CYCLE INTEGRALS OF MODULAR FORMS

DISSERTATION

submitted for the degree of
"Doctor of the Hungarian Academy of Sciences"

Árpád Tóth

Eötvös Loránd University
and
MTA Alfréd Rényi Institute of Mathematics

Budapest

2018

Contents

Introduction

The goal of this dissertation is to describe some new results on the integrals of modular forms along certain closed curves. We will refer to these integrals as cycle integrals of modular forms. When interpreted classically the curves of integration are closed geodesics on the so called modular surface, the quotient of the hyperbolic plane of Bolyai-Lobachevsky by a distinguished discrete group of isometries. In this language the simplest modular forms are eigenfunctions of the Laplace-Beltrami operator. However for certain applications it is better to view them as functions on $\mathrm{PSL}_2(\mathbb{R})$ the isometry group of the hyperbolic plane. In this version modular forms have a natural representation theoretic interpretation. The lift of closed geodesics then become periodic orbits of the geodesic flow, and the integral of modular forms along them give interesting information about these periodic orbits, such as their linking numbers.

For example when the discrete group in question is $\mathrm{SL}_2(\mathbb{Z})$, the right-coset space $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})$ as a 3-manifold is diffeomorphic to the complement of the trefoil knot in S^3, as realized as the link of the surface singularity of $z^2 = w^3$ at the origin. E. Ghys showed that the linking number of this trefoil knot with a periodic orbit of the geodesic flow (called a modular knot in this case) is given by the Rademacher symbol. This symbol is a close relative of the classical Dedekind symbol which arose historically in the computation of cycle integrals of the logarithmic derivative of Dedekind's eta function. In his 2006 ICM talk Ghys brought attention to the intriguing question of understanding linking numbers between modular knots either combinatorially or from an arithmetic point of view. These linking numbers have to be properly interpreted as $H_1(S^3 \setminus \mathrm{Trefoil}) = \mathbb{Z}$. A natural path to take is to consider the symmetrized link of a modular knot, these arise as the union of the periodic orbits corresponding to some $\gamma, \gamma^{-1} \in \mathrm{SL}_2(\mathbb{Z})$. One of the problems considered in this thesis concerns the construction of analogues of Dedekind's eta function whose cycle integrals produce the linking numbers between these symmetrized links. This is Chapter 5 of the present thesis, based on the paper [42].

Another interesting application of cycle integrals presented in this dissertation concerns mock modular forms. Their theory was outlined by Ramanujan in his last letter to Hardy, a few month before his untimely death at age 32. It was only recently in 2002 that Zwegers found an intrinsic description of the elusive idea of what Ramanujan's mock modular forms are. Surprisingly, cycle integrals of $\mathrm{PSL}_2(\mathbb{Z})$-invariant functions with respect to arc length give a natural construction of these objects [40]. This material is presented in Chapter 4.

There is yet another problem that arose recently concerning immersed surfaces on the modular surface whose boundary is one of the closed geodesics. It is natural to expect that the push-forward measure of the natural hyperbolic metric on these surfaces gets equidistributed as the discriminant (a natural invariant of the geodesics) approaches infinity. This problem too leads to cycle integrals, this time of $\mathrm{PSL}_2(\mathbb{Z})$-invariant differential forms, and require a generalization of formulas due to Maass, and Katok-Sarnak [43]. This generalization and the resulting equidistribution results are presented from [43] and they form Chapters 2 and 3 of the thesis.

Finally cycle integrals are closely related to sums of certain exponential sums named after Salie and this is what we consider in the last chapter. This has interesting applications to the equidistribution of the so-called angles of these sums, or equivalently to the distribution of the roots of quadratic congruences to prime moduli. A conceptual background for why one expects such equidistribution is given for example in [112]. The major result of this chapter is from [123], which establishes these equidistribution results.

We will now put the results presented here in the framework of recent advances in automorphic forms. This thesis deals with generalizations of work of Katok, Sarnak, Borcherds, Zagier, Ghys, and many others and leads to some surprising new applications. It is without doubt that for the general public the most exciting new developments in modular form theory are related to Langland's program on the relation between automorphic forms and Galois representations. Instances such as Lafforgue's proof of the Langlands' conjectures for the general linear group $\mathrm{GL}(n, K)$ for function fields or Ngo's proof of the fundamental lemma for general reductive groups received Fields medals. (Lafforgue's work continued earlier research of Drinfeld, another Fields medalist, who treated the case $\mathrm{GL}(2, K)$.) Widely known by the general public is Wiles' work on the modularity of Galois representations associated to elliptic curves that allowed him to prove Fermat's last theorem. Borcherds' achievements for which he too was awarded a Fields medal were more connected to classical modular forms. Another major development is Lindenstrauss proof of the quantum unique ergodicity (QUE) conjecture of Rudnick and Sarnak for arithmetic surfaces for which he received the Fields medal in 2010. Also highly praised is this year's Fields medalist Venkatesh' work that very successfully brought in homogeneous dynamics into the subject. In both Lindenstrauss' and Venkatesh' work the role of number theory, while somewhat hidden, is significant. The research presented in this thesis is in different directions but in the same vain as theirs. Analysis and geometry in the classical sense play a more accentuated role than number theory but arithmetic considerations are crucial in most of the results presented below.

The organization of the dissertation is as follows. Chapter 1 gives a very short introduction to modular forms to set up notation. In Chapter 2 we develop a formalism to deal with cycle integrals of Poincaré series, this formalism will be used on several occasions. The first of these applications is this same chapter's main result about the extension of the Katok-Sarnak formulas taken from [43]. Chapter 3 is about the resulting two-dimensional equidistribution problem from [43]. Chapter 4 gives a brief description of mock modular forms and describes how these can be constructed from cycle integrals [40]. It also includes a construction of certain modular integrals used in connection with linking numbers in the chapter that follows. In that chapter, Chapter 5, we review Ghys' work on the geodesic flow and derive our results from [42] on linking numbers between symmetrized modular knots. Finally we prove the equidistribution of the angles of Salie sums in Chapter 6 [123].

Chapter 1

Background on modular forms

1.1 Hyperbolic geometry in the upper half plane model

One of the major discoveries of János Bolyai [56] was the fact that hyperbolic space contains a surface (the so called parasphere or horopsphere) on which the natural geometry satisfies the axioms of Euclid. Beltrami observed that the projection of a hyperbolic plane to this surface gives models of hyperbolic geometry on an open disc in the Euclidean plane. One of these models is conformal, and can be taken to the upper half plane via Cayley's transformation. This model was rediscovered and popularized by Poincaré while working on Fuchsian funtions, and is usually named the Poincaré upper half plane [121].

Let $\mathcal{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$. The group

$$GL_2^+(\mathbb{R}) = \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : a, b, c, d \in \mathbb{R}, ad - bc > 0\}$$

acts on $\mathbb{C}$ via Möbius transformations, for $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in GL_2^+(\mathbb{R})$

$$gz = \frac{az + b}{cz + d}.$$

If $z \in \mathcal{H}$ then $\operatorname{Im} gz = \det g \frac{\operatorname{Im} z}{|cz+d|^2}$ and so gz is also in H. One checks easily that this is a left action of $GL_2^+(\mathbb{R})$ on $\mathcal{H}$, $(g_1 g_2)(z) = g_1(g_2(z))$. By an easy application of Swartz's lemma one sees that all holomorphic automorphisms of $\mathcal{H}$ are given by such Möbius transformations. These automorphisms as a group are easily seen to be isomorphic to $\mathrm{PSL}_2(\mathbb{R})$.

The metric

$$ds^2 = \frac{1}{y^2}(dx^2 + dy^2)$$

is invariant under the action of $\mathrm{PSL}_2(\mathbb{R})$ (and has constant curvature -1). It follows that the conformal isomorphisms are the orientation preserving isometries.

One verifies easily that vertical lines are geodesics, and then so are their $\mathrm{PSL}_2(\mathbb{R})$-translates, which lead to semi-circles whose center lies on the real line. These are then the hyperbolic lines of the model.

Define the cross ratio of $z_1, z_2, z_3, z_4 \in \mathbb{C}$ by

$$[z_1, z_2, z_3, z_4] = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_2)(z_3 - z_4)}. \tag{1.1.1}$$

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

$$d(z, z^*) = \log |[w, z, z^*, w^*]|, \tag{1.1.2}$$

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. [9]).

If $d(z, w)$ is the distance, then in terms of Cartesian coordinates it is easier to work with

$$\cosh d(z, w) = 1 + 2u(z, w)$$

where

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

The measure associated to the metric is also invariant and is given by

$$\mu = \frac{dxdy}{y^2}$$

The hyperbolic Laplace operator is defined on smooth functions as the operator

$$\Delta = (\Delta_{\mathcal{H}}) = y^2(\partial_x^2 + \partial_y^2) = (y^2 \Delta_{\mathbb{R}^2}).$$

One can show that $-\Delta$ is a non-negative operator on various L^2-spaces (see below), and so the normalization of the eigenvalues is as follows. If

$$\Delta f + \lambda f = 0$$

then we call λ an eigenvalue. With this understanding the eigenvalues that arise for us are non-negative and we can write them using additional parameters:

$$\lambda = \frac{1}{4} + r^2 = s(1 - s).$$

Here $r \in \mathbb{R}$ or $ir \in [-\frac{1}{2}, \frac{1}{2}]$ and $s = \frac{1}{2} + ir$. So each $\lambda \neq \frac{1}{4}$ corresponds to two r values $\pm r$, and two s values $s = \frac{1}{2} \pm ir$.

1.2 $SL_2(\mathbb{Z})$

Let $\Gamma = PSL_2(\mathbb{Z})$, it is a discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$ and so acts totally discontinuously on $\mathcal{H}$. From the Euclidean algorithm or otherwise one shows that $\Gamma = \langle S, T \rangle$, where as Möbius transformations $S(z) = -\frac{1}{z}$ and $T(z) = z + 1$. We will occasionally overload the notation by identifying S and T with the matrices

$$S = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right] \qquad \text{and} \qquad T = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$$

or their image in $\mathrm{PSL}_2(\mathbb{Z})$.

A fundamental domain for Γ is given by $\mathcal{F} = \{z \in H : |\operatorname{Re} z| < 1/2 \text{ and } |z| > 1\}$, see the figure. This means that

1. For each $z \in H$ there is $\gamma \in \Gamma$ such that $\gamma z \in \overline{\mathcal{F}}$.

2. If $z_1, z_2 \in \overline{F}$ and $\gamma \in \Gamma$ are such that $\gamma z_1 = z_2$ then $z_1, z_2 \in \partial\mathcal{F}$, where $\partial\mathcal{F}$ is the boundary of $\mathcal{F}$.

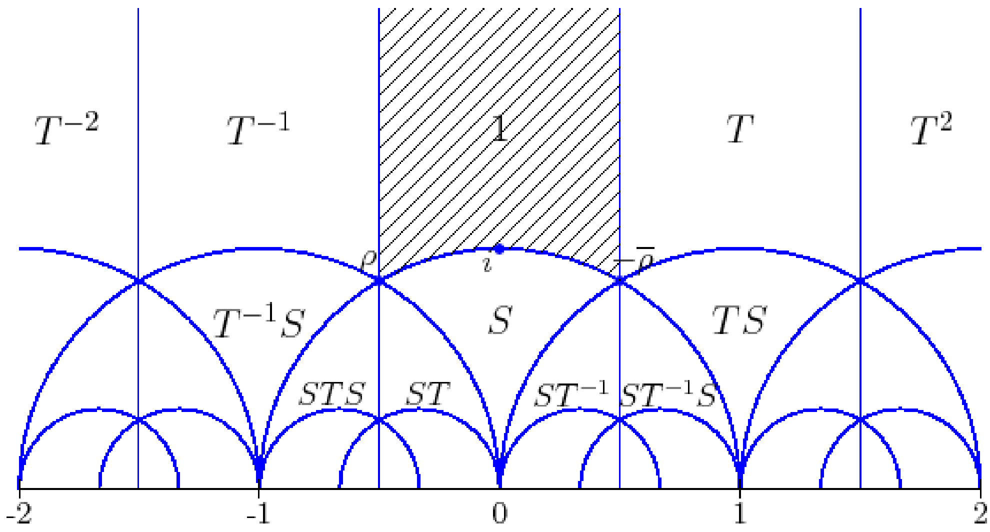

Figure 1.1:

1.3 Invariant functions

The primary object of study in this thesis are functions invariant under Γ: $f(\gamma z) = f(z)$ for all $\gamma \in \Gamma$. These function can be identified with functions $f : \Gamma\backslash\mathcal{H} \to \mathbb{C}$, and for the purposes of spectral theory they can also be identified with functions on $\mathcal{F}$. In what follows we will freely move between these interpretations.

Define now the inner product of two invariant functions via:

$$\langle f, g\rangle = \int_{\Gamma\backslash\mathcal{H}} f\overline{g}d\mu = \int_{\mathcal{F}} f\overline{g}d\mu$$

(This doesn't depend on the choice of the fundamental domain $\mathcal{F}$), and let

$$L^2(\Gamma\backslash\mathcal{H}) = \{f : \langle f, f\rangle < \infty\}.$$

Then Δ is an unbounded symmetric operator with respect to this inner product and the goal is the "spectral decomposition" of Δ on $L^2(\Gamma\backslash\mathcal{H})$.

Fourier expansion

Let

$$\Gamma_\infty = \{\pm \left[\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right] : k \in \mathbb{Z}\}.$$

Because of Γ_∞-invariance, invariant functions have Fourier expansions of the form

$$f(z) = \sum_{n\in\mathbb{Z}} a_n(y)e(nx)$$

where as usual $e(nx) = e^{2\pi inx}$. If $\Delta f + \lambda f = 0$, then by separation of variables the Fourier coefficients will satisfy

$$a_n''(y) + (s(1-s)/y^2 - 4\pi^2n^2)a_n(y) = 0.$$

When $n = 0$ two independent solutions are provided by y^s, y^{1-s}, except at $s = 1/2$, when they are $y^{1/2}, y^{1/2}\log y$. When $\neq 0$, the two independent solutions are given by $\sqrt{y}K_{s-1/2}(y)$

and $\sqrt{y}I_{s-1/2}(y)$ [100, Chapter 10]. Since I_s is exponentially growing as $y \to \infty$, it is clear that for an L^2-eigenfunction we have

$$a_n(y) = 2a(n)\sqrt{y}K_{s-1/2}(2\pi|n|y).$$

for some $a(n) \in \mathbb{C}$. This is one of many possible normalizations for $a(n)$, but will be used in this thesis.

Now the Fourier expansion, which in its general form exist for all invariant functions gives a distinguished subspace given by those forms whose "constant term" $a_0(y)$ is 0 almost everywhere, as a function of y. It is called the space of cusp forms

$$L^2_{cusp} = \{f \in L^2(\Gamma\backslash\mathcal{H}) : \int_0^1 f(x+iy)dx = 0 \text{ (a.e. } y)\}$$

They are easy to characterize in another way that we will now describe.

Incomplete Eisenstein series.

The easiest construction of an invariant function is to average over the group Γ. This can be done for example if we start with a smooth compactly supported function $\kappa : \mathcal{H} \to \mathbb{C}$ and take

$$\sum_{g\in\Gamma} \kappa(\gamma z)$$

(The sum is even locally finite and clearly invariant.)

A variant, which is even more important is that we start with a function ψ that is already Γ_∞ invariant and consider

$$\sum_{g\in\Gamma_\infty\backslash\Gamma} \psi(\gamma z).$$

(This is just like the first construction if we let $\psi(z) = \sum_{n\in\mathbb{Z}} \kappa(z+n)$.)

Here, and in what follows, summation over left, or right coset spaces mean that the sum is over a representative set of each coset. Implicit in these definitions is the (usually trivial) fact that the sum does not depend on this choice of representatives.

In this second construction we can take the Fourier series components of ψ to reduce to the case when $\psi(z) = \phi(y)e(mx)$. These are the Poincaré series that are the main objects of study in what follows.

The simplest case is when $m = 0$ where we define

$$E(z,\phi) = \sum_{g\in\Gamma_\infty\backslash\Gamma} \phi(\mathrm{Im}(\gamma z))$$

and let

$$\mathcal{E} = \text{ closure of } \{E(z,\phi) : \phi \text{ smooth, compactly supported on } \mathbb{R}^+\}$$

Note that in the definition of $E(z,\phi)$, we may replace compactly supported ϕ with functions that satisfy $\phi(y) = O(y^\alpha)$ as $y \to 0^+$, for some $\alpha > 1$. The resulting series are locally uniformly convergent in norm, without any condition on the growth as $y \to \infty$. Such conditions at infinity are required however if one wants $E(z,\phi)$ to be in L^2.

By unfolding the sum that defines $E(z, \phi)$, one gets that u is orthogonal to $E(z, \phi)$, if and only if

$$\int_0^\infty \phi(y) a_0(y) \frac{dy}{y^2} = 0.$$

and so

$$L^2(\Gamma \backslash \mathcal{H}) = \mathcal{E} \oplus L^2_{cusp}.$$

Interestingly, although $\mathcal{E}$ and L^2_{cusp} are defined group theoretically, they have the following spectral description:

- Δ has a pure discrete spectrum on L^2_{cusp}: there is a basis (in the Hilbert space sense) consisting of eigenfunctions. (Called eigenforms.)
- Δ has a pure continuous spectrum on $\mathcal{E}$ except for the constant function which comes as a residue of the Eisenstein series defined below.

1.4 Eisenstein series

The resolution of the space $\mathcal{E}$ of incomplete Eisenstein series is the theory of the Eisenstein series $E(z, s)$ defined for $\mathrm{Re}(s) > 1$ by

$$E(z, s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\mathrm{Im}\, \gamma z)^s = \tfrac{1}{2} (\mathrm{Im}\, z)^s \sum_{\gcd(c,d)=1} |cz + d|^{-2s}, \tag{1.4.1}$$

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. [74])

$$E^*(z, s) = \Lambda(2s) y^s + \Lambda(2-2s) y^{1-s} + 2y^{1/2} \sum_{n \neq 0} |n|^{s-1/2} \sigma_{1-2s}(|n|) K_{s-\frac{1}{2}}(2\pi |n| y) e(nx), \tag{1.4.2}$$

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

$$E^*(z, 1-s) = E^*(s). \tag{1.4.3}$$

Furthermore we have that

$$\mathrm{Res}_{s=1} E^*(z, s) = -\mathrm{Res}_{s=0} E^*(z, s) = \tfrac{1}{2}. \tag{1.4.4}$$

The residue at $s = 1$ gives rise to constant term c_0 in (1.4.5).

Let $\phi : (0, \infty) \to \mathbb{C}$ and consider its Mellin transform:

$$\hat{\phi}(s) = \int_0^\infty \phi(y) y^{-s} \frac{dy}{y}$$

By Mellin inversion, if we choose $\sigma > 1$ where the Eisenstein series converges absolutely we get

$$E(z, \phi) = \frac{1}{2\pi i} \int_{(\sigma)} E(z, s) \hat{\phi}(s) ds.$$

We can move the line of integration to $\sigma = 1/2$. When doing so we pickup a residue at $s = 1$, a constant function. Then we have the following version of the spectral theorem:

Theorem 1.4.1. *If f is a smooth function in $\mathcal{E}$ then*

$$f(z) = c_0 + \frac{1}{4\pi}\int_{-\infty}^{\infty} c(t)E(z, \tfrac{1}{2} + it)dt \tag{1.4.5}$$

where

$$c_0 = \langle f, 1\rangle\langle 1, 1\rangle^{-1} = \frac{3}{\pi}\int_{\mathcal{F}} f(z)d\mu(z)$$

$$c(t) = \langle f, E(z, 1/2 + it)\rangle = \int_{\mathcal{F}} f(z)\overline{E(z, 1/2 + it)}d\mu(z)$$

The complementary statement about the cuspidal part of L^2 is the following

Theorem 1.4.2. *The eigenfunctions u_j of Δ in L^2_{cusp} form a Hilbert-space basis of L^2_{cusp}. The associated eigenvalues λ_j form a discrete set, and $\lambda_j \to \infty$. If $f \in L^2_{cusp}$ is smooth then*

$$f(z) = \sum \frac{\langle f, u_j\rangle}{\langle u_j, u_j\rangle} u_j.$$

1.5 Invariant integral operators and the resolvent kernel

Because of the role it plays in our arguments we outline the proof of the spectral resolution for the cuspidal part. This is merely a sketch, details can be found in [65] and [72].

Let $k(z, w) = \kappa(u(z, w))$ where κ is smooth, compactly supported on $(0, \infty)$, or sufficiently fast decaying at 0 and at ∞. This is a point pair invariant, in the sense that $k(gz, gw) = k(z, w)$, and it follows that $\Delta_z k = \Delta_w k$. Therefore we also have for smooth $f : \mathcal{H} \to \mathbb{C}$ and

$$Lf = \int_{\mathcal{H}} k(z, w)f(w)d\mu(w)$$

that $\Delta Lf = L\Delta f$.

If in the above f is Γ-invariant we can express Lf as

$$Lf = \int_{\mathcal{F}} K(z, w)f(w)d\mu(w)$$

using the kernel

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

This is an integral operator on $L^2(\Gamma\backslash\mathcal{H})$, and maps L^2_{cusp} to itself, since taking the 0-th Fourier coefficient commutes with L.

In general the kernel K is not a bounded function, and the problem of the growth at the cusp is solved by subtracting

$$H(z, w) = \sum_{g\in\Gamma_\infty\backslash\Gamma} h(z, \gamma w)$$

where

$$h(z, w) = \int_{\mathbb{R}} k(z, w + t)dt.$$

$H(z,w)$ is an incomplete Eisenstein series (in w), and so the integral operator with kernel

$$\hat{K}(z,w) = K(z,w) - H(z,w)$$

which is bounded on $\mathcal{F} \times \mathcal{F}$ acts the same way on L^2_{cusp} as L. Therefore L as an operator on L^2_{cusp} is compact and has discrete spectrum. Since it commutes with Δ we almost get that Δ has discrete spectrum in L^2_{cusp}, but this requires the construction of a kernel whose image is dense. This is easiest done via the resolvent.

We will say that λ is in the resolvent set for Δ if there is a bounded operator R, such that $(\Delta + \lambda)Rf = f$ for all f, and $R(\Delta + \lambda)f = f$, whenever defined.

A typical λ is in the resolvent set, when it is not we say λ is in the spectrum. In what follows it is better to use the $\lambda = s(1-s)$ description, we will refer to s as being in the resolvent set, or the spectrum.

To construct $R = R_s$ for any $\operatorname{Re} s > 1$ one looks at geodesic polar coordinates. Start with a function $k_s(u) = k_s(u(z,w))$ such that

$$\Delta_w g_s(u) + s(1-s)k_g(u) = \delta_z$$

where δ_z is Dirac's δ at z. Explicit computations show that g_s has to satisfy

$$u(u+1)g_s''(u) + (2u+1)g_s'(u) + s(1-s)g_s(u) = 0$$

and $g_s(u) \sim C|\log u|$ as $u \to 0^+$. The solutions are standard [100, 48] but their explicit form is not needed for us, only that

$$\sum_{g \in \Gamma} g_s(z, \gamma w)$$

converges to G_s absolutely and locally uniformly on $\operatorname{Re} s > 1$. By a suitable choice, say $s = 2$, this gives that L^2_{cusp} is spanned by the eigenfunctions of $(\Delta - 2)^{-1}$, but then they are eigenfunctions of Δ as well.

To justify the analysis it is convenient to look at $G_a - G_b$ for some $a \neq b, \operatorname{Re} a, \operatorname{Re} b > 1$. Then the singularities of $G_a(z,w), G_b(z,w)$ at $z = w$ cancel each other out, and one may use Hilbert's identity

$$R_a - R_b = (a(1-a) - b(1-b))R_a R_b$$

where $R_s = (\Delta + s(1-s))^{-1}$.

We therefore have

Theorem 1.5.1. *The eigenfunctions u_j of Δ in L^2_{cusp} form a Hilbert-space basis of L^2_{cusp}. The associated eigenvalues from a discrete set, and $\lambda_j \to \infty$. If $f \in L^2_{cusp}$ then*

$$f(z) = \sum_j \frac{\langle f, u_j \rangle}{\langle u_j, u_j \rangle} u_j.$$

1.6 Residues of the Green function

It is easy to compute the spectral resolution of the resolvent kernel, at least heuristically. If

$$G_a(z,w) = c_0 + \sum_{\varphi} c_j(z)\overline{\varphi_j(w)} + \text{ Eisenstein part}$$

then

$$(\Delta + a(1-a)) \int_{\mathcal{F}} G_a(z,w)\varphi_j(w)d\mu(w) = \varphi_j(z)$$

which gives $c_j(z) = \frac{1}{s_j(1-s_j)-a(1-a)}\phi_j(z)$.

Again this can be made precise by considering $G_a - G_b$.

$$G_a(z,w) - G_b(z,w) = \rho(a,b,0) + \frac{1}{4\pi i}\int_{(1/2)} \rho(a,b,s)E(z,s)\overline{E(w,s)}ds +$$
$$\sum_{\varphi} \rho(a,b,s_j)\varphi_j(z)\overline{\varphi_j(w)}.$$

where

$$\rho(a,b,s) = \frac{1}{s(1-s)-a(1-a)} - \frac{1}{s(1-s)-b(1-b)}$$

It is an important fact that ρ as a rational function of s is $O(1/s^4)$. Note that $N(T)$ the number of eigenvalue parameters s_j less than T in magnitude is $O(T^2)$.

From the spectral expansion above we conclude that the Green function has a meromorphic continuation. The two parts behave differently

1. The Eisenstein part has an analytic continuation to $\operatorname{Re} s \in (0,1)$ with no poles.
2. If $s_j = 1/2 + it_j$ is one of the spectral parameters then
$$\operatorname{Res}_{s=s_j} G_s(z,w) = \sum_{\varphi} \langle \varphi,\varphi\rangle^{-1}\varphi(z)\overline{\phi(w)}.$$

We will take the Fourier expansion of $G_s(z,w)$ as in [48]. We have

$$G_s(z,w) = \sqrt{\operatorname{Im} w}\sum_{m\in\mathbb{Z}} F_{-m}(z,s)K_{s-1/2}(2\pi|m|\operatorname{Im} w)e(m\operatorname{Re} w), \tag{1.6.1}$$

where

$$F_m(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} f_m(\gamma z,s), \tag{1.6.2}$$

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

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

By the above we get

Proposition 1.6.1. *For any $m \neq 0$ we have that $F_m(z,s)$ has meromorphic continuation in s to $\operatorname{Re}(s) > 0$ and that*

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

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 (3.2.16).

1.7 Congruence subgroups and cusps

A congruence subgroup is the preimage (under the natural map) of a subgroup of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, for some N. In this thesis we will only use

$$\Gamma_0(q) = \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \Gamma : c \equiv 0 \mod q\},$$

and except for the last chapter even this we will only need for $q = 4$. If Γ' is a finite index

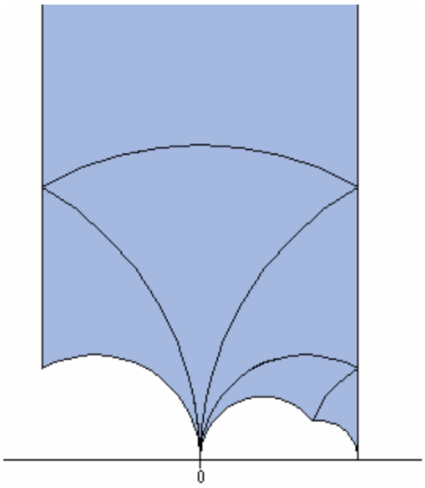

Figure 1.2: A fundamental domain for $\Gamma_0(4)$.

subgroup of Γ it acts on $P^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$ and has finitely many orbits. These orbits of Γ' are called the cusps of Γ'. They are frequently identified with a representative that is in the closure of a fundamental domain in $\mathcal{H}$ when viewed as a subset of $P^1(\mathbb{C})$. If $\mathfrak{c}$ is a cusp, we let $\Gamma'_{\mathfrak{c}}$ be the stabilizer of one of its representatives, (choosing a different representative leads to a Γ' conjugate subgroup). If $\mathfrak{c} = a/c$, $(a, c) = 1$ and $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \Gamma$, then $\mathfrak{c} = \gamma\infty$, and so $\gamma^{-1}\Gamma'_{\mathfrak{c}}\gamma$ fixes ∞, and as such is a finite index subgroup of Γ_∞. We call this index the width of $\mathfrak{c}$.

