$F\left(\frac{-1}{2},\frac{3}{2},1-it,\frac{1}{(1-e^{2R})}\right) = 1 + O(e^{-2R}\min(1,|t|^{-1})).$$

For R small, say less than 1, and t real it is known that, see e.g. [6, Lemma 2.4 (c)],

$$h_R(t) = 2\pi R^2\frac{J_1(Rt)}{Rt}\sqrt{\frac{\sinh R}{R}} + O(R^2\min(R^2,|t|^{-2})).$$

Here J_1 is the Bessel function of order 1. It satisfies

$$2\frac{J_1(x)}{x} = 1_{[0,1]}(x) + O(\min(|x|,|x|^{-3/2})),$$

see [18, B.28, B.35]. It is also convenient to use the uniform bound $h_R(t) = O((R+1)e^{R/2})$, see [6, Lemma 2.4].

The group $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ has no small eigenvalues so we only need to estimate $h^{\pm}(t)$ at $t = i/2$ and for $t \in \mathbb{R}$. We have

$$h^{\pm}(t) = h_{Y\pm\delta}(t)\frac{h_\delta(t)}{4\pi\sinh^2(\delta/2)}.$$

A direct computation [6, Lemma 2.4 (d)] shows that

$$h^{\pm}(i/2) = 2\pi(\cosh(Y\pm\delta)-1)\frac{2\pi(\cosh\delta-1)}{4\pi\sinh^2(\delta/2)} = \pi X + O(1+\delta X). \tag{6.2}$$

To estimate $h^{\pm}(t)$ for t real we combine the bounds above and find

$$\begin{aligned} h^{\pm}(t) &= O\left(\frac{\sqrt{X}}{t^{3/2}}\left(\min(1,(\delta|t|)^{-3/2}) + \min(\delta^2,|t|^{-2})\right)\right) \\ &= O\left(\frac{\sqrt{X}}{t^{3/2}}\left(\min(1,(\delta|t|)^{-3/2})\right)\right), \end{aligned} \tag{6.3}$$

see [26, Eq. (5.5), (5.10)]. Finally

$$h^{\pm}(t) = O(\sqrt{X}\log X), \tag{6.4}$$

where the last bound is uniform for t real.

6.2. **Applying the pre-trace formula.** By the pre-trace formula [18, Theorem 7.4] we have

$$K^{\pm}(z,z,X) = \sum_{t_j} h^{\pm}(t_j)\,|u_j(z)|^2 + \frac{1}{4\pi}\int_{\mathbb{R}} h^{\pm}(t)\,|E(z,1/2+it)|^2\,dt.$$

Using (6.2) we therefore see that

$$\begin{aligned}
&\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left(K^{\pm}(z,z,X)-\frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\right)\\
&\quad=\sum_{t_j\in\mathbb{R}} h^{\pm}(t_j)\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|u_j(z)\right|^2\\
&\qquad+\frac{1}{4\pi}\int_{\mathbb{R}} h^{\pm}(t)\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|E(z,1/2+it)\right|^2 dt+O(1+\delta X)
\end{aligned}$$

(6.5)

$$\begin{aligned}
&\quad=\sum_{t_j} h^{\pm}(t_j)\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f\left|u_j\right|^2 d\mu+Q_c(X,\delta,D)\\
&\qquad+\frac{1}{4\pi}\int_{\mathbb{R}} h^{\pm}(t)\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f\left|E(\cdot,1/2+it)\right|^2 d\mu\, dt+Q_E(X,\delta,D)+O(1+\delta X),
\end{aligned}$$

where

$$Q_c(X,\delta,D)=\sum_{t_j\in\mathbb{R}} h^{\pm}(t_j)\left(\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|u_j(z)\right|^2-\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f\left|u_j\right|^2 d\mu\right),$$

(6.6)

$$\begin{aligned}
Q_E(X,\delta,D)=&\frac{1}{4\pi}\int_{\mathbb{R}} h^{\pm}(t)\left(\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|E(z,1/2+it)\right|^2\right.\\
&\left.\qquad\qquad-\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f\left|E(\cdot,1/2+it)\right|^2 d\mu\right) dt.
\end{aligned}$$

The first and third terms in (6.5) are exactly the expressions that are treated in [26, Sec 6.], where we found – using several deep results from [23], e.g. (1.6) – that the first term is bounded by $O(X^{7/12+\varepsilon})$ as long as we assume that δ tends to zero at least as fast as X^{-c} for some $c>0$. Compare [26, Lemmata 6.2 and 6.3]. The third term is $O(X^{1/2+\varepsilon})$ by [26, Lemma 6.1].

