Annals of Mathematics **184** (2016), 949–990
http://dx.doi.org/10.4007/annals.2016.184.3.8

Geometric invariants for real quadratic fields

By W. Duke, Ö. Imamoḡlu, and Á. Tóth

Abstract

To an ideal class of a real quadratic field we associate a certain surface. This surface, which is a new geometric invariant, has the usual modular closed geodesic as its boundary. Furthermore, its area is determined by the length of an associated backward continued fraction. We study the distribution properties of this surface on average over a genus. In the process we give an extension and refinement of the Katok-Sarnak formula.

1. Introduction

In this paper we will introduce a new geometric invariant associated to a (narrow) ideal class of a real quadratic field. This invariant is a finite area hyperbolic surface[1] *with a boundary* that maps naturally on the modular surface. The boundary is a simple closed geodesic whose image in the modular surface is the usual modular closed geodesic associated to the ideal class. Its length is well known to be expressible in terms of a fundamental unit of the field. The area of the surface is determined by the length of an associated minus (or backward) continued fraction. Since the surface contains more information than the closed geodesic alone, it might be hoped that an investigation of its geometric properties could lead us to a better understanding of the class groups of real quadratic fields. Their mysterious behavior as the discriminant varies has tantalized number theorists since the time of Gauss. One purpose of this work is to initiate this investigation.

W. Duke's research on this paper was partially supported by NSF grant DMS-10-01527.
Ö. Imamoḡlu is supported by SNF grant 200020-144285.
Á. Tóth is supported by OTKA grant NK 104183 and by NKFIH (National Research, Development and Innovation Office) grant ERC_HU_15 118946.

[1]Actually it is a hyperbolic orbifold, or equivalently a hyperbolic Riemann surface with signature, but we will usually abuse terminology and simply call it a surface.

For the present at least, our main goal is to obtain a result about the distribution properties of the surface as it lies in the modular surface. To obtain good results we need to average over a genus of classes. This problem is closely allied with (and in fact completes in a natural way) the problem, introduced in [12], of showing the uniform distribution of the closed geodesics on the modular surface when ordered by their associated discriminant. The analytic approach to the closed geodesic problem leads to estimating the Fourier coefficients of Maass cusp forms of weight 1/2. For the surface problem, this approach also leads to estimating these Fourier coefficients (for different indices), but it requires interesting and nontrivial extensions of formulas of Hecke and Katok-Sarnak. The needed extensions for surface integrals involve genus characters associated to two negative discriminants, as opposed to the closed geodesic case, which leads to two positive discriminants. Discriminants of different signs arise in the case of *CM points*, where the quadratic field is imaginary. Thus points, curves and surfaces associated to quadratic ideal classes all occur in a natural way.

We begin by presenting some background material about real quadratic fields, continued fractions, binary quadratic forms, Fuchsian groups and modular closed geodesics. Then in Section 3 we introduce the surfaces we will study and give some of their properties in Theorem 1. The main uniform distribution result is stated in Section 4 as Theorem 2. Section 5 contains statements of the extended formulas of Hecke (Theorem 3) and Katok-Sarnak (Theorem 4). Assuming these, we prove the main result in Section 6. Theorem 3 is proven in Section 7. The rest of the paper (Sections 8–10) is devoted to the proof of Theorem 4.

Acknowledgements. Duke and Tóth thank the FIM at ETH in Zürich for generous ongoing support of our joint research, including this project. We also thank Alex Kontorovich for his comments, Nickolas Andersen for independently verifying some of our numerical computations and the referee for helpful suggestions.

2. **Preliminaries**

Real quadratic fields. Let $\mathbb{K}/\mathbb{Q}$ be a real quadratic field. Then $\mathbb{K} = \mathbb{Q}(\sqrt{D})$, where $D > 1$ is the discriminant of $\mathbb{K}$. Let $\sigma : \mathbb{K} \to \mathbb{K}$ be the nontrivial Galois automorphism $w \mapsto w^\sigma$, and for $\alpha \in \mathbb{K}$, let $N(\alpha) = \alpha\alpha^\sigma$. Let $\mathrm{Cl}^+(\mathbb{K})$ be the group of fractional ideal classes taken in the narrow sense. Thus two ideals $\mathfrak{a}$ and $\mathfrak{b}$ are in the same narrow class if there is a $\alpha \in \mathbb{K}$ with $N(\alpha) > 0$ so that $\mathfrak{a} = (\alpha)\mathfrak{b}$. Let $h(D) = \#\mathrm{Cl}^+(\mathbb{K})$ be the (narrow) class number and $\varepsilon_D > 1$ be the smallest unit with positive norm in the ring of integers $\mathcal{O}_\mathbb{K}$ of $\mathbb{K}$. We denote by I the principal class and by J the class of the different $(\sqrt{D})$ of K, which coincides with the class of principal ideals (α)

where $N(\alpha) = \alpha\alpha^\sigma < 0$. Then

$$\mathrm{Cl}(\mathbb{K}) = \mathrm{Cl}^+(\mathbb{K})/J$$

is the class group in the wide sense. Clearly $J \neq I$ if and only if $\mathcal{O}_\mathbb{K}$ contains no unit of norm -1. In this case each wide ideal class is the union of two narrow classes, say A and JA. A sufficient condition for $J \neq I$ is that D is divisible by a prime $p \equiv 3 \pmod 4$.

Minus continued fractions. Each ideal class $A \in \mathrm{Cl}^+(\mathbb{K})$ contains fractional ideals of the form $w\mathbb{Z} + \mathbb{Z} \in A$, where $w \in \mathbb{K}$ is such that $w > w^\sigma$. Consider the *minus* (or backward) continued fraction of w:

$$w = [\![a_0, a_1, a_2, \ldots]\!] = a_0 - \cfrac{1}{a_1 - \cfrac{1}{a_2 - \cfrac{1}{a_3 - \cdots}}},$$

where $a_j \in \mathbb{Z}$ with $a_j \geq 2$ for $j \geq 1$. This continued fraction is eventually periodic and has a unique primitive *cycle* $(\!(n_1, \ldots, n_\ell)\!)$ of length ℓ, only defined up to cyclic permutations. Different admissible choices of w lead to the same primitive cycle. The continued fraction is purely periodic precisely when w is *reduced* in the sense that

$$0 < w^\sigma < 1 < w$$

(see [28], [57]). The cycle $(\!(n_1, \ldots, n_\ell)\!)$ characterizes A; it is a *complete* class invariant. The length $\ell = \ell_A$, which is also the number of distinct reduced w, is another invariant as is the sum

$$m = m_A = n_1 + \cdots + n_\ell. \tag{2.1}$$

The cycle of A^{-1} is given by that of A reversed:

$$(\!(n_\ell, \ldots, n_1)\!)\,. \tag{2.2}$$

To see this observe that A^{-1} is represented by $(1/w^\sigma)\mathbb{Z} + \mathbb{Z}$ and by [58, p. 128] the continued fraction of $1/w^\sigma$ has (2.2) as its cycle.

Binary quadratic forms. In place of ideal classes, it is sometimes more convenient to use binary quadratic forms

$$Q(x, y) = [a, b, c] = ax^2 + bxy + cy^2,$$

where $a, b, c \in \mathbb{Z}$ and $D = b^2 - 4ac$. Quadratic forms are especially useful when one wants to consider arbitrary discriminants D. For fundamental D, all quadratic forms are primitive in that $\gcd(a, b, c) = 1$, and we have a simple correspondence between narrow ideal classes of $\mathbb{K}$ and equivalence classes of binary quadratic forms of discriminant D with respect to the usual action of

$\mathrm{PSL}(2,\mathbb{Z})$. This correspondence is induced by $\mathfrak{a} \mapsto Q(x,y)$, where $\mathfrak{a} = w\mathbb{Z} + \mathbb{Z}$ with $w^\sigma < w$ and

$$Q(x,y) = N(x - wy)/N(\mathfrak{a}).$$

The map takes the narrow ideal class of $\mathfrak{a}$ to the Γ-equivalence class of Q. The inverse map is given by $Q(x,y) \mapsto w\mathbb{Z} + \mathbb{Z}$, where

$$w = \frac{-b + \sqrt{D}}{2a},$$

provided we choose Q in its class to have $a > 0$. The following table of correspondences is useful. Suppose that $Q = [a,b,c]$ represents in this way the ideal class A. Then

$$[a,-b,c] \quad \text{represents } A^{-1}, \tag{2.3}$$

$$[-a,b,-c] \quad \text{represents } JA, \tag{2.4}$$

$$[-a,-b,-c] \quad \text{represents } JA^{-1}. \tag{2.5}$$

Fuchsian groups and closed geodesics. Let $\mathcal{H}$ be the upper half plane. As usual, lengths and areas on $\mathcal{H}$ are hyperbolic and determined by the metric and measure

$$ds = y^{-1}|dz| \quad \text{and} \quad d\mu(z) = y^{-2}dxdy,$$

respectively, where $z = x + iy$. Define the cross ratio of $z_1, z_2, z_3, z_4 \in \mathbb{C}$ by

$$[z_1,z_2,z_3,z_4] = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_2)(z_3 - z_4)}. \tag{2.6}$$

A useful formula for the distance between z and z^* in $\mathcal{H}$ is given by

$$d(z,z^*) = \log|[w,z,z^*,w^*]|, \tag{2.7}$$

where $w, w^* \in \mathbb{R}$ are the points where the geodesic arc joining z to z^* intersects $\mathbb{R}$ and where the order in which this arc passes through the points is given by w, z, z^*, w^* (see, e.g., [2]).

Suppose that $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})$ is a nonelementary Fuchsian group. (See [2] for background.) Let Λ be the limit set of Γ. The group Γ is said to be of the first kind when $\Lambda = \mathbb{R}$, otherwise of the second kind. In general, $\mathbb{R} - \Lambda$ is a countable union of mutually disjoint open intervals. Let $\mathcal{N}_\Gamma$ be the intersection of the (non-Euclidean) open half-planes that lie above the geodesics having the same endpoints as these intervals. This $\mathcal{N}_\Gamma$ is called the Nielsen region of Γ. It is shown in [2, Thm. 8.5.2] that $\mathcal{N}_\Gamma$ is the smallest nonempty Γ-invariant open convex subset of $\mathcal{H}$. Clearly $\mathcal{N}_\Gamma = \mathcal{H}$ exactly when Γ is of the first kind.

Suppose now that Γ is finitely generated. Let $\mathcal{H}^*$ be the upper half-plane with all elliptic points of Γ removed. Then $\Gamma\backslash\mathcal{H}^*$ becomes a Riemann surface of genus g with $t < \infty$ conformal disks and finitely many points removed. The group Γ is said to have *signature* $(g; m_1, \dots, m_r; s; t)$, where $m_1, \dots, m_r$ are the orders of the elliptic points and there are s parabolic cusps of $\Gamma\backslash\mathcal{H}^*$. The

boundary circle of each removed disk, assumed positively oriented, is freely homotopic in $\Gamma\backslash\mathcal{H}^*$ to a *unique* closed geodesic (see, e.g., [19, Prop. 1.3]). These geodesics form the boundary of the image in $\Gamma\backslash\mathcal{H}^*$ of (the closure of) the Nielsen region.

Thus $\Gamma\backslash\mathcal{N}_\Gamma$ is a Riemann surface with signature having t geodesic boundary curves, s cusps, and r orbifold points. Let $\mathcal{F}_\Gamma \subset \mathcal{H}$ be a fundamental domain for $\Gamma\backslash\mathcal{N}_\Gamma$. For simplicity, we will identify the surface with $\mathcal{F}_\Gamma$. This should cause no confusion as long as it is understood that for us $\partial\mathcal{F}_\Gamma$ denotes the pre-image of the boundary of the surface as a subset of the boundary of the fundamental domain. In other words, we will not count as part of the boundary of $\mathcal{F}_\Gamma$ those sides of $\overline{\mathcal{F}}_\Gamma$ that are identified by Γ. The Gauss-Bonnet theorem [2, Thm. 10.4.3] gives

$$\tfrac{1}{2\pi}\mathrm{area}(\mathcal{F}_\Gamma) = 2(g-1) + s + t + \sum_{j=1}^{r}\left(1 - \tfrac{1}{m_j}\right). \tag{2.8}$$

Suppose now that $\Gamma = \mathrm{PSL}(2,\mathbb{Z})$ is the usual modular group. As is well known, Γ is generated by

$$S = \pm\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{and} \qquad T = \pm\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and has signature $(0; 2, 3; 1, 0)$. Let $\mathcal{F}$ denote the standard fundamental domain for Γ:

$$\begin{aligned}\mathcal{F} = \{&z \in \mathcal{H}; -1/2 \le \mathrm{Re}\, x \le 0 \ \text{ and } \ |z| \ge 1\} \\ &\cup \{z \in \mathcal{H}; 0 < \mathrm{Re}\, x < 1/2 \ \text{ and } \ |z| > 1\}.\end{aligned}$$

By (2.8) or otherwise we have that $\mathrm{area}(\mathcal{F}) = \frac{\pi}{3}$.

For a fixed narrow ideal class $A \in \mathrm{Cl}^+(\mathbb{K})$ and $\mathfrak{a} = w\mathbb{Z} + \mathbb{Z} \in A$ with $w > w^\sigma$ let $\mathcal{S}_w$ be the geodesic in $\mathcal{H}$ with endpoints w^σ and w. The modular closed geodesic $\mathcal{C}_A$ on $\Gamma\backslash\mathcal{H}$ is defined as follows. Define $\gamma_w = \pm\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, where $a, b, c, d \in \mathbb{Z}$ are determined by

$$\begin{aligned} \varepsilon_D w &= aw + b, \\ \varepsilon_D &= cw + d, \end{aligned} \tag{2.9}$$

with ε_D our unit. Then γ_w is a primitive hyperbolic transformation in Γ with fixed points w^σ and w. Since

$$(cw + d)^{-2} = \varepsilon_D^{-2} < 1,$$

we have that w is the attracting fixed point of γ_w. This induces on the geodesic $\mathcal{S}_w$ a clock-wise orientation. Distinct $\mathfrak{a}$ and w for A induce Γ-conjugate transformations γ_w. If we choose some point z_0 on $\mathcal{S}_w$, then the directed arc on $\mathcal{S}_w$ from z_0 to $\gamma_w(z_0)$, when reduced modulo Γ, is the associated closed geodesic $\mathcal{C}_A$ on $\Gamma\backslash\mathcal{H}$. It is well defined for the class A and gives rise to a unique set of

oriented arcs (which could overlap) in $\mathcal{F}$. We also use $\mathcal{C}_A$ to denote this set of arcs. It is well known and easy to see using (2.9) that

$$\text{length}(\mathcal{C}_A) = 2\log\varepsilon_D. \tag{2.10}$$

For a primitive quadratic form $Q(x,y) = [a',b',c']$ with any nonsquare discriminant $d' > 1$, its group of automorphs in Γ is generated by

$$\gamma_Q = \pm\begin{pmatrix} \frac{t-b'u}{2} & -c'u \\ a'u & \frac{t+b'u}{2} \end{pmatrix}, \tag{2.11}$$

where (t,u) gives the smallest solution with $t,u \geq 1$ to $t^2 - d'u^2 = 4$ (see [44]). If

$$Q(x,y) = N(x-wy)/N(\mathfrak{a})$$

as above, then $\gamma_Q = \gamma_w$ and $\varepsilon_D = \frac{t+u\sqrt{D}}{2}$. Using (2.5) we see that the closed geodesic $\mathcal{C}_{JA^{-1}}$ has the same image as $\mathcal{C}_A$ but with the opposite orientation.

Remark. The arcs of $\mathcal{C}_A$ might retrace back over themselves. When this happens $\mathcal{C}_A$ is said to be *reciprocal*. In terms of the class A, it means that $JA^{-1} = A$ or equivalently $A^2 = J$. Sarnak [46] has given a comprehensive treatment of these remarkable geodesics for arbitrary discriminants.