Some simple observations follow [72] . Let $\mathfrak{c}$ be a cusp of $\Gamma_0(q)$. Then $\mathfrak{c}$ is equivalent to some $\frac{u}{v}$ for which $v|q$. Moreover $\frac{u}{v}$, and $\frac{u'}{v'}$ give rise to the same cusp, if and only if $v = v'$, and $u \equiv u' \bmod (v, q/v)$. The width of the cusp $\frac{u}{v}$ is $\frac{q}{(v^2,q)}$.

For example, when $q = 4$, there are 3 cusps, ∞, 0, and 1/2, of width 1, 4 and 1 see the figure above.

1.8 Half integral weight modular forms

Half integral weight modular forms generalize the (modified) Jacobi theta series,

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

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

$$J(\gamma, z) = \frac{\theta(\gamma z)}{\theta(z)} \quad \text{for } \gamma \in \Gamma_0(4). \tag{1.8.1}$$

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

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

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

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

A Maass form of weight 1/2 for $\Gamma_0(4)$ has weight 1/2, is smooth and satisfies $\Delta_{1/2}F + \lambda F = 0$, where we write $\lambda = \lambda(F) = \frac{1}{4} + (\frac{r}{2})^2$. Such a form has Fourier expansion

$$\psi(z) = \sum_{n\neq 0} b(n) W_{\frac{1}{4}\operatorname{sign} n, \frac{ir}{2}}(4\pi|n|y)e(nx). \tag{1.8.2}$$

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 [48] (see also [106]). It satisfies

$$\big(\Delta_{1/2} + s(1-s)\big) \int_{\Gamma_0(4)\backslash\mathcal{H}} G_{\frac{1}{2}}(z, z'; s)u(z)d\mu(z) = u(z') \tag{1.8.3}$$

for $u \in L^2(\Gamma_0(4)\backslash\mathcal{H}, d\mu)$ with weight 1/2. By Theorem 3.1 of [48] we have the Fourier expansion[1]

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

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

$$F_{1/2,n}(z, s) = \frac{\Gamma(s - \frac{1}{4}\operatorname{sign} n)}{4\pi|n|\Gamma(2s)} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} J(\gamma, z)^{-1} f_{1/2,n}(\gamma z, s) \tag{1.8.4}$$

with

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

As in the weight 0 case, we have the following

Proposition 1.8.1. *$F_{1/2,n}(z, s)$ has a meromorphic continuation to $\operatorname{Re}(s) > 0$ with simple poles at the points $\frac{1}{2} + \frac{ir}{2}$ giving the discrete spectrum of $\Delta_{1/2}$ and that*

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

and

$$\operatorname{Res}_{s=\frac{1}{2}+\frac{ir}{2}}(2s-1)F_{1/2,n}(z, s) = \sum_\psi \overline{b}(n)\psi(z). \tag{1.8.5}$$

Here the sum is over an orthonormal basis $\{\psi\}$ of Maass cusp forms for V_r and $b(n)$ is as 1.8.2.

[1] 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 $k = 1/4$.

1.9 Hecke operators and Shimura-Shintani correspondence

In this section we denote by U the space of cusp forms of weight 0 and U_r the space of cusp forms of eigenvalue $1/4 + r^2$. Similarly V is the space of weight 1/2 cusp forms, and V_r those cusp forms with spectral parameter $1/2 + ir/2$.

For each m one also has the Hecke operator $T(m)$ acting on U [72], these operators commute with each other and Δ, therefore preserving U_r. If φ is also an eigenform of the Hecke operators then one knows that $a(1) \neq 0$, and we may assume that $a(1) = 1$. We will call such a form Hecke-normalized.

It is known that the space U_r has a basis of Hecke-normalized eigenforms $\{\varphi\}$.

Furthermore 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 $\varphi(-\overline{z}) = a(-1)\varphi(z)$. Thus the associated L-function has an Euler product (for $\mathrm{Re}(s) > 1$):

$$L(s;\varphi) = \sum_{n\geq 1} a(n)n^{-s} = \prod_{p \text{ prime}} (1 - a(p)p^{-s} + p^{-2s})^{-1}. \tag{1.9.1}$$

The space V is equipped with Hecke operators $T_{1/2}(m^2)$ [118]. There is an important distinguished subspace of V_r, denoted by V_r^+ and called after Kohnen [82] the *plus space*, that contains those Maass cusp forms $\psi \in V_r$ whose n-th Fourier coefficient $b(n)$ vanishes unless $n \equiv 0, 1 \mod 4$. It is clearly invariant under $\Delta_{1/2}$.

It is shown in [77] that V_r^+ has an orthonormal basis $B_r = \{\psi\}$ consisting of eigenfunction of all Hecke operators $T_{1/2}(p^2)$ where $p > 2$ is prime. Fix such a basis B_r.

Given $\psi \in B_r$ with Fourier expansion

$$\psi(z) = \sum_{n\neq 0} b(n) W_{\frac{1}{4}\operatorname{sign} n, \frac{ir}{2}}(4\pi|n|y)e(nx) \tag{1.9.2}$$

and a fundamental discriminant d with $b(d) \neq 0$ the Hecke relation $T_{1/2}(p^2)\psi = a_\psi(p)\psi$ 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

$$\prod_p (1 - a_\psi(p)p^{-s} + p^{-2s})^{-1} = \sum_{n\geq 1} a_\psi(n)n^{-s} \tag{1.9.3}$$

and let

$$\operatorname{Shim}\psi(z) = y^{1/2} \sum_{n\neq 0} 2a_\psi(|n|)K_{ir}(2\pi|n|y)e(nx). \tag{1.9.4}$$

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

Theorem 1.9.1 (Shimura lift)**.** *If $\psi \in V_r$ is an eigenfunction of all 1/2-weight Hecke operators, then* $\operatorname{Shim}\psi \in U_r$ *and is also an eigenform of all weight 0 Hecke operators.*

1.10 Holomorphic modular forms and functions

Let $\Gamma' \leq \Gamma$ BE a subgroup of finite index, and let $f : \mathcal{H} \to \mathbb{C}$ be a holomorphic function that satisfies

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \tag{1.10.1}$$

whenever $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \Gamma'$. Such a form is called weakly holomorphic. We will now assume that $\Gamma' = \Gamma$. In particular we have $T \in \Gamma'$ and so we have

$$f(z+1) = f(z).$$

It follows from this that there is $\tilde{f} : \{0 < |q| < 1\} \to \mathbb{C}$, holomorphic such that

$$f(z) = \tilde{f}(e^{2\pi iz}) \tag{1.10.2}$$

We will say f is holomorphic (resp. meromorphic) at ∞ if $\tilde{f}$ is holomorphic (resp. meromorphic) at 0. For ease of notation we will denote

$$q = e^{2\pi iz} \tag{1.10.3}$$

so that by (1.10.2) we have

$$f(z) = \sum_n a_n q^n \tag{1.10.4}$$

for some $a_n \in \mathbb{C}$, called the Fourier coefficients of f.

Definition 1.10.1. We let

$$M_k^! = \{f : \mathcal{H} \to \mathbb{C} : f \text{ satisfies 1.10.1}, f(z) = \sum_{n=n_0}^{\infty} a_n q^n, \text{ for some } n_0\}$$

Similarly let

$$M_k = \{f : \mathcal{H} \to \mathbb{C} : f \text{ satisfies 1.10.1}, f(z) = \sum_{n=0}^{\infty} a_n q^n\}$$

and

$$S_k = \{f : \mathcal{H} \to \mathbb{C} : f \text{ satisfies 1.10.1}, f(z) = \sum_{n=1}^{\infty} a_n q^n\}$$

For example let

$$E_4(z) = 1 + 240 \sum_{n=1}^{\infty} \frac{n^3 q^n}{1-q^n}$$

$$E_6(z) = 1 - 504 \sum_{n=1}^{\infty} \frac{n^5 q^n}{1-q^n}$$

and

$$\Delta(z) = \frac{1}{1728}\left(E_4^3(z) - E_6^2(z)\right).$$

Then $E_4 \in M_4, E_6 \in M_6$ and $\Delta \in S_{12}$. Moreover Δ does not vanish on $\mathcal{H}$ and the function

$$j(z) = \frac{E_4^3(z)}{\Delta(z)}$$

is invariant. Clearly it is in $M_0^!$, the space of meromorphic modular functions, and it is well known classically that $M_0^! = \mathbb{C}(j)$. The map $j : \mathcal{H} \to \mathbb{C}$ is conformal except at the orbits of i and $-\frac{1}{2} + \frac{\sqrt{3}}{2}i$. It establishes a bijection between $\Gamma\backslash\mathcal{H}$ and $\mathbb{C}$.

1.11 Closed geodesics

There are three closely related ways to describe closed geodesics. All three will appear in the thesis.

Hyperbolic elements

Let σ be hyperbolic, with fixed points w_1, w_2. The geodesic S_σ connecting the two endpoints becomes a closed geodesic in $\Gamma\backslash\mathcal{H}$. This requires some clarifying because $\Gamma\backslash\mathcal{H}$ is only an orbifold. Since σ is hyperbolic, it has real eigenvalues say $\lambda_1 > \lambda_2$, which we may assume are positive after replacing σ by $-\sigma$ if necessary. Then we have

$$\sigma \begin{bmatrix} w_1 & w_2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} w_1 & w_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

and we can define

$$\sigma(t) = \begin{bmatrix} w_1 & w_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^t & 0 \\ 0 & \lambda_2^t \end{bmatrix} \begin{bmatrix} w_1 & w_2 \\ 1 & 1 \end{bmatrix}^{-1}.$$

If we now fix a point z_0 on S_σ, then the image of the (parametrized) curve $t \to \sigma(t)z_0$ becomes periodic on $\Gamma\backslash\mathcal{H}$ with period 1. This may not be the minimal period, if it is we call σ primitive. Then σ is primitive if and only if it is not a positive power of another hyperbolic element. Since the eigenvalues λ_1, λ_2 are units in the maximal order $\mathcal{O}$ of $K_\sigma = \mathbb{Q}(\sqrt{(a+d)^2-4})$, σ is primitive if and only if the eigenvalues are the totally positive fundamental units $\epsilon_D > 1/\epsilon_D$, where D is the discriminant of $\mathcal{O}$. The above parametrization is not by arc-length, the length of a primitive closed geodesic is easy to compute and is

$$\text{length}(\mathcal{C}_A) = 2\log\epsilon_D. \tag{1.11.1}$$

where ϵ_D is the totally positive fundamental unit in $\mathcal{O}$.

It is easy to see that conjugate hyperbolic elements give the same curve in $\Gamma\backslash\mathcal{H}$.

Real quadratic extensions

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 $w \mapsto w'$ be the non-trivial Galois automorphism of$\mathbb{K}$ and for $\alpha \in \mathbb{K}$ let $N(\alpha) = \alpha\alpha'$. 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 $\epsilon_D > 1$ 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 $N(\alpha) = \alpha\alpha' < 0$. Then

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

is the class group in the wide sense. Clearly $J \neq I$ iff $\mathcal{O}_\mathbb{K}$ 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 $p \equiv 3 \pmod 4$.

For a fixed narrow ideal class $A \in \mathrm{Cl}^+(\mathbb{K})$ and $\mathfrak{a} = w\mathbb{Z} + \mathbb{Z} \in A$ with $w > w'$ let $\mathcal{S}_w$ be the geodesic in $\mathcal{H}$ with endpoints w' and w. The modular closed geodesic $\mathcal{C}_A$ on $\Gamma\backslash\mathcal{H}$

dc_1553_18

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

$$\begin{aligned} \epsilon_D w =& aw + b \\ \epsilon_D =& cw + d, \end{aligned} \tag{1.11.2}$$

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

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

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 (that could overlap) in $\mathcal{F}$. We also use $\mathcal{C}_A$ to denote this set of arcs. Again it is well-known and easy to see using (1.11.2) that

$$\text{length}(\mathcal{C}_A) = 2 \log \epsilon_D. \tag{1.11.3}$$

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

$$[a, -b, c] \text{ represents } A^{-1} \tag{1.11.4}$$
$$[-a, b, -c] \text{ represents } JA \tag{1.11.5}$$
$$[-a, -b, -c] \text{ represents } JA^{-1}. \tag{1.11.6}$$

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

$$\gamma_Q = \pm \begin{bmatrix} \frac{t - b'u}{2} & -c'u \\ a'u & \frac{t + b'u}{2} \end{bmatrix}, \tag{1.11.7}$$

where (t, u) gives the smallest integer solution with $t, u \geq 1$ to

$$t^2 - d'u^2 = 4$$

(see [108]).

If

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

as above then $\gamma_Q = \gamma_w$ and $\varepsilon_D = \frac{t+u\sqrt{D}}{2}$. Therefore the closed geodesic associated to the hyperbolic element γ_Q agrees with the closed geodesic associated to the narrow class A.

Using (1.11.6) we see that the closed geodesic $\mathcal{C}_{JA^{-1}}$ has the same image as $\mathcal{C}_A$ but with the opposite orientation.

It is also possible to describe the primitive quadratic form associated to a hyperbolic element. If $\sigma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is a primitive hyperbolic element, then let $u = \gcd(c, d - a, b)$ and set

$$Q_\sigma(z) = -\frac{c}{u}X^2 + \frac{a-d}{u}XY + \frac{b}{u}Y^2,$$

then Q_σ is primitive and its group of automorphs is generated by $\sigma_{Q_\sigma} = \sigma$.

When $D > 0$ and Q is primitive and $n \in \mathbb{Z}^+$ define $\mathcal{C}_{nQ} = \mathcal{C}_Q$. 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.

Remark 1.11.1. 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 [110] has given a comprehensive treatment of these remarkable geodesics for arbitrary discriminants.

1.12 Genus characters

We need to define genus characters for arbitrary discriminants. We will mainly use the language of binary quadratic forms. 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 $D = d'd$ where d is fundamental we define

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

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

$$S_m(d', d; c) = \sum_{\substack{b (\mathrm{mod}\ c) \\ b^2 \equiv D \ (\mathrm{mod}\ c)}} \chi\left(\left[\tfrac{c}{4}, b, \tfrac{b^2 - D}{c}\right]\right) e\left(\tfrac{2mb}{c}\right). \tag{1.12.1}$$

Clearly

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

We have the identity

$$\chi_d(-Q) = (\operatorname{sgn} d)\chi_d(Q). \tag{1.12.2}$$

A crucial ingredient in what follows is an identity connecting the weight 1/2 Kloosterman sum with $S_m(d,d';c)$ above. In a special case this identity is due to Salié and variants have found many applications in the theory of modular forms. We shall use a general version due essentially to Kohnen [83]. To define the weight 1/2 Kloosterman sum we need an explicit formula for the theta multiplier in $J(\gamma,z)=\theta(\gamma z)/\theta(z)$ introduced above. This may be found in [118, p. 447]. As usual, for non-zero $z\in\mathbb{C}$ and $v\in\mathbb{R}$ we define $z^v=|z|^v\exp(iv\arg z)$ with $\arg z\in(-\pi,\pi]$. We have

$$J(\gamma,z)=(cz+a)^{1/2}\varepsilon_a^{-1}\left(\tfrac{c}{a}\right)\quad\text{for}\quad\gamma=\pm\left[\begin{smallmatrix}*&*\\c&a\end{smallmatrix}\right]\in\Gamma_0(4),$$

where $\left(\frac{c}{a}\right)$ is the extended Kronecker symbol and

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

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

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

be 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

$$K^+(m,n;c)=(1-i)K_{1/2}(m,n;c)\times\begin{cases}1 & \text{if } c/4 \text{ is even}\\ 2 & \text{otherwise.}\end{cases}\tag{1.12.3}$$

It is easily checked that

$$K^+(m,n;c)=K^+(n,m;c)=\overline{K^+(n,m;c)}.\tag{1.12.4}$$

The following identity is proved by a slight modification of the proof given by Kohnen in [83, Prop. 5, p. 259] (see also [35], [75] and [124]).

Proposition 1.12.1. *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*

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

By Möbius inversion in two variables this can be written in the form

$$c^{-1/2}\,K^+(d,m^2D,c)=\sum_{n|\left(m,\frac{c}{4}\right)}\mu(n)\left(\tfrac{D}{n}\right)S_{m/n}\left(d,D;\tfrac{c}{n}\right).\tag{1.12.5}$$

Note that this gives an identity for $K^+(d,d',c)$ for any pair $d,d'\equiv 0,1 \pmod 4$. An immediate consequence of (1.12.5) and the obvious upper bound

$$S_m(d,D;c)\ll_\epsilon c^\epsilon$$

is the upper bound

$$K^+(d, d', c) \ll_\epsilon c^{1/2+\epsilon}, \tag{1.12.6}$$

which holds for any $\epsilon > 0$. Furthermore, since for any $m, n \in \mathbb{Z}$ we have

$$K_{1/2}(m, n; c) = \tfrac{1}{4} K_{1/2}(4m, 4n; 4c),$$

(1.12.6) implies that for any $m, n \in \mathbb{Z}$

$$K_{1/2}(m, n, c) \ll_\epsilon c^{1/2+\epsilon}.$$

This elementary bound correspond to Weil's bound for the ordinary (weight 0) Kloosterman sum

$$K_0(m, n; c) = \sum_{\substack{a (\mathrm{mod}\ c) \\ (a,c)=1}} e\left(\tfrac{ma+n\bar{a}}{c}\right),$$

which states that (see [128], [66, Lemma 2])

$$K_0(m, n; c) \ll_\epsilon (m, n, c)^{1/2} c^{1/2+\epsilon}. \tag{1.12.7}$$

Chapter 2

The Shimura-Shintani correspondence and formulas of Katok-Sarnak type

2.1 Background and statements of results

Correspondences between automorphic forms on different groups have a long and rich history as can be seen in the works of Doi-Naganuma, Shimura, Langlands and many others. The Shimura-Shintani [118, 119] correspondence lead to many applications via the formulas of Waldspurger [125] and Kohnen-Zagier [84]. The analogous formula in case of Maass forms was given by Katok and Sarnak[77]. A particularly important application of such formulas is due to Duke on the equidistribution of CM points and closed geodesics [35]. There are various methods for proving this family of formulas that use either theta-kernels, or spectral methods [11]. Recently an extension of the Katok-Sarnak formula was developed in [41] with applications to a new type of equidistribution result. In this chapter we will present the part of the paper where this extension is proved.

2.1.1 Katok-Sarnak type formulas

Let $u(z) = E(z,s),$. We recall 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

$$\Lambda(s,\chi_d) = \pi^{-s/2}\Gamma(\tfrac{s+\mathfrak{a}}{2})|d|^{s/2}L(s,\chi_d). \tag{2.1.1}$$

Theorem 2.1.1. *For the genus character χ associated to $D = d'd$ and $\mathrm{Re}(s) = \frac{1}{2}$ we have*

$$\Lambda(s,\chi_{d'})\Lambda(s,\chi_d) = \sum_{Q\in}\chi(Q)\begin{cases}\int_{C_Q} i\partial_z E^*(z,s)dz & \text{if } d',d<0\\ \int_{C_Q} E^*(z,s)y^{-1}|dz| & \text{if } d',d>0\\ 2\sqrt{\pi}\omega_D^{-1}\, E^*(z_A,s) & \text{if } d'd<0.\end{cases}$$

This formula is due to Hecke except when $d',d<0$. Theorem 2.1.1 can be expressed in terms of Maass forms of weight 1/2.

Set for fundamental d

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

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 [40, Proposition 2 p.959] that

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

has weight 1/2 for $\Gamma_0(4)$. The idea behind this example originates in the papers of H. Cohen [27] and Goldfeld and Hoffstein [51]. See also [117], [34].

The formula

$$\Lambda(s, \chi_{d'}) \Lambda(s, \chi_d) = 2\sqrt{\pi} |D|^{3/4} b(d', s) b(d, s) \tag{2.1.2}$$

in connection with Theorem 2.1.1 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 [77], together with [8], gives the case $d = 1$ in the following.[1]

Theorem 2.1.2 ([43]). *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*

$$12\sqrt{\pi} |D|^{\frac{3}{4}} b(d') \bar{b}(d) = \langle \varphi, \varphi \rangle^{-1} \sum_{Q \in} \chi(Q) \begin{cases} \int_{C_Q} i \partial_z \varphi(z) z & \text{if } d', d < 0 \\ \int_{C_Q} \varphi(z) y^{-1} |dz| & \text{if } d', d > 0 \\ 2\sqrt{\pi}\, \omega_D^{-1}\, \varphi(z_Q) & \text{if } d'd < 0, \end{cases} \tag{2.1.3}$$

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*

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

[1] Except that when $d' < 0$ we get in (2.1.3) on the RHS $2\sqrt{\pi}\,\omega_D^{-1}$ instead of their $(2\sqrt{\pi}\,\omega_D)^{-1}$.

Remarks. The $\sqrt{\pi}$ in (2.1.2) and (2.1.3) is an artifact of the normalization of the Whittaker function. Also, if we choose F in Theorem 2.1.2 so that $\langle F, F\rangle = 6$, which is the index of $\Gamma_0(4)$ in Γ, then we get 2 in the LHS of (2.1.3) instead of 12, which matches the Eisenstein series case (2.1.2). Perhaps not coincidentally,

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

and by [26] 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 [77] and in general by Baruch and Mao [8]. Here we quote their result in our context. Under the same assumptions as in Theorem 2.1.2 we have

$$12\pi|d||b(d)|^2 = \langle\varphi, \varphi\rangle^{-1}\Gamma(\tfrac{1}{2} + \tfrac{ir}{2} - \tfrac{\operatorname{sign} d}{4})\Gamma(\tfrac{1}{2} - \tfrac{ir}{2} - \tfrac{\operatorname{sign} d}{4})L(\tfrac{1}{2}, \varphi, \chi_d), \tag{2.1.4}$$

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 sub-convexity bound, this time for a twisted L-function.

Results like Theorem 2.1.2 and (2.1.4) have a long history, especially in the holomorphic case. Some important early papers are those by Kohnen and Zagier [84], Shintani [119] and Waldspurger [125]. All of these relied on the fundamental paper of Shimura [118].

Examples

It is interesting to evaluate numerically some examples of Theorem 2.1.2. This is possible thanks to computations done by Strömberg [122]. 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 $r/2 = 6.889875675\ldots$. 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 [14]. The first six values to twelve places are given in Table 2.1.

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 2.2.

Let us illustrate Theorem 2.1.2 in a few cases. Consider first the quadratic field $\mathbb{Q}(\sqrt{3})$, for which $D = 12 = 4 \cdot 3$. There are 2 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 $D = (-3)(-4)$

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

dc_1553_18

Table 2.1: Hecke eigenvalues

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

Table 2.2: Weight 1/2 coefficients

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

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 [14].

2.2 Proofs

2.2.1 Maass forms and the resolvent kernel

Our proof of Theorem 2.1.2 is similar in spirit to that of Hecke's for the Eisenstein series case. We will employ resolvent kernels for the Laplacian 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 [48]. Other references include Hejhal [65] and Roelcke [106].

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

$$F_m(z, s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} f_m(\gamma z, s), \tag{2.2.1}$$

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 [98] and Niebur [99], is a Γ-invariant eigenfunction of Δ:

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

As in Proposition 1.8.1 we will get to the Maass cusp forms through residues of $F_m(z,s)$.

Proposition 2.2.1. *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 (3.2.16).

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 [48],[40])

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

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

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

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 as outlined in Section 1.1.8. By Theorem 3.1 of [48] we have the Fourier expansion[2]

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

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

$$F_{1/2,n}(z,s) = \tfrac{\Gamma(s-\frac{1}{4}\,\mathrm{sign}\, n)}{4\pi|n|\Gamma(2s)} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} J(\gamma,z)^{-1}f_{1/2,n}(\gamma z,s) \tag{2.2.2}$$

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

with

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

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

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

and

$$\operatorname{Res}_{s=\frac{1}{2}+\frac{ir}{2}}(2s-1)F_{1/2,n}(z,s) = \sum_{\psi} \overline{b}(n)\psi(z). \tag{2.2.3}$$

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

$$\psi(z) = \sum_{n\neq 0} b(n)W_{\frac{1}{4}\operatorname{sign} n, \frac{ir}{2}}(4\pi|n|y)e(nx). \tag{2.2.4}$$

2.2.2 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 (1.8.4) the projection operator $\mathrm{pr}^+ : V_r \to V_r^+$ defined by $\mathrm{pr}^+ = \frac{2}{3}WU + \frac{1}{3}$, where[3]

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

We will need an expansion of each of the Fourier coefficients of $\mathrm{pr}^+F_{1/2,m}(z,s)$ 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 [118]) 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 $m, n \in \mathbb{Z}$

$$K_{1/2}(m,n;c) = \sum_{a \pmod c} \left(\tfrac{c}{a}\right)\varepsilon_a e\left(\tfrac{ma+n\overline{a}}{c}\right) \tag{2.2.5}$$

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

$$K^+(m,n;c) = K^+(n,m;c) = \overline{K^+(n,m;c)}. \tag{2.2.6}$$

[3]The constant $\sqrt{2}$ which is not present in [40] is due to the factor $y^{1/4}$ that comes from our normalization.

It follows[4] from [40, Proposition 2 p.959] that for $\mathrm{Re}(s) > 1$ and d any non-zero integer with $d \equiv 0, 1 \pmod 4$ we have

$$\mathrm{pr}^+ F_{1/2,d}(z,s) = \frac{2}{3}\frac{\Gamma(s - \frac{\mathrm{sign}\, d}{4})}{4\pi|d|\Gamma(2s)} M_{\frac{1}{4}\mathrm{sign}\, d, s-\frac{1}{2}}(4\pi|d|y)e(dx) + \sum_{n\equiv 0,1(4)} \Phi^+(n,d;s) W_{\frac{1}{4}\mathrm{sign}\, n, s-\frac{1}{2}}(4\pi|n|y)e(nx) \tag{2.2.7}$$

where for $n \neq 0$ we have

$$\Phi^+(n,d;s) = \frac{1}{|nd|^{\frac{1}{2}}} \frac{\Gamma(s-\frac{\mathrm{sign}\, n}{4})\Gamma(s-\frac{\mathrm{sign}\, d}{4})}{3\sqrt{\pi}\, 2^{2-2s}\, \Gamma(2s-\frac{1}{2})} \sum_{\substack{c\equiv 0(4)\\ c>0}} \frac{K^+(n,d;c)}{c} \begin{cases} I_{2s-1}\Big(\frac{4\pi\sqrt{|nd|}}{c}\Big) & \text{if } nd < 0 \\ J_{2s-1}\Big(\frac{4\pi\sqrt{|nd|}}{c}\Big) & \text{if } nd > 0. \end{cases} \tag{2.2.8}$$

As in [48, 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 (2.2.7) and (1.8.5) the following residue formula.

Theorem 2.2.2. *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 (1.8.2).

2.2.3 Cycle integrals of Poincaré series

We next give an identity from which the extended Katok–Sarnak formula will be derived. Our main source is [40], 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 non-fundamental discriminants.

As further preparation for the proof of Theorem 3.2.3, in this section we will compute the cycle integrals of certain general Poincaré series, which we will then specialize. This will be used both here and in Chapter 5. To begin we need to make some elementary observations about cycle integrals. For $Q \in \mathcal{Q}_d$ with $d > 0$ not a square let S_Q be the oriented semi-circle defined by

$$a|z|^2 + b\,\mathrm{Re}\, z + c = 0, \tag{2.2.9}$$

directed counterclockwise if $a > 0$ and clockwise if $a < 0$. Clearly

$$S_{gQ} = gS_Q, \tag{2.2.10}$$

for any $g \in \Gamma$. Given $z \in S_Q$ let C_Q be the directed arc on S_Q from z to $g_Q z$, where g_Q was defined in (1.11). It can easily be checked that C_Q has the same orientation as S_Q. It is convenient to define

$$dz_Q = \frac{\sqrt{d}\, dz}{Q(z,1)}. \tag{2.2.11}$$

[4]There is a typo in (2.19) of [40]. It should read $P_d^+(z,s) = \frac{3}{2}\mathrm{pr}^+(P_d(z,s))$.