The second term $Q_c(X,\delta,D)$ in (6.5) is where we need bounds on the sup-norm. We first notice that we have the trivial bound

$$\begin{aligned}
\sum_{T\leq t_j\leq 2T}\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|u_j(z)\right|^2&=\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\sum_{T\leq t_j\leq 2T}\left|u_j(z)\right|^2\\
&\ll_f\left\|\sum_{T\leq t_j\leq 2T}\left|u_j(z)\right|^2\right\|_\infty.
\end{aligned}$$

Since $\int_{\Gamma\backslash\mathbb{H}} f(z)\left|u_j(z)\right|^2 d\mu(z)\ll_f 1$ we easily find from Proposition 3.1 and Weyl's law (1.5) that

$$\sum_{T\leq t_j\leq 2T}\left(\frac{1}{h(D)}\sum_{z\in\Lambda_D} f(z)\left|u_j(z)\right|^2-\frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}\int_{\Gamma\backslash\mathbb{H}} f\left|u_j\right|^2 d\mu\right)=O_f(T^2).$$

Interpolating this – using $\min(B,C) \le B^a C^{1-a}$ for $0 \le a \le 1$ – with the bound we get from bounding each term using Theorem 5.2 and applying Theorem 1.5 we find that for any $0 \le a \le 1$

$$\sum_{T \le t_j \le 2T} \left(\frac{1}{h(D)} \sum_{z \in \Lambda_D} f(z)\, |u_j(z)|^2 - \frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} \int_{\Gamma\backslash\mathbb{H}} f\, |u_j|^2 \, d\mu \right) = O_f(T^{2+a(3/2-1/8)+\varepsilon} D^{-a/12+\varepsilon})).$$

Using (6.4) for a bounded set of t_j's, (6.3) in the ranges $1 \le |t_j| < \delta^{-1}$ and $|t_j| \ge \delta^{-1}$, dyadic decomposition, and the estimate above, we find that the quantity $Q_c(X,\delta,D)$ in (6.6) is $O(X^{1/2}\delta^{-(1/2+a(3/2-1/8)+\varepsilon)} D^{-a/12+\varepsilon})$, for any $0 \le a \le 1$ satisfying $a(3/2-1/8)+\varepsilon < 1$.

The strategy for bounding the fourth term of (6.5) is basically the same. We use Proposition 3.3, Lemma 5.3, Theorem 5.2 and (1.7) to see that

$$\int_T^{2T} \left(\frac{1}{h(D)} \sum_{z \in \Lambda_D} f(z)\, |E(z,1/2+it)|^2 - \frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} \int_{\Gamma\backslash\mathbb{H}} f\, |E(\cdot,1/2+it)|^2 \, d\mu \right) dt$$

is $O_f(\min(T^2, T^{5/2-1/8+\varepsilon} D^{-1/12+\varepsilon}) = O(T^{2+(1/2-1/8)b+\varepsilon} D^{-b/12+\varepsilon})$ for any $0 \le b \le 1$. Doing dyadic decomposition we find that the fourth quantity $Q_E(X,\delta,D)$ is $O(X^{1/2}\delta^{-(1/2+b(1/2-1/8)+\varepsilon)} D^{-b/12+\varepsilon})$. Chosing $a = b$ we see that this term is smaller than the third.