3. **Hyperbolic surfaces**

The basic object we will study is a certain hyperbolic surface with boundary associated to A. This surface is built out of the cycle $((n_1,\dots,n_\ell))$ of A. For each class A, choose *once and for all* a fixed $w\mathbb{Z}+\mathbb{Z} \in A$ with w reduced, hence a fixed ℓ-tuple $(n_1,\dots n_\ell)$. For each $k = 1,\dots \ell$, define the elliptic element of order 2 in Γ:

$$S_k = T^{(n_1+\cdots n_k)}ST^{-(n_1+\cdots n_k)}. \tag{3.1}$$

Consider the subgroup of the modular group

$$\Gamma_A = \langle S_1, S_2, \dots, S_\ell, T^m\rangle = \langle S, S_1, \dots, S_{\ell-1}, T^m\rangle, \tag{3.2}$$

where m was defined in (2.1). We will show below in Theorem 1 that Γ_A is an infinite index (i.e., *thin*) subgroup of Γ, hence a Fuchsian group of the second kind. A different choice of $w\mathbb{Z}+\mathbb{Z} \in A$ with reduced w leads to a conjugate subgroup Γ_A in Γ, in fact conjugate by a translation. In case $\ell = 1$ we have that $\Gamma_A = \langle S, T^{n_1}\rangle$, which is among those studied by Hecke [26].

Let $\mathcal{N}_A = \mathcal{N}_{\Gamma_A}$ be the Nielsen region of Γ_A and $\mathcal{F}_A = \mathcal{F}_{\Gamma_A}$ the associated surface. Before giving its properties, it is useful to see some examples.

Example. Consider the quadratic field $\mathbb{Q}(\sqrt{7})$, for which $D = 28 = 4 \cdot 7$. There are two classes: the principal class I with associated cycle $((3, 6))$ and J with cycle $((3, 3, 2, 2, 2))$.

The fundamental norm one unit is $\varepsilon_{28} = 8 + 3\sqrt{7}$. The class I contains

$$(\tfrac{3+\sqrt{7}}{2})\mathbb{Z} + \mathbb{Z}$$

with reduced $w = \frac{3+\sqrt{7}}{2} = [\![\overline{3, 6}]\!]$. A fundamental domain for the Fuchsian group of the second kind

$$\Gamma_I = \langle S, T^3 S T^{-3}, T^9 \rangle$$

is indicated in Figure 1. It has signature $(0; 2, 2; 1, 1)$. The surface $\mathcal{F}_I$ is depicted in Figure 2 and is bounded from below by the simple closed geodesic $\partial\mathcal{F}_I$ consisting of the two large circular arcs. The length of $\partial\mathcal{F}_I$ is $2\log(8+3\sqrt{7})$ and the area of $\mathcal{F}_I$ is 2π. Another depiction is in Figure 3, where the two distinguished points are the points of order 2 and segments connect them to the boundary geodesic.

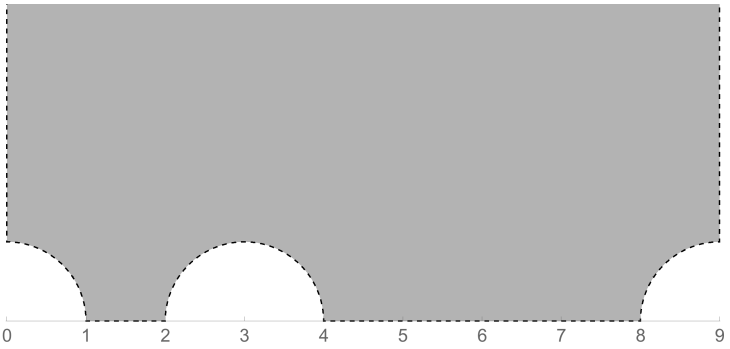

Figure 1. Fundamental domain for Γ_I when $d = 28$.

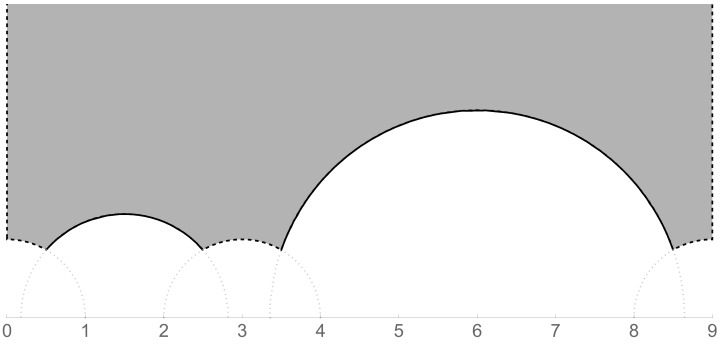

Figure 2. The surface $\mathcal{F}_I$.

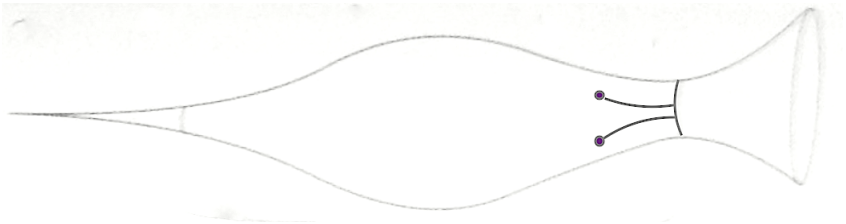

Figure 3. The surface $\mathcal{F}_I$.

The other class J contains the ideal $(\frac{5+\sqrt{7}}{3})\mathbb{Z}+\mathbb{Z}$ with reduced

$$\tfrac{5+\sqrt{7}}{3} = [\![\overline{3,3,2,2,2}]\!].$$

A fundamental domain for the Fuchsian group of the second kind

$$\Gamma_J = \langle S, T^3ST^{-3}, T^6ST^{-6}, T^8ST^{-8}, T^{10}ST^{-10}, T^{12}\rangle$$

is indicated in Figure 4. It has signature $(0;2,2,2,2,2,2;1,1)$. The surface $\mathcal{F}_J$ is pictured in Figure 5. It has area 5π. The closed geodesic that bounds $\mathcal{F}_J$ also has length $2\log(8+3\sqrt{7})$.

When either surface $\mathcal{F}_I$ or $\mathcal{F}_J$ is mapped to $\mathcal{F}$ we obtain overlapping polygons and the image of their boundaries are the closed geodesics $\mathcal{C}_I$ and $\mathcal{C}_J$, which have the same image as sets but with opposite orientations. This is depicted in Figure 6.

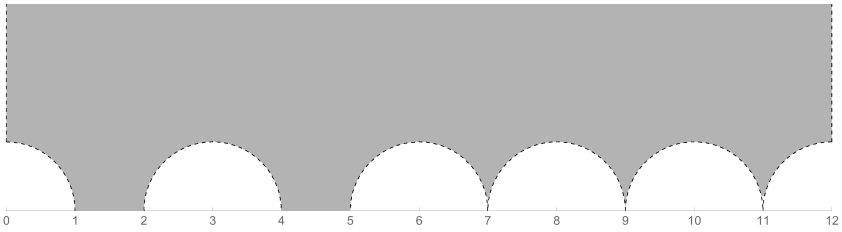

Figure 4. Fundamental domain for Γ_J in case $d = 28$.

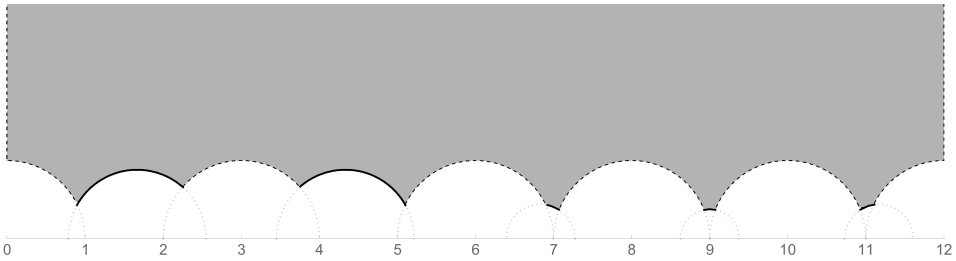

Figure 5. The surface $\mathcal{F}_J$.

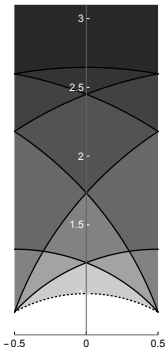

Figure 6. Projection of $\mathcal{F}_I$ and $\partial\mathcal{F}_I$ to the modular surface.

THEOREM 1. *The group Γ_A defined in* (3.2) *is Fuchsian of the second kind with signature*

$$(0; \underbrace{2, \dots, 2}_{\ell \text{ times}}; 1; 1).$$

The hyperbolic Riemann surface $\mathcal{F}_A$ thus has genus 0, *contains ℓ points of order* 2 *and has one cusp and one boundary component. The boundary $\partial\mathcal{F}_A$ is a simple closed geodesic whose image in $\mathcal{F}$ is $\mathcal{C}_A$. We have*

$$\text{length}(\partial\mathcal{F}_A) = 2\log\epsilon_D \quad \text{and} \quad \text{area}(\mathcal{F}_A) = \pi\ell_A. \tag{3.3}$$

The conformal class of $\mathcal{F}_A$ determines A.

Proof. The first two statements of Theorem 1 follow easily from an examination of the fundamental domain for

$$\Gamma_A = \langle S, S_1, \dots, S_{\ell-1}, T^m \rangle$$

constructed as in the examples above. That this construction is valid is an easy consequence of the Poincaré theorem for fundamental polygons [42] (see also [39]). It also follows that Γ_A is isomorphic to the free product

$$\mathbb{Z} * \underbrace{\mathbb{Z}/2\mathbb{Z} * \cdots * \mathbb{Z}/2\mathbb{Z}}_{\ell \text{ times}}.$$

Note that the unique boundary circle of $\Gamma_A\backslash\mathbb{H}^*$ can be visualized by identifying endpoints of the intervals on $\mathbb{R}$ bounding the fundamental domain using elliptic elements and the translation of Γ_A.

We next show that the boundary component of $\mathcal{F}_A$ is a simple closed geodesic whose image in $\Gamma\backslash\mathcal{H}$ is $\mathcal{C}_A$. Recall that we have fixed a choice of reduced w for each ideal class A. Using the minus continued fraction of w we have by [33] that for γ_w from (2.9),

$$\gamma_w = S_1 S_2 \cdots S_\ell T^m, \tag{3.4}$$

where S_k is given in (3.1) and m in (2.1). In particular,

$$\gamma_w \in \Gamma_A = \langle S_1, S_2, \dots, S_\ell, T^m \rangle.$$

Set $w_1 = w$ and for $k = 2, \dots \ell$ set $w_k = S_{k-1} \cdots S_1(w)$. By the construction of the minus continued fraction expansion we have that

(3.5)
$$n_1 + \cdots + n_{k-1} < w_k^\sigma < n_1 + \cdots + n_{k-1} + 1 \le n_1 + \cdots + n_k - 1 < w_k < n_1 + \cdots + n_k$$

for $k = 1, \dots \ell$ where, when $k = 1$, the first two inequalities mean that $0 < w^\sigma < 1$. Therefore the intersection of the unit circle with the geodesic in $\mathcal{H}$ with endpoints w^σ and w exists and defines a point z. We have by (3.4) that

$$\gamma_w(z) = \gamma(z) = S_1 S_2 \cdots S_\ell T^m(z),$$

and so

$$T^{-m} S_\ell \cdots S_2 S_1 \gamma(z) = z. \tag{3.6}$$

The circular arc from z to $\gamma(z)$ will intersect the circle with equation $(x - n_1)^2 + y^2 = 1$ at some z^* since by (3.5) we have that $n_1 - 1 < w < n_1$. The image of the arc from z^* to $\gamma(z)$ under S_1 covers part of the geodesic joining w_2^σ to w_2. Again the excess arc from $S_1(z^*)$ to $S_1\gamma(z)$ will intersect the circle $(x - n_1 - n_2)^2 + y^2 = 1$ at some z^{**} since $n_1 + n_2 - 1 < w_2 < n_1 + n_2$. Using now S_2 we can map the new excess arc from $S_2(z^{**})$ to $S_2 S_1 \gamma(z)$. We can repeat this process of cutting off arcs until we have applied S_ℓ. Now observe that by (3.6), upon application of T^{-m}, we have returned to z. Since the maps are orientation preserving isometries, we see that the bounding geodesic arcs piece together to give exactly one copy of $\mathcal{C}_A$, known to have length $2 \log \varepsilon_D$. See Figure 7 for an illustration of this argument when $w = \frac{3+\sqrt{7}}{2}$ from our first example above.

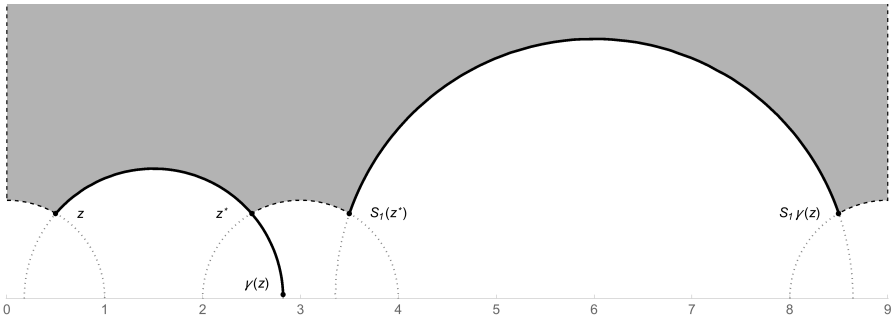

Figure 7. Cutting up $\partial \mathcal{F}_A$.

It is easily seen that the constructed geodesic is freely homotopic to the boundary circle of $\Gamma_A\backslash\mathcal{H}^*$ and hence by uniqueness is the boundary curve of $\mathcal{F}_A$. Furthermore, $\mathcal{F}_A$ can be identified with the intersection of the fundamental domain for Γ_A with the region above this boundary curve.

The fact that the area of $\mathcal{F}_A$ is $\pi\ell_A$ is an immediate consequence of (2.8).

Finally we must show that the conformal type of $\mathcal{F}_A$ determines A to complete the proof of Theorem 1. We will do this by demonstrating that this conformal type determines the cycle $((n_1,\dots n_\ell))$. By the above construction of $\mathcal{F}_A$, each elliptic fixed point in $\mathcal{F}_A$ determines a unique point on the boundary geodesic that is closest to it. The boundary geodesic (which is simple and oriented) determines an ordering of these points, which is unique up to cyclic permutations. This determines an ordering of the elliptic fixed points. Using (2.7) we compute the cycle of hyperbolic distances between successive fixed points of $S_0, S_1, \dots S_\ell$ in $\mathcal{H}$. This is given by $(V(n_1), V(n_2), \dots V(n_\ell))$, where $V(x)$ is the monotone increasing function

$$V(x) = \log\left(\frac{x}{2}\left(\sqrt{x^2+4}+x\right)+1\right).$$

The cycle of distances is a conformal invariant since these distances and the orientation of the boundary geodesic are preserved under conformal equivalence. The cycle of distances clearly determines the cycle $((n_1,\dots n_\ell))$ since V is monotone increasing.

This completes the proof of Theorem 1. □

4. **Uniform distribution**

In this section we state the main result of this paper. To obtain satisfactory results about the uniform distribution of $\mathcal{F}_A$, we average over a genus of ideal classes of $\mathbb{K}$. A genus is an element of the group of genera, which is (isomorphic to) the quotient group

$$\mathrm{Gen}(\mathbb{K}) = \mathrm{Cl}^+(\mathbb{K})/(\mathrm{Cl}^+(\mathbb{K}))^2. \tag{4.1}$$

It is classical that $\mathrm{Gen}(\mathbb{K}) \cong (\mathbb{Z}/2\mathbb{Z})^{\omega(D)-1}$ so if G_D is a genus in $\mathrm{Cl}^+(\mathbb{K})$, then

$$\#G_D = 2^{1-\omega(D)}h(D), \tag{4.2}$$

where $\omega(D)$ is the number of distinct prime factors of D.