On the geodesic corresponding to Q we have the $dz_Q = \frac{|dz|}{y} = ds$. If $z' = gz$ for some $g \in \Gamma$ we have

$$dz'_{gQ} = dz_Q. \tag{2.2.12}$$

For any Γ–invariant function f on $\mathcal{H}$ the integral $\int_{C_Q} f(z)dz_Q$ is both independent of $z \in S_Q$ and is a class invariant. This is an immediate consequence of the following lemma that expresses this cycle integral as a sum of integrals over arcs in a fixed fundamental domain for Γ. This lemma will also be used in Chapter 5 as well. Let $\mathcal{F}$ be the standard fundamental domain for Γ

$$\mathcal{F} = \{z \in \mathcal{H}; -\tfrac{1}{2} \leq \operatorname{Re} z \leq 0, |z| \geq 1\} \cup \{z \in \mathcal{H}; 0 < \operatorname{Re} z < \tfrac{1}{2}, |z| > 1\}.$$

Lemma 2.2.3. *Let $Q \in \mathcal{Q}_d$ be a form with $d > 0$ not a square and $\mathcal{F}' = g\mathcal{F}$ be the image of $\mathcal{F}$ under any fixed $g \in \Gamma$. Suppose that f is Γ-invariant and continuous on S_Q. Then for any $z \in S_Q$ we have*

$$\int_{C_Q} f(z)dz_Q = \sum_{q \in (Q)} \int_{S_q \cap \mathcal{F}'} f(z)dz_q, \tag{2.2.13}$$

where (Q) denotes the class of Q.

Proof. Let $\tilde{f}(z) = f(z)$ if $z \in \mathcal{F}'$ and $\tilde{f}(z) = 0$ otherwise, so $f(z) = \sum_{g \in \Gamma} \tilde{f}(gz)$ with only a discrete set of exceptions. Thus

$$\int_{C_Q} f(z)dz_Q = \int_{C_Q} \sum_{g \in \Gamma} \tilde{f}(gz)dz_Q = \sum_{g \in \Gamma/\Gamma_Q} \sum_{\sigma \in \Gamma_Q} \int_{C_Q} \tilde{f}(g\sigma z)dz_Q = \sum_{g \in \Gamma/\Gamma_Q} \int_{S_Q} \tilde{f}(gz)dz_Q.$$

Take gz as a new variable. By (2.2.10) and (2.2.12) we get

$$\int_{C_Q} f(z)dz_Q = \sum_{g \in \Gamma/\Gamma_Q} \int_{S_{gQ}} \tilde{f}(z)dz_{gQ},$$

which immediately yields (2.2.13). □

The general Poincaré series are built from a test function $\phi : \mathbb{R}^+ \to \mathbb{C}$ assumed to be smooth and to satisfy $\phi(y) = O_\varepsilon(y^{1+\varepsilon})$, for any $\varepsilon > 0$. For any $m \in \mathbb{Z}$ let

$$G_m(z, \phi) = \sum_{g \in \Gamma_\infty \backslash \Gamma} e\left(m \operatorname{Re} gz\right) \phi(\operatorname{Im} gz). \tag{2.2.14}$$

This sum converges uniformly on compacta and defines a smooth Γ-invariant function on $\mathcal{H}$. We will express its cycle integrals in terms of the sum $S_m(d, d'; c)$ from (1.12.1). Define for $t > 0$ the integral transform.

$$\Phi_m(t) = \int_0^\pi \cos(2\pi m t \cos\theta)\phi(t \sin\theta)\frac{d\theta}{\sin\theta}.$$

For ϕ as above we see that this integral converges absolutely and that $\Phi_m(t) = O_\epsilon(t^{1+\epsilon})$. As we have seen, we may assume without loss that $d, D > 0$.

dc_1553_18

Proposition 2.2.4. *Suppose that $d, d' > 0$ with $D = dd'$ not a square. Then for all $m \in \mathbb{Z}$*

$$\sum_{Q \in \Gamma \backslash \mathcal{Q}_{dd'}} \chi(Q) \int_{C_Q} G_m(z,\phi) dz_Q = \sum_{0<c\equiv 0(4)} S_m(d,d';c) \Phi_m\big(\tfrac{2\sqrt{dd'}}{c}\big).$$

Proof. For each Q, interchanging the sum defining G_m and the integral yields

$$\int_{C_Q} G_m(z,\phi) dz_Q = \sum_{g\in\Gamma_\infty\backslash\Gamma} \int_{C_Q} \mathrm{e}(m \operatorname{Re} gz)\phi(\operatorname{Im} gz) dz_Q. \tag{2.2.15}$$

Now Γ_Q, the group of automorphs of Q, acts freely on $\Gamma_\infty\backslash\Gamma$ so we have that

$$\sum_{g\in\Gamma_\infty\backslash\Gamma} \int_{C_Q} \mathrm{e}(m \operatorname{Re} gz)\phi(\operatorname{Im} gz) dz_Q =$$
$$\sum_{g\in\Gamma_\infty\backslash\Gamma/\Gamma_Q} \sum_{\sigma\in\Gamma_Q} \int_{C_Q} \mathrm{e}(m \operatorname{Re} g\sigma z)\phi(\operatorname{Im} g\sigma z) dz_Q =$$
$$\sum_{g\in\Gamma_\infty\backslash\Gamma/\Gamma_Q} \int_{S_Q} \mathrm{e}(m \operatorname{Re} gz)\phi(\operatorname{Im} gz) dz_Q.$$

Applying (2.2.12) and (2.2.10) in the last expression, we get from (2.2.15) that

$$\int_{C_Q} G_m(z,\phi) dz_Q = \sum_{g\in\Gamma_\infty\backslash\Gamma/\Gamma_Q} \int_{S_{gQ}} \mathrm{e}(m \operatorname{Re} z)\phi(\operatorname{Im} z) dz_{gQ}$$

and hence that

$$\sum_{Q\in\Gamma\backslash\mathcal{Q}_{dD}} \chi(Q) \int_{C_Q} G_m(z,\phi) dz_Q = \sum_{Q\in\Gamma_\infty\backslash\mathcal{Q}_{dD}} \chi(Q) \int_{S_Q} \mathrm{e}(m \operatorname{Re} z)\phi(\operatorname{Im} z) dz_Q.$$

We now need to parameterize the cycle explicitly. Let

$$z_Q = \frac{-b}{2a} + \frac{i\sqrt{d}}{2|a|}, \tag{2.2.16}$$

which is easily seen to be the apex of the circle S_Q. We can parameterize S_Q by $\theta \in (0,\pi)$ via

$$z = \begin{cases} \operatorname{Re} z_Q + e^{i\theta} \operatorname{Im} z_Q & \text{if } a > 0 \\ \operatorname{Re} z_Q - e^{-i\theta} \operatorname{Im} z_Q & \text{if } a < 0. \end{cases}$$

With this parameterization we find that

$$Q(z,1) = \frac{d}{4a} \cdot \begin{cases} e^{2i\theta} - 1 & \text{if } a > 0 \\ e^{-2i\theta} - 1 & \text{if } a < 0 \end{cases}$$

and hence that $dz_Q = d\theta/\sin\theta$. If $\chi(-Q) = -\chi(Q)$ the integrals cancel each other. When $\chi(Q) = \chi(-Q)$ we arrive at the identity

$$\sum_{Q\in\Gamma\backslash\mathcal{Q}_{dD}} \chi(Q) \int_{C_Q} G_m(z,\phi) dz_Q = 2 \sum_{Q\in\Gamma_\infty\backslash\mathcal{Q}_{dD}^+} \chi(Q)\, \mathrm{e}(m \operatorname{Re} z_Q)\Phi_m(\operatorname{Im} z_Q).$$

The proof of Proposition 2.2.4 is thus reduced to the following lemma. □

Lemma 2.2.5. *Let ϕ be as above and suppose that $dd' = D$ is not a square. Then for all $m \in \mathbb{Z}$ we have the identity*

$$\sum_{\Gamma_\infty\backslash\mathcal{Q}_D^+} \chi(Q)\,\mathrm{e}(m\,\mathrm{Re}\,z_Q)\phi(\mathrm{Im}\,z_Q) = \tfrac{1}{2}\sum_{0<c\equiv 0(4)} S_m(d,d';c)\phi\Big(\tfrac{2\sqrt{|D|}}{c}\Big),$$

where z_Q is defined in (2.2.16).

Proof. Under the growth condition on ϕ both series are absolutely convergent, and can be rearranged at will. Consider the left hand side. For $g = \pm\left(\begin{smallmatrix}1 & k\\ 0 & 1\end{smallmatrix}\right) \in \Gamma_\infty$ and $Q = [a,b,c] \in \mathcal{Q}_D$, $gQ = [a, b-2ka, *]$ and so the map

$$[a,b,c] \mapsto (a, b \bmod 2a)$$

is Γ_∞-invariant. Thus

$$\sum_{\Gamma_\infty\backslash\mathcal{Q}_D^+} \chi(Q)e(m\,\mathrm{Re}\,z_Q)\phi(\mathrm{Im}\,z_Q) = \sum_{a=1}^{\infty}\phi\left(\frac{\sqrt{|D|}}{2a}\right)\sum_{b\,(2a)}\chi([a,b,\tfrac{b^2-D}{4a}])\,\mathrm{e}(-\tfrac{mb}{2a}).$$

The sum in b is restricted to those values for which $\frac{b^2-D}{4a}$ is an integer. This happens exactly when $b^2 \equiv D \pmod{4a}$. Thus the inner sum is

$$\sum_{\substack{b\,(2a)\\ b^2\equiv D\,(4a)}} \chi([a,b,\tfrac{b^2-D}{4a}])e(-\tfrac{mb}{2a}) = \tfrac{1}{2}\sum_{\substack{b\,(4a)\\ b^2\equiv D\,(4a)}} \chi([a,b,\tfrac{b^2-D}{4a}])\,\mathrm{e}(-\tfrac{2mb}{4a}) = \tfrac{1}{2}S_m(d,d';4a).$$

Replace $4a$ by c to finish the proof. □

We remark that the positive definite version of Lemma 2.2.4 is following well-known formula for $dd' = D < 0$:

$$\sum_{Q\in\Gamma\backslash\mathcal{Q}_D} w_Q^{-1}\chi(Q)G_m(z,\phi) = \tfrac{1}{2}\sum_{0<c\equiv 0(4)} S_m(d,d';c)\phi\Big(\tfrac{2\sqrt{|dD|}}{c}\Big). \tag{2.2.17}$$

This formula is an immediate consequence of Lemma 2.2.5.

The following is the weight 2 analog of Proposition 2.2.4. Because the proof requires minor modifications of the proof presented above for Proposition 2.2.4 it will be omitted.

Proposition 2.2.6. *Suppose that $d', d < 0$ and that $dd' = D$ is not a square. Then for all $m \in \mathbb{Z}$*

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

where

$$\Psi_m(t) = it\int_0^\pi e(mt\cos\theta)\phi(t\sin\theta)e^{i\theta}d\theta \tag{2.2.18}$$

The following result together with Propositions 1.6.1 and 2.2.2, will be used to derive the extended Katok-Sarnak formula. The first and second parts follow directly from [40], but we include them here for the sake of completeness. Recall that F_m was defined in (1.6.2) and Φ^+ in (2.2.8).

Theorem 2.2.7 ([43]). *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*

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

$$\sum_{Q \in \Gamma\backslash \mathcal{Q}_D} \chi(Q) \begin{cases} 2\sqrt{\pi}\omega_Q^{-1} F_m(z_Q, s) & \text{if } d'd < 0, \\ \int_{\mathcal{C}_Q} F_m(z,s) y^{-1} |dz| & \text{if } d', d > 0, \\ \int_{\mathcal{C}_Q} i\, \partial_z F_m(z,s) dz & \text{if } d', d < 0. \end{cases} \qquad (2.2.19)$$

Lemma 2.2.8. *For $m \neq 0$, $d'd < 0$ and $\mathrm{Re}(s) > 1$ we have*

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

Proof. This follows directly from the above. See also [40, Prop.4 p.970]. □

Similarly we have for the second case the following.

Lemma 2.2.9. *For $m \neq 0$, $d', d > 0$ with $d'd$ not a square and $\mathrm{Re}(s) > 1$ we have*

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

Proof. Note that $\frac{\sqrt{D}}{Q(z,1)} dz = y^{-1}|dz|$ on $\mathcal{C}_Q$. Therefore Proposition 2.2.4 is applicable. The result follows from AppendixA.2 where the associated transform Φ is evaluated for the Poincaré series above. □

The third case requires some new computations.

Lemma 2.2.10. *For $m \neq 0$, $d', d < 0$ with $d'd$ not a square and $\mathrm{Re}(s) > 1$ we have*

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

Proof. Now (1.6.2) and a calculation using differentiation formulas for the Whittaker functions in [93, 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

$$f_{2,m}(z) = -s|m|^{-1/2}(2\pi y)^{-1} \tfrac{\Gamma(s)}{\Gamma(2s)} M_{\mathrm{sgn}(m), s-1/2}(4\pi|m|y) e(mx). \qquad (2.2.20)$$

We can now apply Proposition 2.2.6 for the Poincaré series

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

formed by $f(z) = e(m\,\mathrm{Re}\, z)\phi(\mathrm{Im}\, z)$ where

$$\phi(t) = -s|m|^{-1/2}(2\pi y)^{-1}\tfrac{\Gamma(s)}{\Gamma(2s)}M_{\mathrm{sgn}(m),s-1/2}(4\pi|m|y).$$

Recall that in this case

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

where

$$\Psi_m(t) = it\int_0^\pi e(mt\cos\theta)\phi(t\sin\theta)e^{i\theta}d\theta \tag{2.2.21}$$

The proof of the theorem is now reduced to the following lemma about special functions.

Lemma 2.2.11. *For $\mu\in\mathbb{C}$, $t>0$ and $Re(s)>0$*

$$\int_0^\pi e^{\pm i(tcos\theta+\mu\theta)}M_{\mu,s-1/2}(2t\sin\theta)\frac{d\theta}{\sin\theta} = G(s,\mu)t^{1/2}J_{s-1/2}(t) \tag{2.2.22}$$

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 Appendix A.3. □

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 [83, Prop. 5, p. 259] (see also [35], [75] and [124]).

Lemma 2.2.12. *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*

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

Proceeding as in [40], Proposition 2.2.7 follows from Lemmas 2.2.9,2.2.10 and 2.2.12. □

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 [114] (see also [55]). This might make some of the calculations somewhat simpler but that would not give an exact formula like we obtain in Proposition 2.2.7. 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 [40] for the first two cases of Proposition 2.2.7. The last case can also be applied in this way. It is also possible to prove Theorem 2.1.1 by these methods.

2.2.4 Proof of Theorem 2.1.2

Recall the plus space V_r^+ of Maass cusp forms of weight 1/2 defined in Section 2.2.1 above. It is shown in [77] 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

$$\psi(z) = \sum_{n\neq 0} b(n) W_{\frac{1}{4}\operatorname{sgn} n, \frac{ir}{2}}(4\pi|n|y)e(nx) \tag{2.2.23}$$

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

$$\prod_p (1 - a_\psi(p)p^{-s} + p^{-2s})^{-1} = \sum_{n\geq 1} a_\psi(n)n^{-s} \tag{2.2.24}$$

and let

$$\operatorname{Shim}\psi(z) = y^{1/2}\sum_{n\neq 0} 2a_\psi(|n|)K_{ir}(2\pi|n|y)e(nx). \tag{2.2.25}$$

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

It is convenient to define

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

Theorem 2.1.2 follows easily from the next Proposition.

Proposition 2.2.13. *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 (1.9.2) 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*

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

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

$$12\pi^{\frac{1}{2}}D^{\frac{3}{4}}\sum_{\psi\in B_r} b(d')\overline{b}(d)a_\psi(m) = \sum_\varphi a(m)\mathrm{T}r_{d,d'}(\varphi), \tag{2.2.26}$$

where φ is summed over all Hecke–Maass cusp forms with Laplace eigenvalue $\frac{1}{2}+r^2$. We have from Propositions 1.6.1 and 2.2.7 that for every $m \neq 0$

$$6\pi^{\frac{1}{2}}D^{\frac{3}{4}}|m|\sum_{\substack{n|m\\ n>0}} n^{-\frac{3}{2}}\left(\tfrac{d}{n}\right)\operatorname{Res}_{s=\frac{1}{2}+ir}(2s-1)\Phi^+\left(d', \tfrac{m^2d}{n^2}; \tfrac{s}{2}+\tfrac{1}{4}\right) = \sum_\varphi 2a(m)\mathrm{T}r_{d,d'}(\varphi). \tag{2.2.27}$$

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 2.2.2 gives

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

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 (1.8.2). By (2.2.27) we get

$$24\pi^{\frac{1}{2}}D^{\frac{3}{4}}m\sum_{\psi\in B_r}b(d')\sum_{\substack{n|m\\n>0}}n^{-\frac{3}{2}}\left(\tfrac{d}{n}\right)\overline{b}\left(\tfrac{m^2d}{n^2}\right)=\sum_{\varphi}2a(m)\mathrm{Tr}_{d,d'}(\varphi),$$

and we obtain (2.2.26) 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 (2.2.26) and (1.9.4) that

$$12\pi^{1/2}D^{3/4}\sum_{\psi}b(d')\overline{b}(d)\mathrm{Shim}(\psi)=\sum_{\varphi}\mathrm{Tr}_{d,d'}(\varphi)\varphi. \tag{2.2.28}$$

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 p.129 of Biró in [11], one can conclude that $\mathrm{Shim}(\psi)$ is a weight 0 Maass form with eigenvalue $\frac{1}{2}+r^2$ and by (1.9.3) it is some φ. This leads to

$$12\pi^{1/2}D^{3/4}\sum_{\varphi}\sum_{\mathrm{Shim}(\psi)=\varphi}b(d')\overline{b}(d)\varphi=\sum_{\varphi}\mathrm{Tr}_{d,d'}(\varphi)\varphi.$$

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

$$12\pi^{1/2}D^{3/4}\sum_{\mathrm{Shim}(\psi)=\varphi}b(d')\overline{b}(d)=\mathrm{Tr}_{d,d'}(\varphi).$$

Finally, it is known (see [8, Theorem 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 2.2.13.

□

Remarks. Some of our arguments in the proof of Theorem 2.1.2 are quite similar in spirit to those of Biró in [11], 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 [92], was first used by Siegel to study indefinite quadratic forms. It would be interesting to apply this method to give our extension.

Chapter 3

A new geometric invariant for real quadratic fields

The purpose of this chapter is to give a geometric interpretation and related applications of the new case in the extension of Katok-Sarnak formula.

3.1 Background and statements of results

3.1.1 Fuchsian groups

Suppose that $\Gamma \subset \mathrm{PSL}(2,\mathbb{R})$ is a non-elementary Fuchsian group (see [9] 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 [9, Thm 8.5.2] that $\mathcal{N}_\Gamma$ is the smallest non-empty 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* $(g; m_1, \ldots, m_r; s; t)$ where $m_1, \ldots, m_r$ 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 is freely homotopic in $\Gamma\backslash\mathcal{H}^*$ to a *unique* un-oriented closed geodesic (see e.g. [47, Prop. 1.3]). These geodesics are the image in $\Gamma\backslash\mathcal{H}^*$ of the boundary 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 boundary of the surface (as a subset of $\mathcal{H}$) and not of the fundamental domain as a subset of $\mathcal{H}$. 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 [9, Thm 10.4.3] gives

$$\tfrac{1}{2\pi}\mathrm{area}(\mathcal{F}_\Gamma) = 2(g-1) + s + t + \sum_{j=1}^{r}\left(1 - \tfrac{1}{m_j}\right). \tag{3.1.1}$$

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

$$S = \pm \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \qquad \text{and} \qquad T = \pm \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$$

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

$$\mathcal{F} = \{z \in \mathcal{H}; -1/2 \leq \operatorname{Re} x \leq 0 \text{ and } |z| \geq 1\} \cup \{z \in \mathcal{H}; 0 < \operatorname{Re} x < 1/2 \text{ and } |z| > 1\}.$$

By (3.1.1) or otherwise we have that $\operatorname{area}(\mathcal{F}) = \frac{\pi}{3}$.

An interesting question whether a closed geodesic is the boundary of an immersed surface in $\Gamma\backslash\mathcal{H}$. We will show that this the case, if we are not looking at immersed disks, immersions of a disk from which a point os removed.

Theorem 3.1.1. *There is a Fuchsian group of the second kind with signature*

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

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 will also prove that these immersed surfaces when averaged over genera become equidistributed as the discriminant approaches ∞.

3.2 Proofs

3.2.1 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

$$0 < w'igma < 1 < w$$

(see [61], [132]). 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

$$m = m_A = n_1 + \cdots + n_\ell. \tag{3.2.1}$$

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

$$(\!(n_\ell, \ldots, n_1)\!)\,. \tag{3.2.2}$$

To see this observe that A^{-1} is represented by $(1/w'igma)\mathbb{Z} + \mathbb{Z}$ and by [135, p.128] the continued fraction of $1/w'igma$ has (3.2.2) as its cycle.

3.2.2 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 $((n_1, \ldots, n_\ell))$ 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 $(n_1, \ldots n_\ell)$. For each $k = 1, \ldots \ell$ define the elliptic element of order 2 in Γ:

$$S_k = T^{(n_1+\cdots n_k)} S T^{-(n_1+\cdots n_k)}. \tag{3.2.3}$$

Consider the subgroup of the modular group

$$\Gamma_A = \langle S_1, S_2, \ldots, S_\ell, T^m \rangle = \langle S, S_1, \ldots, S_{\ell-1}, T^m \rangle, \tag{3.2.4}$$

where m was defined in (3.2.1). We will show below in Theorem 4.1.1 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 [64].

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 2 classes: the principal class I with associated cycle $((3, 6))$ and J with cycle $((3, 3, 2, 2, 2))$.

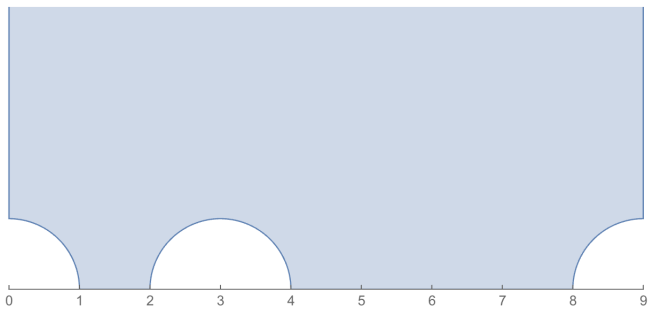

Figure 3.1: Fundamental Domain for Γ_I when $d = 28$.

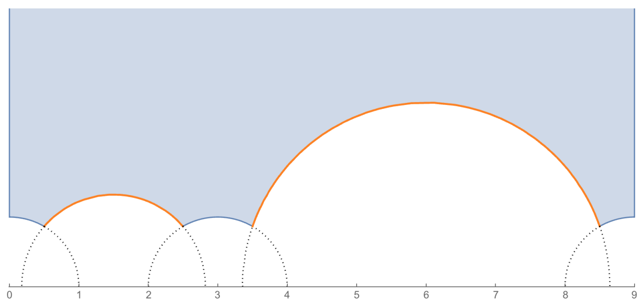

Figure 3.2: The Surface $\mathcal{F}_I$.

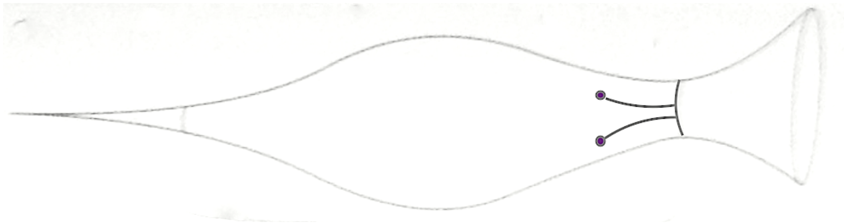

Figure 3.3: The Surface $\mathcal{F}_I$.

The fundamental norm one unit is $\epsilon_{28} = 8 + 3\sqrt{7}$. 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

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

is indicated in Figure 3.1. It has signature $(0; 2, 2; 1, 1)$. The surface $\mathcal{F}_I$ is depicted in Figure 3.2 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 3.3, where the two distinguished points are the points of order 2 and segments connect them to the boundary geodesic.

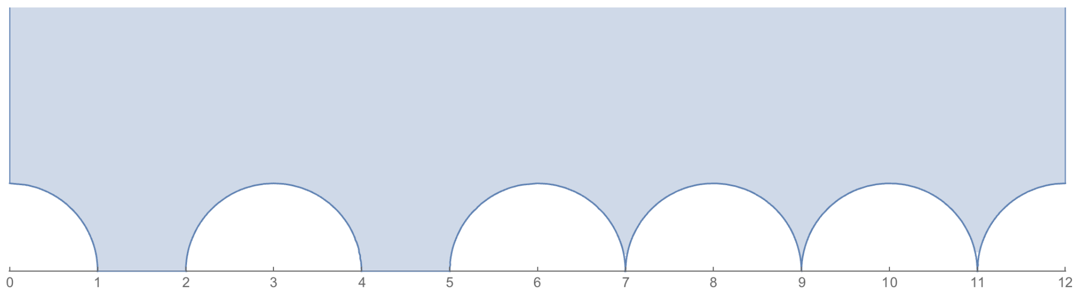

Figure 3.4: Fundamental Domain for Γ_J in case $d = 28$.

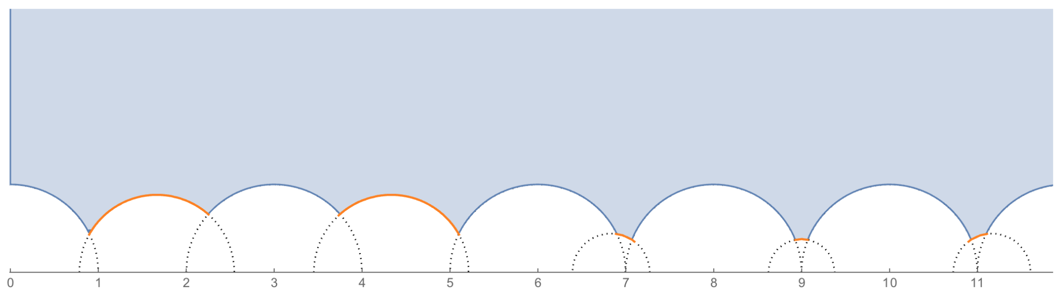

Figure 3.5: The Surface $\mathcal{F}_J$.

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 3.4. It has signature $(0; 2, 2, 2, 2, 2; 1, 1)$. The surface $\mathcal{F}_J$ is pictured in Figure 3.5. 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 3.6.

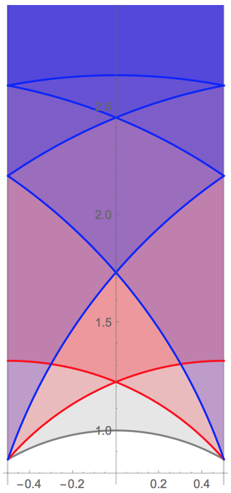

Figure 3.6: Projection of $\mathcal{F}_I$ and $\partial\mathcal{F}_I$ to the modular surface.