Putting everything together we find that

$$\frac{1}{h(D)} \sum_{z \in \Lambda_D} f(z) \left(K^{\pm}(z,z,X) - \frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} \right) = O(\delta X + X^{7/12+\varepsilon} + X^{1/2}\delta^{-(1/2+a(3/2-1/8)+\varepsilon)} D^{-a/12+\varepsilon}).$$

Choosing a minimal ($a = 0$) and balancing error terms we recover Selberg's bound $O(X^{2/3+\varepsilon})$ with no saving due to the averaging in D. If we choose a as large as allowed, i.e. close to $1/(3/2-1/8) = 8/11$, the error is

$$O(\delta X + X^{7/12+\varepsilon} + X^{1/2}\delta^{-(3/2+\varepsilon)} D^{-a/12+\varepsilon}).$$

To balance the first and third term we choose $\delta = X^{-1/5} D^{-4/165}$, which gives Theorem 1.1.

Proof of Theorem 1.3. We assume (C1), (C2), and (C3). Using Remark 2.3 and [26, Remark 6.4] and the same technique as above, we find that for any $0 \le a < 1$

$$\frac{1}{h(D)} \sum_{z \in \Lambda_D} f(z) \left(K^{\pm}(z,z,X) - \frac{\pi X}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} \right) = O(\delta X + X^{1/2+\varepsilon} + X^{1/2}\delta^{-(1/2+a(3/2-1/2)+\varepsilon)} D^{-a/4+\varepsilon}).$$

We choose a close to 1 and $\delta = X^{-1/5} D^{-1/10}$ to get the result. □

References

[1] Joseph Bernstein and Andre Reznikov. Estimates of automorphic functions. *Mosc. Math. J.*, 4(1):19–37, 310, 2004. 13

[2] Valentin Blomer. Epstein zeta-functions, subconvexity, and the purity conjecture. *arXiv:1602.02326*, 2016. 3, 4

[3] Valentin Blomer, Gergely Harcos, and Philippe Michel. Bounds for modular L-functions in the level aspect. *Ann. Sci. École Norm. Sup. (4)*, 40(5):697–740, 2007. 6

[4] Valentin Blomer and Roman Holowinsky. Bounding sup-norms of cusp forms of large level. *Invent. Math.*, 179(3):645–681, 2010. 6

[5] Valentin Blomer and Péter Maga. Subconvexity for sup-norms of cusp forms on PGL(n). *Selecta Math. (N.S.)*, 22(3):1269–1287, 2016. 6

[6] Fernando Chamizo. Some applications of large sieve in Riemann surfaces. *Acta Arith.*, 77(4):315–337, 1996. 16, 17

[7] Soumya Das and Jyoti Sengupta. L^∞ norms of holomorphic modular forms in the case of compact quotient. *Forum Math.*, 27(4):1987–2001, 2015. 6

[8] Harold Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1980. Revised by Hugh L. Montgomery. 2

[9] William Duke. Hyperbolic distribution problems and half-integral weight Maass forms. *Invent. Math.*, 92(1):73–90, 1988. 2, 4

[10] William Duke, John B. Friedlander, and Henryk Iwaniec. The subconvexity problem for Artin L-functions. *Invent. Math.*, 149(3):489–577, 2002. 6

[11] John D. Fay. Fourier coefficients of the resolvent for a Fuchsian group. *J. Reine Angew. Math.*, 293/294:143–203, 1977. 9, 11, 13

[12] Anton Good. Beiträge zur Theorie der Dirichletreihen, die Spitzenformen zugeordnet sind. *J. Number Theory*, 13(1):18–65, 1981. 10

[13] Dennis A. Hejhal. *The Selberg trace formula for* $\mathrm{PSL}(2, \mathbf{R})$. *Vol. 2*, volume 1001 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983. 14

[14] Jeffrey Hoffstein and Paul Lockhart. Coefficients of Maass forms and the Siegel zero. *Ann. of Math. (2)*, 140(1):161–181, 1994. With an appendix by Dorian Goldfeld, Jeffrey Hoffstein and Daniel Lieman. 5, 6

[15] Aleksandar Ivić. On sums of Hecke series in short intervals. *J. Théor. Nombres Bordeaux*, 13(2):453–468, 2001. 5

[16] Henryk Iwaniec. Fourier coefficients of modular forms of half-integral weight. *Invent. Math.*, 87(2):385–401, 1987. 4