THEOREM 2. *Suppose that for each positive fundamental discriminant $D>1$, we choose a genus $G_D \in \mathrm{Gen}(\mathbb{K})$. Let Ω be an open disc contained in the fundamental domain $\mathcal{F}$ for $\Gamma = \mathrm{PSL}(2,\mathbb{Z})$, and let $\Gamma\Omega$ be its orbit under the action of Γ. We have*

$$\tfrac{\pi}{3}\sum_{A\in G_D}\mathrm{area}(\mathcal{F}_A\cap\Gamma\Omega) \sim \mathrm{area}(\Omega)\sum_{A\in G_D}\mathrm{area}(\mathcal{F}_A), \tag{4.3}$$

as $D\to\infty$ through fundamental discriminants.

In view of Theorem 1, the uniform distribution of closed geodesics proven in [12] (generalized to genera) can be stated in the following form:[2]

$$\frac{\pi}{3}\sum_{A\in G_D}\operatorname{length}(\partial\mathcal{F}_A\cap\Gamma\Omega)\sim\operatorname{area}(\Omega)\sum_{A\in G_D}\operatorname{length}(\partial\mathcal{F}_A) \tag{4.4}$$

as $D\to\infty$ through fundamental discriminants.

The statement of (4.4) given in [12] has averaging over the entire class group. Unlike (4.4), (4.3) is actually trivial when one averages over the whole group since we get an even covering in that case and the $\sim$ can be replaced by equality. The reason is that $\mathcal{F}_A$ and $\mathcal{F}_{JA^{-1}}$ are complementary in that their union covers $\mathcal{F}$ evenly and the images of their boundary geodesics are the same as sets but with opposite orientations. For instance, the surfaces $\mathcal{F}_I$ and $\mathcal{F}_J$ are complementary. In general, (4.3) is trivial when J is in the principal genus. This happens if and only if D is not divisible by any primes $p\equiv 3 \pmod 4$ or, equivalently, when D is the sum of two squares (see, e.g., [24, Prop. 3.1]). In particular, for any class A that satisfies $A^2=J$, so that $\mathcal{C}_A$ is reciprocal, we have that $\mathcal{F}_A$ covers $\mathcal{F}$ evenly.

An interesting special case for which (4.3) is nontrivial is when $D=4p$ where $p\equiv 3 \pmod 4$ is prime. The case $p=7$ was illustrated above. There are exactly two genera, one containing I and the other containing J. Cohen and Lenstra [9] have conjectured that I and J are the only classes in their respective genera for $>75\%$ of such p. This happens exactly when $\mathbb{K}$ has wide class number one. Suppose that arbitrarily large such p exist. Then Theorem 2 and (4.4) imply that as $p\to\infty$ through such p, we have that

$$\frac{\operatorname{area}(\mathcal{F}_I\cap\Gamma\Omega)}{\operatorname{area}(\mathcal{F}_I)}\sim\frac{\operatorname{area}(\Omega)}{\operatorname{area}(\mathcal{F})}\quad\text{and}\quad\frac{\operatorname{length}(\partial\mathcal{F}_I\cap\Gamma\Omega)}{\operatorname{length}(\partial\mathcal{F}_I)}\sim\frac{\operatorname{area}(\Omega)}{\operatorname{area}(\mathcal{F})}.$$

Remarks. Since $\mathcal{F}_I$ and $\mathcal{F}_J$ are complementary, their distribution properties are directly related. A pretty class number formula of Hirzebruch and Zagier [28] (see also [57]) states that for such $p>3$,

$$\ell_J-\ell_I=3h(-p),$$

where $h(-p)$ is the class number of the imaginary quadratic field $\mathbb{Q}(\sqrt{-p})$. Upon using that $\operatorname{area}(\mathcal{F}_A)=\pi\ell_A$, this is equivalent to the area formula

$$\operatorname{area}(\mathcal{F}_J)-\operatorname{area}(\mathcal{F}_I)=3\pi h(-p).$$

There is a third hyperbolic distribution problem, one associated to imaginary quadratic fields. For $\mathbb{K}=\mathbb{Q}(\sqrt{D})$ with $D<0$, we may again associate to each ideal class A a geometric object, a CM point we denote by $z_A\in\mathcal{F}$ where

[2]Recall our convention concerning $\partial\mathcal{F}_A$.

$z_A\mathbb{Z}+\mathbb{Z}\in A$. Choose for each D a genus G_D, noting that (4.1) and (4.2) are valid for $D<0$. Then by [12] generalized to genera we have that

$$\tfrac{\pi}{3}\,\#\{A\in G_D \mid z_A\in\Omega\}\sim \operatorname{area}(\Omega)\,\#G_D \tag{4.5}$$

as $D\to-\infty$ through fundamental discriminants.

5. **The analytic approach**

Here we give a brief review of the analytic method and then state the extensions of formulas of Hecke and Katok-Sarnak that we will use to prove Theorem 2. Since it creates no new difficulties, we will allow both positive and negative D and set things up so that only obvious modifications are needed to prove the other two uniform distribution results (4.4) and (4.5). The analytic approach that we follow is based on the spectral theory of the Laplacian for automorphic forms and strong sub-convexity estimates for L-values, or equivalently nontrivial estimates of Fourier coefficients of modular forms of half-integral weight. Standard references for this section are Hejhal's book [27], the book of Iwaniec [30] and that of Iwaniec and Kowalski [31]. Some other related distribution problems are treated in Sarnak's book [45].

In this paper we will make use of many standard special functions, including the Bessel functions I_s, J_s, K_s and the Whittaker functions $M_{r,s}, W_{r,s}$. Some standard references for their properties are [38] and [55].

Spectral expansion. First we review the spectral expansion. The initial idea is to employ hyperbolic Weyl integrals, which are analogous to the usual Weyl sums used in proving the uniform distribution of sequences of points on a circle. One approximates the characteristic function of $\Gamma\Omega$ from above and from below by smooth Γ-invariant functions with compact support. If $f:\mathcal{H}\to\mathbb{R}^+$ is such a function, we expand it spectrally:

$$f(z)=c_0+\frac{1}{4\pi}\int_{-\infty}^{\infty}c(t)E(z,\tfrac12+it)dt+\sum_{\varphi}c(\varphi)\langle\varphi,\varphi\rangle^{-1}\varphi(z), \tag{5.1}$$

where $\langle\varphi,\varphi\rangle=\int_{\mathcal{F}}|\varphi(x)|^2d\mu(z)$. Here $E(z,s)$ is the Eisenstein series of weight 0 given for $\operatorname{Re}(s)>1$ by

$$E(z,s)=\sum_{\gamma\in\Gamma_\infty\backslash\Gamma}(\operatorname{Im}\gamma z)^s=\tfrac12(\operatorname{Im}z)^s\sum_{\gcd(c,d)=1}|cz+d|^{-2s}, \tag{5.2}$$

where Γ_∞ is the subgroup of Γ generated by T. Clearly $E(z,s)$ is an eigenfunction of

$$\Delta=-y^{-2}(\partial_x^2+\partial_y^2)$$

with eigenvalue $\lambda = s(1-s)$. If we define $E^*(z,s) = \Lambda(2s)E(z,s)$, the Fourier expansion of $E^*(z,s)$ is given by (see, e.g., [31])

$$E^*(z,s) = \Lambda(2s)y^s + \Lambda(2-2s)y^{1-s} + 2y^{1/2}\sum_{n\neq 0}|n|^{s-1/2}\sigma_{1-2s}(|n|)K_{s-\frac{1}{2}}(2\pi|n|y)e(nx), \tag{5.3}$$

where $\Lambda(s) = \pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)$. Then $E^*(z,s)$ is entire except at $s = 0, 1$, where it has simple poles and satisfies the functional equation

$$E^*(z,1-s) = E^*(s). \tag{5.4}$$

Furthermore, we have that

$$\mathrm{Res}_{s=1}E^*(z,s) = -\mathrm{Res}_{s=0}E^*(z,s) = \tfrac{1}{2}. \tag{5.5}$$

The residue at $s = 1$ gives rise to constant term c_0 in (5.1).

The second sum in (5.1) is over the countably infinite set of Hecke-Maass cusp forms φ. Like the Eisenstein series, these are Maass forms in that they are Γ-invariant eigenfunctions of Δ with $\Delta\varphi = \lambda\varphi$, where we express the eigenvalue uniquely as

$$\lambda = \lambda(\varphi) = \tfrac{1}{4} + r^2 \tag{5.6}$$

and choose $r \geq 0$. Being a Hecke-Maass cusp form means that, in addition, φ is an eigenfunction of all the Hecke operators, that $\|\varphi\|^2 = \langle\varphi,\varphi\rangle < \infty$ and that the constant term in its Fourier expansion at $i\infty$ is zero. We can and always will normalize such a Hecke-Maass cusp form φ so that this Fourier expansion has the form[3]

$$\varphi(z) = 2y^{1/2}\sum_{m\neq 0}a(m)K_{ir}(2\pi|m|y)e(mx), \tag{5.7}$$

where $a(1) = 1$. We can also assume that

$$a(-n) = a(-1)a(n) = \pm a(n).$$

If $a(-1) = 1$, we say that φ is even, otherwise odd since $\varphi(-\overline{z}) = a(-1)\varphi(z)$ or equivalently $\overline{\varphi}(z) = a(-1)\varphi(z)$. Thus the associated L-function has a Euler product (for $\mathrm{Re}(s) > 1$):

$$L(s;\varphi) = \sum_{n\geq 1}a(n)n^{-s} = \prod_{p\text{ prime}}(1 - a(p)p^{-s} + p^{-2s})^{-1}. \tag{5.8}$$

Furthermore, its completion

$$\Lambda(s;\varphi) = \pi^{-s}\Gamma(\tfrac{s+ir+\epsilon}{2})\Gamma(\tfrac{s-ir+\epsilon}{2})L(s;\varphi) \tag{5.9}$$

[3]Note the 2 in front!

is entire and satisfies the functional equation $\Lambda(s;\varphi) = (-1)^{\epsilon}\Lambda(1-s;\varphi)$, where $\epsilon = \frac{1-a(-1)}{2}$.

Remark. Note that the Eisenstein series is also an even Hecke eigenform and that its associated L-function

$$L(s;t) = \sum_{n\geq 1} n^{it}\sigma_{-2it}(m)m^{-s} = \zeta(s+it)\zeta(s-it),$$

defined for a fixed t, satisfies $\Lambda(s,t) = \pi^{-s}\Gamma(\frac{s+it}{2})\Gamma(\frac{s-it}{2})L(s,t) = \Lambda(1-s,t)$. Unlike $\Lambda(s;\varphi)$, it has poles, reflecting the fact that $E(z,s)$ is not a cusp form.

Weyl's law gives that as $x \to \infty$,

$$\#\{\varphi; \lambda(\varphi) \leq x\} \sim \frac{x}{12}. \tag{5.10}$$

The first five values of λ to five decimal places (see [5]) are

$$91.14134,\ \ 148.43213,\ \ 190.13154,\ \ 206.41679,\ \ 260.68740. \tag{5.11}$$

It appears to be likely that each λ is simple, but this is open. The eigenvalues in (5.11) all belong to odd forms except the third.

For our f, the spectral expansion (5.1) converges uniformly on compact subsets of $\mathcal{H}$.

Hyperbolic Weyl integrals. The Weyl integrals give the remainder terms in the asymptotics and are of two types depending on whether they come from the Eisenstein series or the Hecke-Maass cusp forms. Let $u(z)$ denote either $E(z,s)$ for $\mathrm{Re}(s) = 1/2$ or $\langle\varphi,\varphi\rangle^{-1}\varphi(z)$. Note that $E(z,s)$ is absolutely integrable over $\mathcal{F}_A$ for $\mathrm{Re}(s) = 1/2$ by (5.3). To pick out genera we need genus characters, or what is the same thing, real characters of $\mathrm{Cl}^+(\mathbb{K})$. These are in one-to-one correspondence with factorizations $D = d'd$, where d', d are fundamental discriminants. See Section 7 for more information about the genus characters. Given such a χ define

$$\mathrm{Weyl}(u,\chi) = \sum_{A\in \mathrm{Cl}^+(\mathbb{K})} \chi(A) \begin{cases} \frac{\lambda}{2}\int_{\mathcal{F}_A} u(z)d\mu(z) & \text{if } d', d < 0, \\ \int_{\partial\mathcal{F}_A} u(z)y^{-1}|dz| & \text{if } d', d > 0, \\ \frac{1}{\omega_D}\, u(z_A) & \text{if } d'd < 0. \end{cases} \tag{5.12}$$

Here $\omega_D = 1$ except that $\omega_{-3} = 3$ and $\omega_{-4} = 2$.

To prove uniform distribution by the analytic method we need estimates for $\mathrm{Weyl}(u,\chi)$ for real χ that are nontrivial in the D-aspect and uniform (but weak) in the spectral aspect. This is enough since the Weyl integral in (5.12) in the first case is zero when $d', d > 0$ as is that in the second case when $d', d < 0$.

When $u(z) = E(z,s)$, we apply a version of a classical formula of Hecke. Let $L(s,\chi_d)$ be the Dirichlet L-function with character given by the Kronecker

symbol $\chi_d(\cdot) = \left(\frac{d}{\cdot}\right)$, and for $\alpha = \frac{1}{2}(1 - \operatorname{sign} d)$ define the completed L-function

$$\Lambda(s, \chi_d) = \pi^{-s/2}\Gamma(\tfrac{s+\alpha}{2})|d|^{s/2}L(s, \chi_d). \tag{5.13}$$

THEOREM 3. *For the genus character χ associated to $D = d'd$ and* $\operatorname{Re}(s) = \frac{1}{2}$, *we have*

$$\Lambda(s, \chi_{d'})\Lambda(s, \chi_d) = \sum_{A \in \mathrm{Cl}^+(\mathbb{K})} \chi(A) \begin{cases} \frac{\lambda}{2} \int_{\mathcal{F}_A} E^*(z, s) d\mu(z) & \text{if } d', d < 0, \\ \int_{\partial \mathcal{F}_A} E^*(z, s) y^{-1} |dz| & \text{if } d', d > 0, \\ 2\sqrt{\pi}\omega_D^{-1}\ E^*(z_A, s) & \text{if } d'd < 0. \end{cases}$$

This formula, due to Hecke except when $d', d < 0$, reduces the problem of estimating nontrivially the Weyl integrals for Eisenstein series to obtaining a sub-convexity bound for Dirichlet L-functions in the conductor aspect. Of course, this is one of the first such to be done and is a famous result of Burgess.