Theorem 3.2.1. *The group Γ_A defined in (3.2.4) 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

$$\text{length}(\partial\mathcal{F}_A) = 2\log\epsilon_D \qquad \text{and} \qquad \text{area}(\mathcal{F}_A) = \pi\ell_A. \tag{3.2.5}$$

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

Proof. The first two statements of Theorem 3.2.1 follow easily from an examination of the fundamental domain for

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

constructed like in the examples above. That this construction is valid is an easy consequence of the Poincaré theorem for fundamental polygons [104] (see also [94]). 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 $\Gamma_A\backslash\mathcal{H}^*$ 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 $\mathcal{C}_A$. Using the minus continued fraction of w we have by [76] that for γ_w from (1.11.2)

$$\gamma_w = S_1 S_2 \cdots S_\ell T^m, \tag{3.2.6}$$

where S_k is given in (3.2.3) and m in (3.2.1). In particular,

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

Recall that we have fixed a choice of reduced w for each ideal class A. Consider the point z, the intersection of the unit circle with the geodesic in $\mathcal{H}$ with endpoints $w'igma$ and w, which exists since w is reduced. We have by (3.2.6) that

$$\gamma_w(z) = \gamma(z) = S_1 S_2 \cdots S_\ell T^m(z), \quad \text{so}$$

$$T^{-m} S_\ell \cdots S_2 S_1 \gamma(z) = z. \tag{3.2.7}$$

The circular arc from z to $\gamma(z)$ will intersect the circle with equation $(x-n_1)^2+y^2=1$ at some z^*, say, since by the construction of the backward continued fraction expansion we have that $n_1 - 1 < w < n_1$. The image of the arc from z^* to $\gamma(z)$ under S_1 covers another part of the boundary of $\mathcal{F}_A$. 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 again $n_1+n_2-1 < S_1(w) < n_1+n_2$. Using now S_2 we can map the new excess arc from $S_2(z^{**})$ to $S_2S_1\gamma(z)$. We can repeat this process of cutting off arcs until we have applied S_ℓ. Now observe that by (3.2.7), upon application of T^{-m}, we have returned to z. Since the maps are isometries we see the bounding geodesic arcs piece together to give exactly one copy of $\mathcal{C}_A$, known to have length $2\log\epsilon_D$.

See Figure 3.7 for an illustration of the proof when $w = \frac{3+\sqrt{7}}{2}$ from our first example above. Here $\gamma = T^3\,S\,T^6\,S = \pm\left(\begin{smallmatrix}17 & -3\\ 6 & -1\end{smallmatrix}\right)$ while $z^* = \frac{5+\sqrt{3}i}{2}$, $S_1(z^*) = \frac{7+\sqrt{3}i}{2}$ and $S_1\gamma(z) = \frac{17+\sqrt{3}i}{2}$.

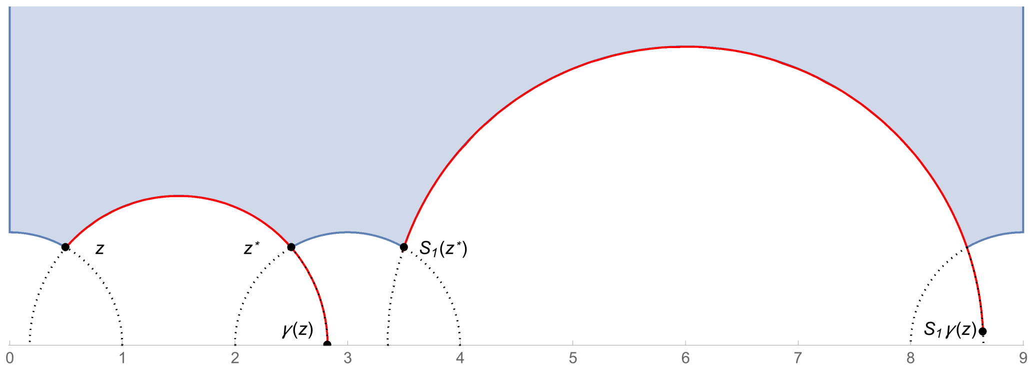

Figure 3.7: Cutting up $\partial\mathcal{F}_A$.

Clearly 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 $\mathcal{F}_A$.

The fact that the area of $\mathcal{F}_A$ is $\pi\ell_A$ is an immediate consequence of (3.1.1).

Finally we must show that the conformal type of $\mathcal{F}_A$ determines A to complete the proof of Theorem 3.2.1. 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 (1.1.2) 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 3.2.1. □

Remark 3.2.2. Although this is not needed for us, note that the shaded region in the figure can be described as the convex polygon U that is the intersection of the strip $0 \le \operatorname{Re} z \le m$ with the closed hyperbolic half-planes that lie above the geodesics whose endpoints are $S_k \dots S_1 w'igma$ and $S_k \dots S_1 w$, and the closed hyperbolic half-planes that lie above the geodesics given by $T^{n_1+\dots+n_k}C$, where C is the unit semi-circle. It is easy to see that U is contained in the closure of Nielsen region of Γ (since the semi-circles with endpoints $S_k \dots S_1 w'igma$ and $S_k \dots S_1 w$ are). The image of U in $\Gamma_A \backslash \mathcal{H}$ is then part of the closure of $\Gamma \backslash \mathcal{N}_A$ with the same boundary, and so must equal to it, (since the latter is path-connected). This shows that the shaded region in the figure is the intersection of the (closure of the) Nielsen region with the (closure of the) fundamental domain for Γ_A.

3.2.3 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

$$\operatorname{Gen}(\mathbb{K}) = \operatorname{Cl}^+(\mathbb{K})/(\operatorname{Cl}^+(\mathbb{K}))^2. \tag{3.2.8}$$

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

$$\#G_D = 2^{1-\omega(D)} h(D), \tag{3.2.9}$$

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

Theorem 3.2.3. *Suppose that for each positive fundamental discriminant $D > 1$ we choose a genus $G_D \in \operatorname{Gen}(\mathbb{K})$. Let Ω be an open disc contained in the fundamental domain $\mathcal{F}$ for $\Gamma = \operatorname{PSL}(2,\mathbb{Z})$ and let $\Gamma\Omega$ be its orbit under the action of Γ. We have*

$$\tfrac{\pi}{3} \sum_{A \in G_D} \operatorname{area}(\mathcal{F}_A \cap \Gamma\Omega) \sim \operatorname{area}(\Omega) \sum_{A \in G_D} \operatorname{area}(\mathcal{F}_A), \tag{3.2.10}$$

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

In view of Theorem 3.2.1, the uniform distribution of closed geodesics proven in [35] (generalized to genera) can be stated in the following form[1]:

$$\tfrac{\pi}{3}\sum_{A\in G_D}\mathrm{length}(\partial\mathcal{F}_A\cap\Gamma\Omega)\sim\mathrm{area}(\Omega)\sum_{A\in G_D}\mathrm{length}(\partial\mathcal{F}_A) \tag{3.2.11}$$

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

The statement of (3.2.11) given in [35] has averaging over the entire class group. Unlike (3.2.11), (3.2.10) 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, (3.2.10) 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. [59, 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 (3.2.10) is non-trivial 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 [28] 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 3.2.3 and (3.2.11) imply that as $p\to\infty$ through such p we have that

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

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 [61] (see also [131]) 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 $\mathrm{area}(\mathcal{F}_A)=\pi\ell_A$, this is equivalent to the area formula

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

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 $z_A\mathbb{Z}+\mathbb{Z}\in A$. Choose for each D a genus G_D, noting that (3.2.8) and (3.2.9) are valid for $D<0$. Then by [35] generalized to genera we have that

$$\tfrac{\pi}{3}\,\#\{A\in G_D\mid z_A\in\Omega\}\sim\mathrm{area}(\Omega)\,\#G_D \tag{3.2.12}$$

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

3.2.4 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 3.2.3. Since it creates no

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

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 (3.2.11) and (3.2.12). 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 non-trivial estimates of Fourier coefficients of modular forms of half-integral weight. Standard references for this section are Hejhal's book [65], the book of Iwaniec [72] and that of Iwaniec and Kowalski [74]. Some other related distribution problems are treated in Sarnak's book [111].

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 [93] and [126].

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:

$$f(z) = c_0 + \frac{1}{4\pi}\int_{-\infty}^{\infty} c(t)E(z, \tfrac{1}{2} + it)dt + \sum_{\varphi} c(\varphi)\langle\varphi, \varphi\rangle^{-1}\varphi(z), \tag{3.2.13}$$

where $\langle\varphi, \varphi\rangle = \int_{\mathcal{F}} |\varphi(x)|^2 d\mu(z)$. Here $E(z, s)$ is the Eisenstein series as in 1.4.1 of weight 0 given for $\mathrm{Re}(s) > 1$ by

$$E(z, s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (\mathrm{Im}\,\gamma z)^s = \tfrac{1}{2}(\mathrm{Im}\,z)^s \sum_{\gcd(c,d)=1} |cz + d|^{-2s}, \tag{3.2.14}$$

where Γ_∞ is the subgroup of Γ generated by T.

The second sum in (1.4.5) 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

$$\lambda = \lambda(\varphi) = \tfrac{1}{4} + r^2 \tag{3.2.15}$$

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[2]

$$\varphi(z) = 2y^{1/2} \sum_{m\neq 0} a(m)K_{ir}(2\pi|m|y)e(mx), \tag{3.2.16}$$

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 an Euler product (for $\mathrm{Re}(s) > 1$):

$$L(s; \varphi) = \sum_{n\geq 1} a(n)n^{-s} = \prod_{p \text{ prime}} (1 - a(p)p^{-s} + p^{-2s})^{-1}. \tag{3.2.17}$$

[2]Note the 2 in front!

Furthermore, its completion

$$\Lambda(s;\varphi) = \pi^{-s}\Gamma(\tfrac{s+ir+\epsilon}{2})\Gamma(\tfrac{s-ir+\epsilon}{2})L(s;\varphi), \tag{3.2.18}$$

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

Remark 3.2.4. 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 $x \to \infty$

$$\#\{\varphi; \lambda(\varphi) \leq x\} \sim \frac{x}{12}. \tag{3.2.19}$$

The first five values of λ to five decimal places (see [14]) are

$$91.14134,\quad 148.43213,\quad 190.13154,\quad 206.41679,\quad 260.68740. \tag{3.2.20}$$

It appears to be likely that each λ is simple but this is open. The eigenvalues in (3.2.20), all belong to odd forms except the third.

For our f the spectral expansion (1.4.5) 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 (1.4.2). 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 1.12 for more information about the genus characters. Given such a χ define

$$\mathrm{Weyl}(u,\chi) = \sum_{A\in \mathrm{Cl}^+(\mathbb{K})} \chi(A) \begin{cases} \frac{\lambda}{2}\int_{\mathcal{F}_A} u(z)d\mu(z) & \text{if } d', d < 0 \\ \int_{\partial\mathcal{F}_A} u(z)y^{-1}|dz| & \text{if } d', d > 0 \\ \frac{1}{\omega_D}\, u(z_A) & \text{if } d'd < 0. \end{cases} \tag{3.2.21}$$

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 non-trivial in the D-aspect and uniform (but weak) in the spectral aspect. This is enough since the Weyl integral in (3.2.21) in the first case is zero when $d', d > 0$ as is that in the second case when $d', d < 0$.

3.2.5 Proof of Theorem 3.2.3

To get the surface case of Theorem 2.1.1 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 3.2.5. *For u as in (3.2.21) we have*

$$\tfrac{\lambda}{2}\int_{\mathcal{F}_A} u(z)d\mu(z) = \int_{\mathcal{C}_A} i\,\partial_z u(z)dz. \tag{3.2.22}$$

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 (3.2.1).

Lemma 3.2.6. *Suppose that $F(z)$ is any real analytic Γ_A-invariant function on $\mathcal{H}$ that satisfies*

$$\Delta F = -y^2(F_{xx} + F_{yy}) = s(1-s)F \tag{3.2.23}$$

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

$$\tfrac{s(1-s)}{2}\int_{\mathcal{F}_A} F(z)d\mu(z) = \int_{\partial\mathcal{F}_A} i\,\partial_z F(z)dz. \tag{3.2.24}$$

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 $dz\,d\overline{z} = -2idx\,dy$, we have by (3.2.23)

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

By our growth assumption on F we get (3.2.24) by letting $Y \to \infty$. □

To deduce Lemma 3.2.5, note that both $E(z,s)$ and $\varphi(z)$ satisfy (3.2.23) 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 (1.4.2). 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 $\mathcal{C}_A$.

We now show how to deduce Theorem 3.2.3 (and (3.2.11) and (3.2.12)) from Theorems 2.1.1 and 2.1.2. 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 (1.4.5). It is a little more complicated to obtain a lower bound for the main term in (3.2.10) 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 (3.2.9) that

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

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 3.2.7. *For any $\epsilon > 0$ we have that*

$$\sum_{A \in G_D} \mathrm{area}(\mathcal{F}_A) \gg_\epsilon D^{1/2-\epsilon}. \tag{3.2.25}$$

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

Proof. We have by Theorem 3.2.1 that

$$\sum_{A \in G_D} \mathrm{area}(\mathcal{F}_A) = \sum_{A \in G_D} \ell_A. \tag{3.2.26}$$

We have the identity (see [135, p.167] or [135, p.138])

$$\prod_{w \text{ reduced}} w = \epsilon_D. \tag{3.2.27}$$

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

$$\sqrt{D} \geq \tfrac{\sqrt{D}}{a} = w - w'igma > w - 1.$$

We conclude that $w < \sqrt{D} + 1$, so (3.2.27) easily implies that

$$\ell_A > \frac{\log \epsilon_D}{\log(\sqrt{D} + 1)}. \tag{3.2.28}$$

Using (3.2.26), (3.2.28), (3.2.9) and Siegel's theorem (see [30]), we derive (3.2.25). $\square$

Remark 3.2.8. It is also possible to give an upper bound for ℓ_A. For example, Eichler [44] gave a general argument that yields for the modular group that

$$\ell_A < c \log \epsilon_D$$

for an explicit c.

We now turn to estimating the Weyl integrals.

Proposition 3.2.9. *There is a constant $C > 0$ such that for any $\epsilon > 0$ we have*

$$\mathrm{Weyl}(E(\cdot, s), \chi) \ll_\epsilon |s|^C |D|^{7/16+\epsilon} \tag{3.2.29}$$

$$\mathrm{Weyl}(\langle \varphi, \varphi \rangle^{-1} \varphi, \chi) \ll_\epsilon r^C |D|^{13/28+\epsilon} \tag{3.2.30}$$

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

Proof. By Theorem 2.1.1 and standard estimates for the gamma function quotient and for $\zeta(2s)$, we have for $\mathrm{Re}(s) = 1/2$ that

$$\mathrm{Weyl}(s, \chi) \ll_\epsilon |s|^C |L(s, \chi_{d'}) L(s, \chi_d)| D^{1/4+\epsilon}. \tag{3.2.31}$$

Thus (3.2.29) now follows from the subconvexity bound of Burgess [23] made uniform in s (see [74, Theorem 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 (3.2.30) of Proposition 3.2.9 follows straight from Theorem 2.1.2 and Theorem 5 of [35]. □

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

$$\int_{\mathcal{F}_{A^{-1}}} \varphi(z) d\mu(z) = - \int_{\mathcal{F}_A} \varphi(z) d\mu(z). \tag{3.2.32}$$

To get (3.2.32) observe that by (3.2.2) 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 3.2.1. The cases $d', d > 0$ and $d'd < 0$ are handled similarly by using (1.11.4).

Theorem 3.2.3 follows from Propositions 3.2.7 and 3.2.9 and the fact that the spectral coefficients in (1.4.5) satisfy

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

for any $A > 0$ and by the Weyl law (3.2.19). See e.g. [72].

Remarks. There has been a lot of progress on subconvexity estimates since the paper [35] that we quote was written. We were content to use the result of [35] here since *any* strong non-trivial 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 [71], techniques for dealing directly with the L-functions were developed in a series of papers starting with [38]. See also [39]. Currently the best known subconvexity bound for the L-functions (2.1.4) was obtained in the breakthrough paper [29] of Conrey and Iwaniec, which gives the exponent $1/3 + \epsilon$ of $|D|$ in both estimates of Proposition 3.2.9 but under the technical assumption that D is odd. This result was improved by Young [130], who gives the same value $1/3 + \epsilon$ for the exponent of C in these estimates. See also the paper of Blomer and Harcos [12]. 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 [45] of Einsiedler, Lindenstrauss, Michel and Venkatesh.

Chapter 4

Mock modular forms and cycle integrals

4.1 Background and statements of results

One of the early appearance of modular forms is the generating series

$$\sum_{n=1}^{\infty} p(n)q^n$$

of the partition function $p(n)$. With a slight modification we see that

$$q^{-1/24}\sum_{n=1}^{\infty} p(n)q^n = q^{-1/24}\prod_{n=1}^{\infty}\frac{1}{1-q^n} = \frac{1}{\eta(z)}.$$

Here $q = e^{2\pi iz}$ and $\eta(z)$ is Dedekind's function, which is a modular form of weight 1/2. An interesting identity follows from considering the Durfee square of the Ferrers diagram of the partition. This is the maximal square that can be fit in the Ferrers diagram.

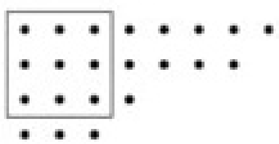

Figure 4.1: The Durfee square of the partition $3 + 4 + 7 + 8 = 22$

If we let $b_m(n)$ be the number of partitions of n with a Durfee square of size $m \times m$, then

$$\sum_{n=1}^{\infty} b_m(n)q^n = \frac{q^{m^2}}{(1-q)^2(1-q^2)^2...(1-q^m)^2}$$

since $\frac{1}{(1-q)(1-q^2)...(1-q^m)}$ is the closed form of the generating series of the number of partitions of n into parts not exceeding m. Putting the above together we get the Eulerian series identity

$$\sum_{n=1}^{\infty} p(n)q^n = 1 + \sum_{m=1}^{\infty}\frac{q^{m^2}}{(1-q)^2(1-q^2)^2...(1-q^m)^2}$$

The modularity of η allowed Ramanujan and Hardy to derive [60], via the circle method, the asymptotic expression

$$p(n) \sim \frac{e^{\pi\sqrt{2n/3}}}{4n\sqrt{3}}. \tag{4.1.1}$$

Mock modular forms, introduced by Ramanujan in his last letter to Hardy, play an important role in extending such considerations to more constrained partitions. Consider for example the q-series

$$f(z) = q^{-1/24} \sum_{n \geq 0} \frac{q^{n^2}}{(1+q)^2 \cdots (1+q^n)^2} \qquad \big(q = e(z) = e^{2\pi i z}, z \in \mathcal{H}\big).$$

Up to the factor $q^{-1/24}$ this is one of Ramanujan's original mock theta functions, and is related to the number of partitions of even ranks. Ramanujan's original definition of what a mock modular forms should be was somewhat vague, his examples are holomorphic functions that have at rational numbers the same asymptotics as (meromorphic) modular forms. This is clearly motivated by the desire to extend the circle method to such restricted partitions.

Ramanujan's ideas, while motivating, were hard to conceptualize, despite work by many great mathematicians, including Watson, Selberg, Andrews and others [3, 113, 127]. This has changed recently, due to the discovery of Zwegers [136, 137] that Ramanujan's mock theta functions of weight 1/2 can be completed to become modular by the addition of a certain non-holomorphic function on the upper half plane $\mathcal{H}$. This complement is associated to a modular form of weight 3/2, the *shadow* of the mock theta function. To illustrate on the above example $f(z) = q^{-1/24} \sum_{n \geq 0} \frac{q^{n^2}}{(1+q)^2 \cdots (1+q^n)^2}$, the shadow of f is the weight 3/2 cusp form (a unary theta series)

$$g(z) = \sum_{n \in 1+6\mathbb{Z}} n\, q^{n^2/24}.$$

The Eichler integral of g is

$$g^*(z) = \sum_{n \in 1+6\mathbb{Z}} \operatorname{sgn}(n)\, \beta\big(\tfrac{n^2 y}{6}\big)\, q^{-n^2/24} \qquad \big(y = \operatorname{Im} z\big).$$

Here $\beta(x)$ is defined for $x > 0$ in terms of the complementary error function and the standard incomplete gamma function by

$$\beta(x) = \operatorname{erfc}(\sqrt{\pi x}) = \frac{1}{\sqrt{\pi}} \Gamma(\tfrac{1}{2}, \pi x), \quad \text{where} \quad \Gamma(s,x) = \int_x^\infty t^s e^{-t}\, \tfrac{dt}{t}. \tag{4.1.2}$$

Observe that the Fourier expansion of the non-holomorphic Eichler integral $g^*(z)$ mirrors that of $g(z)$.

It is proved in [137] that the completion

$$\hat{f}(z) = f(z) + g^*(z)$$

transforms like a modular form of weight 1/2 for $\Gamma(2)$, the well known congruence subgroup of $\Gamma = PSL(2, \mathbb{Z})$.

The appearance of the Eichler integral is best explained via the ξ operator introduced in [20]. Let $f : \mathcal{H} \to \mathbb{C}$ and define

$$\xi_k(f) = 2iy^k \overline{\tfrac{\partial f}{\partial \bar{z}}}.$$

It is easily checked that

$$\xi_k\big((cz+d)^{-k} f(gz)\big) = (cz+d)^{k-2}(\xi_k f)(gz)$$

for any $g = \pm\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{PSL}(2,\mathbb{R})$. Thus if $f(z)$ has weight k for Γ then $\xi_k f$ has weight $2-k$ and $\xi_k f = 0$ if and only if f is holomorphic. The weight k Laplacian, first introduced by Maass, can be conveniently defined by

$$\Delta_k = -\xi_{2-k} \circ \xi_k \tag{4.1.3}$$

If f is a real analytic function on $\mathcal{H}$ of weight k for Γ that is harmonic on $\mathcal{H}$ in the sense that

$$\Delta_k f = 0$$

then f will have a Fourier expansion at $i\infty$ each of whose terms has at most linearly exponential growth. Such an f is called harmonic weak Maass form if it has only finitely many such growing terms. The space of all such forms is denoted by $H_k^!$. It is clear that the space of weakly holomorphic modular forms $M_k^!$ is a subset of $H_k^!$. It follows easily from its general properties that ξ_k maps $H_k^!$ to $M_{2-k}^!$ with kernel $M_k^!$.

In addition to leading to a number of new results about mock theta functions, the work of Zwegers has stimulated the study of other kinds of mock modular forms as well (see [101] and [133] for surveys on some of these developments). For example g^* satisfies

$$\xi_{1/2} g^*(z) = \frac{\sqrt{6}}{3} \sum_{n \in 1+6\mathbb{Z}} q^{n^2/24}$$

and the right hand side of of the above identity is a a cusp form of weight $3/2$, (that above we called, following Zagier, the shadow of f). An explicit determination of the transformation formula allowed Bringmann and Ono to derive an exact formula of Rademacher-type for the coefficients of Ramanujan's mock theta function $f(q)$. This formula refines Ramanujan's first order asymptotic (4.1.1), and had been conjectured by Andrews and Dragonette (see [36] for references). Combined with Rademacher's formula it gives exact formulas for the number of partitions with even (resp. odd) ranks. The work of Zwegers also stimulated numerous other results about mock theta functions. See the surveys [36, 101, 133].

4.1.1 Cycle integrals of the j-function and mock modular forms

In this section we will consider mock modular forms of weight $1/2$ for $\Gamma_0(4)$. In some sense this is the simplest case, but has not been treated before our work in [40] because the associated shadows, if not zero, cannot be cusp forms. We will show that they are nevertheless quite interesting, and have remarkable connections with cycle integrals of the modular j-function and modular integrals having rational period functions. First let us define mock modular forms precisely in this context. Let

$$\theta(z) = 1 + 2\,q + 2\,q^4 + 2\,q^9 + 2\,q^{16} + \cdots$$

be the Jacobi theta series, which is a modular form of weight 1/2 for $\Gamma_0(4)$. Set

$$J(\gamma, z) = \theta(\gamma z)/\theta(z) \quad \text{for } \gamma \in \Gamma_0(4). \tag{4.1.4}$$

For $k \in 1/2 + \mathbb{Z}$ say that f defined on $\mathcal{H}$ has weight k for $\Gamma_0(4)$ (or simply has weight k) if

$$f(\gamma z) = J(\gamma, z)^{2k} f(z) \qquad \text{for all} \qquad \gamma \in \Gamma_0(4).$$

Let $M_k^!$ be the space comprising functions holomorphic on $\mathcal{H}$ of weight k for $\Gamma_0(4)$ whose Fourier coefficients $a(n)$ in the expansion $f(z) = \sum_n a(n) q^n$ are supported on integers $n \geq n_0$, for some $n_0 \in \mathbb{Z}$, with $(-1)^{k-1/2} n \equiv 0, 1 \pmod 4$.

Specializing now to the case of weight 1/2, let $\mathcal{E}(z)$ be the entire function given by any of the following formulas

$$\mathcal{E}(z) = \int_0^1 e^{-\pi z u^2} du = \frac{\operatorname{erf}(\sqrt{\pi z})}{2\sqrt{z}} = \sum_{n=0}^{\infty} \frac{(-\pi z)^n}{(2n+1)n!}. \tag{4.1.5}$$

For any $g(z) = \sum_n b_n \, q^n \in M_{3/2}^!$ we define the non-holomorphic Eichler integral of g by

$$g^*(z) = -4\sqrt{y} \sum_{n \leq 0} b_n \, \mathcal{E}(4ny) \, q^{-n} + \sum_{n>0} \frac{b_n}{\sqrt{n}} \beta(4ny) q^{-n}. \tag{4.1.6}$$

where $\beta(x)$ is as in (4.1.2). Let $f(z) = \sum_n a_n q^n$ be holomorphic on $\mathcal{H}$ and such that its coefficients a_n are supported on integers $n_0 < n$, $n_0 \in \mathbb{Z}$, and also satisfying $n \equiv 0, 1 \pmod 4$. We will say that $f(z)$ is a mock modular form of weight 1/2 for $\Gamma_0(4)$ if there exists a $g \in M_{3/2}^!$, its shadow, so that

$$\hat{f}(z) = f(z) + g^*(z)$$

has weight 1/2 for $\Gamma_0(4)$. Denote by $\mathbb{M}_{1/2}$ the space of all mock modular forms of weight 1/2 for $\Gamma_0(4)$. Obviously $M_{1/2}^! \subset \mathbb{M}_{1/2}$ but it is not at all clear that there are any non-modular mock modular forms.

Nevertheless they exist as I showed with Duke and Imamoglu in [40]. In fact they are closely related to the work of Borcherds and Zagier on traces of singular moduli of the classical j-function

$$j(z) = q^{-1} + 744 + 196884 \, q + \cdots .$$

It is well-known and easily shown that $\mathbb{C}[j]$, has a unique basis $\{j_m\}_{m \geq 0}$ whose members are of the form $j_m(z) = q^{-m} + \emptyset(q)$. For example

$$j_0 = 1, \quad j_1 = j - 744, \quad j_2 = j^2 - 1488j + 159768, \ldots . \tag{4.1.7}$$

Here $j_1(z)$ is the normalized Hauptmodule for Γ. In this paper, unless otherwise specified, d is assumed to be an integer $d \equiv 0, 1(\mod 4)$ and is called a discriminant if $d \neq 0$. For each discriminant D let $\mathcal{Q}_D$ be the set of integral binary quadratic forms of discriminant D that are positive definite if $D < 0$. The forms are acted on as usual by Γ, resulting in finitely many classes $\Gamma \backslash \mathcal{Q}_D$. Let Γ_Q be the group of automorphs of Q (see section 1.11 for more details).

Suppose that $D<0$. For $Q\in\mathcal{Q}_D$ and z_Q a root of Q in $\mathcal{H}$, the numbers $j_1(z_Q)$ are known by the classical theory of complex multiplication to form a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-invariant set of algebraic integers, so that their weighted sum

$$\mathrm{Tr}_D(j_1)=\sum_{Q\in\Gamma\backslash\mathcal{Q}_D}|\Gamma_Q|^{-1}j_1(z_Q) \tag{4.1.8}$$

is an integer. A beautiful theorem of Zagier [134] asserts that these integers give the Fourier coefficients of a weight 3/2 weakly holomorphic form in $\mathbf{T}_-(z)\in M^!_{3/2}$:

$$\begin{aligned}\mathbf{T}_-(z)&=-q^{-1}+2+\sum_{D\le 0}\mathrm{Tr}_D(j_1)|D|q^{|D|} \\ &=-q^{-1}+2-248\,q^3+492\,q^4-4119\,q^7+7256\,q^8+\cdots.\end{aligned} \tag{4.1.9}$$

A natural question is whether one can give a similar statement for the numbers $\mathrm{Tr}_D(j_1)$ defined for non-square $D>0$ by

$$\mathrm{Tr}_D(j_1)=\tfrac{1}{2\pi}\sum_{Q\in\Gamma\backslash\mathcal{Q}_D}\int_{C_Q}j_1(z)\tfrac{dz}{Q(z,1)}. \tag{4.1.10}$$

Here C_Q is any smooth curve from any $z\in\mathcal{H}$ to $g_Q z$, where g_Q is a certain distinguished generator of the infinite cyclic group Γ_Q of automorphs of Q. Note: $\mathrm{Tr}_D(j_1)$ is well-defined.