[17] Henryk Iwaniec. Small eigenvalues of Laplacian for $\Gamma_0(N)$. *Acta Arith.*, 56(1):65–82, 1990. 6

[18] Henryk Iwaniec. *Spectral methods of automorphic forms*, volume 53 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2002. 2, 8, 17

[19] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004. 4

[20] Henryk Iwaniec and Peter Sarnak. L^∞ norms of eigenfunctions of arithmetic surfaces. *Ann. of Math. (2)*, 141(2):301–320, 1995. 3, 5, 6

[21] Hervé Jacquet and Nan Chen. Positivity of quadratic base change L-functions. *Bull. Soc. Math. France*, 129(1):33–90, 2001. 5

[22] Junehyuk Jung. Quantitative quantum ergodicity and the nodal domains of Hecke-Maass cusp forms. *Comm. Math. Phys.*, 348(2):603–653, 2016. 3, 4

[23] Wen Zhi Luo and Peter Sarnak. Quantum ergodicity of eigenfunctions on $\mathrm{PSL}_2(\mathbf{Z})\backslash\mathbf{H}^2$. *Inst. Hautes Études Sci. Publ. Math.*, (81):207–237, 1995. 3, 7, 18

[24] Simon Marshall. Sup norms of Maass forms on semisimple groups. *arXiv:1405.7033*, 2014. 6

[25] Yiannis N. Petridis. On squares of eigenfunctions for the hyperbolic plane and a new bound on certain L-series. *Internat. Math. Res. Notices*, (3):111–127 (electronic), 1995. 13

[26] Yiannis N. Petridis and Morten S. Risager. Local average in hyperbolic lattice point counting, with an Appendix by Niko Laaksonen. *Mathematische Zeitschrift*, pages 1–26, 2016. 2, 3, 16, 17, 18, 19

[27] Ralph Phillips and Zeév Rudnick. The circle problem in the hyperbolic plane. *J. Funct. Anal.*, 121(1):78–116, 1994. 1

[28] Walter Roelcke. Das Eigenwertproblem der automorphen Formen in der hyperbolischen Ebene. i, ii. Math. Ann. **167** (1966), 292–337; ibid **168** (1966), 261–324., 1966. 11, 12

[29] Abhishek Saha. Hybrid sup-norm bounds for Maass newforms of powerful level. *arXiv:1509.07489 (To appear in Algebra and Number Theory)*, 2015. 6

[30] Peter Sarnak. Integrals of products of eigenfunctions. *Internat. Math. Res. Notices*, (6):251 ff., approx. 10 pp. (electronic), 1994. 13

[31] Peter Sarnak. Letter to Cathleen Morawetz. August 2004. 6, 7, 13

[32] Andreas Seeger and Christopher D. Sogge. Bounds for eigenfunctions of differential operators. *Indiana Univ. Math. J.*, 38(3):669–682, 1989. 5

[33] Atle Selberg. *Göttingen lecture notes in Collected papers. Vol. I.* Springer-Verlag, Berlin, 1989. With a foreword by K. Chandrasekharan. 14

[34] Nicolas Templier. Hybrid sup-norm bounds for Hecke–Maass cusp forms. *J. Eur. Math. Soc. (JEMS)*, 17(8):2069–2082, 2015. 6

[35] Edward C. Titchmarsh. *The theory of the Riemann zeta-function.* The Clarendon Press Oxford University Press, New York, second edition, 1986. Edited and with a preface by D. R. Heath-Brown. 5

[36] Matthew P. Young. Weyl-type hybrid subconvexity bounds for twisted L-functions and Heegner points on shrinking sets. *arXiv:1405.5457 (To appear in Journal Eur. Math. Soc.)*, 2014. 3, 4

[37] Matthew P. Young. A note on the sup norm of Eisenstein series. *arXiv:1504.03272 (To appear in Quarterly Journal of Mathematics)*, 2015. 3

Department of Mathematics, University College London, Gower Street, London WC1E 6BT, United Kingdom

E-mail address: `i.petridis@ucl.ac.uk`

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

E-mail address: `risager@math.ku.dk`