Theorem 3 can be expressed in terms of Maass forms of weight 1/2. These generalize the (modified) Jacobi theta series,

$$\theta(z) = \operatorname{Im}(z)^{1/4} \sum_{n \in \mathbb{Z}} e(n^2 z),$$

which is a modular form of weight 1/2 for $\Gamma_0(4)$. Set

$$J(\gamma, z) = \frac{\theta(\gamma z)}{\theta(z)} \quad \text{for } \gamma \in \Gamma_0(4). \tag{5.14}$$

Say F defined on $\mathcal{H}$ has weight 1/2 for $\Gamma_0(4)$ if

$$F(\gamma z) = J(\gamma, z) F(z) \qquad \text{for all} \qquad \gamma \in \Gamma_0(4).$$

For fundamental d, set

$$b(d, s) = (4\pi)^{-1/4} |d|^{-3/4} \Lambda(s, \chi_d),$$

and define $b(dm^2, s)$ for $m \in \mathbb{Z}^+$ by means of the Shimura relation

$$m \sum_{\substack{n|m \\ n>0}} n^{-\frac{3}{2}} \left(\tfrac{d}{n}\right) b\left(\tfrac{m^2 d}{n^2}, s\right) = m^{s-1/2} \sigma_{1-2s}(m) b(d, s).$$

Then it follows from [16, Prop. 2 p. 959] that

$$\begin{aligned} E^*_{1/2}(z, s) = {} & \Lambda(2s) 2^s y^{\frac{s}{2} + \frac{1}{4}} + \Lambda(2 - 2s) 2^{1-s} y^{\frac{3}{4} - \frac{s}{2}} \\ & + \sum_{\substack{n \equiv 0, 1 \pmod 4 \\ n \neq 0}} b(n, s) W_{\frac{1}{4} \operatorname{sgn} n, \frac{s}{2} - \frac{1}{4}}(4\pi |n| y) e(nx) \end{aligned}$$

has weight 1/2 for $\Gamma_0(4)$. The idea behind this example originates in the papers of H. Cohen [8] and Goldfeld and Hoffstein [21]. See also [50], [15].

The formula

$$\Lambda(s, \chi_{d'})\Lambda(s, \chi_d) = 2\sqrt{\pi} |D|^{3/4} b(d', s) b(d, s) \tag{5.15}$$

in connection with Theorem 3 hints strongly as to what should take place for cusp forms; this is the extension (and refinement) of the formula of Katok-Sarnak mentioned earlier. Their result from [34], together with [1], gives the case $d = 1$ in the following.[4]

THEOREM 4. *Let*

$$\varphi(z) = 2y^{1/2} \sum_{n\neq 0} a(n) K_{ir}(2\pi|n|y) e(nx)$$

be a fixed even Hecke-Maass cusp form for Γ. *Then there exists a unique nonzero* $F(z)$ *with weight* $1/2$ *for* $\Gamma_0(4)$ *with Fourier expansion*

$$F(z) = \sum_{\substack{n\equiv 0,1 (\mathrm{mod}\ 4)\\ n\neq 0}} b(n) W_{\frac{1}{4}\operatorname{sgn} n, \frac{ir}{2}}(4\pi|n|y) e(nx),$$

such that for any pair of co-prime fundamental discriminants d' *and* d, *we have* (5.16)

$$12\sqrt{\pi}|D|^{\frac{3}{4}}\, b(d')\bar{b}(d) = \langle \varphi, \varphi\rangle^{-1} \sum_{A\in \mathrm{Cl}^+(\mathbb{K})} \chi(A) \begin{cases} \frac{\lambda}{2}\int_{\mathcal{F}_A} \varphi(z) d\mu(z) & \text{if } d', d < 0,\\ \int_{\partial \mathcal{F}_A} \varphi(z) y^{-1} |dz| & \text{if } d', d > 0,\\ 2\sqrt{\pi}\,\omega_D^{-1}\ \varphi(z_A) & \text{if } d'd < 0, \end{cases}$$

where χ *is the genus character associated to* $D = d'd$. *Here*

$$\langle F, F\rangle = \int_{\Gamma_0(4)\backslash \mathcal{H}} |F|^2 d\mu = 1$$

and the value of $b(n)$ *for a general discriminant* $n = dm^2$ *for* $m \in \mathbb{Z}^+$ *is determined by means of the Shimura relation*

$$m \sum_{\substack{n|m\\ n>0}} n^{-\frac{3}{2}} \left(\frac{d}{n}\right) b\Big(\frac{m^2 d}{n^2}\Big) = a(m) b(d).$$

Remarks. The $\sqrt{\pi}$ in (5.15) and (5.16) is an artifact of the normalization of the Whittaker function. Also, if we choose F in Theorem 4 so that $\langle F, F\rangle = 6$, which is the index of $\Gamma_0(4)$ in Γ, then we get 2 in the left-hand side of (5.16) instead of 12, which matches the Eisenstein series case (5.15). Perhaps not coincidentally,

$$\mathrm{Res}_{s=1} E^*_{1/2}(z, s) = \tfrac{1}{2}\theta(z),$$

and by [7] we have $\langle \frac{1}{2}\theta(z), \frac{1}{2}\theta(z)\rangle = 6$.

It is also possible to evaluate $|b(d)|^2$. When $d = 1$, this was done in [34] and in general by Baruch and Mao [1]. Here we quote their result in our

[4]Except that when $d' < 0$, in (5.16) we get $2\sqrt{\pi}\,\omega_D^{-1}$ on the right-hand side instead of their $(2\sqrt{\pi}\,\omega_D)^{-1}$.

context. Under the same assumptions as in Theorem 4 we have

$$12\pi|d||b(d)|^2 = \langle\varphi,\varphi\rangle^{-1}\Gamma(\tfrac{1}{2}+\tfrac{ir}{2}-\tfrac{\operatorname{sign} d}{4})\Gamma(\tfrac{1}{2}-\tfrac{ir}{2}-\tfrac{\operatorname{sign} d}{4})L(\tfrac{1}{2},\varphi,\chi_d), \tag{5.17}$$

where

$$L(s,\varphi,\chi_d) = \sum_{n\geq 1}\chi_d(n)a(n)n^{-s}.$$

Hence in the cuspidal case our problem also reduces to obtaining a subconvexity bound, this time for a twisted L-function.

Results like Theorem 4 and (5.17) have a long history, especially in the holomorphic case. Some important early papers are those by Kohnen and Zagier [36], Shintani [49] and Waldspurger [54]. All of these relied on the fundamental paper of Shimura [48].

Examples. It is interesting to evaluate numerically some examples of Theorem 4. This is possible thanks to computations done by Strömberg [52]. Note that half-integral weight Fourier coefficients, even in the holomorphic case, are notoriously difficult to compute.

For example, for $\varphi(z)$, we take the first occurring even Hecke-Maass form with eigenvalue

$$\lambda = 190.13154731 = \cdots = \tfrac{1}{2}+r^2,$$

where $r/2 = 6.889875675\cdots$. We have

$$\langle\varphi,\varphi\rangle = 7.26300636\times 10^{-19}.$$

A large number of Hecke eigenvalues for this φ are given (approximately, but with great accuracy) in the accompanying files of the paper of Booker, Strömbergsson and Venkatesh [5]. The first six values to twelve places are given in Table 1.

p	$a(p)$
2	1.549304477941
3	0.246899772453
5	0.737060385348
7	-0.261420075765
11	-0.953564652617
13	0.278827029162

Table 1. Hecke eigenvalues

A few values of $b(d)$ for fundamental d (except for $d=1$, which we computed independently) are computed from Strömberg's Table 5 and given in our Table 2.

$d > 0$	$b(d)$	$d < 0$	$b(d)$
1	10894.40532	-3	6404.69711
5	894.31877	-4	11927.63292
8	2191.95607	-7	8495.02618
12	-1298.74136	-8	-4512.60385

Table 2. Weight 1/2 coefficients

Let us illustrate Theorem 4 in a few cases. Consider first the quadratic field $\mathbb{Q}(\sqrt{3})$, for which $D = 12 = 4 \cdot 3$. There are two classes: the principal class I with associated cycle $(\!(4)\!)$ and J with cycle $(\!(2,3)\!)$. For $D = 12 = (1)(12)$,

$$12^{7/4}\sqrt{\pi}\; b(1)b(12) = 2\langle\varphi,\varphi\rangle^{-1}\int_{\partial\mathcal{F}_I}\varphi(z)y^{-1}|dz| = -1.94029\times 10^9$$

and for $D = (-3)(-4)$,

$$12^{7/4}\sqrt{\pi}\; b(-3)b(-4) = \lambda\langle\varphi,\varphi\rangle^{-1}\int_{\mathcal{F}_I}\varphi(z)d\mu(z) = 1.04759\times 10^{10}.$$

Two examples when $D < 0$: $D = (1)(-3)$,

$$18\; 3^{3/4}\; b(1)b(-3) = \langle\varphi,\varphi\rangle^{-1}\varphi(\tfrac{1+\sqrt{-3}}{2}) = 2.86296\times 10^9$$

and $D = (1)(-4)$,

$$12\; 4^{3/4}\; b(1)b(-4) = \langle\varphi,\varphi\rangle^{-1}\varphi(i) = 4.41046\times 10^9.$$

In these examples the integrals and special values were computed by approximating φ by its Fourier expansion and using the Fourier coefficients given in the files accompanying [5].

6. **Proof of Theorem 2**

We now show how to deduce Theorem 2 (and (4.4) and (4.5)) from Theorems 3 and 4. In order to show that we actually have an asymptotic formula we need a lower bound for the main term that is larger than the remainder terms. The main term comes from the constant c_0 in the spectral expansion (5.1). It is a little more complicated to obtain a lower bound for the main term in (4.3) than the corresponding bounds for geodesics or CM points, which we get almost directly from Siegel's theorem. For the geodesic case, we have by the class number formula and (4.2) that

$$\sum_{A\in G_D}\operatorname{length}(\partial\mathcal{F}_A) = 2^{2-\omega(D)}h(D)\log\varepsilon_D.$$

Similarly, when $D < 0$ we have

$$\#G_D = 2^{1-\omega(D)}h(D).$$

By Siegel's theorem we obtain that the main term in either case is $\gg_\epsilon |D|^{1/2-\epsilon}$ for any $\epsilon > 0$, where the implied constant is not effective.

Unlike the lengths of the closed geodesics, the areas of the surfaces $\mathcal{F}_A$ are not the same for different A! Still, we have the needed lower bound.

PROPOSITION 1. *For any $\epsilon > 0$, we have that*

$$\sum_{A \in G_D} \operatorname{area}(\mathcal{F}_A) \gg_\epsilon D^{1/2-\epsilon}. \tag{6.1}$$

The implied constant is not effectively computable for a given ϵ.

Proof. We have by Theorem 1 that

$$\sum_{A \in G_D} \operatorname{area}(\mathcal{F}_A) = \sum_{A \in G_D} \ell_A. \tag{6.2}$$

We have the identity (see [58, p. 167] or [58, p. 138])

$$\prod_{w \text{ reduced}} w = \varepsilon_D. \tag{6.3}$$

Now for a reduced w, there are $a, b, c \in \mathbb{Z}$ with $D = b^2 - 4ac$ and

$$a, c > 0 \text{ and } a + b + c < 0$$

so that $w = \frac{-b+\sqrt{D}}{2a}$. Thus

$$\sqrt{D} \geq \tfrac{\sqrt{D}}{a} = w - w^\sigma > w - 1.$$

We conclude that $w < \sqrt{D} + 1$, so (6.3) easily implies that

$$\ell_A > \frac{\log \varepsilon_D}{\log(\sqrt{D}+1)}. \tag{6.4}$$

Using (6.2), (6.4), (4.2) and Siegel's theorem (see [11]), we derive (6.1). □

Remark. It is also possible to give an upper bound for ℓ_A. For example, Eichler [17] gave a general argument that yields for the modular group that

$$\ell_A < c \log \varepsilon_D$$

for an explicit c.

We now turn to estimating the Weyl integrals.

PROPOSITION 2. *There is a constant $C > 0$ such that for any $\epsilon > 0$, we have*

$$\operatorname{Weyl}(E(\cdot, s), \chi) \ll_\epsilon |s|^C |D|^{7/16+\epsilon}, \tag{6.5}$$

$$\operatorname{Weyl}(\langle \varphi, \varphi \rangle^{-1} \varphi, \chi) \ll_\epsilon r^C |D|^{13/28+\epsilon}, \tag{6.6}$$

where $\operatorname{Re}(s) = 1/2$ and φ is any even Hecke–Maass cusp form with Laplace eigenvalue $\frac{1}{4} + r^2$.

Proof. By Theorem 3 and standard estimates for the gamma function quotient and for $\zeta(2s)$, we have for $\mathrm{Re}(s) = 1/2$ that

$$\mathrm{Weyl}(s,\chi) \ll_\epsilon |s|^C |L(s,\chi_{d'})L(s,\chi_d)| D^{1/4+\epsilon}. \tag{6.7}$$

Thus (6.5) now follows from the subconvexity bound of Burgess [6] made uniform in s (see [31, Thm. 12.9 p. 329]): for any $\epsilon > 0$, we have

$$L(s,\chi_d) \ll |s||d|^{3/16+\epsilon},$$

where the implied constant depends only on ϵ.

Part (6.6) of Proposition 2 follows straight from Theorem 4 and Theorem 5 of [12]. □

To see that it is enough to restrict to even Maass cusp forms observe that for an odd form all the Weyl integrals are identically zero. To see this first observe that $\chi(A) = \chi(A^{-1})$. There is a symmetry under $A \to A^{-1}$ of all the geometric objects that forces the corresponding sum of integrals for A and A^{-1} to cancel for an odd form φ. For example, when $d', d < 0$ we have that

$$\int_{\mathcal{F}_{A^{-1}}} \varphi(z) d\mu(z) = -\int_{\mathcal{F}_A} \varphi(z) d\mu(z). \tag{6.8}$$

To get (6.8) observe that by (2.2) the cycle for A^{-1} is that for A reversed. This has the effect of making a left translate by T^{-m_A} of the fundamental domain $\mathcal{F}_{A^{-1}}$ a mirror image in the imaginary axis of $\mathcal{F}_A$. Here we are using the fundamental domains constructed in the proof of Theorem 1. The cases $d', d > 0$ and $d'd < 0$ are handled similarly by using (2.3).

Theorem 2 follows from Propositions 1 and 2 and the fact that the spectral coefficients in (5.1) satisfy

$$c(t) \ll |s|^{-A} \qquad \text{and} \qquad c(\varphi) \ll |r|^{-A}$$

for any $A > 0$ and by the Weyl law (5.10); see, e.g., [30].

Remarks. There has been a lot of progress on subconvexity estimates since the paper [12] that we quote was written. We were content to use the result of [12] here since *any* strong nontrivial estimate is enough to get the uniform distribution results. By "strong" we mean a power savings in the exponent, and this is required due to our use of Siegel's theorem for the main term.

After the fundamental paper of Iwaniec [29], techniques for dealing directly with the L-functions were developed in a series of papers starting with [13]. See also [14]. Currently the best known subconvexity bound for the L-functions (5.17) was obtained in the breakthrough paper [10] of Conrey and Iwaniec, which gives the exponent $1/3 + \epsilon$ of $|D|$ in both estimates of Proposition 2 but under the technical assumption that D is odd. This result was improved by Young [56], who gives the same value $1/3 + \epsilon$ for the exponent of C in these

estimates. See also the paper of Blomer and Harcos [4]. Although we have not pursued this here, such explicit hybrid estimates would allow one to improve the ranges of certain parameters in the distribution results.

In a different direction, it would be interesting to see if the methods of arithmetic ergodic theory could be applied here along the lines of the paper [18] of Einsiedler, Lindenstrauss, Michel and Venkatesh.