In [40] we proved that the generating function

$$\mathbf{T}_+(z)=\sum_{d>0}\mathrm{Tr}_d(j_1)q^d \tag{4.1.11}$$

(with a suitable definition of $\mathrm{Tr}_D(j_1)$ when D is a perfect square) defines a mock modular form of weight 1/2 for $\Gamma_0(4)$ with shadow $\mathbf{T}_-(z)$ from (4.1.9).

Theorem 4.1.1 ([40])**.** *The function $\widehat{\mathbf{T}}_+(z)$ on $\mathcal{H}$ defined by*

$$\begin{aligned}\widehat{\mathbf{T}}_+(z)&=\mathbf{T}_+(z)+\mathbf{T}^*_-(z)\\ &=\sum_{D>0}\mathrm{Tr}_D(j_1)\,q^D+4\sqrt{y}\,\mathcal{E}(-4y)\,q-8\sqrt{y}+\sum_{D<0}\frac{\mathrm{Tr}_D(j_1)}{\sqrt{|D|}}\,\beta(4|D|\,y)\,q^D\end{aligned}$$

has weight 1/2 for $\Gamma_0(4)$.

Zagier [134] showed that $g_1(z)=\mathbf{T}_-(z)$ from (4.1.9) is the first member of a basis $\{g_D\}_{0<D\equiv 0,1(4)}$ for $M^!_{3/2}$, where for each $D>0$ the function $g_D(z)$ is uniquely determined by having a q-expansion of the form[1]

$$g_D(z)=-q^{-D}+\sum_{\substack{n\le 0\\ n\equiv 0,1\ (\mathrm{mod}\ 4)}}a(D,n)q^{|n|}. \tag{4.1.12}$$

We define $a(D,n)=0$ unless $d,n\equiv 0,1\pmod 4$.

For $D\le 0$ consider the "dual" form

$$f_D(z)=q^D+\sum_{n>0}a(n,D)q^n. \tag{4.1.13}$$

[1]This is the negative of the $g_D(z)$ defined in [134].

As was shown in [134], the set $\{f_D\}_{D\le 0}$ coincides with the basis given by Borcherds [15] for $M^!_{1/2}$. Thus $f_0(z) = \theta(z)$ and the first few terms of the next function are

$$f_{-3}(z) = q^{-3} - 248q + 26752q^4 - 85995q^5 + 1707264q^8 + \cdots .$$

One of the main results in [40] is that Borcherds' basis extends naturally to a basis for $\mathbb{M}_{1/2}$. The construction of this extension relies heavily on the spectral theory of Maass forms.

Theorem 4.1.2 ([40]). *For each $D > 0$ there is a unique mock modular form $f_D(z) \in \mathbb{M}_{1/2}$ with shadow $g_D(z)$ having a Fourier expansion of the form*

$$f_D(z) = \sum_{n>0} a(n, D)q^n. \tag{4.1.14}$$

These Fourier coefficients $a(n, D)$ satisfy $a(n, D) = a(D, n)$. The set $\{f_D\}_{d\equiv 0,1(\bmod 4)}$ gives a basis for $\mathbb{M}_{1/2}$.

We have thus defined $a(n, d)$ for all d, n with $n > 0$. We use them to evaluate certain twisted traces, which we now define. Suppose that $d > 0$ is a fundamental discriminant. There is a function $\chi_d : \mathcal{Q}_{dd'} \to \{-1, 1\}$ that restricts to a real character (a genus character) on the group of primitive classes and can be used to define a general twisted trace for $dd' = D$ not a square by

$$\mathrm{Tr}_{d,d'}(j_m) = \begin{cases} \frac{1}{\sqrt{D}} \sum \chi_d(Q)|\Gamma_Q|^{-1} j_m(z_Q), & \text{if } dd' < 0\ ; \\ \frac{1}{2\pi} \sum \chi_d(Q) \int_{C_Q} j_m(z) \frac{dz}{Q(z,1)} & \text{if } dd' > 0, \end{cases} \tag{4.1.15}$$

each sum being over $Q \in \Gamma\backslash\mathcal{Q}_D$.

In the paper [40] we established the following evaluation, which generalizes a well-known result of Zagier [134, (25)] to include positive d.

Theorem 4.1.3 ([40]). *Let $a(n, d)$ be the mock modular coefficients defined in (4.1.13) and (4.1.14). Suppose that $m \ge 1$. For $d' \equiv 0, 1 \pmod 4$ and fundamental $d > 0$ with dd' not a square we have*

$$\mathrm{Tr}_{d,d'}(j_m) = \sum_{n|m} \left(\frac{d'}{m/n}\right) n\, a(n^2 d', d). \tag{4.1.16}$$

Together with Theorem 4.1.2, Theorem 4.1.3 implies Theorem 3.2.1 after we define $\mathrm{Tr}_D(j_1)$ to be equal to $a(D, 1)$ when D is a perfect square.

In particular, for non-square dd' with $d > 0$, Theorem 4.1.3 gives

$$a(d, d') = \begin{cases} \frac{1}{\sqrt{d}} \sum \chi(Q)|\Gamma_Q|^{-1} j_1(z_Q), & \text{if } dd' < 0; \\ \frac{1}{2\pi} \sum \chi(Q) \int_{C_Q} j_1(z) \frac{dz}{Q(z,1)}, & \text{if } dd' > 0, \end{cases} \tag{4.1.17}$$

where each sum is over $Q \in \Gamma\backslash\mathcal{Q}_{dD}$.

The proof in [40] that we reproduce below uses Poincaré series and a Kloosterman sum identity that generalizes a well-known result of Salié.

Concerning the case $m = 0$, there exists an interesting "second order" mock modular form $\mathbf{Z}_+(z)$ of weight 1/2 that is almost, but not quite, in $\mathbb{M}_{1/2}$ with Fourier expansion

$$\mathbf{Z}_+(z) = \sum_{d>0} \mathrm{Tr}_d(1)\, q^d. \tag{4.1.18}$$

Here $\mathrm{Tr}_d(1)$ must be defined suitably for square d while for $d > 1$ a fundamental discriminant we have

$$\mathrm{Tr}_d(1) = \pi^{-1} d^{-1/2}\, h(d) \log \epsilon_d,$$

where $h(d)$ is the narrow class number of $\mathbb{Q}(\sqrt{d})$ and ϵ_d is its smallest unit > 1 of norm 1. A (generalized) shadow of $\mathbf{Z}_+(z)$ is the completion of the mock modular form $\mathbf{Z}_-(z)$ of weight 3/2 with shadow $\theta(z)$ discovered by Zagier in 1975 [131] (see also [62]) whose Fourier expansion is

$$\mathbf{Z}_-(z) = \sum_{d \leq 0} \mathrm{Tr}_d(1) q^{|d|}. \tag{4.1.19}$$

Here for any $d \leq 0$ we have that $\mathrm{Tr}_d(1) = H(|d|)$, the usual Hurwitz class number, whose first few values are given by

$$H(0) = -\tfrac{1}{12}, \quad H(3) = \tfrac{1}{3}, \quad H(4) = \tfrac{1}{2}, \quad H(7) = 1, \ \ldots.$$

The completion of $\mathbf{Z}_-(z)$, which has weight 3/2 for $\Gamma_0(4)$ is given by

$$\widehat{\mathbf{Z}}_-(z) = \mathbf{Z}_-(z) + \frac{1}{16\pi} \sum_{n \in \mathbb{Z}} \Gamma(-\tfrac{1}{2}, 4\pi n^2 y) q^{-n^2}. \tag{4.1.20}$$

Define for $y > 0$ the special function

$$\alpha(y) = \frac{\sqrt{y}}{4\pi} \int_0^\infty t^{-1/2} \log(1+t) e^{-\pi y t} dt.$$

The next result shows that $\mathbf{Z}_+(z)$ from (4.1.18) has $\widehat{\mathbf{Z}}_-(z)$ as a generalized shadow.

Theorem 4.1.4 ([40])**.** *The function $\widehat{\mathbf{Z}}_+(z)$ whose Fourier expansion is given by*

$$\widehat{\mathbf{Z}}_+(z) = \sum_{d>0} \mathrm{Tr}_d(1)\, q^d + \frac{\sqrt{y}}{3} + \sum_{d<0} \frac{\mathrm{Tr}_d(1)}{\sqrt{|d|}} \beta(4|d|y) q^d + \sum_{n \neq 0} \alpha(4n^2 y) q^{n^2} - \tfrac{1}{4\pi} \log y, \tag{4.1.21}$$

has weight 1/2 for $\Gamma_0(4)$.

The automorphic nature of $\widehat{\mathbf{Z}}_+(z)$ gives some reason to hope that there might be a connection between the cycle integrals of j and abelian extensions of real quadratic fields.

4.1.2 Modular integrals with rational period functions

The last result we quote form [40] concerns an unexpected connection between mock modular forms of weight 1/2 and modular integrals having rational period functions. Define for each $D \equiv 0, 1(\mod 4)$

$$F_D(z) = -\mathrm{Tr}_D(1) - \sum_{m \geq 1} \Big(\sum_{n|m} n\, a(n^2, D) \Big) q^m. \tag{4.1.22}$$

Note that $F_D(z)$ is the derivative of the formal Shimura lift of f_D. When $D < 0$ Borcherds showed that F_D is a meromorphic modular form of weight 2 for Γ having a simple pole with residue $|\Gamma_Q|^{-1}$ at each point $z_Q \in \mathcal{H}$ of discriminant d. Thus one has corresponding properties of the infinite product

$$q^{-\mathrm{Tr}_D(1)} \prod_{m\geq 1}(1-q^m)^{a(m^2,D)}.$$

In case $D=0$ one finds that this product is $\Delta(z)^{1/12}$, and we have that

$$F_0(z) = \tfrac{1}{12} - 2\sum_{n\geq 1}\sigma(m)q^m = \tfrac{1}{12}E_2(z).$$

This is a holomorphic modular integral of weight 2 with a rational period function:

$$F_0(z) - z^{-2}F_0(-\tfrac{1}{z}) = -\tfrac{1}{2\pi i}\, z^{-1}.$$

Another important result of [40] is the following

Theorem 4.1.5 ([40]). *For each $D>0$ not a square the function F_D defined in (4.1.22) is a holomorphic modular integral of weight 2 with a rational period function:*

$$F_D(z) - z^{-2}F_D(-\tfrac{1}{z}) = \frac{1}{\pi}\sum_{\substack{c<0<a\\ b^2-4ac=D}}(az^2+bz+c)^{-1}. \tag{4.1.23}$$

The Fourier expansion of $F_D(z)$ can be expressed in the form

$$F_d(z) = -\sum_{m\geq 0}\mathrm{Tr}_D(j_m)\, q^m.$$

Note that the period function has simple poles at certain real quadratic integers of discriminant D, in analogy to the behavior of $F_D(z)$ when $D<0$. The existence of a holomorphic F satisfying (4.1.23) with growth conditions was proved by Knopp [85], [86]. He used a certain Poincaré series built out of cocycles, however, it appears to be very difficult to extract explicit information about F from this construction. At the end of their paper [25], Choie and Zagier raised the problem of explicit construction of a modular integral with a given rational period function. Parson [102] gave a more direct construction in weights $k>2$ using series of the form

$$\sum_{a>0}(az^2+bz+c)^{-k/2},$$

which are partial versions of certain hyperbolic Poincaré series studied by Zagier, but they do not converge when $k=2$. In any case, the expression of the Fourier coefficients as sums of cycle integrals is not immediate from this construction, although it is possible to deduce such expressions this way, at least in higher weights, using methods from [40].

For the rational period functions that occur in (4.1.23) the modular integral given by $F_D(z)$ also gives a real quadratic analogue of (the logarithmic derivative of) the Borcherds product. Also note that the function F_D is closely related to the modular integrals considered in the next chapter. The underlying technical achievement of that chapter is the general transformation formula for these functions under an arbitrary modular substitution just as in the case of the Dedekind η function, Dedekind's main result is the determination of the transformation formula of $\log\eta$ under all modular substitutions, and not just inversion.

4.2 Proofs

4.2.1 Weakly harmonic modular forms

We begin by proving Theorem 4.1.2 using the theory of weakly harmonic forms. Set $k \in 1/2 + \mathbb{Z}$. If f of weight k for $\Gamma_0(4)$ is smooth, for example, it will have a Fourier expansion in each cusp. For the cusp at $i\infty$ we have the Fourier expansion

$$f(z) = \sum_n a(n; y)e(nx) \tag{4.2.1}$$

which, if f is holomorphic, has $a(n; y) = a(n)e(niy)$. Set

$$f^{\mathrm{e}}(z) = \sum_{n\equiv 0(2)} a(n; \tfrac{y}{4})e(\tfrac{nx}{4}) \quad \text{and} \quad f^{\mathrm{o}}(z) = \sum_{n\equiv 1(2)} a(n; \tfrac{y}{4})e(\tfrac{n}{8})e(\tfrac{nx}{4}). \tag{4.2.2}$$

Suppose that the Fourier coefficients $a(n; y)$ satisfy the *plus space condition*, meaning that they vanish unless $(-1)^{k-1/2}n \equiv 0, 1 \pmod 4$. An easy extension of arguments given in [84, p.190] shows that such an f satisfies

$$(\tfrac{2z}{i})^{-k} f(-\tfrac{1}{4z}) = \alpha f^{\mathrm{e}}(z) \qquad \text{and} \qquad (\tfrac{2z+1}{i})^{-k} f(\tfrac{z}{2z+1}) = \alpha f^{\mathrm{o}}(z) \tag{4.2.3}$$

where

$$\alpha = (-1)^{\lfloor \frac{2k+1}{4} \rfloor} 2^{-k+\frac{1}{2}}.$$

In particular, the behavior of such an f at the cusps 0 and 1/2 is determined by that at $i\infty$. Thus to check that such a form is weakly holomorphic, meaning it is holomorphic on $\mathcal{H}$ and meromorphic in the cusps, one only needs look at the Fourier expansion at $i\infty$, as we have done in the Introduction. As there, let $M_k^!$ denote the space of all such forms. Let $M_k^+ \subset M_k^!$ denote the subspace of holomorphic forms (having no pole in the cusps) and $S_k^+ \subset M_k^+$ the subspace of cusp forms (having zeros there).

Consider the Maass-type differential operator ξ_k defined for any $k \in \mathbb{R}$ through its action on a differentiable function f on $\mathcal{H}$ by

$$\xi_k(f) = 2iy^k \overline{\tfrac{\partial f}{\partial \bar{z}}}.$$

This operator is studied in detail in [20]. It is easily checked that

$$\xi_k\big((\gamma z + \delta)^{-k} f(gz)\big) = (\gamma z + \delta)^{k-2} (\xi_k f)(gz)$$

for any $g \in \mathrm{PSL}(2, \mathbb{R})$. Thus if $f(z)$ has weight k for $\Gamma_0(4)$ then $\xi_k f$ has weight $2-k$ and $\xi_k f = 0$ if and only if f is holomorphic. Also ξ_k preserves the plus space condition. The weight k Laplacian can be conveniently defined by

$$\Delta_k = -\xi_{2-k} \circ \xi_k.$$

Specializing now to $k = 1/2$, suppose that h is a real analytic function on $\mathcal{H}$ of weight 1/2 for $\Gamma_0(4)$ that is harmonic on $\mathcal{H}$ in the sense that

$$\Delta_{1/2} h = 0. \tag{4.2.4}$$

By separation of variables every such h has a (unique) Fourier expansion in the cusp at ∞ of the form

$$h(z) = \sum_n b(n)\mathcal{M}_n(y)e(nx) + \sum_n a(n)\,\mathcal{W}_n(y)e(nx). \tag{4.2.5}$$

The functions $\mathcal{W}_n(y)$ and $\mathcal{M}_n(y)$ in the Fourier expansion (4.2.5) are defined in terms of the functions $\beta(x)$ and $\mathcal{E}(z)$ from (4.1.2) and (4.1.5) by

$$\mathcal{W}_n(y) = e^{-2\pi ny}\begin{cases}|n|^{-\frac{1}{2}}\,\beta(4|n|y) & \text{if } n<0,\\ -4y^{\frac{1}{2}} & \text{if } n=0,\\ n^{-\frac{1}{2}} & \text{if } n>0,\end{cases} \tag{4.2.6}$$

$$\mathcal{M}_n(y) = e^{-2\pi ny}\begin{cases}1-\beta(4|n|\,y) & \text{if } n<0,\\ 1 & \text{if } n=0,\\ 4(ny)^{\frac{1}{2}}\,\mathcal{E}(-4ny) & \text{if } n>0.\end{cases} \tag{4.2.7}$$

We remark that $\mathcal{W}_n(y)$ and $\mathcal{M}_n(y)$ are special cases of Whittaker functions, (see (4.2.15)) and we use the notation $\mathcal{W}$ and $\mathcal{M}$ to suggest this relation. More importantly, the definitions (4.2.6) and (4.2.7) make possible the complete symmetry of the Fourier coefficients of the basis to be given in the next Proposition. It becomes clear after working with them that one can define the normalization for the Fourier coefficients in different reasonable ways, each with advantages and disadvantages. Note that the function $\mathcal{W}_n(y)$ is exponentially decaying while $\mathcal{M}_n(y)$ is exponentially growing in y (see (A.1.4)).

Let $H^!_{1/2}$ denote the space of all real analytic functions on $\mathcal{H}$ of weight 1/2 for $\Gamma_0(4)$ that satisfy (4.2.4), whose Fourier coefficients at ∞ are supported on integers n with $n \equiv 0, 1 (\mathrm{mod} 4)$ and that have only finitely many non-zero coefficients $b(n)$. As before this is enough to control bad behavior in the other cusps. We will call such an $h \in H^!_{1/2}$ weakly harmonic.[2] This space was identified by Bruinier and Funke [19] as being interesting arithmetically. It follows easily from its general properties that $\xi_{1/2}$ maps $H^!_{1/2}$ to $M^!_{3/2}$ with kernel $M^!_{1/2}$. This is also directly visible after a calculation from (4.2.7) and (4.2.6) yields the formulas

$$\xi_{1/2}\big(\mathcal{M}_n(y)e(nx)\big) = 2|n|^{\frac{1}{2}}q^{-n} \qquad \xi_{1/2}\big(\mathcal{W}_n(y)e(nx)\big) = \begin{cases}0 & \text{if } n>0\\ -2q^{|n|} & \text{if } n\le 0.\end{cases} \tag{4.2.8}$$

Given h in (4.2.5) with $b(n) = 0$ for all n, we infer that $\xi_{1/2}h \in M^+_{3/2} = \{0\}$. This implies that $h \in S^+_{1/2} = \{0\}$, and proves the following uniqueness result.

Lemma 4.2.1. *If $h \in H^!_{1/2}$ has Fourier expansion*

$$h(z) = \sum_n b(n)\mathcal{M}_n(y)e(nx) + \sum_n a(n)\,\mathcal{W}_n(y)e(nx), \tag{4.2.9}$$

then $h = 0$ if and only if $b(n) = 0$ for all $n \equiv 0, 1 \pmod 4$.

[2] The definition of harmonic weak Maass forms, for example as given in [21] and elsewhere, is more restrictive and does not apply to the non-holomorphic $h \in H^!_{1/2}$, so we use the terminology *weakly harmonic* to avoid confusion.

It is now easy to explain the relation between mock modular forms and weakly harmonic ones (c.f. [133]). It follows easily from (4.2.6), (4.2.7) and (4.2.8), or directly, that for $g(z) \in M^!_{3/2}$

$$\xi_{1/2}\, g^*(z) = -2\, g(z),$$

where $g^*(z)$ was defined in (4.1.6). As a consequence we see that if $f \in \mathbb{M}_{1/2}$ and if $\hat{f} = f + g^*$ is its completion, then $\hat{f} \in H^!_{1/2}$ since $\xi_{1/2}\ \hat{f}(z) = -2g(z)$ so $\Delta_{1/2}\hat{f} = 0$. Also $\hat{f}(z)$ satisfies the plus space condition. In fact it is easy to see that $f \mapsto \hat{f}$ defines an isomorphism from $\mathbb{M}_{1/2}$ to $H^!_{1/2}$. Given $h \in H^!_{1/2}$ let $g = -\frac{1}{2}\,\xi_{1/2}(h) \in M^!_{3/2}$ and take $h^+ = h - g^*$. It is easily checked that $h \mapsto h^+$ gives the inverse of $f \mapsto \hat{f}$. Call h^+ the *holomorphic part* of h. In terms of the Fourier expansion (4.2.5)

$$h^+(z) = \sum_{n\le 0} b(n)q^n + \sum_{n>0} a(n)n^{-1/2}q^n. \tag{4.2.10}$$

The next result gives one natural basis for $H^!_{1/2}$.

Proposition 4.2.2. *For each $d \equiv 0, 1 \pmod 4$ there is a unique $h_d \in H^!_{1/2}$ with Fourier expansion of the form*

$$h_d(z) = \mathcal{M}_d(y)e(dx) + \sum_{n\equiv 0,1(4)} a_d(n)\mathcal{W}_n(y)e(nx). \tag{4.2.11}$$

The set $\{h_d\}_{d\equiv 0,1\,(4)}$ forms a basis for $H^!_{1/2}$. The coefficients $a_d(n)$ satisfy the symmetry relation

$$a_d(n) = a_n(d) \tag{4.2.12}$$

for all integers $n, d \equiv 0, 1 \pmod 4$. When $d > 0$ we have

$$\xi_{1/2}\, h_d(z) = -2d^{\frac{1}{2}}\, g_d(z), \tag{4.2.13}$$

where $g_d \in M^!_{3/2}$ has Fourier expansion given in (4.1.12).

Theorem 4.1.2 is an immediate consequence of this proposition. We see that for $d \le 0$ we have that $h_d = f_d$ from (4.1.13) and $a(n, d) = n^{-1/2}a_d(n)$ unless $n = d < 0$, in which case $a_d(d) = |d|^{1/2}$. If $d > 0$ let $f_d(z) = \sum_{n>0} a(n, d)q^n$ be the holomorphic part of $d^{-1/2}\, h_d$. This gives the $f_d(z)$ from Theorem 4.1.2 and we find that for $n > 0$ we have

$$a(n, d) = (dn)^{-1/2}a_d(n). \tag{4.2.14}$$

We remark that the fact we quoted from [134] that $\{f_d\}_{d\le 0}$ from (4.1.13) gives the Borcherds basis for $M^!_{1/2}$ also follows from the symmetry relation (4.2.12) and (4.2.13) of Proposition 4.2.2.

We now turn to the construction of h_d. We will give a uniform construction using Poincaré series. Due to some delicate convergence issues that arise from this approach, we will define them through analytic continuation. For fixed s with $\mathrm{Re}(s) > 1/2$ and $n \in \mathbb{Z}$ let

$$\mathcal{M}_n(y, s) = \begin{cases} \Gamma(2s)^{-1}(4\pi|n|y)^{-\frac{1}{4}} M_{\frac{1}{4}\operatorname{sgn} n, s-\frac{1}{2}}(4\pi|n|y) & \text{if } n \neq 0, \\ y^{s-\frac{1}{4}} & \text{if } n = 0 \end{cases}$$

$$\mathcal{W}_n(y, s) = \begin{cases} |n|^{-\frac{3}{4}}\Gamma(s + \frac{\operatorname{sgn} n}{4})^{-1}(4\pi y)^{-\frac{1}{4}} W_{\frac{1}{4}\operatorname{sgn} n, s-\frac{1}{2}}(4\pi|n|y), & \text{if } n \neq 0 \\ \frac{2^{2s-\frac{1}{2}}}{(2s-1)\Gamma(2s-1/2)}\, y^{\frac{3}{4}-s}, & \text{if } n = 0, \end{cases}$$

where M and W are the usual Whittaker functions (see Appendix A.1). By (A.1.6) and (A.1.7), for $n \neq 0$ we have that

$$\mathcal{M}_n(y) = \mathcal{M}_n(y, 3/4) \text{ and } \mathcal{W}_n(y) = \mathcal{W}_n(y, 3/4) \tag{4.2.15}$$

where $\mathcal{M}_n(y)$ and $\mathcal{W}_n(y)$ were given in (4.2.7) and (4.2.6). However, $\mathcal{M}_0(y) = \mathcal{W}_0(y, 3/4)$ and $\mathcal{W}_0(y) = \mathcal{M}_0(y, 3/4)$.[3] We also need the usual I and J-Bessel functions, defined for fixed ν and $y > 0$ by (see e.g. [89])

$$I_\nu(y) = \sum_{k=0}^{\infty} \frac{(y/2)^{\nu+2k}}{k!\,\Gamma(\nu+k+1)} \qquad \text{and} \qquad J_\nu(y) = \sum_{k=0}^{\infty} \frac{(-1)^k (y/2)^{\nu+2k}}{k!\,\Gamma(\nu+k+1)}. \tag{4.2.16}$$

For $m \in \mathbb{Z}$ let

$$\psi_m(z, s) = \mathcal{M}_m(y, s) e(mx). \tag{4.2.17}$$

It follows from (A.1.3) and (4.1.3) that

$$\Delta_{1/2}\psi_m(z, s) = (s - \tfrac{1}{4})(\tfrac{3}{4} - s)\psi_m(z, s).$$

Define the Poincaré series

$$P_m(z, s) = \sum_{g \in \Gamma_\infty \backslash \Gamma_0(4)} j(g, z)^{-1} \psi_m(gz, s),$$

where Γ_∞ is the subgroup of translations. By (A.1.5) this series converges absolutely and uniformly on compacta for $\operatorname{Re} s > 1$. The function $P_0(z, s)$ is the usual weight $1/2$ Eisenstein series. It is clear that for $\operatorname{Re}(s) > 1$ and any m the function $P_m(z, s)$ has weight $1/2$ and that P_m satisfies

$$\Delta_{1/2} P_m(z, s) = (s - \tfrac{1}{4})(\tfrac{3}{4} - s) P_m(z, s).$$

As in [83], in order to get forms whose Fourier expansions are supported on $n \equiv 0, 1 \pmod 4$ we will employ the projection operator $\mathrm{pr}^+ = \frac{2}{3}(U_4 \circ W_4) + \frac{1}{3}$, where

$$(U_4 f)(z) = \tfrac{1}{4} \sum_{\nu \bmod 4} f\left(\tfrac{z+\nu}{4}\right) \qquad \text{and} \qquad (W_4 f)(z) = \left(\tfrac{2z}{i}\right)^{-1/2} f(-1/4z).$$

For each $d \equiv 0, 1 \pmod 4$ and $\operatorname{Re}(s) > 1$ define

$$P_d^+(z, s) = \mathrm{pr}^+(P_d(z, s)). \tag{4.2.18}$$

Proposition 4.2.3. *For any $d \equiv 0, 1 \pmod 4$ and $\operatorname{Re}(s) > 1$ the function $P_d^+(z, s)$ has weight 1/2 and satisfies*

$$\Delta_{1/2} P_d^+(z, s) = (s - \tfrac{1}{4})(\tfrac{3}{4} - s) P_d^+(z, s).$$

[3] This notational switching is inessential but gives a cleaner statement of Proposition 4.2.2 and some other results.

dc_1553_18

Its Fourier expansion is given by

$$P_d^+(z,s) = \mathcal{M}_d(y,s)e(dx) + \sum_{n\equiv 0,1(4)} b_d(n,s)\mathcal{W}_n(y,s)e(nx), \quad \textit{where} \tag{4.2.19}$$

$$b_d(n,s) = \sum_{0<c\equiv 0(4)} K^+(d,n;c) \times \begin{cases} 2^{\frac{1}{2}}\pi|dn|^{\frac{1}{4}}c^{-1}I_{2s-1}\left(\frac{4\pi\sqrt{|dn|}}{c}\right) & \textit{if } dn<0; \\ 2^{\frac{1}{2}}\pi|dn|^{\frac{1}{4}}c^{-1}J_{2s-1}\left(\frac{4\pi\sqrt{|dn|}}{c}\right) & \textit{if } dn>0, \\ \pi^{s+\frac{1}{4}}|d+n|^{s-\frac{1}{4}}c^{-2s} & \textit{if } dn=0,\ d+n\neq 0, \\ 2^{\frac{1}{2}-2s}\pi^{\frac{1}{2}}\Gamma(2s)c^{-2s} & \textit{if } d=n=0, \end{cases} \tag{4.2.20}$$

where $K^+(d,n;c)$ is the modified Kloosterman sum defined in (1.12.3) below. The sum defining each $b_d(n,s)$ is absolutely convergent.

Proof. The first statement is clear. So is the last statement using the trivial bound for $K^+(d,n;c)$ and the definitions (4.2.16).

For the calculation of the Fourier expansion we employ the following lemma, whose proof is standard and follows from an application of Poisson summation using an integral formula found in [48, p. 176]. See [81, Lemma 2, p. 20] or [83] for the prototype result.

Lemma 4.2.4. *Let $\left[\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right] \in SL(2,\mathbb{R})$ have $c>0$ and suppose that* $\mathrm{Re}(s)>1/2$. *Then for ψ_m defined in (4.2.17) with any $m\in\mathbb{Z}$, we have*

$$\sum_{r\in Z}(c(z+r)+d)^{-1/2}\psi_m\left(\frac{a(z+r)+b}{c(z+r)+d}, s\right) = 2\pi i^{-1/2}\sum_{n\in\mathbb{Z}} e\left(\tfrac{am+nd}{c}\right)\mathcal{W}_n(y,s)e(nx)$$

$$\times \begin{cases} c^{-1}|mn|^{\frac{1}{4}}J_{2s-1}(4\pi\sqrt{|mn|}c^{-1}) & \textit{if } mn>0 \\ c^{-1}|mn|^{\frac{1}{4}}I_{2s-1}(4\pi\sqrt{|mn|}c^{-1}) & \textit{if } mn<0 \\ 2^{-\frac{1}{2}}\pi^{s-\frac{3}{4}}c^{-2s}|m+n|^{s-\frac{1}{4}} & \textit{if } mn=0,\ m+n\neq 0, \\ \pi^{-\frac{1}{2}}(2c)^{-2s}\Gamma(2s) & \textit{if } m=n=0, \end{cases}$$

where both sides of the identity converge uniformly on compact subsets of $\mathcal{H}$.

With this Lemma, the computation of the Fourier coefficients parallels so closely that given in [81, pp. 18–27] in the holomorphic case that we will omit the details. □

It is a well-known consequence of the theory of the resolvent kernel that $P_d(z,s)$ has an analytic continuation in s to $\mathrm{Re}(s)>1/2$ except for possibly finitely many simple poles in $(1/2,1)$. These poles may only occur at points of the discrete spectrum of $\Delta_{1/2}$ on the Hilbert space consisting of weight $1/2$ functions f on $\mathcal{H}$ that satisfy

$$\int_{\Gamma\backslash\mathcal{H}} |f(z)|^2 y\, d\mu < \infty \qquad (d\mu = y^{-2}dx\, dy),$$

and this space contains the residues.[4] It is easily seen from (4.2.18) that $P_d^+(z,s)$ also has an analytic continuation to $\mathrm{Re}(s)>1/2$ with at most finitely many simple poles in $(1/2,1)$.

[4] See [48, p.179] and its references, especially [106] and [46]. A very clear treatment when the weight is 0 and the multiplier is trivial is given in [98]. In particular, see Satz 6.8 p.60; the case of weight 1/2 is similar.

Actually, such poles can only occur in $(\frac{1}{2}, \frac{3}{4}]$, since by (1.12.6) and (4.2.16) the series in (4.2.20) giving the Fourier coefficient $b_d(n, s)$ converges absolutely for $\mathrm{Re}(s) > 3/4$. Thus for $\mathrm{Re}(s) > 1/2$ away from these poles the function $P_d^+(z, s)$ has weight 1/2 and satisfies

$$\Delta_{1/2} P_d^+(z, s) = (s - \tfrac{1}{4})(\tfrac{3}{4} - s) P_d^+(z, s).$$

Furthermore a residue at $s = 3/4$ is a weight 1/2 weakly harmonic form $f \in H_{1/2}^!$. In fact, the Fourier expansion of f can be obtained from that of P_d^+ in (4.2.19) by taking residues term by term, a process that is easily justified using the integral representations for the Fourier coefficients since the convergence is uniform on compacta. This shows that the Fourier expansion of f is supported on n with $n \equiv 0, 1 \pmod 4$ and that it can have no exponentially growing terms. Another way to see these facts is to observe that f is the projection of the residue of P_d, which comes from the discrete spectrum. Thus by Lemma 4.2.1 applied to $f - b(0)\theta$, we obtain the following result.

Lemma 4.2.5. *For each d and each $z \in \mathcal{H}$ the function $P_d^+(z, s)$ has an analytic continuation around $s = 3/4$ with at most a simple pole there with residue*

$$\mathrm{res}_{s=3/4} P_d^+(z, s) = \rho_d \, \theta(z), \tag{4.2.21}$$

where $\rho_d \in \mathbb{C}$.

When $d = 0$ this result is well-known. In fact, $b_0(n, s)$ can be computed in terms of Dirichlet L-functions. We have the following formulas (see e.g. [69]).

Lemma 4.2.6. *For $m \in \mathbb{Z}^+$ and D a fundamental discriminant we have that*

$$\sum_{n|m} \left(\tfrac{D}{n}\right) b_0\left(\tfrac{Dm^2}{n^2}, s\right) = 2^{2-4s} \pi^{s+\frac{1}{4}} m^{\frac{3}{2}-2s} |D|^{s-\frac{1}{4}} \sigma_{4s-2}(m) \frac{L_D(2s - \frac{1}{2})}{\zeta(4s-1)} \quad \textit{and} \tag{4.2.22}$$

$$b_0(0, s) = \pi^{\frac{1}{2}} 2^{\frac{5}{2}-6s} \Gamma(2s) \frac{\zeta(4s-2)}{\zeta(4s-1)}, \tag{4.2.23}$$

where L_D is the Dirichlet L-function.

By Möbius inversion, (4.2.22) gives for $m \neq 0$ the identity

$$b_0(Dm^2, s) = 2^{2-4s} \pi^{s+\frac{1}{4}} |D|^{s-\frac{1}{4}} \frac{L_D(2s - 1/2)}{\zeta(4s-1)} \sum_{n|m} \mu(m/n) \left(\tfrac{D}{m/n}\right) n^{\frac{3}{2}-2s} \sigma_{4s-2}(n). \tag{4.2.24}$$

This yields a direct proof of Lemma 4.2.5 in case $d = 0$. Since $b_0(d, s) = b_d(0, s)$, which is clear from (1.12.4) and (4.2.20), a calculation using (4.2.24) and the (4.2.23) also gives the constant ρ_d in (4.2.21):

$$\rho_d = \begin{cases} \frac{3}{4\pi} & \text{if } d = 0, \\ \frac{6}{\pi}\sqrt{d} & \text{if } d \text{ is a non-zero square}, \\ 0 & \text{otherwise}. \end{cases} \tag{4.2.25}$$

We are finally ready to define the basis functions h_d. For $d \neq 0$ let

$$h_d(z, s) = P_d^+(z, s) - \frac{b_d(0, s)}{b_0(0, s)} P_0^+(z, s). \tag{4.2.26}$$

It has the Fourier expansion

$$h_d(z,s) = \mathcal{M}_d(y,s)e(dx) - \frac{b_d(0,s)}{b_0(0,s)}y^{s-\frac{1}{4}} + \sum_{0\neq n\equiv 0,1(4)} a_d(n,s)\mathcal{W}_n(y,s)e(nx), \quad \text{where} \tag{4.2.27}$$

$$a_d(n,s) = b_d(n,s) - \frac{b_d(0,s)b_0(n,s)}{b_0(0,s)}. \tag{4.2.28}$$

Lemma 4.2.7. *For each nonzero $d \equiv 0,1 \pmod 4$ the function $h_d(z,s)$ defined in (4.2.26) has an analytic continuation to $s = 3/4$ and*

$$h_d(z,3/4) = h_d(z) \in H^!_{1/2}.$$

The Fourier expansion of each such h_d at ∞ has the form (4.2.11), where for each nonzero $n \equiv 0,1 \pmod 4$ we have

$$a_d(n) = \lim_{s\to 3/4^+} a_d(n,s).$$

Furthermore, $a_d(0) = 2\sqrt{d}$ if d is a square and $a_d(0) = 0$ otherwise.

Proof. Observe that $h_d(z,s)$ defined in (4.2.26) is holomorphic at $s = 3/4$, since otherwise by Proposition 4.2.5 it would have as residue there a nonzero multiple of $\theta(z)$, which cannot happen since (4.2.27) does not yield the constant term in θ. From (4.2.27) its Fourier expansion is given by

$$h_d(z) = \mathcal{M}_d(y)e(dx) + \sum_{n\equiv 0,1(4)} a_d(n)\mathcal{W}_n(y)e(nx),$$

where $a_d(n) = \lim_{s\to 3/4^+} a_d(n,s)$ for $n \neq 0$ and, after recalling the definition of $\mathcal{W}_0(y)$ from (4.2.6), we have that

$$a_d(0) = \lim_{s\to 3/4^+} \frac{b_d(0,s)}{4b_0(0,s)}. \tag{4.2.29}$$

Here again we use the integral representations for the Fourier coefficients and the fact that $h_d(z,s) \to h_d(z)$ uniformly on compacta as $s \to 3/4^+$. Thus $h_d \in H^!_{1/2}$ for all $d \neq 0$. The last statement of Lemma 4.2.7 can easily be obtained from (4.2.29), (4.2.24) and (4.2.23). $\square$

Continuing with the proof of Proposition 4.2.2, we next show that the symmetry relation (4.2.12) holds. By (1.12.4) and (4.2.20) we have that $b_d(n,s) = b_n(d,s)$, hence by (4.2.28)

$$a_d(n,s) = a_n(d,s). \tag{4.2.30}$$

Now (4.2.12) follows from Lemma 4.2.7 and (4.2.30), where we use that $h_0 = \theta$ in case $nd = 0$. Note that $a_0(0) = 0$. A direct calculation using (4.2.8) together with (4.2.12) yields (4.2.13). This completes the proof of Proposition 4.2.2 and hence of Theorem 4.1.2.

4.2.2 Cycle integrals of Poincaré series

As further preparation for the proof of Theorem 4.1.3, in this section we will compute the cycle integrals of certain general Poincaré series, which we will then specialize in order to treat j_m. To begin we need to make some elementary observations about cycle integrals.

As in (2.2.9) for $Q \in \mathcal{Q}_d$ with $d > 0$ not a square let S_Q be the oriented semi-circle defined by

$$a|z|^2 + b\,\mathrm{Re}\,z + c = 0, \tag{4.2.31}$$

directed counterclockwise if $a > 0$ and clockwise if $a < 0$. Clearly

$$S_{gQ} = gS_Q, \tag{4.2.32}$$

for any $g \in \Gamma$. Given $z \in S_Q$ let C_Q be the directed arc on S_Q from z to $g_Q z$, where g_Q was defined in Section 1.11. It can easily be checked that C_Q has the same orientation as S_Q. It is convenient to define

$$dz_Q = \frac{\sqrt{d}\,dz}{Q(z,1)}. \tag{4.2.33}$$

If $z' = gz$ for some $g \in \Gamma$ we have

$$dz'_{gQ} = dz_Q. \tag{4.2.34}$$

For any Γ–invariant function f on $\mathcal{H}$ the integral $\int_{C_Q} f(z)dz_Q$ is both independent of $z \in S_Q$ and is a class invariant.

Now we will specialize the Poincaré series G_m from (2.2.14) and construct the modular functions j_m. Let $G_m(z,s) = G_m(z,\phi_{m,s})$, where

$$\phi_{m,s}(y) = \begin{cases} y^s & \text{if } m = 0 \\ 2\pi|m|^{\frac{1}{2}}y^{\frac{1}{2}}I_{s-\frac{1}{2}}(2\pi|m|y) & \text{if } m \neq 0, \end{cases}$$

with $I_{s-1/2}$ the Bessel function as before. The resulting Γ-invariant function satisfies

$$\triangle_0 G_m(z,s) = s(1-s)G_m(z,s).$$

The function G_0 is the usual Eisenstein series while G_m for $m \neq 0$ was studied by Neunhöffer [98] and Niebur [99], among others. The required analytic properties of $G_m(z,s)$ in s are most easily obtained from their Fourier expansions. For the Eisenstein series we have the well known formulas (see e.g. [74])

$$G_0(z,s) = y^s + c_0(0,s)y^{1-s} + \sum_{n\neq 0} c_0(n,s)K_{s-\frac{1}{2}}(2\pi|n|y)e(nx),$$

where $K_{s-\frac{1}{2}}$ is the K-Bessel function (see e.g. [89]),

$$c_0(0,s) = \frac{\xi(2s-1)}{\xi(2s)} \qquad \text{and for } n \neq 0 \qquad c_0(n,s) = \frac{2y^{1/2}}{\xi(2s)}|n|^{s-1/2}\sigma_{1-2s}(|n|),$$

with $\xi(s) = \pi^{-\frac{s}{2}}\Gamma(s/2)\zeta(s)$. For $m \neq 0$ the Fourier expansion of G_m can be found in [48], and is given by

$$\begin{aligned} G_m(z,s) = 2\pi|m|^{\frac{1}{2}}y^{\frac{1}{2}}I_{s-\frac{1}{2}}(2\pi|m|y)e(mx) + c_m(0,s)y^{1-s} \\ +4\pi|m|^{1/2}y^{1/2}\sum_{n\neq 0}|n|^{1/2}c_m(n,s)K_{s-\frac{1}{2}}(2\pi|n|y)e(nx), \end{aligned}$$

where

$$c_m(0,s) = \frac{4\pi|m|^{1-s}\sigma_{2s-1}(|m|)}{(2s-1)\xi(2s)}$$

and

$$c_m(n;s) = \sum_{c>0} c^{-1} K_0(m,n;c) \cdot \begin{cases} I_{2s-1}(4\pi\sqrt{|mn|}\,c^{-1}) & \text{if } mn < 0 \\ J_{2s-1}(4\pi\sqrt{|mn|}\,c^{-1}) & \text{if } mn > 0. \end{cases}$$

Define for $m \in \mathbb{Z}^+$ and $\mathrm{Re}(s) > 1$

$$j_m(z,s) = G_{-m}(z,s) - \frac{2m^{1-s}\sigma_{2s-1}(m)}{\pi^{-(s+\frac{1}{2})}\Gamma(s+\frac{1}{2})\zeta(2s-1)} G_0(z,s). \tag{4.2.35}$$

It follows from its Fourier expansion, Weil's bound (1.12.7) and (4.2.16) that $j_m(z,s)$ has an analytic continuation to $\mathrm{Re}(s) > 3/4$. Furthermore, since a bounded harmonic function is constant, for $m \in \mathbb{Z}^+$ we have

$$j_m(z,1) = j_m(z), \tag{4.2.36}$$

where j_m was defined above (4.1.7) (c.f. [99]). Alternatively, we could apply the theory of the resolvent kernel in weight 0 to get the analytic continuation of $j_m(z,s)$ up to $\mathrm{Re}\, s > 1/2$.

In view of (4.2.36), in order to compute the traces of $j_m(z,s)$ it is enough to compute them for $G_m(z,s)$. We have the following identities, which are known when $m = 0$ (Dirichlet/Hecke) and when $dD < 0$ (see e.g. [37], [35], [21]). For the convenience of the reader we will give a uniform proof.

Proposition 4.2.8. *Let $\mathrm{Re}(s) > 1$ and $m \in \mathbb{Z}$. Suppose that d and D are not both negative and that dD is not a square. Then, when $dD < 0$ we have*

$$\sum_{Q \in \Gamma\backslash Q_{dD}} \frac{\chi(Q)}{w_Q} G_m(z_Q, s) = \begin{cases} \sqrt{2}\pi|m|^{\frac{1}{2}}\,|dD|^{\frac{1}{4}} \sum\limits_{c\equiv 0(4)} \frac{S_m(d,D;c)}{c^{1/2}} I_{s-\frac{1}{2}}\left(\frac{4\pi\sqrt{m^2|dD|}}{c}\right) & \text{if } m \neq 0 \\ 2^{s-1}\,|dD|^{\frac{s}{2}} \sum\limits_{c\equiv 0(4)} \frac{S_0(d,D;c)}{c^s} & \text{if } m = 0, \end{cases}$$

while when $dD > 0$ we have

$$\sum_{Q \in \Gamma\backslash Q_{dD}} \frac{\chi(Q)}{B(s)} \int_{C_Q} G_m(z,s)dz_Q = \begin{cases} \sqrt{2}\pi|m|^{\frac{1}{2}}\,|dD|^{\frac{1}{4}} \sum\limits_{c\equiv 0(4)} \frac{S_m(d,D;c)}{c^{1/2}} J_{s-\frac{1}{2}}\left(\frac{4\pi\sqrt{m^2|dD|}}{c}\right) & \text{if } m \neq 0 \\ 2^{s-1}\,|dD|^{\frac{s}{2}} \sum\limits_{c\equiv 0(4)} \frac{S_0(d,D;c)}{c^s} & \text{if } m = 0, \end{cases}$$

where $B(s) = 2^s\Gamma(\frac{s}{2})^2/\Gamma(s)$.

Proof. By (2.2.17) the proof of Proposition 4.2.8 reduces to the case $dD > 0$. Applying Lemma 2.2.4 when $m = 0$ we use the well-known evaluation

$$\int_0^\pi (\sin\theta)^{s-1} d\theta = 2^{s-1}\frac{\Gamma(\frac{s}{2})^2}{\Gamma(s)}.$$

When $m \neq 0$ we need the following not-so-well-known evaluation to finish the proof. □

Lemma 4.2.9. *For $\mathrm{Re}(s) > 0$ we have*

$$\int_0^\pi \cos(t\cos\theta) I_{s-\frac{1}{2}}(t\sin\theta)\frac{d\theta}{(\sin\theta)^{1/2}} = 2^{s-1}\frac{\Gamma(\frac{s}{2})^2}{\Gamma(s)} J_{s-1/2}(t).$$

Proof. See Appendix A.2. □

4.2.3 The traces in terms of Fourier coefficients

In this section we complete the proofs of Theorems 4.1.3 and 4.1.4. We need to express the traces of j_m in terms of the Fourier coefficients of our basis h_d. This is first done for $j_m(z,s)$ with $\mathrm{Re}(s)>1$ by applying Proposition 1.12.1 to transform the sum of exponential sums in Proposition 4.2.8 into a sum of Kloosterman sums, which is then related to the coefficients of $h_d(z,s)$. The method of using Kloosterman sums in this way was first applied by Zagier [131] to base change, then by Kohnen [83] to the Shimura lift and more recently to weakly holomorphic forms in [18], [35], [75] and [21].

Theorem 4.1.3 follows from Lemma 4.2.7, (4.2.36) and the next result by taking the limit as $s\to 1^+$ of both sides of (4.2.37). Also we use the relationship between $a(n,d)$ and $a_d(n)$ given in and above equation (4.2.14). We remark that we actually get a slightly more general result than Theorem 4.1.3 in that we may allow $D<0$, but the general result is best left in terms of the coefficients $a_d(n)$.

Proposition 4.2.10. *Let $m\in\mathbb{Z}^+$ and $\mathrm{Re}(s)>1$. Suppose that d and D are not both negative and that dD is not a square. Then*

$$\sum_{n|m}\left(\tfrac{D}{n}\right)a_d\left(\tfrac{m^2D}{n^2},\tfrac{s}{2}+\tfrac14\right)=\begin{cases}\sum_Q \chi(Q)w_Q^{-1}j_m(z_Q,s) & \text{if } dD<0,\\ \\ B(s)^{-1}\sum_Q\chi(Q)\int_{C_Q}j_m(z,s)dz_Q & \text{if } dD>0,\end{cases}\tag{4.2.37}$$

where each sum on the right hand side is over $Q\in\Gamma\backslash\mathcal{Q}_{dD}$.

Proof. It is convenient to set for any $m\in\mathbb{Z}$

$$T_m(s)=\begin{cases}\sum_Q \chi(Q)w_Q^{-1}G_m(z_Q,s) & \text{if } dD<0,\\ B(s)^{-1}\sum_Q\chi(Q)\int_{C_Q}G_m(z,s)dz_Q & \text{if } dD>0,\end{cases}$$

where each sum is over $Q\in\Gamma\backslash\mathcal{Q}_{dD}$. By Propositions 4.2.8 and 1.12.1 we have for $m\neq 0$ and $\mathrm{Re}(s)>1$ that

$$T_m(s)=\pi|2m|^{\frac12}\,|dD|^{\frac14}\sum_{n|m}\left(\tfrac{D}{n}\right)n^{-\frac12}\sum_{c\equiv 0(4)}c^{-1}K^+\left(d,\tfrac{m^2D}{n^2};c\right)\cdot\begin{cases}I_{s-\frac12}\left(\frac{4\pi}{c}\sqrt{\frac{m^2}{n^2}\,|Dd|}\right) & \text{if } dD<0,\\ J_{s-\frac12}\left(\frac{4\pi}{c}\sqrt{\frac{m^2}{n^2}\,|Dd|}\right) & \text{if } dD>0,\end{cases}$$

while when $m=0$ we have

$$T_0(s)=2^{s-1}\,|dD|^{\frac{s}{2}}\,L_D(s)\sum_{c\equiv 0(4)}c^{-s-1/2}K^+\big(d,0;c\big).$$

Thus by (4.2.20) of Proposition 4.2.3 we derive that

$$T_m(s)=\begin{cases}\sum_{n|m}\left(\tfrac{D}{n}\right)b_d\left(\tfrac{m^2D}{n^2},\tfrac{s}{2}+\tfrac14\right), & \text{if } m\neq 0\\ \\ 2^{s-1}\pi^{-\frac{s+1}{2}}\,|D|^{\frac{s}{2}}\,L_D(s)b_d(0,\tfrac{s}{2}+\tfrac14) & \text{if } m=0.\end{cases}\tag{4.2.38}$$

In view of (4.2.35), in order to prove Proposition 4.2.10 it is enough to show that

$$\sum_{n|m}\left(\tfrac{D}{n}\right)a_d\left(\tfrac{m^2D}{n^2},\tfrac{s}{2}+\tfrac14\right)=T_m(s)-\frac{2m^{1-s}\sigma_{2s-1}(m)}{\pi^{-(s+\frac12)}\Gamma(s+\frac12)\zeta(2s-1)}T_0(s).\tag{4.2.39}$$

By (4.2.28) and the first formula of (4.2.38) the left hand side of (4.2.39) is

$$T_m(s) - \frac{b_d(0, \frac{s}{2}+\frac{1}{4})}{b_0(0, \frac{s}{2}+\frac{1}{4})} \sum_{n|m} \left(\tfrac{D}{n}\right) b_0(\tfrac{m^2D}{n^2}, \tfrac{s}{2}+\tfrac{1}{4}).$$

Hence by the second formula of (4.2.38) we are reduced to showing that

$$b_0(0, \tfrac{s}{2}+\tfrac{1}{4})^{-1} \sum_{n|m} \left(\tfrac{D}{n}\right) b_0(\tfrac{m^2D}{n^2}, \tfrac{s}{2}+\tfrac{1}{4}) = \frac{2^s \pi^{s/2} |D|^{s/2} m^{1-s} \sigma_{2s-1}(m) L_D(s)}{\Gamma(s+\frac{1}{2})\zeta(2s-1)},$$

which follows by Lemma 4.2.6. This finishes the proof of Proposition 4.2.10, hence of Theorem 4.1.3. □

We now give a quick proof of Theorem 4.1.4. By (4.2.25) we have

$$\operatorname{Res}_{s=\frac{3}{4}} P_0^+(z,s) = \tfrac{3}{4\pi}\theta(z).$$

The function $\widehat{\mathbf{Z}}_+(z)$ can now be defined through the limit formula[5]

$$\widehat{\mathbf{Z}}_+(z) = \tfrac{1}{3} \lim_{s\to\frac{3}{4}} \left(P_0^+(z,s) - \frac{\frac{3}{4\pi}\theta(z)}{s-3/4} \right). \tag{4.2.40}$$

It follows from (4.2.40) that $\widehat{\mathbf{Z}}_+(z)$ has weight $1/2$ and satisfies

$$\Delta_{1/2}(\widehat{\mathbf{Z}}_+) = -\tfrac{1}{8\pi}\theta. \tag{4.2.41}$$

Finally, using (4.2.38) when $m=0$ and the fact that $G_0(z,s)$ has a simple pole at $s=1$ with residue $3/\pi$, one shows that $\widehat{\mathbf{Z}}_+(z)$ has a Fourier expansion of the form (4.1.21).

The statement that $\widehat{\mathbf{Z}}_+(z)$ has generalized shadow $\widehat{\mathbf{Z}}_-(z)$ from (4.1.20) can now be made precise since it follows from (4.2.41) and the easily established identity

$$\xi_{3/2}\,\widehat{\mathbf{Z}}_= - \tfrac{1}{4\pi}\theta,$$

that

$$\xi_{1/2}\,\widehat{\mathbf{Z}}_+ = -2\widehat{\mathbf{Z}}_-.$$

4.2.4 Rational period functions

We now prove Theorem 4.1.5. First we give a rough bound for the traces in terms of m when $d>0$ is not a square that is sufficient to show that F_d is holomorphic in $\mathcal{H}$.

Proposition 4.2.11. *For $d>0$ not a square and $m \in \mathbb{Z}^+$ we have for all $\epsilon>0$ that*

$$\mathrm{Tr}_d(j_m) \ll_{d,\epsilon} m^{5/4+\epsilon}.$$

[5]We remark that a similar limit formula was considered in [34].