7. **Proof of Theorem 3**

To prove Theorem 3, we begin by giving some background on Hecke's L-functions. The zeta function for an ideal class A is given for $\mathrm{Re}(s) > 1$ by

$$\zeta_A(s) = \sum_{\mathfrak{a} \in A} N(\mathfrak{a})^{-s},$$

where $\mathfrak{a}$ runs over all *integral* ideals in A. When $\mathbb{K} = \mathbb{Q}(\sqrt{D})$ with $D < 0$, we have

$$\pi^{-s}\Gamma(s)\zeta_A(s) = \tfrac{2^s}{\omega_D}|D|^{-\frac{s}{2}}\, E^*(z_A, s). \tag{7.1}$$

In order to generalize this in the presence of an infinite groups of units, Hecke invented his famous trick of dividing out the action of the unit group on generators of principal ideals. For real quadratic fields, this procedure amounts to either integrating the Eisenstein series with respect to arc length over the associated closed geodesic $\mathcal{C}_A$ or integrating its derivative over this geodesic. He showed

$$\pi^{-s}\Gamma(\tfrac{s}{2})^2 D^{s/2}\big(\zeta_A(s) + \zeta_{JA}(s)\big) = 2\int_{\mathcal{C}_A} E^*(z,s) y^{-1} |dz| \tag{7.2}$$

and

$$\pi^{-s}\Gamma(\tfrac{s+1}{2})^2 D^{s/2}\big(\zeta_A(s) - \zeta_{JA}(s)\big) = 2\int_{\mathcal{C}_A} i\,\partial_z E^*(z,s)\, dz. \tag{7.3}$$

Hecke's L-function for a character χ of $\mathrm{Cl}^+(\mathbb{K})$ is given for $\mathrm{Re}(s) > 1$ by

$$L(s,\chi) = \sum_{\mathfrak{a}} \chi(\mathfrak{a}) N(\mathfrak{a})^{-s} = \prod_{\mathfrak{p}} (1 - N(\mathfrak{p})^{-s})^{-1},$$

where $\mathfrak{a}$ runs over all integral ideals in $\mathbb{K}$ and $\mathfrak{p}$ over all (finite) primes. Clearly

$$L(s,\chi) = \sum_{A} \chi(A)\zeta_A(s).$$

When $D < 0$, the completed L-function is

$$\Lambda(s,\chi) = \pi^{-s}\Gamma(\tfrac{s}{2})\Gamma(\tfrac{s+1}{2})|D|^{s/2}L(s,\chi) = 2^{1-s}\pi^{1/2-s}|D|^{s/2}\Gamma(s)L(s,\chi), \tag{7.4}$$

while when $D > 0$, it is

$$\Lambda(s,\chi) = \pi^{-s}\Gamma(\tfrac{s+\alpha}{2})\Gamma(\tfrac{s+\alpha}{2})D^{s/2}L(s,\chi),$$

where

$$\alpha = \alpha_\chi = \tfrac{1}{2}(1 - \chi(J)). \tag{7.5}$$

Then from (7.1) and (7.4) we have when $D < 0$ that

$$\Lambda(s, \chi) = \tfrac{2\sqrt{\pi}}{\omega_D} \sum_{A \in \mathrm{Cl}^+(\mathbb{K})} \chi(A) E^*(z_A, s), \tag{7.6}$$

while from (7.2) and (7.3) we have when $D > 0$ that

$$\Lambda(s, \chi) = \sum_{A \in \mathrm{Cl}^+(\mathbb{K})} \chi(A) \begin{cases} \int_{\mathcal{C}_A} E^*(z, s) y^{-1} |dz| & \text{if } \chi(J) = 1, \\ \int_{\mathcal{C}_A} i\, \partial_z E^*(z, s)\, dz & \text{if } \chi(J) = -1. \end{cases} \tag{7.7}$$

These equations show that the analytic properties of $\Lambda(s, \chi)$ are inherited from those of $E^*(z, s)$. For instance, (5.4) gives the functional equation

$$\Lambda(1 - s, \chi) = \Lambda(s, \chi).$$

When $\chi(J) = 1$, the narrow class character χ descends to a character of the wide class group, so these formulas are usually given in terms of the wide class group $\mathrm{Cl}(\mathbb{K})$ in this case. When $\chi(J) = -1$, the character is a norm class character (see [24], [25]).

Genus characters. The characters of $\mathrm{Cl}^+(\mathbb{K})$ that we need are the genus characters, which are the same as the real characters. See [51] for their basic theory. The genus characters are in one-to-one correspondence with decompositions $D = d'd$, where d and d' are fundamental discriminants. Given such a decomposition, the associated character χ is defined for a prime ideal $\mathfrak{p}$ of $\mathbb{K}$ by

$$\chi(\mathfrak{p}) = \begin{cases} \left(\frac{d}{N(\mathfrak{p})}\right) & \text{if } \mathfrak{p} \nmid d, \\ \left(\frac{d'}{N(\mathfrak{p})}\right) & \text{if } \mathfrak{p} \nmid d', \end{cases}$$

where $\left(\frac{d}{\cdot}\right)$ is the Kronecker symbol. From now on by χ we shall always mean, unless otherwise specified, the genus character associated to $D = d'd$. Later we need a generalization of χ to some cases where d is fundamental but d', hence D, need not be.

We have that $\chi(J) = \operatorname{sign} d = \operatorname{sign} d'$, and we also have Kronecker's decomposition $L(s, \chi) = L(s, \chi_{d'}) L(s, \chi_d)$. Equivalently,

$$\Lambda(s, \chi) = \Lambda(s, \chi_{d'}) \Lambda(s, \chi_d).$$

By (7.6) we have when $D < 0$ that

$$\Lambda(s, \chi_{d'}) \Lambda(s, \chi_d) = \tfrac{2\sqrt{\pi}}{\omega_D} \sum_{A \in \mathrm{Cl}^+(\mathbb{K})} \chi(A) E^*(z_A, s),$$

while by (7.7) we get when $D > 0$ that

$$\Lambda(s,\chi_{d'})\Lambda(s,\chi_d) = \sum_{A\in \mathrm{Cl}^+(\mathbb{K})} \chi(A) \begin{cases} \int_{\mathcal{C}_A} E^*(z,s) y^{-1} |dz| & \text{if } d', d > 0, \\ \int_{\mathcal{C}_A} i\,\partial_z E^*(z,s)\, dz & \text{if } d', d < 0. \end{cases} \tag{7.8}$$

Stokes' theorem. To get the surface case of Theorem 3 we need to express the Weyl surface integrals in terms of cycle integrals. Of course, the main tool for this is Stokes' theorem. We do the cusp form case at the same time.

LEMMA 1. *For u as in* (5.12), *we have*

$$\tfrac{\lambda}{2} \int_{\mathcal{F}_A} u(z) d\mu(z) = \int_{\mathcal{C}_A} i\,\partial_z u(z) dz. \tag{7.9}$$

By an integral over $\mathcal{C}_A$ we always mean the integral from $z_0 \in \mathcal{S}_w$ to $\gamma_w(z_0) \in \mathcal{S}_w$ along the arc on $\mathcal{S}_w$, assuming that the integral is independent of z_0.

A little more generally we have the following lemma. Recall that m was defined in (2.1).

LEMMA 2. *Suppose that $F(z)$ is any real analytic Γ_A-invariant function on $\mathcal{H}$ that satisfies*

$$\Delta F = -y^2(F_{xx} + F_{yy}) = s(1-s)F \tag{7.10}$$

and the growth condition $\int_0^m \partial_z F(x+iY)dx = o(1)$ as $Y \to \infty$. Then we have

$$\tfrac{s(1-s)}{2} \int_{\mathcal{F}_A} F(z) d\mu(z) = \int_{\partial \mathcal{F}_A} i\,\partial_z F(z) dz. \tag{7.11}$$

Proof. By Stokes' theorem we have

$$\int_{\mathcal{F}_A(Y)} \partial_{\overline{z}}(\partial_z F(z)) dz\, d\overline{z} = -\int_{\partial \mathcal{F}_A} \partial_z F(z) dz + \int_0^m \partial_z F(x+iY) dx,$$

where $\mathcal{F}_A(Y) = \{z \in \mathcal{F}_A; \mathrm{Im}(z) < Y\}$. Using that $dz\, d\overline{z} = -2i dx\, dy$, by (7.10) we have

$$\partial_{\overline{z}} \partial_z F(z) dz\, d\overline{z} = \tfrac{i}{2} s(1-s) F(z) d\mu(z).$$

By our growth assumption on F we get (7.11) by letting $Y \to \infty$. □

To deduce Lemma 1, note that both $E(z,s)$ and $\varphi(z)$ satisfy (7.10) and that the growth condition for φ is clear while that for $E(z,s)$ when $\mathrm{Re}(s) = 1/2$ follows from its Fourier expansion (5.3). Finally, since both $\varphi(z)$ and $E(z,s)$ are Γ-invariant we may replace the integrals over $\partial \mathcal{F}_A$ by integrals over $\mathcal{C}_A$.

Applying Lemma 1 when $u(z) = E(z,s)$ we finish the proof of Theorem 3.

8. **Maass forms and the resolvent kernel**

Our proof of Theorem 4 is similar in spirit to that of Hecke's for the Eisenstein series case. We will employ resolvent kernels for the Laplacians of weight 0 and weight 1/2. The residue of such a resolvent at a spectral point gives the reproducing kernel for the associated eigenspace. Our principal reference here is the paper of Fay [20]. Other references include Hejhal [27] and Roelcke [43].

We begin with the case of Maass cusp forms of weight 0 for Γ. For $\mathrm{Re}(s)>1$, consider the Poincaré series

$$F_m(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} f_m(\gamma z, s), \tag{8.1}$$

where $f_0(z,s) = y^s$ and for $m \neq 0$,

$$f_m(z,s) = y^{1/2} I_{s-1/2}(2\pi|m|y)e(mx) = \tfrac{|m|^{-1/2}}{2\pi}\tfrac{\Gamma(s)}{\Gamma(2s)} M_{0,s-\frac{1}{2}}(4\pi|m|y)e(mx).$$

The function $F_m(z,s)$, which was first studied by Neunhöffer [40] and Niebur [41], is a Γ-invariant eigenfunction of Δ:

$$\Delta F_m(z,s) = s(1-s)F_m(z,s).$$

We will get to the Maass cusp forms through residues of $F_m(z,s)$.

Proposition 3. *For any $m \neq 0$, we have that $F_m(z,s)$ has meromorphic continuation in s to $\mathrm{Re}(s) > 0$ and that*

$$\mathrm{Res}_{s=\frac{1}{2}+ir}(2s-1)F_m(z,s) = \sum_{\varphi}\langle\varphi,\varphi\rangle^{-1}2a(m)\varphi(z),$$

where the (finite) sum is over all Hecke-Maass cusp forms φ with Laplace eigenvalue $\frac{1}{4}+r^2$ and $a(m)$ is defined in (5.7).

Proof. For $u \in L^2(\mathcal{F}, d\mu)$, the resolvent kernel $G(z,z';s)$ for Δ satisfies

$$\big(\Delta - s(1-s)\big)\int_{\mathcal{F}} G(z,z';s)u(z)d\mu(z) = u(z'). \tag{8.2}$$

The function $F_m(z,s)$ occurs in the Fourier expansion of $G(z,z';s)$, which is given by

$$G(z,z';s) = \sqrt{y'}\sum_{m\in\mathbb{Z}} F_{-m}(z,s)K_{s-1/2}(2\pi|m|y')e(mx'), \tag{8.3}$$

valid when $y' > y$. This follows from Theorem 3.1 on page 173 of [20]. The meromorphic continuation of the resolvent kernel in s is proven in [27] (see also [43]). If $s = \frac{1}{2}+ir$ is such that the eigenspace with eigenvalue $\frac{1}{4}+r^2$

is nonempty and $\{u\}$ is an orthonormal basis for this eigenspace, then the integral kernel

$$H(z,z',s) = G(z,z';s) - \frac{1}{\frac{1}{4}+r^2-s(1-s)}\sum \overline{u(z)}u(z') \tag{8.4}$$

has a holomorphic extension to $s=\frac{1}{2}+ir$. This is a consequence of the fact that it represents the resolvent on the orthogonal complement of this eigenspace and vanishes identically on the eigenspace itself.

Now let $u=\frac{\varphi}{\|\varphi\|}$, where φ is a Hecke-Maass cusp form. The set $\{u\}$ of all such is an orthonormal basis for our eigenspace. Then by (8.3) and (8.4) we have for $0<y<y'$ that

$$\int_{-1/2}^{1/2} H(z,z',s)e(-mx')dx' = \sqrt{y'}F_{-m}(z,s)K_{s-1/2}(2\pi|m|y') \\ -\frac{1}{\frac{1}{4}+r^2-s(1-s)}\sum \langle\varphi,\varphi\rangle^{-1}\overline{\varphi(z)}2a(m)K_{ir}(2\pi|m|y'),$$

after referring to (5.7). Since the left-hand side is holomorphic at $s=\frac{1}{2}+ir$, we get that

$$\mathrm{Res}_{s=\frac{1}{2}+ir}(2s-1)F_{-m}(z,s) = \sum_{\varphi}\langle\varphi,\varphi\rangle^{-1}2a(m)\overline{\varphi(z)}.$$

Replacing m with $-m$ and noting that $a(-m)\overline{\varphi(z)} = a(m)\varphi(z)$ from below (5.7) we get Proposition 3. □

For comparison with the weight $1/2$ case that we will treat next, it is instructive to carry the analysis one step further. The Fourier expansion of $F_m(z,s)$ is given by (see [20],[16])

$$F_m(z,s) = f_m(z,s) + \tfrac{2|m|^{1/2-s}\sigma_{2s-1}(|m|)}{(2s-1)\Lambda(2s)}y^{1-s}] \\ + 2y^{1/2}\sum_{n\neq 0}\Phi(m,n;s)K_{s-\frac{1}{2}}(2\pi|n|y)e(nx),$$

where for $\mathrm{Re}(s)>1$,

$$\Phi(m,n;s) = \sum_{c>0} c^{-1}K(m,n;c)\cdot\begin{cases} I_{2s-1}(4\pi\sqrt{|mn|}\,c^{-1}) & \text{if } mn<0,\\ J_{2s-1}(4\pi\sqrt{|mn|}\,c^{-1}) & \text{if } mn>0.\end{cases}$$

Here $K(m,n;c)$ is the Kloosterman sum

$$K(m,n;c) = \sum_{\substack{a \,(\mathrm{mod}\, c)\\ (a,c)=1}} e\big(\tfrac{ma+n\bar{a}}{c}\big).$$

It follows that for fixed m, n with $mn \neq 0$, the function $\Phi(m, n; s)$ has meromorphic continuation to $\mathrm{Re}(s) > 0$ and

$$\mathrm{Res}_{s=\frac{1}{2}+ir}(2s-1)\Phi(-m,n;s) = 2\sum_{\varphi}\langle\varphi,\varphi\rangle^{-1}a(m)a(n),$$

where the sum is over all Hecke-Maass cusp forms φ for Γ with eigenvalue $\frac{1}{4}+r^2$.

There is a parallel (yet more intricate) development for Maass forms of weight $1/2$. The invariant Laplace operator of weight $1/2$ is given by

$$\Delta_{1/2} = -y^2(\partial_x^2+\partial_y^2) + \tfrac{1}{2}iy\partial_x.$$

This means that for all $\gamma \in \Gamma_0(4)$, we have

$$(\Delta_{1/2}F)(\gamma z) = J(\gamma, z)\Delta_{1/2}F(z).$$

A Maass form of weight $1/2$ for $\Gamma_0(4)$ has weight $1/2$, is smooth and satisfies $\Delta_{1/2}F = \lambda F$, where we write $\lambda = \lambda(F) = \frac{1}{4} + (\frac{r}{2})^2$. Usually we also require some growth conditions as well in the three cusps of $\Gamma_0(4)$. In particular, a Maass cusp form F is in $L^2(\Gamma_0(4)\backslash\mathcal{H}, d\mu)$ and has the further property that its zeroth Fourier coefficient in each cusp vanishes.