Proof. It follows from [67, Thm 1. p. 110] that for fixed d not a square and $x > 0$, we have for all $\epsilon > 0$ that
$$\sum_{0<n<x} S_m(d,1;4n) \ll_{d,\epsilon} (mx)^\epsilon (m^{5/4} + x^{3/4}), \tag{4.2.42}$$
after replacing d by $4d$ if necessary. For $1 < s < 2$ we have by (1.12) and the well-known bound (see e.g. [89, pp. 122])
$$J_v(y) \ll_\nu y^{-1/2}$$
that
$$\sum_{0<n\le m} S_m(d,1;4n)\sqrt{\tfrac{m}{n}}\, J_{s-\frac{1}{2}}\left(\pi\sqrt{|d|}\tfrac{m}{n}\right) \ll_{d,\epsilon} m^{1+\epsilon}.$$
By (4.2.16) we have for $x > m$
$$\sum_{m<n<x} S_m(d,1;4n)\sqrt{\tfrac{m}{n}}\, J_{s-\frac{1}{2}}\left(\pi\sqrt{|d|}\tfrac{m}{n}\right) \ll_{d,\epsilon} m^s \left|\sum_{m<n<x} S_m(d,1;4n) n^{-s}\right| + m^{1+\epsilon}.$$
Summation by parts and (4.2.42) give
$$m^s \sum_{m<n<x} S_m(d,1;4n) n^{-s} \ll_{d,\epsilon} m^{5/4+\epsilon}.$$
Now Proposition 4.2.11 follows by Proposition 4.2.8 and (4.2.35) by taking $s \to 1^+$ in the resulting uniform inequality
$$\sum_{Q\in\Gamma\backslash\mathcal{Q}_d} \int_{C_Q} j_m(z,s)\,dz_Q \ll_{d,\epsilon} m^{5/4+\epsilon}$$
and using (4.2.36). $\square$

It follows from Theorem 4.1.3 and Proposition 4.2.11 that the function F_d defined in (4.1.22) for $d > 0$ not a square can be represented by the series
$$F_d(z) = -\sum_{m\ge 0} \operatorname{Tr}_d(j_m)\, q^m, \tag{4.2.43}$$
which gives a holomorphic function on $\mathcal{H}$. The basis $\{j_m\}_{m\ge 0}$ has a generating function that goes back to Faber (see e.g. [4]):
$$\sum_{m\ge 0} j_m(z) q^m = \frac{j'(z)}{j(z) - j(z)}, \quad \text{where} \quad j'(z) = \frac{1}{2\pi i}\frac{d\,j}{d\,z}. \tag{4.2.44}$$
Note that this formal series converges when $\operatorname{Im}(z) > \operatorname{Im}(z)$ and that for fixed z not a zero of j' it has a simple pole at $z = z$ with residue $(2\pi i)^{-1}$. It follows from (4.2.44) and (4.2.43) that for $\operatorname{Im}(z)$ sufficiently large we have
$$F_d(z) = \frac{1}{2\pi} \sum_{Q\in\Gamma\backslash\mathcal{Q}_d} \int_{C_Q} \frac{j'(z)}{j(z)-j(z)} \frac{dz}{Q(z,1)}, \tag{4.2.45}$$
where we take for C_Q an arc on S_Q, the semi-circle defined in (2.2.9). Let $\mathcal{F}' = -\mathcal{F}^{-1}$ be the image of the standard fundamental domain under inversion $z \mapsto -1/z$. By (4.2.45)

and Lemma 2.2.3 applied to each class of $\mathcal{Q}_d$ and to each fundamental domain $\mathcal{F}$ and $\mathcal{F}'$, we can write

$$F_d(z) = \frac{1}{4\pi} \sum_{Q \in \mathcal{Q}_d} \left(\int_{S_Q \cap \mathcal{F}} \frac{j'(z)}{j(z) - j(z)} \frac{dz}{Q(z,1)} + \int_{S_Q \cap \mathcal{F}'} \frac{j'(z)}{j(z) - j(z)} \frac{dz}{Q(z,1)} \right).$$

Now it is easily seen that each of these integrals is invariant under $Q \mapsto -Q$, so we may restrict the sum to $\mathcal{Q}_d^+$, giving

$$F_d(z) = \frac{1}{2\pi} \sum_{Q \in \mathcal{Q}_d^+} \left(\int_{S_Q \cap \mathcal{F}} \frac{j'(z)}{j(z) - j(z)} \frac{dz}{Q(z,1)} + \int_{S_Q \cap \mathcal{F}'} \frac{j'(z)}{j(z) - j(z)} \frac{dz}{Q(z,1)} \right). \quad (4.2.46)$$

Recall from [25] that an indefinite quadratic form $Q = [a, b, c]$ is called *simple* if $c < 0 < a$. It is easily seen that $Q \in \mathcal{Q}_d$ is simple if and only if $Q \in \mathcal{Q}_d^+$ and S_Q intersects $\mathcal{F}'' = \mathcal{F} \cup \mathcal{F}'$. For simple Q let $A_Q = S_Q \cap \mathcal{F}''$ be the arc in $\mathcal{F}''$ oriented from right to left. Clearly A_Q must connect the two "vertical" sides of $\mathcal{F}''$.[6] Thus from (4.2.46) we obtain the identity

$$F_d(z) = \frac{1}{2\pi} \sum_{\substack{Q \text{ simple} \\ b^2 - 4ac = d}} \int_{A_Q} \frac{j'(z)}{j(z) - j(z)} \frac{dz}{Q(z,1)}.$$

Now we deform each arc A_Q in the sum of integrals to B_Q, which is within $\mathcal{F}''$ and has the same endpoints as A_Q, but travels above z. By evaluating each resulting residue at z, we get the formula

$$F_d(z) = \frac{1}{2\pi} \sum_{\substack{Q \text{ simple} \\ b^2 - 4ac = d}} \int_{B_Q} \frac{j'(z)}{j(z) - j(z)} \frac{dz}{Q(z,1)} + \frac{1}{2\pi} \sum_{\substack{Q \text{ simple} \\ b^2 - 4ac = d}} Q(z,1)^{-1},$$

which is also valid at $-1/z$. A simple calculation now shows that (4.1.23) holds in a neighborhood of z, hence for all $z \in \mathcal{H}$. Thus Theorem 4.1.5 follows.

Finally, for fixed $m \in \mathbb{Z}^+$ the inequality (4.2.42) can be used to show that the series in Proposition 4.2.8 converges when $s = 1$. They yield the formula upon using the elementary evaluation

$$J_{1/2}(y) = \sqrt{\tfrac{2}{\pi y}} \sin y.$$

[6]For example, when $d = 12$ the simple forms are $[1, 0, -3], [1, -2, -2], [1, 2, -2], [3, 0, -1], [2, 2, -1], [2, -2, -1]$. A diagram showing the corresponding arcs A_Q in this case is given in Figure 1.

Chapter 5

Modular cocycles and linking numbers

5.1 Background and statements of results

5.1.1 The Rademacher symbol as a linking number

The Dedekind η-function, Dedekind's symbol and the Rademacher symbol

Let $G = \mathrm{SL}(2,\mathbb{R})$ and $\Gamma = \mathrm{SL}(2,\mathbb{Z})$. The homogeneous space $\Gamma\backslash G$ is diffeomorphic to the 3-manifold $\mathcal{M}$ the complement of a trefoil knot in the 3-sphere S^3. (This is due to Quillen, see below.) Recall the diagonal geodesic flow on $\Gamma\backslash G$. Suppose that $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ is a primitive hyperbolic element with an eigenvalue $\epsilon > 1$. Fix a $g \in G$ so that $g^{-1}\gamma g = \left(\begin{smallmatrix} \epsilon & 0 \\ 0 & 1/\epsilon \end{smallmatrix}\right)$. Then

$$\Gamma g \mapsto \Gamma g \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix}$$

where $t \in [0, \log \epsilon]$ gives a primitive oriented closed orbit in $\Gamma\backslash G$ which depends only on the conjugacy class of γ. The image of this orbit in M is a modular knot. Below we will review Ghys [52] beautiful result that the linking number of this knot with the trefoil (with some orientation) is given by the Rademacher symbol

$$\Psi(\gamma) = \Phi(\gamma) - 3\,\mathrm{sign}(c(a+d)). \tag{5.1.1}$$

Here $\Phi(\gamma)$ is the Dedekind symbol defined for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ by

$$\Phi(\gamma) = \begin{cases} \frac{b}{d} & \text{if } c = 0 \\ \frac{a+d}{c} - 12\,\mathrm{sign}\, c \cdot s(a,c) & \text{if } c \neq 0, \end{cases} \tag{5.1.2}$$

where $s(a,c)$ is the Dedekind sum defined for $\gcd(a,c) = 1$, $c \neq 0$ by

$$s(a,c) = \sum_{n=1}^{|c|-1} \left(\left(\frac{n}{c}\right)\right) \left(\left(\frac{na}{c}\right)\right). \tag{5.1.3}$$

As usual, $((x)) = 0$ if $x \in \mathbb{Z}$ and otherwise $((x)) = x - \lfloor x \rfloor - 1/2$.

dc_1553_18

The Rademacher symbol defined for all $\gamma \in \Gamma$ by (5.1.1) is a conjugacy class invariant [105] and, for γ hyperbolic, it is the homogenization of the Dedekind symbol $\Phi(\gamma)$ [7] [24]. More precisely,

$$\Psi(\gamma) = \lim_{n\to\infty} \frac{\Phi(\gamma^n)}{n} \tag{5.1.4}$$

In addition to its role here, the Dedekind sum $s(a,c)$ occurs in surprisingly diverse contexts (see e.g. [6], [105], [79]). Among its many properties we note here only the famous reciprocity formula for $a, c > 0$

$$s(a,c) - s(-c,a) = \tfrac{1}{12}\left(\tfrac{a^2+c^2+1}{ac}\right) - \tfrac{1}{4}. \tag{5.1.5}$$

The Dedekind symbol arose in Dedekind's [31] evaluation of the transformation law for the logarithm of

$$\eta(z) = q^{1/24} \prod_{m\geq 1} (1-q^m)$$

or equivalently that of $\log \Delta$, where

$$\Delta(z) = q \prod_{m\geq 1} (1-q^m)^{24}.$$

Here as usual $q = e(z) = e^{2\pi i z}$ for $z \in \mathcal{H}$. Thus for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ we have

$$\log \Delta(\gamma z) - \log \Delta(z) = 6 \log(-(cz+d)^2) + 2\pi i \Phi(\gamma), \tag{5.1.6}$$

where $\Phi(\gamma)$ is given by the formula (5.1.2) and where we choose $\arg(-(cz+d)^2) \in (-\pi, \pi)$.

Quillen's identification

Let E_4, E_6 be the Eisenstein series of weight 4 and 6. They provide a realization of $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})$ as follows. Any function of weight k can be lifted to $\mathrm{SL}_2(\mathbb{R})$ via

$$\tilde{f}(g) = f(gi) j_k(g, i)$$

where $j_k(g,z) = (cz+d)^{-k}$. We will first map $\mathrm{SL}_2(\mathbb{R})$ into $\mathbb{C}^2$ via

$$F : g \mapsto (\tilde{E}_6(g), \tilde{E}_4(g)).$$

Since $E_4^3(z) - E_6^2(z)$ does not vanish on $\mathcal{H}$ the image of F avoids the set $V = \{(z,w) : z^2 - w^3 = 0\}$. We can act by $\mathbb{R}_+^*$ on $\mathbb{C}^2$ via

$$(z,w) \mapsto \lambda \cdot (z,w) = (\lambda^3 z, \lambda^2 w).$$

Then it is clear that if $(z,w) \in V$ then so is $\lambda \cdot (z,w)$, and conversely if $(z,w) \notin V$ then $\lambda \cdot (z,w) \notin V$.

We need the following

1. if $g_1, g_2 \in \mathrm{SL}_2(\mathbb{R})$ and $\exists \lambda > 0$ such that

$$\lambda \cdot F(g_1) = F(g_2)$$

then $\exists \gamma \in \Gamma$ such that $\gamma g_1 = g_2$.

2. For $g \in \mathrm{SL}_2(\mathbb{R})$ $\exists g \in G$, $\lambda > 0$ such that $(z, w) = \lambda \cdot F(g)$ satisfies $|z|^2 + |w|^2 = 2$. This λ depends smoothly on g.

3. If $(z, w) \in \mathbb{C}^2$ is such that $|z|^2 + |w|^2 = 2$ and $z^3 - w^2 \neq 0$ then $\exists g \in G$, $\lambda > 0$ such that
$$(z, w) = \lambda \cdot F(g).$$

These are standard and follow from properties of the modular invariant j, see the appendix.

Instead of S^3 we will use $X = \{(z, w) : |z|^2 + |w|^2 = 2\}$ which is clearly homeomorphic to S^3 but is a more convenient normalization.

By the above for each $(z, w) \in X \setminus V$ there is a unique $g \in \Gamma\backslash G$, and a unique $\lambda > 0$ such that
$$(z, w) = \lambda \cdot F(g)$$
and this maps $\Gamma\backslash G$ one-to-one and onto $X \setminus V$. One easily checks that this is a homeomorphism ($g \mapsto \lambda(g)$ is continuous).

It remains to identify $X \cap V$. Assume $(z, w) \in X$ and $z^2 = w^3$. Then
$$|w|^3 + |w|^2 = 2$$
which is only possible if $|w| = 1$, in which case we also have $|z| = 1$. Therefore
$$X \cap V = \{(z, w) \in S^1 : z^2 = w^3\}.$$

This set is clearly homeomorphic to S^1 via $e^{i\theta} \mapsto (e^{3i\theta}, e^{2i\theta})$ and so is a knot. It is not difficult to construct a knot diagram that identifies this knot with the trefoil, or alternatively we may just take this construction as the definition of the trefoil knot.

The geodesic flow. Periodic orbits.

Let $x \in \Gamma\backslash\mathrm{SL}_2(\mathbb{R})$ and consider the map $\phi_t : \Gamma\backslash\mathrm{SL}_2(\mathbb{R}) \to \Gamma\backslash\mathrm{SL}_2(\mathbb{R})$
$$x \mapsto x\phi(t)$$
where
$$\phi(t) = \left[\begin{smallmatrix} e^t & 0 \\ 0 & e^{-t} \end{smallmatrix}\right].$$

It is easy to see that $\phi_{t+s} = \phi_t \circ \phi_s$, the resulting flow is called the geodesic flow. When $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$ is identified with unit tangent bundle with $\Gamma\backslash\mathcal{H}$ this is the geodesic flow in the geometric sense, the flow induced by geodesics.

The periodic orbits of ϕ_t correspond to conjugacy classes of primitive hyperbolic conjugacy classes $\{\gamma\}$ in Γ. Here hyperbolic means that γ have distinct real eigenvalues, and primitive refers to non-divisibility, there is no $\sigma \in \Gamma$ such that $\sigma^n = \gamma$ for some n.

To make the relation between hyperbolic elements and periodic orbits explicit note that if $\gamma \in \Gamma$ has $\operatorname{tr}\gamma > 2$ and fixed points $w' < w$ then both γ and γ^{-1} are diagonalized by $M = \frac{1}{\sqrt{w-w'}}\left[\begin{smallmatrix} w & w' \\ 1 & 1 \end{smallmatrix}\right]$. By replacing γ with γ^{-1} if necessary we may assume that
$$\gamma M = M \begin{bmatrix} \varepsilon & 0 \\ 0 & 1/\varepsilon \end{bmatrix}$$
where $\varepsilon > 1$. When $a + d > 2$ this is equivalent to $c > 0$. Both
$$\tilde{\gamma}_+(t) = M\phi(t) \text{ and } \tilde{\gamma}_-(t) = MS\phi(t)$$
are periodic orbits of the geodesic flow $g \mapsto g\phi(t)$ on $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$.

dc_1553_18

Ghys' theorem

As above define $\tilde{\Delta} : SL_2(\mathbb{R}) \to \mathbb{C}$ by

$$\tilde{\Delta}(g) = \Delta(gi) j_{12}(g,i)$$

where for $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{SL}_2(R)$

$$j_{12}(g,z) = (cz+d)^{-12}.$$

Similar lifts $\tilde{E}_4, \tilde{E}_6$ of the classical Eisenstein series E_4 and E_6 give an embedding of $\Gamma\backslash \mathrm{SL}_2(\mathbb{R})$ into $\mathbb{C}^2$. The 3-manifold $\{(\tilde{E}_4(g), \tilde{E}_6(g) : g \in \mathrm{SL}_2(\mathbb{R})\}$ is disjoint from the hypersurface $\mathcal{V} = \{(z,w) : z^3 = w^2\}$ and is easily seen to be homeomorphic to the complement of $\mathcal{V} \cap S^3$, the trefoil knot, in S^3. Let $\gamma \in \mathrm{SL}_2(\mathbb{R})$ be hyperbolic, with $\operatorname{tr}\gamma > 2$. We are looking for the linking number of the closed periodic orbit $\tilde{\gamma}_+$ with the trefoil (after the above identification). Since $\tilde{E}_4{}^3 - \tilde{E}_6{}^2 = \tilde{\Delta}$, a general topological argument shows that this linking number is the same as the winding number of $\tilde{\Delta}(\tilde{\gamma}_+(t))$ around 0. This in turn can be computed as follows

$$2\pi i \operatorname{ind}(\tilde{\Delta}(\tilde{\gamma}_+(t)), 0) = \int_{\tilde{\gamma}_+} \frac{d\tilde{\Delta}}{\tilde{\Delta}} = \int_{\tilde{\gamma}_+} \frac{d\Delta}{\Delta} + \int_{\tilde{\gamma}_+} \frac{dj_{12}}{j_{12}}.$$

The first integral can be evaluated from the transformation formula of $\log\Delta$ from $\tilde{\gamma}_+(0)i = Mi = z_0$ to $\tilde{\gamma}_+(\log\varepsilon)i = \gamma z_0$

$$\log\Delta(\gamma z_0) - \log\Delta(z_0) = 12\log\left(\frac{cz_0+d}{i\operatorname{sign}c}\right) + 2\pi i\Phi(\gamma)$$

with $\Phi(\gamma)$ as in (5.1.2). (See [105] equation (60) on page 49.)

Similarly the value of the second integral is $12\log(cz_0+d)$ and the linking number of the closed orbit of a hyperbolic γ is given by

$$\frac{6}{\pi i}\left(\log\left(\frac{cz_0+d}{i\operatorname{sign}c}\right) - \log(cz_0+d)\right) + \Phi(\gamma)$$

Finally for $\operatorname{Im} z_0 > 0$

$$\frac{6}{\pi i}\left(\log\left(\frac{cz_0+d}{i\operatorname{sign}c}\right) - \log(cz_0+d)\right) = -3\operatorname{sign}c$$

leading to Ghys' theorem.

5.1.2 Generalizations of the Dedekind symbol

At the end of his paper Ghys mentions the problem of interpreting the linking number between two modular knots. In this section we will approach this question by giving an appropriate generalization of the Dedekind symbol. To do this we give an equivalent but slightly different approach to the above results about the Dedekind symbol: it arises as a limiting value of the weight 0 cocycle whose derivative is $\frac{12c}{cz+d}$. This limiting value is an integer and its homogenization is also an integer that gives the linking number with the trefoil.

To put this into perspective, let $\mathcal{P}$ be the space of holomorphic functions f on $\mathcal{H}$ such that $f(z) \ll y^{\alpha} + y^{-\alpha}$ for some α depending on f. For any integer $k \in 2\mathbb{Z}$, $\gamma \in \Gamma$ acts on

$\mathcal{P}$ by the usual slash action defined via $f|_k\gamma = (cz+d)^{-k}f(\gamma z)$. A 1-cocycle of weight k for Γ with coefficients in $\mathcal{P}$ is a map $\Gamma \to \mathcal{P}$ given by $\gamma \mapsto r(\gamma, z)$ with

$$r(\sigma\gamma, z) = r(\sigma, z)|_k\gamma + r(\gamma, z)$$

for all $\gamma, \sigma \in \Gamma$. Now given a 1-cocycle $r(\gamma, z)$ of weight 2 for Γ there will be a unique 1-cocycle $R(\gamma, z)$ of weight 0 for Γ such that

$$\frac{d}{dz}R(\gamma, z) = r(\gamma, z), \tag{5.1.7}$$

the uniqueness following from the fact that $H^1(\Gamma, \mathbb{C}) = \{0\}$. We call $R(\gamma, z)$ the primitive of $r(\gamma, z)$.

The weight 2 cocycle relevant to the Dedekind sum is given for $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ by

$$r(\gamma, z) = \frac{12c}{cz+d}$$

which, up to a constant, appears in the transformation formula of the weight 2 Eisenstein series $E_2(z)$ (and which therefore equals a multiple of $\Delta'(z)/\Delta(z)$). It follows from (5.1.6) that the primitive for this cocycle is

$$R(\gamma, z) = 6\log(-(cz+d)^2) + 2\pi i\Phi(\gamma),$$

provided $c \neq 0$, from which we have the limit formula for $\Phi(\gamma)$ in (5.1.2):

$$\Phi(\gamma) = \tfrac{1}{2\pi}\lim_{y\to\infty} \operatorname{Im} R(\gamma, iy). \tag{5.1.8}$$

As an attempt to generalize the linking number formula of Ghys to two closed orbits, we will associate to any conjugacy class $\mathcal{C}$ of hyperbolic $\sigma \in \Gamma$ with $\operatorname{tr}\sigma > 2$ the weight two 1-cocycle defined for $c \neq 0$ and $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ by

$$r_{\mathcal{C}}(\gamma, z) := \varepsilon_{\mathcal{C}} \sum \frac{1}{z-w} - \frac{1}{z-w'}, \tag{5.1.9}$$

where the sum is over the fixed points w', w of $\sigma \in \mathcal{C}$, satisfying $w' < -d/c < w$ and

$$\varepsilon_{\mathcal{C}} = \begin{cases} 1 & \text{if } \sigma \not\sim \sigma^{-1} \\ 2 & \text{if } \sigma \sim \sigma^{-1} \end{cases}. \tag{5.1.10}$$

If $c = 0$ we let $r_{\mathcal{C}}(\gamma, z) = 0$.

With Duke and Imamoglu we proved the following theorems.

Theorem 5.1.1. *Let $r_{\mathcal{C}}(\gamma, z)$ be defined as in (5.1.9). Then $r_{\mathcal{C}}(\gamma, z)$ is a weight 2 cocycle for Γ.*

Let $\mathcal{C}$ be a primitive conjugacy class of Γ. Theorem 5.1.1 has an immediate corollary for the functional equation of the function

$$F_{\mathcal{C}}(z) = \sum_{m=0}^{\infty} a_{\mathcal{C}}(m)e^{2\pi imz}$$

where the coefficient $a_{\mathcal{C}}(n)$ is given by the cycle integral

$$a_{\mathcal{C}}(m) = \sqrt{D'} \int_{z_0}^{\sigma z_0} j_m(z) \tfrac{dz}{Q_\sigma(z)}. \tag{5.1.11}$$

Here $\sigma = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) \in \mathcal{C}$ is primitive and we set $Q_\sigma(z) = c'z^2 + (d' - a')z - b'$ and $D' = (a' + d')^2 - 4$. The path of integration can be taken as any path from z_0 to σz_0. In particular, if λ is the eigenvalue > 1 of σ^2 then

$$a_{\mathcal{C}}(0) = \log \lambda,$$

assuming that $\operatorname{tr} \sigma > 2$.

Theorem 5.1.2. *For the function $F_{\mathcal{C}}$ defined above satisfies*

$$F_{\mathcal{C}}(\gamma z)(cz + d)^{-2} - F_{\mathcal{C}}(z) = r_{\mathcal{C}}(\gamma, z).$$

This is extends Theorem 4.1.5 to an arbitrary element of Γ, to a slightly more general function then F_d (which arises as the sum of certain $F_{\mathcal{C}}$-s.)

Concerning the generalization of Dedekind's symbol let

$$G_{\mathcal{C}}(z) = a_{\mathcal{C}}(0)z + \frac{1}{2\pi i} \sum_{m=1}^{\infty} \frac{a_{\mathcal{C}}(m)}{m} e^{2\pi i m z}$$

be a primitive of $F_{\mathcal{C}}(z)$ and $R_{\mathcal{C}}(\gamma, z)$ be the unique primitive of $r_{\mathcal{C}}(\gamma, z)$ which is a 0-cycle, so that we have

$$G_{\mathcal{C}}(\gamma z) - G_{\mathcal{C}}(z) = \int_z^{\gamma z} F_{\mathcal{C}}(w)dw = R_{\mathcal{C}}(\gamma, z).$$

Next we define the Dedekind symbol for $\mathcal{C}$ and any $\gamma \in \Gamma$ by

$$\Phi_{\mathcal{C}}(\gamma) = \tfrac{2}{\pi} \lim_{y \to \infty} \operatorname{Im} R_{\mathcal{C}}(\gamma, iy) \tag{5.1.12}$$

We also extended (5.1.8) form $\log \eta$ to the functions $F_{\mathcal{C}}$.

Theorem 5.1.3. $\Phi_{\mathcal{C}}(\gamma)$ *exists and is an integer.*

5.1.3 Linking number between symmetric modular links

In order to define the linking number of two cycles in a manifold we must assume that they are each homologous to 0 and that they don't intersect. For two orbits as above one can either fill in the trefoil appropriately to get S^3, as is done in [53], or restrict attention to orbits that are null-homologous as in [32]. It is not immediately clear what role modular forms may play in the first approach as the SL_2-geometry is then lost. In the second course one may use a theorem that goes back to Birkhoff to show that the link determined by a primitive hyperbolic element and its inverse is null-homologous in $\mathcal{M}$. Our main result with Duke and Imamoglu is that for two such distinct symmetric links, denoted also by $\mathcal{C}_\gamma$, and $\mathcal{C}_\sigma$, their linking number $Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma)$ is given by the homogenization of $\Phi_{\mathcal{C}_\sigma}$. More precisely

Theorem 5.1.4. *If $\mathcal{C}$ is a hyperbolic conjugacy class then*

$$\lim_{n\to\infty} \frac{\Phi_{\mathcal{C}_\sigma}(\gamma^n)}{n}.$$

exists.

The value of the limit, called the homogenization of $\Phi_{\mathcal{C}}$ will be denoted by $\Psi_{\mathcal{C}_\sigma}(\gamma)$. Recall that Ghys result interprets the homogenization of the Dedekind symbol as a linking number. The natural extension of this to $\Psi_{\mathcal{C}}$ is the following

Theorem 5.1.5. *Let $\mathcal{C}_\sigma$ and $\mathcal{C}_\gamma$ denote also the links associated to two different primitive conjugacy classes. Then*

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = \Psi_{\mathcal{C}_\sigma}(\gamma)$$

5.1.4 Reciprocity

With Duke and Imamoglu I also gave an expression for $\Phi_{\mathcal{C}}(\gamma)$ in terms of a special value of a certain Dirichlet series that has some properties analogous to the Dedekind sum $s(a, c)$ from (5.1.3), including the reciprocity formula (5.1.5). That something like this might be possible is indicated by the fact that for the Dirichlet series

$$L(s, a/c) = \sum_{n\geq 1} \sigma(n) e(\tfrac{a}{c} n) n^{-s},$$

where $\sigma(n)$ is the usual divisor sum, we have the limit formula

$$s(a, c) = \tfrac{1}{2\pi i} \lim_{s\to 1} \left[L(s, \tfrac{a}{c}) + \tfrac{1}{2s-2} \right], \tag{5.1.13}$$

assuming $c > 0$.

The Dirichlet series associated to the cocycles of Theorem 5.1.1 are given explicitly as follows. For each $m \geq 0$ let j_m be the unique modular function holomorphic on $\mathcal{H}$ whose Fourier expansion begins

$$j_m(z) = q^{-m} + \emptyset(q)$$

and define for $\alpha \in \mathbb{Q}$ the Dirichlet series

$$L_{\mathcal{C}}(s, \mathfrak{a}) = \sum_{n\geq 1} a_{\mathcal{C}}(n) e(n\mathfrak{a}) n^{-s}, \tag{5.1.14}$$

where the coefficient $a_{\mathcal{C}}(n)$ is given by the cycle integral

$$a_{\mathcal{C}}(m) = \sqrt{D'} \int_{z_0}^{\sigma z_0} j_m(z) \tfrac{dz}{Q_\sigma(z)}. \tag{5.1.15}$$

as in (5.1.11).

Our next theorem gives the connection of this Dirichlet series to $\Phi_{\mathcal{C}}(\gamma)$.

Theorem 5.1.6. *Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ with $c \neq 0$ and and $L_{\mathcal{C}}(s, a/c)$ be the Dirichlet series as in (5.1.14). Then $L_{\mathcal{C}}(s, a/c)$ converges for $\mathrm{Re}(s) > 9/4$, has a meromorphic continuation to $s > 0$ and is holomorphic at $s = 1$. Moreover*

$$\Phi_{\mathcal{C}}(\gamma) = -\tfrac{1}{\pi^2} \operatorname{Re} L_{\mathcal{C}}(1, a/c). \tag{5.1.16}$$

It is interesting that $\Phi_{\mathcal{C}}(\gamma)$ depends only on $a/c \mod 1$. Furthermore, we have the following reciprocity formula:

For $z_i \in \mathbb{C} \cup \{\infty\}$, let

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

denote the cross ratio. We assume that $(a, c) = 1$ and $ac \neq 0$. Then

$$\frac{1}{i\pi}\left[L_{\mathcal{C}}(1, a/c) - L_{\mathcal{C}}(1, -c/a)\right] = -2\left(\tfrac{a^2+c^2+1}{ac}\right)\log\lambda + \varepsilon_{\mathcal{C}}\sum_{w'<0<w}\log[\tfrac{a}{c}, w, w', -\tfrac{c}{a}] \quad (5.1.17)$$

Here we interpret the imaginary part of the logarithm of a negative real number to be π.