The resolvent kernel $G_{1/2}(z, z'; s)$ for $\Delta_{1/2}$ in this case was also studied by Fay [20] (see also [43]). It satisfies

$$\big(\Delta_{1/2} - s(1-s)\big)\int_{\Gamma_0(4)\backslash\mathcal{H}} G_{\frac{1}{2}}(z,z';s)u(z)d\mu(z) = u(z') \tag{8.5}$$

for $u \in L^2(\Gamma_0(4)\backslash\mathcal{H}, d\mu)$ with weight $1/2$. By Theorem 3.1 of [20] we have the Fourier expansion[5]

$$G_{1/2}(z', z; s) = \sum_n F_{1/2,n}(z,s)W_{\frac{1}{4}\operatorname{sign} n, s-\frac{1}{2}}(4\pi|n|y')e(-nx')$$

valid for $\mathrm{Im}\, z' > \mathrm{Im}\, z$, where for $n \neq 0$ and $\mathrm{Re}(s) > 1$,

$$F_{1/2,n}(z,s) = \frac{\Gamma(s-\frac{1}{4}\operatorname{sign} n)}{4\pi|n|\Gamma(2s)}\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} J(\gamma,z)^{-1}f_{1/2,n}(\gamma z, s) \tag{8.6}$$

with

$$f_{1/2,n}(z,s) = M_{\frac{1}{4}\operatorname{sign} n, s-\frac{1}{2}}(4\pi|n|\,\mathrm{Im}\, z)e(n\,\mathrm{Re}\, z).$$

As above it follows that $F_{1/2,n}(z,s)$ has a meromorphic continuation to $\mathrm{Re}(s) > 0$ with simple poles at the points $\frac{1}{2}+\frac{ir}{2}$ giving the discrete spectrum of $\Delta_{1/2}$ and that

$$\mathrm{Res}_{s=\frac{1}{2}+\frac{ir}{2}}(2s-1)G_{1/2}(z',z) = \sum \overline{\psi}(z')\psi(z)$$

[5] Note that in the notation of Fay, $F_{1/2,n}(z,s) = -\overline{F}_n(z,\overline{s})$. The minus sign comes from his definition of $\Delta_{1/2}$. We are also using his (38), which gives $G_{1/2}(z,z';s) = \overline{G}_{1/2}(z',z;\overline{s})$. Observe as well that for weight $1/2$, his $k = 1/4$.

and

$$\text{Res}_{s=\frac{1}{2}+\frac{ir}{2}}(2s-1)F_{1/2,n}(z,s)=\sum_{\psi}\overline{b}(n)\psi(z). \tag{8.7}$$

Here the sum is over an orthonormal basis $\{\psi\}$ of Maass cusp forms for V_r and $b(n)$ is defined by

$$\psi(z)=\sum_{n\neq 0}b(n)W_{\frac{1}{4}\operatorname{sign} n,\frac{ir}{2}}(4\pi|n|y)e(nx). \tag{8.8}$$

Plus space. There is an important distinguished subspace of V_r, denoted by V_r^+ and called after Kohnen the *plus space*, that contains those Maass cusp forms $\psi\in V_r$ whose n-th Fourier coefficient $b(n)$ vanishes unless $n\equiv 0,1 \pmod 4$. It is clearly invariant under $\Delta_{1/2}$. We shall apply to $F_{1/2,n}(z,s)$ from (8.6) the projection operator $\text{pr}^+ : V_r\to V_r^+$ defined by $\text{pr}^+=\frac{2}{3}WU+\frac{1}{3}$, where[6]

$$U\psi(z)=\tfrac{\sqrt{2}}{4}\sum_{\nu=0}^{3}\psi(\tfrac{z+\nu}{4}) \quad \text{and} \quad W\psi(z)=e^{\frac{i\pi}{4}}\left(\tfrac{z}{|z|}\right)^{-\frac{1}{2}}\psi(-\tfrac{1}{4z}).$$

We will need an expansion of each of the Fourier coefficients of $\text{pr}^+F_{1/2,m}(z,s)$ when $m\equiv 0,1 \pmod 4$. These involve certain Kloosterman sums of weight $1/2$ that we now recall. Let $\left(\frac{c}{a}\right)$ be the extended Kronecker symbol (see [48]), and set

$$\varepsilon_a=\begin{cases}1 & \text{if } a\equiv 1 \pmod 4,\\ i & \text{if } a\equiv 3 \pmod 4.\end{cases}$$

Then for $c\in\mathbb{Z}^+$ with $c\equiv 0 \pmod 4$ and $m,n\in\mathbb{Z}$,

$$K_{1/2}(m,n;c)=\sum_{a(\text{mod } c)}\left(\tfrac{c}{a}\right)\varepsilon_a e\left(\tfrac{ma+n\overline{a}}{c}\right)$$

defines the weight $1/2$ Kloosterman sum. Here $\overline{a}\in\mathbb{Z}$ satisfies $a\overline{a}\equiv 1 \pmod c$. It is convenient to define the modified Kloosterman sum

$$K^+(m,n;c)=(1-i)K_{1/2}(m,n;c)\times\begin{cases}1 & \text{if } c/4 \text{ is even},\\ 2 & \text{otherwise.}\end{cases}$$

It is easily checked that

$$K^+(m,n;c)=K^+(n,m;c)=\overline{K^+(n,m;c)}. \tag{8.9}$$

[6]The constant $\sqrt{2}$ that is not present in [16] is due to the factor $y^{1/4}$ that comes from our normalization.

It follows[7] from [16, Prop. 2, p. 959] that for $\mathrm{Re}(s) > 1$ and d any nonzero integer with $d \equiv 0, 1 \pmod 4$, we have

$$\mathrm{pr}^+ F_{1/2,d}(z,s) = \frac{2}{3}\frac{\Gamma(s-\frac{\operatorname{sign} d}{4})}{4\pi|d|\Gamma(2s)} M_{\frac{1}{4}\operatorname{sign} d, s-\frac{1}{2}}(4\pi|d|y)e(dx) + \sum_{n\equiv 0,1(4)} \Phi^+(n,d;s) W_{\frac{1}{4}\operatorname{sign} n, s-\frac{1}{2}}(4\pi|n|y)e(nx), \tag{8.10}$$

where for $n \neq 0$, we have

$$\Phi^+(n,d;s) = \frac{1}{|nd|^{\frac{1}{2}}} \frac{\Gamma(s-\frac{\operatorname{sign} n}{4})\Gamma(s-\frac{\operatorname{sign} d}{4})}{3\sqrt{\pi}\, 2^{2-2s}\, \Gamma(2s-\frac{1}{2})} \cdot \sum_{\substack{c\equiv 0(4)\\ c>0}} \frac{K^+(n,d;c)}{c} \begin{cases} I_{2s-1}\Big(\frac{4\pi\sqrt{|nd|}}{c}\Big) & \text{if } nd<0, \\ J_{2s-1}\Big(\frac{4\pi\sqrt{|nd|}}{c}\Big) & \text{if } nd>0. \end{cases} \tag{8.11}$$

As in [20, Cor. 3.6 p. 178] we have that $\Phi^+(n,d;s)$ has a meromorphic continuation to all s, and it is now straightforward to get from (8.10) and (8.7) the following residue formula.

PROPOSITION 4. *For fixed discriminants d', d, the function $\Phi^+(d',d;s)$ has meromorphic continuation to* $\mathrm{Re}(s) > 0$ *and*

$$\mathrm{Res}_{s=\frac{1}{2}+\frac{ir}{2}}(2s-1)\Phi^+(d',d;s) = \sum_{\psi} b(d')\bar{b}(d),$$

where the sum is over an orthonormal basis of cusp forms ψ for V_r^+ and $b(d)$ is the Fourier coefficient of ψ as in (8.8).

9. **Cycle integrals of Poincaré series**

We next give an identity from which the extended Katok–Sarnak formula will be derived. Our main source is [16], where other relevant references are also given. As in the previous section, we will deal with general discriminants. This causes no essential new difficulties and makes it easier to quote some of our previous results. It also makes it clear how one could approach our main theorem for nonfundamental discriminants. Thus we need to define genus characters for arbitrary discriminants. We will use the language of binary quadratic forms. As in [16], let $\mathcal{Q}_D$ be the set of Q with discriminant D that are positive definite when $D < 0$. For $Q = [a,b,c]$ with discriminant $D = d'd$

[7]There is a typo in (2.19) of [16]. It should read $P_d^+(\tau,s) = \frac{3}{2}\mathrm{pr}^+(P_d(\tau,s))$.

where d is fundamental, we define

$$\chi(Q) = \begin{cases} \left(\frac{d}{m}\right) & \text{if } (a,b,c,d) = 1 \text{ where } Q \text{ represents } m \text{ and } (m,d) = 1, \\ 0, & \text{if } (a,b,c,d) > 1. \end{cases}$$

In case $D > 0$ is fundamental and Q is any binary quadratic form associated to $\mathfrak{a} \in A$ as in Section 2, we have that $\chi(Q) = \chi(A)$. When $D > 0$ and Q is primitive and $n \in \mathbb{Z}^+$, define $\mathcal{C}_{nQ} = \mathcal{C}_Q$ as in [16] using (2.11).[8] When $D < 0$, let $z_Q = \frac{-b+\sqrt{D}}{2a} \in \mathcal{H}$ if $Q = [a,b,c]$ and let ω_Q be the number of automorphs of Q in Γ.

The following result together with Propositions 3 and 4, will be used to derive the extended Katok-Sarnak formula. The first and second parts follow directly from [16], but we include them here for the sake of completeness. Recall that F_m was defined in (8.1) and Φ^+ in (8.11).

PROPOSITION 5. *Let $m \neq 0$ and $\mathrm{Re}(s) > 1$. Suppose that d is a fundamental discriminant and that d' is any discriminant such that $D = d'd$ is not a square. Then*

$$6\pi^{\frac{1}{2}}|D|^{\frac{3}{4}}|m| \sum_{\substack{n|m \\ n>0}} n^{-\frac{3}{2}} \left(\tfrac{d}{n}\right) \Phi^+\left(d', \tfrac{m^2 d}{n^2}; \tfrac{s}{2} + \tfrac{1}{4}\right)$$

$$= \sum_{Q \in \Gamma\backslash\mathcal{Q}_D} \chi(Q) \begin{cases} 2\sqrt{\pi}\omega_Q^{-1} F_m(z_Q, s) & \text{if } d'd < 0, \\ \int_{\mathcal{C}_Q} F_m(z,s) y^{-1} |dz| & \text{if } d', d > 0, \\ \int_{\mathcal{C}_Q} i\, \partial_z F_m(z,s) dz & \text{if } d', d < 0. \end{cases}$$

We assume that d is a fundamental discriminant and that $D = d'd$. We need an associated exponential sum, defined for $c \equiv 0 \pmod 4$ by

$$S_m(d', d; c) = \sum_{\substack{b \pmod c \\ b^2 \equiv D \pmod c}} \chi\left(\left[\tfrac{c}{4}, b, \tfrac{b^2 - D}{c}\right]\right) e\left(\tfrac{2mb}{c}\right). \tag{9.1}$$

Clearly

$$S_{-m}(d', d; c) = \overline{S_m(d', d; c)} = S_m(d', d; c).$$

LEMMA 3. *For $m \neq 0$, $d'd < 0$ and $\mathrm{Re}(s) > 1$, we have*

$$\sum_{Q \in \Gamma\backslash\mathcal{Q}_D} \chi(Q)\omega_Q^{-1} F_m(z_Q, s) = 2^{-1/2}|D|^{1/4} \sum_{0 < c \equiv 0(4)} \frac{S_m(d', d; c)}{\sqrt{c}} I_{s-1/2}\left(4\pi|m|\tfrac{\sqrt{D}}{c}\right).$$

Proof. This follows directly from [16, Prop. 4 p. 970]. □

Similarly we have for the second case the following.

[8] Actually, we use $g_Q = \gamma_Q^{-1}$ in [16].

LEMMA 4. *For $m \neq 0$, $d', d > 0$ with $d'd$ not a square and $\mathrm{Re}(s) > 1$, we have*

$$\sum_{Q \in \Gamma \backslash \mathcal{Q}_D} \chi(Q) \int_{\mathcal{C}_Q} F_m(z,s) y^{-1} |dz| = 2^{s-1/2} \frac{\Gamma(\frac{s}{2})^2}{\Gamma(s)} D^{1/4} \sum_{0<c \equiv 0(4)} \frac{S_m(d',d;c)}{\sqrt{c}} J_{s-1/2}(4\pi |m| \tfrac{\sqrt{D}}{c}).$$

Proof. Again this follows directly from [16, Prop.4]. Note that $\frac{\sqrt{D}}{Q(z,1)} dz = y^{-1}|dz|$ on $\mathcal{C}_Q$. □

The third case requires some new computations.

LEMMA 5. *For $m \neq 0$, $d', d < 0$ with $d'd$ not a square and $\mathrm{Re}(s) > 1$, we have*

$$\sum_{Q \in \Gamma \backslash \mathcal{Q}_D} \chi(Q) \int_{\mathcal{C}_Q} i\partial_z F_m(z,s) dz = 2^{s-1/2} \frac{\Gamma(\frac{s+1}{2})^2}{\Gamma(s)} D^{1/4} \sum_{0<c \equiv 0(4)} \frac{S_m(d',d;c)}{\sqrt{c}} J_{s-1/2}(4\pi |m| \tfrac{\sqrt{D}}{c}).$$

Proof. Now (8.1) and a calculation using differentiation formulas for the Whittaker functions in [38, p.302] gives for that for $\mathrm{Re}(s) > 1$,

$$-2i\partial_z F_m(z,s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} f_{2,m}(\gamma z, s) \frac{d(\gamma z)}{dz},$$

where

$$f_{2,m}(z) = -s|m|^{-1/2} (2\pi y)^{-1} \tfrac{\Gamma(s)}{\Gamma(2s)} M_{\mathrm{sgn}(m), s-1/2}(4\pi |m| y) e(mx). \tag{9.2}$$

The proof proceeds in a very similar way as for Lemma 4 except that we need the analogs of two technical lemmas in [16].

For a smooth function $\phi(y)$ that satisfies $\phi(y) \ll y^{1+\varepsilon}$ as $y \to 0$, let $f(\tau) = e(m \,\mathrm{Re}\, \tau) \phi(\mathrm{Im}\, \tau)$ and

$$P_m(\tau, \phi) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} f(\gamma \tau) \frac{d(\gamma z)}{dz}.$$

Then the following is the weight 2 analog of Lemma 7 in [16].

LEMMA 6. *Suppose that $d', d < 0$ and that $dd' = D$ is not a square. Then for all $m \in \mathbb{Z}$,*

$$\sum_{Q \in \Gamma \backslash \mathcal{Q}_D} \chi(Q) \int_{C_Q} P_m(\tau, \phi) d\tau = \sum_{0<c \equiv 0(4)} S_m(d, d'; c) \Phi_m\big(\tfrac{2\sqrt{D}}{c}\big),$$

where

$$\Phi_m(t) = it \int_0^\pi e(mt\cos\theta)\phi(t\sin\theta)e^{i\theta}d\theta. \tag{9.3}$$

Taking $\phi(t) = -s|m|^{-1/2}(2\pi y)^{-1}\frac{\Gamma(s)}{\Gamma(2s)}M_{\mathrm{sgn}(m),s-1/2}(4\pi|m|y)$ in (9.3) reduces the proof of Lemma 5 to the the following lemma about special functions.

LEMMA 7. *For $\mu \in \mathbb{C}$, $t > 0$ and $\mathrm{Re}(s) > 0$,*

$$\int_0^\pi e^{\pm i(t\cos\theta+\mu\theta)}M_{\mu,s-1/2}(2t\sin\theta)\frac{d\theta}{\sin\theta} = G(s,\mu)t^{1/2}J_{s-1/2}(t), \tag{9.4}$$

where

$$G(s,\mu) = e(\pm\mu/4)(2\pi)^{3/2}\frac{2^{-s}\Gamma(2s)}{\Gamma(\frac{s+1+\mu}{2})\Gamma(\frac{s+1-\mu}{2})}.$$

Proof. See the appendix. □

The following identity, which allows us to relate the cycle integrals to the spectral coefficients, is proved by a slight modification of the proof given by Kohnen in [35, Prop. 5, p. 259] (see also [12], [32] and [53]).

LEMMA 8. *For positive $c \equiv 0 \pmod 4$, $d, m \in \mathbb{Z}$ with $d' \equiv 0, 1 \pmod 4$ and d a fundamental discriminant, we have*

$$S_m(d', d; c) = \sum_{n|\left(m,\frac{c}{4}\right)} \left(\tfrac{d}{n}\right)\sqrt{\tfrac{n}{c}}\, K^+\left(d', \tfrac{m^2 d}{n^2}; \tfrac{c}{n}\right).$$

Proceeding as in [16], Proposition 5 follows from Lemmas 4,5 and 8. □

Remarks. For the purpose of proving the extended Katok–Sarnak formula by the method of spectral residues we actually have many choices of Poincaré series to use since we can add a holomorphic form without changing the residues. Thus we could employ the Poincaré series originally used by Selberg [47] (see also [22]). This might make some of the calculations somewhat simpler, but that would not give an exact formula as we obtain in Proposition 5. One advantage of an exact formula is that we can also use it to show that cycle integrals of modular functions give weight 1/2 weak Maass forms. This was done in [16] for the first two cases of Proposition 5. The last case can also be applied in this way. It is also possible to prove Theorem 3 by these methods.

10. **Proof of Theorem 4**

Recall the plus space V_r^+ of Maass cusp forms of weight 1/2 defined in Section 8 above. It is shown in [34] that V_r^+ has an orthonormal basis $B_r = \{\psi\}$ consisting of eigenfunction of all Hecke operators T_{p^2} where $p > 2$ is prime.

Fix such a basis B_r. Given $\psi \in B_r$ with Fourier expansion

$$\psi(z) = \sum_{n\neq 0} b(n) W_{\frac{1}{4}\operatorname{sgn} n, \frac{ir}{2}}(4\pi|n|y)e(nx) \tag{10.1}$$

and a fundamental discriminant d with $b(d) \neq 0$, the Hecke relation $T_{p^2}\psi = a_\psi(p)\psi$ implies that

$$L_d(s+\tfrac{1}{2}) \sum_{n\geq 1} b(dn^2) n^{-s+1} = b(d) \prod_p (1 - a_\psi(p)p^{-s} + p^{-2s})^{-1}.$$

Define the numbers $a_\psi(n)$ via

$$\prod_p (1 - a_\psi(p)p^{-s} + p^{-2s})^{-1} = \sum_{n\geq 1} a_\psi(n) n^{-s} \tag{10.2}$$

and let

$$\operatorname{Shim}\psi(z) = y^{1/2} \sum_{n\neq 0} 2a_\psi(|n|) K_{ir}(2\pi|n|y)e(nx). \tag{10.3}$$

Note that for some d, we must have that $b(d) \neq 0$ so that this is always defined.

It is convenient to define

$$T(\varphi, \chi) = \frac{1}{\langle \varphi, \varphi\rangle} \sum_{Q\in \Gamma\backslash \mathcal{Q}_D} \chi(Q) \begin{cases} 2\sqrt{\pi}\omega_Q^{-1}\varphi(z_Q) & \text{if } d'd < 0, \\ \int_{\mathcal{C}_Q} \varphi(z) y^{-1}|dz| & \text{if } d', d > 0, \\ \int_{\mathcal{C}_Q} i\,\partial_z\varphi(z) dz & \text{if } d', d < 0. \end{cases}$$

Theorem 4 follows easily from the next proposition together with Lemma 1.

Proposition 6. *For any even Hecke-Maass cusp form φ for Γ with Laplace eigenvalue $\frac{1}{2} + r^2$, there is a unique $\psi \in B_r$ with Fourier expansion given in* (10.1) *so that $\varphi = \operatorname{Shim}\psi$ and such that for d a fundamental discriminant and d' any discriminant such that $D = d'd$ is not a square, we have*

$$T(\varphi, \chi) = 12\pi^{\frac{1}{2}} D^{\frac{3}{4}} b(d')\bar{b}(d),$$

where χ is the genus character associated to the factorization $D = d'd$.

Proof. Let $m > 0$, and suppose that $D = d'd > 1$ where d is fundamental. First we will show that

$$12\pi^{\frac{1}{2}} D^{\frac{3}{4}} \sum_{\psi\in B_r} b(d')\bar{b}(d) a_\psi(m) = \sum_\varphi a(m) T(\varphi, \chi), \tag{10.4}$$

where φ is summed over all Hecke–Maass cusp forms with Laplace eigenvalue $\frac{1}{2} + r^2$. We have from Propositions 3 and 5 that for every $m \neq 0$,

$$\begin{aligned} 6\pi^{\frac{1}{2}} D^{\frac{3}{4}} |m| \sum_{\substack{n|m \\ n>0}} n^{-\frac{3}{2}} \left(\tfrac{d}{n}\right) \operatorname{Res}_{s=\frac{1}{2}+ir} (2s-1)\Phi^+\left(d', \tfrac{m^2 d}{n^2}; \tfrac{s}{2} + \tfrac{1}{4}\right) \\ = \sum_\varphi 2a(m) T(\varphi, \chi). \end{aligned} \tag{10.5}$$

Observe that

$$\begin{aligned}\mathrm{Res}_{s=\frac{1}{2}+ir}(2s-1)\Phi^+\left(d',d;\tfrac{s}{2}+\tfrac{1}{4}\right) &= 2ir\lim_{s\to\frac{1}{2}+ir}\left(s-(\tfrac{1}{2}+ir)\right)\Phi^+\left(d',d;\tfrac{s}{2}+\tfrac{1}{4}\right)\\ \text{which, setting } s=2w-\tfrac{1}{2}, \quad &= 4ir\lim_{w\to\frac{1}{2}+\frac{ir}{2}}\left(w-(\tfrac{1}{2}+\tfrac{ir}{2})\right)\Phi^+\left(d',d;w\right)\\ &= 4\mathrm{Res}_{s=\frac{1}{2}+\frac{ir}{2}}(2s-1)\Phi^+\left(d',d;s\right).\end{aligned}$$

Therefore, Proposition 4 gives

$$\mathrm{Res}_{s=\frac{1}{2}+ir}(2s-1)\Phi^+\left(d',\tfrac{m^2d}{n^2};\tfrac{s}{2}+\tfrac{1}{4}\right)=4\sum_{\psi}b(d')\bar{b}(\tfrac{m^2d}{n^2}),$$

where the sum is over an orthonormal basis of cusp forms $\{\psi\}$ for V_r^+ and $b(d)$ is the Fourier coefficient of ψ as in (8.8). By (10.5) we get

$$24\pi^{\frac{1}{2}}D^{\frac{3}{4}}m\sum_{\psi\in B_r}b(d')\sum_{\substack{n|m\\ n>0}}n^{-\frac{3}{2}}\left(\tfrac{d}{n}\right)\bar{b}\left(\tfrac{m^2d}{n^2}\right)=\sum_{\varphi}2a(m)T(\varphi,\chi),$$

and we obtain (10.4) by using the Hecke relation

$$m\sum_{\substack{n|m\\ n>0}}n^{-\frac{3}{2}}\left(\tfrac{d}{n}\right)b\left(\tfrac{m^2d}{n^2}\right)=a_\psi(m)b(d).$$

It follows from (10.4) and (10.3) that

$$12\pi^{1/2}D^{3/4}\sum_{\psi}b(d')\bar{b}(d)\mathrm{Shim}(\psi)=\sum_{\varphi}T(\varphi,\chi)\varphi. \tag{10.6}$$

This identity is valid for all discriminants d,d' where d is fundamental, and dd' is not a square. As in the proof of Theorem 1 on page 129 of Biró in [3], one can conclude that $\mathrm{Shim}(\psi)$ is a weight 0 Maass form with eigenvalue $\frac{1}{2}+r^2$ and by (10.2) it is some φ. This leads to

$$12\pi^{1/2}D^{3/4}\sum_{\varphi}\sum_{\mathrm{Shim}(\psi)=\varphi}b(d')\bar{b}(d)\varphi=\sum_{\varphi}T(\varphi,\chi)\varphi.$$

The linear independence of the Maass forms φ now gives the following version of the proposition:

$$12\pi^{1/2}D^{3/4}\sum_{\mathrm{Shim}(\psi)=\varphi}b(d')\bar{b}(d)=T(\varphi,\chi).$$

Finally, it is known (see [1, Thm. 1.2]) that $\psi\mapsto\varphi=\mathrm{Shim}(\psi)$ gives a bijection between B_r and the even Hecke-Maass cusp forms φ with Laplace eigenvalue $\frac{1}{2}+r^2$, thus finishing the proof of Proposition 6. □

Remarks. Some of our arguments in the proof of Theorem 4 are quite similar in spirit to those of Biró in [3], who applies the Kuznetsov formula to prove a generalization of the Katok-Sarnak formula to general levels, but still for only positive discriminants d.

The method employed by Katok-Sarnak to prove their formula is based on a theta correspondence. This idea, which is a refinement of that introduced by Maass [37], was first used by Siegel to study indefinite quadratic forms. It would be interesting to apply this method to give our extension.

Appendix A. An integral

In this appendix we give a proof of the following integral formula, which was given in Lemma 7. For $\mu \in \mathbb{C}$, $t > 0$ and $\mathrm{Re}(s) > 0$,

$$\int_0^\pi e^{\pm i(t\cos\theta+\mu\theta)} M_{\mu,s-1/2}(2t\sin\theta)\frac{d\theta}{\sin\theta} = G(s,\mu)t^{1/2}J_{s-1/2}(t), \tag{A.1}$$

where

$$G(s,\mu) = e(\pm\mu/4)(2\pi)^{3/2}\frac{2^{-s}\Gamma(2s)}{\Gamma(\frac{s+1+\mu}{2})\Gamma(\frac{s+1-\mu}{2})}.$$

Proof. To prove the lemma we will restrict to the case when the signs in (A.1) are positive since the formula with negative signs follows by complex conjugation. To prove the lemma we will prove that both sides of (A.1) satisfy the same order differential equation and that the Taylor series coefficients of both sides agree up to order 2.

Let $\lambda = s - 1/2$ and $g(t) = t^{1/2}J_\lambda(t)$. A simple computation shows that

$$t^{3/2}\left[g''(t) + (1 + (1/4 - \lambda^2)/t^2)g(t)\right] = t^2J''_\lambda(t) + tJ'_\lambda(t) + (t^2 - \lambda^2)J_\lambda(t) = 0.$$

Hence we want to show that the left-hand side of (A.1) also satisfies

$$f''(t) + (1 + (1/4 - \lambda^2)/t^2)f(t) = 0.$$

Factoring out the t-dependent part we need to compute

$$h''(t) + \left(1 + \frac{1/4 - \lambda^2}{t^2}\right)h(t)$$

for $h(t) = e^{i(t\cos\theta)}M_{\mu,\lambda}(2t\sin\theta)$.

The fact that the Whittaker function $M_{\mu,\lambda}$ satisfies the differential equation

$$M''_{\mu,\lambda}(2t\sin\theta) = \left(\frac{1}{4} - \frac{\mu}{2t\sin\theta} - \frac{1/4 - \lambda^2}{4t^2\sin^2\theta}\right)M_{\mu,\lambda}(2t\sin\theta)$$

and

$$\begin{aligned} h''(t) = {} & [-\cos^2\theta M_{\mu,\lambda}(2t\sin\theta) + 4\sin^2\theta M''_{\mu,\lambda}(2t\sin\theta)]e^{i(t\cos\theta)} \\ & + 4i\cos\theta\sin\theta M'_{\mu,\lambda}(2t\sin\theta)e^{i(t\cos\theta)} \end{aligned}$$

leads to

$$h''(t)+\left(1-\tfrac{\lambda^2-1/4}{t^2}\right)h(t) \\ =\left(2\sin^2\theta-\tfrac{2\mu\sin\theta}{t}\right)h(t)+2i\sin 2\theta e^{\pm i(t\cos\theta)}M'_{\mu,\lambda}(2t\sin\theta).$$

Using this last equation gives for the integral in (A.1)

$$\left(\frac{d^2}{dt^2}+\left(1+\frac{1/4-\lambda^2}{t^2}\right)\right)\int_0^\pi e^{i(t\cos\theta+\mu\theta)}M_{\mu,\lambda}(2t\sin\theta)\frac{d\theta}{\sin\theta} \\ =\int_0^\pi\left(2\sin\theta-\frac{2\mu}{t}\right)h(t)e^{i\mu\theta}d\theta+2i\int_0^\pi 2\cos\theta e^{i(t\cos\theta)+i\mu\theta}M'_{\mu,\lambda}(2t\sin\theta)d\theta.$$

Now we use $\frac{d}{d\theta}M_{\mu,\lambda}(2t\sin\theta)=M'_{\mu,\lambda}(2t\sin\theta)2t\cos\theta$ and integration by parts to get

$$2i\int_0^\pi e^{i(t\cos\theta+\mu\theta)}M'_{\mu,\lambda}(2t\sin\theta)2\cos\theta d\theta \\ =-\frac{2i}{t}\int_0^\pi\frac{d}{d\theta}\left(e^{i(t\cos\theta+\mu\theta)}\right)M_{\mu,\lambda}(2t\sin\theta)d\theta$$

as $\left[e^{i(t\cos\theta+\mu\theta)}M_{\mu,\lambda}(2t\sin\theta)\right]_0^\pi=0$. Finally, since

$$-\frac{2i}{t}\int_0^\pi\frac{d}{d\theta}\left(e^{i(t\cos\theta+\mu\theta)}\right)M_{\mu,\lambda}(2t\sin\theta)d\theta=\int_0^\pi\left(\frac{2\mu}{t}-2\sin\theta\right)h(t)e^{i\mu\theta}d\theta,$$

we have

$$\left(\frac{d^2}{dt^2}+\left(1+\frac{1/4-\lambda^2}{t^2}\right)\right)\int_0^\pi e^{i(t\cos\theta+\mu\theta)}M_{\mu,\lambda}(2t\sin\theta)\frac{d\theta}{\sin\theta}=0.$$

This proves that both sides of (A.1) satisfy the same differential equation.

To prove the lemma we still need to check the Taylor coefficients. To this end we use the Taylor expansions of the exponential function and of the Whittaker function, namely,

$$M_{\mu,s-1/2}(x)=e^{-x/2}x^s\sum_{n=0}^\infty\frac{(s-\mu)_n}{(2s)_n}\frac{x^n}{n!}.$$

Then

$$\int_0^\pi e^{i(t\cos\theta+\mu\theta)}M_{\mu,s-1/2}(2t\sin\theta)\frac{d\theta}{\sin\theta} \\ =\sum_{n=0}^\infty\sum_{m=0}^\infty\frac{(s-\mu)_n}{(2s)_n n!}(2t)^{n+s}\frac{(it)^m}{m!}\int_0^\pi e^{i(m+\mu)\theta}(\sin\theta)^{n+s-1}d\theta.$$

Using the integral formula (see [23, p. 511, 3.892(1)]),

$$\int_0^\pi e^{i\beta x}\sin^{\nu-1}x dx=\frac{\pi e^{i\pi\beta/2}\Gamma(\nu)}{\Gamma(\frac{\nu+\beta+1}{2})\Gamma(\frac{\nu-\beta+1}{2})},$$

and $(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$ gives

$$\begin{aligned}(A.2)\qquad &\int_0^\pi e^{i(t\cos\theta+\mu\theta)} M_{\mu,s-1/2}(2t\sin\theta)\frac{d\theta}{\sin\theta}\\ &\quad = (2\pi)e(\mu/4)\frac{\Gamma(2s)}{\Gamma(s-\mu)}\\ &\qquad \times\sum_{\ell=0}^{\infty}\sum_{m+n=\ell}\frac{(-1)^m\Gamma(s-\mu+n)\Gamma(s+n)}{m!n!\Gamma(2s+n)\Gamma(\frac{n+s+m+\mu+1}{2})\Gamma(\frac{n+s-m-\mu+1}{2})}t^{s+\ell}.\end{aligned}$$

On the other hand, using the Taylor expansion

$$t^{1/2}J_{s-1/2}(t) = \sum_{r=0}^{\infty}\frac{(-1)^r(t/2)^{s+2r}}{r!\Gamma(s+1/2+r)}.$$

gives for the right-hand side of (A.1)

(A.3)

$$G(s,\mu)t^{1/2}J_{s-1/2}(t) = \frac{(\pi)^{3/2}e(\mu/4)2^{2-2s}\Gamma(2s)}{\Gamma(\frac{s+1+\mu}{2})\Gamma(\frac{s+1-\mu}{2})}\sum_{r=0}^{\infty}\frac{(-1)^r2^{-2r}}{r!\Gamma(s+1/2+r)}t^{s+2r}$$

A straightforward calculation shows that the coefficients of t^s, t^{s+1} and t^{s+2} in (A.2) and (A.3) match, which is more than what is needed to finish the proof of the lemma. □

References

[1] E. M. Baruch and Z. Mao, A generalized Kohnen-Zagier formula for Maass forms, *J. Lond. Math. Soc.* **82** (2010), 1–16. MR 2669637. Zbl 1217.11047. http://dx.doi.org/10.1112/jlms/jdq009.

[2] A. F. Beardon, *The Geometry of Discrete Groups*, *Grad. Texts in Math.* **91**, Springer-Verlag, New York, 1983. MR 0698777. Zbl 0528.30001. http://dx.doi.org/10.1007/978-1-4612-1146-4.

[3] A. Biró, Cycle integrals of Maass forms of weight 0 and Fourier coefficients of Maass forms of weight 1/2, *Acta Arith.* **94** (2000), 103–152. MR 1779112. Zbl 0969.11017. Available at http://matwbn.icm.edu.pl/ksiazki/aa/aa94/aa9421.pdf.

[4] V. Blomer and G. Harcos, Hybrid bounds for twisted L-functions, *J. Reine Angew. Math.* **621** (2008), 53–79. MR 2431250. Zbl 1193.11044. http://dx.doi.org/10.1515/CRELLE.2008.058.

[5] A. R. Booker, A. Strömbergsson, and A. Venkatesh, Effective computation of Maass cusp forms, *Int. Math. Res. Not.* **2006** (2006), Art. ID 71281, 34. MR 2249995. Zbl 1154.11018. http://dx.doi.org/10.1155/IMRN/2006/71281.

[6] D. A. Burgess, On character sums and L-series, *Proc. London Math. Soc.* **12** (1962), 193–206. MR 0132733. Zbl 0106.04004. http://dx.doi.org/10.1112/plms/s3-12.1.193.

[7] F. L. Chiera, On Petersson products of not necessarily cuspidal modular forms, *J. Number Theory* **122** (2007), 13–24. MR 2287108. Zbl 1118.11021. http://dx.doi.org/10.1016/j.jnt.2006.03.003.

[8] H. Cohen, Sums involving the values at negative integers of L functions of quadratic characters, in *Séminaire de Théorie des Nombres*, 1974–1975 (Univ. Bordeaux I, Talence, Exp. No. 3), Centre Nat. Recherche Sci., Talence, 1975, 21 pp. Lab. Théorie des Nombres. MR 0389784. Zbl 0316.10015.

[9] H. Cohen and H. W. Lenstra, Jr., Heuristics on class groups, in *Number Theory* (New York, 1982), *Lecture Notes in Math.* **1052**, Springer-Verlag, New York, 1984, pp. 26–36. MR 0750661. Zbl 0532.12008. http://dx.doi.org/10.1007/BFb0071539.

[10] J. B. Conrey and H. Iwaniec, The cubic moment of central values of automorphic L-functions, *Ann. of Math.* **151** (2000), 1175–1216. MR 1779567. Zbl 0973.11056. http://dx.doi.org/10.2307/121132.

[11] H. Davenport, *Multiplicative Number Theory*, third ed., *Grad. Texts in Math.* **74**, Springer-Verlag, New York, 2000, revised and with a preface by Hugh L. Montgomery. MR 1790423. Zbl 1002.11001.

[12] W. Duke, Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.* **92** (1988), 73–90. MR 0931205. Zbl 0628.10029. http://dx.doi.org/10.1007/BF01393993.

[13] W. Duke, J. Friedlander, and H. Iwaniec, Bounds for automorphic L-functions, *Invent. Math.* **112** (1993), 1–8. MR 1207474. Zbl 0765.11038. http://dx.doi.org/10.1007/BF01232422.

[14] W. Duke, J. B. Friedlander, and H. Iwaniec, Weyl sums for quadratic roots, *Int. Math. Res. Not.* **2012** (2012), 2493–2549. MR 2926988. Zbl 1300.11086. http://dx.doi.org/10.1093/imrn/rnr112.

[15] W. Duke and Ö. Imamoḡlu, A converse theorem and the Saito-Kurokawa lift, *Internat. Math. Res. Notices* (1996), 347–355. MR 1389957. Zbl 0849.11039. http://dx.doi.org/10.1155/S1073792896000220.

[16] W. Duke, Ö. Imamoḡlu, and Á. Tóth, Cycle integrals of the j-function and mock modular forms, *Ann. of Math.* **173** (2011), 947–981. MR 2776366. Zbl 1270.11044. http://dx.doi.org/10.4007/annals.2011.173.2.8.

[17] M. Eichler, Grenzkreisgruppen und kettenbruchartige Algorithmen, *Acta Arith.* **11** (1965), 169–180. MR 0228436. Zbl 0148.32503. Available at http://eudml.org/doc/204752.

[18] M. Einsiedler, E. Lindenstrauss, P. Michel, and A. Venkatesh, The distribution of closed geodesics on the modular surface, and Duke's theorem, *Enseign. Math.* **58** (2012), 249–313. MR 3058601. Zbl 1312.37032. http://dx.doi.org/10.4171/LEM/58-3-2.

[19] B. Farb and D. Margalit, *A Primer on Mapping Class Groups*, *Princeton Math. Ser.* **49**, Princeton Univ. Press, Princeton, NJ, 2012. MR 2850125. Zbl 1245.57002.

[20] J. D. Fay, Fourier coefficients of the resolvent for a Fuchsian group, *J. Reine Angew. Math.* **293/294** (1977), 143–203. MR 0506038. Zbl 0352.30012. http://dx.doi.org/10.1515/crll.1977.293-294.143.

[21] D. Goldfeld and J. Hoffstein, Eisenstein series of $\frac{1}{2}$-integral weight and the mean value of real Dirichlet L-series, *Invent. Math.* **80** (1985), 185–208. MR 0788407. Zbl 0564.10043. http://dx.doi.org/10.1007/BF01388603.

[22] D. Goldfeld and P. Sarnak, Sums of Kloosterman sums, *Invent. Math.* **71** (1983), 243–250. MR 0689644. Zbl 0507.10029. http://dx.doi.org/10.1007/BF01389098.

[23] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, eighth ed., Elsevier/Academic Press, Amsterdam, 2015, translated from the Russian, translation edited and with a preface by Daniel Zwillinger and Victor Moll, revised from the seventh edition [MR 2360010]. MR 3307944. Zbl 1300.65001.

[24] W. F. Hammond, The modular groups of Hilbert and Siegel, *Amer. J. Math.* **88** (1966), 497–516. MR 0201397. Zbl 0144.34103. http://dx.doi.org/10.2307/2373204.

[25] W. F. Hammond and F. Hirzebruch, L-series, modular imbeddings, and signatures, *Math. Ann.* **204** (1973), 263–270. MR 0374038. Zbl 0245.10017. http://dx.doi.org/10.1007/BF01354577.

[26] E. Hecke, *Lectures on Dirichlet Series, Modular Functions and Quadratic Forms*, Vandenhoeck & Ruprecht, Göttingen, 1983, edited by Bruno Schoeneberg, with the collaboration of Wilhelm Maak. MR 0693092. Zbl 0507.10015.

[27] D. A. Hejhal, *The Selberg Trace Formula for* $\mathrm{PSL}(2,\mathbf{R})$. *Vol.* 2, *Lecture Notes in Math.* **1001**, Springer-Verlag, New York, 1983. MR 0711197. Zbl 0543.10020. http://dx.doi.org/10.1007/BFb0061302.

[28] F. E. P. Hirzebruch, Hilbert modular surfaces, *Enseignement Math.* **19** (1973), 183–281. MR 0393045. Zbl 0285.14007.

[29] H. Iwaniec, Fourier coefficients of modular forms of half-integral weight, *Invent. Math.* **87** (1987), 385–401. MR 0870736. Zbl 0606.10017. http://dx.doi.org/10.1007/BF01389423.

[30] H. Iwaniec, *Spectral Methods of Automorphic Forms*, second ed., *Grad. Stud. Math.* **53**, Amer. Math. Soc., Providence, RI, 2002. MR 1942691. Zbl 1006.11024. http://dx.doi.org/10.1090/gsm/053.

[31] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, *Amer. Math. Soc. Colloq. Publ.* **53**, Amer. Math. Soc., Providence, RI, 2004. MR 2061214. Zbl 1059.11001. http://dx.doi.org/10.1090/coll/053.

[32] P. Jenkins, Kloosterman sums and traces of singular moduli, *J. Number Theory* **117** (2006), 301–314. MR 2213767. Zbl 1177.11038. http://dx.doi.org/10.1016/j.jnt.2005.06.009.

[33] S. Katok, Coding of closed geodesics after Gauss and Morse, *Geom. Dedicata* **63** (1996), 123–145. MR 1413625. Zbl 0884.20030. http://dx.doi.org/10.1007/BF00148213.

[34] S. Katok and P. Sarnak, Heegner points, cycles and Maass forms, *Israel J. Math.* **84** (1993), 193–227. MR 1244668. Zbl 0787.11016. http://dx.doi.org/10.1007/BF02761700.

[35] W. Kohnen, Fourier coefficients of modular forms of half-integral weight, *Math. Ann.* **271** (1985), 237–268. MR 0783554. Zbl 0542.10018. http://dx.doi.org/10.1007/BF01455989.

[36] W. Kohnen and D. Zagier, Values of L-series of modular forms at the center of the critical strip, *Invent. Math.* **64** (1981), 175–198. MR 0629468. Zbl 0468.10015. http://dx.doi.org/10.1007/BF01389166.

[37] H. Maass, Über die räumliche Verteilung der Punkte in Gittern mit indefiniter Metrik, *Math. Ann.* **138** (1959), 287–315. MR 0120224. Zbl 0089.06102. http://dx.doi.org/10.1007/BF01344150.

[38] W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics*, *Grundlehren Math. Wiss.* **52**, Springer-Verlag, New York, 1966. MR 0232968. Zbl 0143.08502.

[39] B. Maskit, On Poincaré's theorem for fundamental polygons, *Advances in Math.* **7** (1971), 219–230. MR 0297997. Zbl 0223.30008. http://dx.doi.org/10.1016/S0001-8708(71)80003-8.

[40] H. Neunhöffer, Über die analytische Fortsetzung von Poincaréreihen, *S.-B. Heidelberger Akad. Wiss. Math.-Natur. Kl.* (1973), 33–90. MR 0352007. Zbl 0272.10015.

[41] D. Niebur, A class of nonanalytic automorphic functions, *Nagoya Math. J.* **52** (1973), 133–145. MR 0337788. Zbl 0288.10010. Available at http://projecteuclid.org/euclid.nmj/1118794882.

[42] H. Poincaré, Théorie des groupes fuchsiens, *Acta Math.* **1** (1882), 1–76. MR 1554574. JFM 14.0338.01. http://dx.doi.org/10.1007/BF02391835.

[43] W. Roelcke, Das Eigenwertproblem der automorphen Formen in der hyperbolischen Ebene. I, II. (German), *Math. Ann.* **167** (1966), 292–337, ibid. **168** (1966), 261–324 [http://dx.doi.org/10.1007/BF01361556]. MR 0243062. Zbl 0152.07705. http://dx.doi.org/10.1007/BF01364540.

[44] P. Sarnak, Class numbers of indefinite binary quadratic forms, *J. Number Theory* **15** (1982), 229–247. MR 0675187. Zbl 0499.10021. http://dx.doi.org/10.1016/0022-314X(82)90028-2.

[45] P. Sarnak, *Some Applications of Modular Forms*, *Cambridge Tracts in Math.* **99**, Cambridge Univ. Press, Cambridge, 1990. MR 1102679. Zbl 0721.11015. http://dx.doi.org/10.1017/CBO9780511895593.

[46] P. Sarnak, Reciprocal geodesics, in *Analytic Number Theory*, *Clay Math. Proc.* **7**, Amer. Math. Soc., Providence, RI, 2007, pp. 217–237. MR 2362203. Zbl 1198.11039.

[47] A. Selberg, On the estimation of Fourier coefficients of modular forms, in *Proc. Sympos. Pure Math., Vol. VIII*, Amer. Math. Soc., Providence, R.I., 1965, pp. 1–15. MR 0182610. Zbl 0142.33903.

[48] G. Shimura, On modular forms of half integral weight, *Ann. of Math.* **97** (1973), 440–481. MR 0332663. Zbl 0266.10022. http://dx.doi.org/10.2307/1970831.

[49] T. Shintani, On construction of holomorphic cusp forms of half integral weight, *Nagoya Math. J.* **58** (1975), 83–126. MR 0389772. Zbl 0316.10016. Available at http://projecteuclid.org/euclid.nmj/1118795445.

[50] C. L. Siegel, Die Funktionalgleichungen einiger Dirichletscher Reihen, *Math. Z.* **63** (1956), 363–373. MR 0074533. Zbl 0070.07702. http://dx.doi.org/10.1007/BF01187948.

[51] C. L. Siegel, *Advanced Analytic Number Theory*, second ed., *Tata Instit. Fundamental Research Studies in Mathematics* no. 9, Tata Inst. Fund. Res., Bombay, 1980. MR 0659851. Zbl 0478.10001.

[52] F. Strömberg, Computation of Maass waveforms with nontrivial multiplier systems, *Math. Comp.* **77** (2008), 2375–2416. MR 2429890. Zbl 1260.11039. http://dx.doi.org/10.1090/S0025-5718-08-02129-7.

[53] Á. Tóth, On the evaluation of Salié sums, *Proc. Amer. Math. Soc.* **133** (2005), 643–645. MR 2113909. Zbl 1092.11036. http://dx.doi.org/10.1090/S0002-9939-04-07768-8.

[54] J.-L. Waldspurger, Sur les coefficients de Fourier des formes modulaires de poids demi-entier, *J. Math. Pures Appl.* **60** (1981), 375–484. MR 0646366. Zbl 0431.10015.

[55] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions, reprint of the fourth (1927) edition, *Cambridge Math. Lib.*, Cambridge Univ. Press, Cambridge, 1996. MR 1424469. Zbl 0951.30002. http://dx.doi.org/10.1017/CBO9780511608759.

[56] M. Young, Weyl-type hybrid subconvexity bounds for twisted l-functions and Heegner points on shrinking sets, to appear *J. Eur. Math. Soc.*

[57] D. Zagier, A Kronecker limit formula for real quadratic fields, *Math. Ann.* **213** (1975), 153–184. MR 0366877. Zbl 0283.12004. http://dx.doi.org/10.1007/BF01343950.

[58] D. Zagier, *Zetafunktionen und Quadratische Körper*, Springer-Verlag, New York, 1981, Eine Einführung in die höhere Zahlentheorie. [An introduction to higher number theory], Hochschultext. [University Text]. MR 0631688. Zbl 0459.10001.

(Received: October 19, 2015)
(Revised: July 15, 2016)

University of California, Los Angeles, CA
E-mail: wdduke@ucla.edu
http://www.math.ucla.edu/~wdduke/

Institute of Mathematics, ETH, Zürich, Switzerland
E-mail: ozlem@math.ethz.ch
https://www.math.ethz.ch/research/number-theory-geometry/ozlem-imamoglu.html

Eötvös Loránd University, Budapest, Hungary
E-mail: toth@cs.elte.hu
http://www.cs.elte.hu/u/toth?lang=en