Note that (5.1.17) is in some sense analogous to (5.1.5) and allows for a fast calculation of $L_{\mathcal{C}}(1, a/c)$. Now we restrict ourselves to the imaginary part of $R_{\mathcal{C}}(\gamma, z)$. Recall that

$$\Phi_{\mathcal{C}}(\gamma) = \tfrac{2}{\pi}\lim_{y\to\infty}\operatorname{Im} R_{\mathcal{C}}(\gamma, iy) = -\tfrac{1}{\pi^2}\operatorname{Re} L_{\mathcal{C}}(1, a/c).$$

We start with

Theorem 5.1.7. *Let $\gamma \in \Gamma$ be a hyperbolic element. Then $\Phi_{\mathcal{C}}(\gamma) = -|I_{\mathcal{C}}(-d/c, i\infty)|$ where $|I_{\mathcal{C}}(-d/c, i\infty)|$ counts the number of intersections of image of the half line $\{-d/c + it : t > 0\}$ and the closed geodesic associated to $\mathcal{C}$.*

Note that since $L_{\mathcal{C}}(1, a/c)$ depends only on $a/c \mod 1$, so does $\Phi_{\mathcal{C}}(\gamma)$ and hence for $c \neq 0$ we can write $\Phi_{\mathcal{C}}(a/c) = \Phi_{\mathcal{C}}(\gamma)$. The following theorem is an analogue of Dedekind's reciprocity formula. It allows for, via Euclid's algorithm, a quick computation of $\Phi_{\mathcal{C}}(\gamma)$.

Theorem 5.1.8. *Let $\mathcal{C}$ be a hyperbolic conjugacy class and $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. For $ac \neq 0$ we have*

$$\Phi_{\mathcal{C}}(a/c) = \Phi_{\mathcal{C}}(-c/a) + \frac{\epsilon_{\mathcal{C}}}{2}\sum_{w'<0<w}(1 - \operatorname{sign}[\tfrac{a}{c}, w, w', -\tfrac{c}{a}]) \quad (5.1.18)$$

5.2 Proofs

5.2.1 Dirichlet series associated to weight 2 cocycles

Recall that a (strongly) parabolic cocycle of weight k for Γ with coefficients in $\mathcal{P}$ is a map $\Gamma \to \mathcal{P}$ given by $\gamma \mapsto r(\gamma, z)$ with

$$r(\sigma\gamma, z) = r(\sigma, z)|_k\gamma + r(\gamma, z)$$

for all $\gamma, \sigma \in \Gamma$ which also satisfies $r(T, z) \equiv 0$.

It follows from a more general result of Knopp [88] that given a parabolic cocycle $r(\gamma, z)$ for Γ there is $F(z) = \sum_{n\geq 0} a_n e^{2\pi i n z}$ with $a_n \ll n^C$ for some $C > 0$ such that $\forall \gamma \in \Gamma$,

$$F|_k\gamma(z) = F(z) + r(\gamma, z).$$

The function $F(z)$ is called the modular integral associated to $r(\gamma, z)$. We now restrict ourselves to the case of $k = 2$ and let $r(\gamma, z)$ be a cocycle of weight 2. We associate to $r(\gamma, z)$ and its modular integral a Dirichlet series

$$L(F, s, a/c) = \sum_{n\geq 1} a_n e(\frac{an}{c}) n^{-s}.$$

dc_1553_18

In this section we will first prove a general theorem giving the relation of the special value of $L(F,1,a/c)$ to the unique weight 0 cocycle $R(\gamma,z)$ which satisfies $R'(\gamma,z)=r(\gamma,z)$.

This is based on the fact the function $G(z)=a_0z+\sum_{n>0}\frac{a_n}{2\pi in}e^{2\pi inz}$ is a primitive of $F(z)$ and satisfies $\frac{d}{dz}\left(G(\gamma z)-G(z)\right)=r(\gamma,z)$. This gives a relation between $R(\gamma,z)$ and $\int_z^{\gamma z}(F(w)-a_0)dw$, which in turn expresses $\lim_{y\to\infty}R(\gamma,iy)$ in terms of the "period-integral" $\int_{a/c}^{i\infty}(F(w)-a_0)dw$. If F were a weight 2 cusp form, it is well known that this period integral is expressible in terms of the central value of a twisted Dirichlet series of F. The next theorem shows the case of modular integrals is similar. More precisely we have the following theorem.

Theorem 5.2.1. *Let $r(\gamma,z)\in\mathcal{P}$ be a cocycle of weight 2 and $F(z)=\sum_{n\geq 0}a_nq^n$ be its modular integral. Assume that $a_n\ll n^\alpha$ for some $\alpha>0$. For $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$, let*

$$\Lambda(s,\tfrac{a}{c})=\Lambda(F,s,\tfrac{a}{c})=\left(\tfrac{2\pi}{c}\right)^{-s}\Gamma(s)\sum_{n\geq 1}a_ne(\frac{an}{c})n^{-s} \tag{5.2.1}$$

and

$$H(s,\tfrac{a}{c})=\Lambda(s,\tfrac{a}{c})+\int_1^\infty r(\gamma,-d/c+it/c)t^{1-s}dt+\frac{a_0}{s}-\frac{a_0}{2-s}. \tag{5.2.2}$$

Then $H(s,\frac{a}{c})$ is entire and satisfies the functional equation $H(s,\frac{a}{c})=-H(2-s,\frac{-d}{c})$. Moreover if

$$R(\gamma,z)=\tfrac{-i}{c}H(1,\tfrac{a}{c})+\int_{-\frac{d}{c}+\frac{i}{c}}^{z}r(\gamma,w)dw+a_0\left(\tfrac{a+d}{c}\right) \tag{5.2.3}$$

Then $R(\gamma,z)$ is the weight zero cocycle such that $R'(\gamma,z)=r(\gamma,z)$.

Proof. Let $z_t=\frac{-d}{c}+\frac{i}{ct}$ so that $\gamma z_t=\frac{a}{c}+\frac{it}{c}$ and $cz_t+d=i/t$. Then

$$\begin{aligned}\Lambda(s,a/c)&=\int_0^\infty(F(\gamma z_t)-a_0)t^{s-1}dt\\&=\int_0^1(F(\gamma z_t)-a_0)t^{s-1}dt+\int_1^\infty(F(\gamma z_t)-a_0)t^{s-1}dt\\&=-\frac{a_0}{s}-\int_1^\infty F(\gamma z_{1/t})(it)^{-2}t^{1-s}dt+\int_1^\infty(F(\gamma z_t)-a_0)t^{s-1}dt\\&=-\frac{a_0}{s}-\int_1^\infty[F(z_{1/t})+r(\gamma,z_{1/t})]t^{1-s}dt+\int_1^\infty(F(\gamma z_t)-a_0)t^{s-1}dt\\&=-\frac{a_0}{s}+\frac{a_0}{2-s}-\int_1^\infty r(\gamma,z_{1/t})t^{1-s}dt\\&\quad-\int_1^\infty(F(z_{1/t})-a_0)t^{1-s}dt+\int_1^\infty(F(\gamma z_t)-a_0)t^{s-1}dt\end{aligned}$$

Hence

$$\begin{aligned}H(s,\tfrac{a}{c})=&\Lambda(s,\tfrac{a}{c})+\int_1^\infty r(\gamma,-d/c+it/c)t^{1-s}dt+\frac{a_0}{s}-\frac{a_0}{2-s}\\=&-\int_1^\infty(F(z_{1/t})-a_0)t^{1-s}dt+\int_1^\infty(F(\gamma z_t)-a_0)t^{s-1}dt\end{aligned} \tag{5.2.4}$$

Both integrals in (5.2.4) converge for all $s \in \mathbb{C}$ due to the exponential decay of the integrands proving the analytic continuation of $H(s, \frac{a}{c})$ to the whole complex plane. The functional equation $H(s, a/c) = -H(2-s, -d/c)$ also follows easily from (5.2.4) since $z_{1/t} = \frac{-d}{c} + \frac{it}{c}$ and $\gamma z_t = \frac{a}{c} + \frac{it}{c}$.

We next take the limit $s \to 1$ to get

$$\begin{aligned} H(1, \tfrac{a}{c}) &= -\frac{c}{i}\int_{z_1}^{i\infty}(F(z)-a_0)dz + \frac{c}{i}\int_{\gamma z_1}^{i\infty}(F(z)-a_0)dz \\ &= -\frac{c}{i}\left(G(\gamma z_1) - G(z_1) - a_0\left(\frac{a+d}{c}\right)\right) \end{aligned}$$

where $G(z) = a_0 z + \sum_{n\geq 1}\frac{a_n}{2\pi i n}q^n$. Since $G'(z) = F(z)$,

$$G(\gamma z) - G(z) = \int_{z_1}^{z} r(\gamma, w)dw + \Phi(\gamma)$$

with $\Phi(\gamma) = (G(\gamma z_1) - G(z_1))$.

Hence

$$R(\gamma, z) = \int_{z_1}^{z} r(\gamma, w)dw + (G(\gamma z_1) - G(z_1)) = G(\gamma z) - G(z)$$

is a cocycle being the boundary of a function G. This finishes the proof of the theorem since clearly $R'(\gamma, z) = r(\gamma, z)$. □

As an immediate corollary of Theorem 5.2.1 we prove the limit formula (5.1.13) for the classical Dedekind sums defined as in (5.1.3).

Corollary 5.2.2. *Let $s(a, c)$ be the Dedekind sum and*

$$L(s, a/c) = \sum_{n\geq 1}\sigma(n)e(\tfrac{a}{c}n)n^{-s},$$

Then

$$s(a, c) = \tfrac{1}{2\pi i}\lim_{s\to 1}\left[L(s, \tfrac{a}{c}) + \tfrac{1}{2s-2}\right].$$

Proof. We apply Theorem 5.2.1 in the case of Eisenstein series $F(z) = E_2(z) = 1 - 24\sum\sigma(n)q^n$ and its cocycle $r(\gamma, z) = \frac{6}{\pi i}\frac{c}{cz+d}$, so that $L(F, s, a/c) = -24L(s, a/c)$. For simplicity assume $c > 0$. As a primitive of $r(\gamma, z)$ we choose $\frac{6}{\pi i}\log\left(\frac{cz+d}{i}\right)$. Using (5.2.2) and (5.2.3) we have

$$R(\gamma, z) = \lim_{s\to 1}\left[-\frac{24}{2\pi i}L(s, a/c) - \frac{6}{\pi i}\frac{1}{s-1}\right] + \frac{6}{\pi i}\int_{-d/c+i/c}^{z}\frac{c}{cw+d}dw + \left(\frac{a+d}{c}\right) \tag{5.2.5}$$

$$= \frac{12}{2\pi i}\log\frac{cz+d}{i} + \Phi(\gamma) \tag{5.2.6}$$

where

$$\Phi(\gamma) = \lim_{s\to 1}\left[-\frac{12}{\pi i}L(s, a/c) - \frac{6}{\pi i}\frac{1}{s-1}\right] + \left(\frac{a+d}{c}\right)$$

The limit formula (5.1.13) now follows from Dedekind's formula (5.1.2) for $\Phi(\gamma)$. □

5.2.2 Weight 2 rational cocycles for the modular group

In this section we restrict ourselves to cocycles of weight 2 which are rational functions. The simplest example is $r(\gamma, z) = 12c/(cz + d)$ whose poles are in $\mathbb{Q}$. In the case $r(\gamma, z)$ is a rational cocycle whose poles are not rational it is known that $r(S, z)$ can be written as a finite linear combination of functions of the form

$$\sqrt{D} \sum_{AC<0} \frac{\operatorname{sign} A}{Az^2 + Bz + C} \tag{5.2.7}$$

where $Q(X, Y) = AX^2 + BXY + CY^2$ runs over quadratic forms in the class $\mathcal{C}$ (see [5, 25, 102, 103]). Rational period functions were introduced by Knopp in the 1970s [86, 87] who showed using results from [85] that they have modular integrals. His construction arises from a meromorphic Poincaré series formed out of cocycles and is very difficult to compute (see also [44]). On the other hand in [40] and [41] certain explicit modular integrals were constructed whose Fourier coefficients are given by cycle integrals of weakly holomorphic forms. These functions are parametrized by classes of indefinite quadratic forms $\mathcal{C}$ and are given by the Fourier expansion

$$F_{\mathcal{C}}(z) = \sum_{m \geq 0} a_{\mathcal{C}}(m) e(mz). \tag{5.2.8}$$

with

$$a_{\mathcal{C}}(m) = \sqrt{D} \int_{z_0}^{\sigma z_0} j_m(z) \tfrac{dz}{Q(z)}. \tag{5.2.9}$$

Here j_m is the unique modular function whose Fourier expansion has the form $q^{-m} + O(q)$, Q is any quadratic form in the class $\mathcal{C}$, $\sigma = \sigma_Q$ is a distinguished generator of the group of automorphs of Q. The value of the integral is independent of the path and the point $z_0 \in \mathcal{H}$. In [40] it is shown that the function $F_{\mathcal{C}}$ arises from the cycle integral of the Green function $\frac{j'(z)}{j(z)-j(w)}$. The cycle integral of this Green function is modular but with jump singularities along the geodesic. $F_{\mathcal{C}}$ is then the analytic continuation from the connected component of the cusp. It is holomorphic, but no longer invariant.

The association $Q \mapsto \sigma_Q$ sets up a bijection between elements of the class $\mathcal{C}$ of the quadratic form Q and the conjugacy class of σ_Q, which by abuse of notation will also be denoted by $\mathcal{C}$. Since it is more convenient for us to express our results in terms of the hyperbolic conjugacy class, we briefly recall this correspondence. If $Q(X, Y) = AX^2 + BXY + CY^2$ has discriminant $D = B^2 - 4AC$, and t, u are the smallest positive solutions of Pell's equation $t^2 - Du^2 = 4$ then

$$\sigma_Q = \begin{bmatrix} \frac{t+Bu}{2} & Cu \\ -Au & \frac{t-Bu}{2} \end{bmatrix}.$$

Conversely if $\sigma = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) \in \mathcal{C}$ is a primitive hyperbolic element and we set $Q_\sigma(z) = (c'X^2 + (d' - a')XY - b'Y^2)$, and $Q = \frac{-1}{u} Q_\sigma$ with $u = \gcd(c', d' - a', b')$ then $\sigma_Q = \sigma$. It follows that with $D' = (a' + d')^2 - 4$ we also have

$$a_{\mathcal{C}}(m) = \sqrt{D'} \int_{z_0}^{\sigma z_0} j_m(z) \tfrac{dz}{Q_\sigma(z)}$$

as in (5.1.11).

As in [40] one can show that $a_{\mathcal{C}}(m) \ll m^{5/4+\epsilon}$ for any $\epsilon > 0$ and $F_{\mathcal{C}}$ satisfies the transformation property

$$z^{-2}F_{\mathcal{C}}(-1/z) - F_{\mathcal{C}}(z) = \varepsilon_{\mathcal{C}} \sum_{w'_Q<0<w_Q} \frac{1}{z-w} - \frac{1}{z-w'}. \tag{5.2.10}$$

Note that the rational function on the right hand side above is the same as in (5.2.7).

Here for $Q \in \mathcal{C}$, $w'_Q < w_Q$ are the two roots of $Q(t,1) = 0$. If $\sigma = \sigma_Q$ then these are also the fixed points $w'_\sigma < w_\sigma$ of σ, and $\varepsilon_{\mathcal{C}}$ is defined as in (5.1.10).

If $\mathcal{C}$ denotes the class of Q or the class of the hyperbolic element σ_Q we let

$$\mathcal{W}_{\mathcal{C}} = \big\{(w'_Q, w_Q) : Q \in \mathcal{C}\big\} = \{(w'_\sigma, w_\sigma) : \sigma \in \mathcal{C}\} \tag{5.2.11}$$

the ordered pairs of roots of $Q \in \mathcal{C}$ or equivalently the fixed points of σ.

For a fixed $\gamma \in SL_2(\mathbb{Z})$, we let as in (5.1.9),

$$r_{\mathcal{C}}(\gamma, z) := \varepsilon_{\mathcal{C}} \sum \frac{1}{z-w} - \frac{1}{z-w'}$$

where the sum is over $(w', w) \in \mathcal{W}_{\mathcal{C}}$, satisfying $w' < -d/c < w$ if $c \neq 0$ and $r_{\mathcal{C}}(\gamma, z) \equiv 0$ otherwise.

Remark 5.2.3. Although the set $\mathcal{W}_{\mathcal{C}}$ is infinite, the sum defining $r_{\mathcal{C}}(\gamma, z)$ is finite. To see this note that in the case that $-d/c$ is an integer the number of terms is the same as the number of quadratic forms $[A, B, C]$ for which $AC < 0$. Otherwise consider a form $[A, B, C]$ satisfying $\frac{-B-\sqrt{D}}{2A} < \frac{-d}{c} < \frac{-B+\sqrt{D}}{2A}$, then the form $[A, cB, c^2C]$ has discriminant c^2D and its roots are separated by $-d$, and integer.

For later use we give another description of $r_{\mathcal{C}}(\gamma, z)$. For $\sigma \in \mathcal{C}$ a fixed hyperbolic element, let $w_\sigma, w_{\sigma'}$ be its two fixed points, $\Gamma_\sigma = \{g \in \Gamma : g^{-1}\sigma g = \sigma\}$, and S_σ be the semicircle whose endpoints are w_σ and w'_σ. Let $\partial\mathcal{H} = \mathbb{R} \cup i\infty$ and $\overline{\mathcal{H}} = \mathcal{H} \cup \partial\mathcal{H}$.

For $z_1, z_2 \in \overline{\mathcal{H}}$ we denote the geodesic segment joining z_1 and z_2 by ℓ_{z_1,z_2}. Let

$$I_{\mathcal{C}}(z_1, z_2) = \{\alpha \in \Gamma/\Gamma_\sigma : \alpha S_\sigma \text{ intersects } \ell_{z_1,z_2}\}. \tag{5.2.12}$$

and let $|I_{\mathcal{C}}(z_1, z_2)|$ denote the cardinality of $I_{\mathcal{C}}(z_1, z_2)$.

Note that if we define the net of σ, $\mathcal{N}_\sigma$ as the preimage of the closed geodesic associated to σ in $\mathcal{H}$ so that

$$\mathcal{N}_\sigma := \bigcup_{g\in\Gamma} gS_\sigma = \bigcup_{g\in\Gamma} S_{g^{-1}\sigma g}, \tag{5.2.13}$$

then $|I_{\mathcal{C}}(\alpha, \beta)|$ counts the number of intersections of the geodesic segment $\ell_{\alpha,\beta}$ with the semicircles in $\mathcal{N}_\sigma$, the net of σ. Moreover $\mathcal{W}_{\mathcal{C}}$ is simply the set of end points of the geodesics in the net $\mathcal{N}_\sigma$.

With the above notation we also have

$$r_{\mathcal{C}}(\gamma, z) = \sum_{\alpha \in I_{\mathcal{C}}(-d/c, i\infty)} \operatorname{sign}(\alpha w_\sigma - \alpha w'_\sigma) \left(\frac{1}{z - \alpha w_\sigma} - \frac{1}{z - \alpha w'_\sigma} \right) \tag{5.2.14}$$

Theorem 5.2.4. *For any $\gamma, \sigma \in \Gamma$, with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$*

$$r_{\mathcal{C}}(\sigma\gamma, z) = r_{\mathcal{C}}(\sigma, \gamma z)(cz+d)^{-2} + r_{\mathcal{C}}(\gamma, z) \tag{5.2.15}$$

Proof. To ease the notation the dependence on $\mathcal{C}$, which is fixed, is suppressed. As usual let T and S denote the two generators of Γ corresponding to the translation $z \to z+1$ and the inversion $z \to -1/z$ respectively. First note that $r(T\gamma, z) = r(\gamma, z)$. Hence if we prove

$$r(S\gamma, z) = r(S, \gamma z)(cz+d)^{-2} + r(\gamma, z) \tag{5.2.16}$$

the proposition follows by induction on the word length expressing σ in terms of the generators S and T. Recall that for $z, w \in \mathbb{C}$ and $\gamma \in \Gamma$

$$\frac{w-w'}{(\gamma z - w)(\gamma z - w')}(cz+d)^{-2} = \frac{\gamma^{-1}w - \gamma^{-1}w'}{(z-\gamma^{-1}w)(z-\gamma^{-1}w')} \tag{5.2.17}$$

Since $S\gamma = \left[\begin{smallmatrix} -c & -d \\ a & b \end{smallmatrix}\right]$ to prove (5.2.16), using (5.2.17) we have to prove that

$$\sum_{w'<-b/a<w} \left(\frac{1}{z-w} - \frac{1}{z-w'}\right) - \sum_{w'<-d/c<w} \left(\frac{1}{z-w} - \frac{1}{z-w'}\right) = \\ \sum_{w'<0<w} \left(\frac{1}{z-\gamma^{-1}w} - \frac{1}{z-\gamma^{-1}w'}\right) \tag{5.2.18}$$

all sums over pairs $(w', w) \in \mathcal{W}$.

Assume first that $ac > 0$ so that $-d/c < -b/a$. On the left hand side of (5.2.18) we have

$$\sum_{-d/c<w'<-b/a<w} \left(\frac{1}{z-w} - \frac{1}{z-w'}\right) - \sum_{w'<-d/c<w<-b/a} \left(\frac{1}{z-w} - \frac{1}{z-w'}\right) \tag{5.2.19}$$

On the other hand we can write for the right hand side of (5.2.18)

$$\sum_{w'<0<w} \left(\frac{1}{z-\gamma^{-1}w} - \frac{1}{z-\gamma^{-1}w'}\right) = \\ \sum_{w'<0<w<a/c} \left(\frac{1}{z-\gamma^{-1}w} - \frac{1}{z-\gamma^{-1}w'}\right) + \sum_{w'<0<a/c<w} \left(\frac{1}{z-\gamma^{-1}w} - \frac{1}{z-\gamma^{-1}w'}\right) \tag{5.2.20}$$

Now note that

$$\gamma^{-1}z = -\frac{d}{c} - \frac{1}{c^2(z-a/c)}$$

and the function $x \to -\frac{d}{c} - \frac{1}{c^2(x-a/c)}$ is monotonic for $x \in (-\infty, a/c)$ and also for $x \in (a/c, \infty)$.

It follows that for $w' < 0 < w < a/c$,

$$-d/c < \gamma^{-1}w' < -b/a < \gamma^{-1}w \tag{5.2.21}$$

and similarly that for $w' < 0 < a/c < w$

$$\gamma^{-1}w < -d/c < \gamma^{-1}w' < -b/a. \tag{5.2.22}$$

Using (5.2.21) and (5.2.22) in (5.2.20) we get that

$$\begin{aligned}
&\sum_{w'<0<w}\left(\frac{1}{z-\gamma^{-1}w}-\frac{1}{z-\gamma^{-1}w'}\right)\\
&\quad=\sum_{w'<0<w<a/c}\left(\frac{1}{z-\gamma^{-1}w}-\frac{1}{z-\gamma^{-1}w'}\right)+\sum_{w'<0<a/c<w}\left(\frac{1}{z-\gamma^{-1}w}-\frac{1}{z-\gamma^{-1}w'}\right)\\
&\quad=\sum_{-d/c<w'<-b/a<w}\left(\frac{1}{z-w}-\frac{1}{z-w'}\right)-\sum_{w'<-d/c<w<-b/a}\left(\frac{1}{z-w}-\frac{1}{z-w'}\right)
\end{aligned}\tag{5.2.23}$$

This proves (5.2.18) when $ac>0$. The case $ac<0$ follows in the same manner. The case $ac=0$ can be checked easily since $c=0$ corresponds to $\gamma=T^m$ whereas $a=0$ rises from $\gamma=ST^m$.

This proves Theorem 5.2.4 and hence also Theorem 5.1.1 from the introduction. □

Extending our earlier work we show that

Theorem 5.2.5. *For any hyperbolic conjugacy class $\mathcal{C}$ the function $F_{\mathcal{C}}(z)$ is holomorphic on $\mathcal{H}$ and satisfies*

$$(cz+d)^{-2}F_{\mathcal{C}}(\gamma z)=F_{\mathcal{C}}(z)+r_{\mathcal{C}}(\gamma,z)\tag{5.2.24}$$

Proof. The claim is trivial for T and has been established for the generator $S=\left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right]$ in [41]. It is possible to give a proof of the general case along the lines of the proof of (5.2.10) given in [41]. However the algebraic proof above already established that the rational function $r_{\mathcal{C}}(\gamma,z)$ defined in (5.1.9) is a weight 2 cocycle. Since it agrees with the cocycle associated to $F_{\mathcal{C}}(z)$ for the generators $\gamma=S=\left[\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right]$ and $T=\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]$ the difference is a 1-cocycle that vanishes on both S and T and so must vanish identically. □

5.2.3 The Dirichlet Series associated to $F_{\mathcal{C}}(z)$

Guided by the example of the Eisenstein series $E_2(z)$ and its primitive $\log\Delta(z)$, it is natural to study a primitive of a general modular integral, and the associated weight zero cocycle that appears in its transformation property.

We look at this problem in the case of the function $F_{\mathcal{C}}(z)$ and determine the unique weight 0 primitive $R_{\mathcal{C}}(\gamma,z)$ of the cocycles $r_{\mathcal{C}}(\gamma,z)$ in terms of the special values of the Dirichlet series $L(F_{\mathcal{C}},s,a/c)$.

The next theorem and its corollary, which are based on Theorem 5.2.1, proves Theorem 5.1.6 from the introduction.

Theorem 5.2.6. *Let $F_{\mathcal{C}}(z)$ be the modular integral in (5.2.8) and $L_{\mathcal{C}}(s,a/c):=L(F_{\mathcal{C}},s,a/c)$ be its associated Dirichlet series. Then $L_{\mathcal{C}}(s,a/c)$ converges for $\mathrm{Re}(s)>9/4$, has a meromorphic continuation to $s>0$ and is holomorphic at $s=1$. Moreover if $R_{\mathcal{C}}(\gamma,z)$ is the unique weight 0 cocycle such that $R'_{\mathcal{C}}(\gamma,z)=r_{\mathcal{C}}(\gamma,z)$ then*

$$R_{\mathcal{C}}(\gamma,z)=\varepsilon_{\mathcal{C}}\sum_{w<\frac{-d}{c}<w'}\log(z-w)-\log(z-w')+\tfrac{1}{2\pi i}L_{\mathcal{C}}(1,a/c)+a_{\mathcal{C}}(0)\left(\frac{a+d}{c}\right)\tag{5.2.25}$$

Proof. The convergence of $L_{\mathcal{C}}(s,a/c)$ for $\mathrm{Re}(s) > 9/4$ follows from the bound $a_{\mathcal{C}}(m) \ll m^{5/4+\epsilon}$ which was proved in Proposition 6 of [40].

To prove (5.2.25), in Theorem 5.2.1 we let $r(\gamma,z) = r_{\mathcal{C}}(\gamma,z) = \varepsilon_{\mathcal{C}} \sum_{w'<-d/c<w} \frac{1}{z-w} - \frac{1}{z-w'}$. As a primitive of $r_{\mathcal{C}}(\gamma,z)$ we choose

$$\varepsilon_{\mathcal{C}} \sum_{w'<-d/c<w} \log(z-w) - \log(z-w').$$

Once again using (5.2.2) and (5.2.3) we have

$$\begin{aligned} R_{\mathcal{C}}(\gamma,z) =& \frac{-i}{c} \lim_{s\to 1} \left[\left(\frac{2\pi}{c}\right)^{-s} \Gamma(s) L_{\mathcal{C}}(s,a/c) + \int_1^\infty r_{\mathcal{C}}(\gamma, -d/c + it/c) t^{1-s} dt \right] \\ &+ \int_{z_1}^{z} r_{\mathcal{C}}(\gamma,w)dw + a_{\mathcal{C}}(0)\left(\frac{a+d}{c}\right) \end{aligned} \tag{5.2.26}$$

where $z_1 = -d/c + i/c$.

Contrary to the case of E_2, the Dirichlet series $L_{\mathcal{C}}(s,a/c)$ has no pole at $s=1$. This is due to the fact that at $s=1$ the first integral in (5.2.26) has the finite value

$$\varepsilon_{\mathcal{C}} \sum_{w'<-d/c<w} \log(z_1-w) - \log(z_1-w').$$

To finish the proof of Theorem 5.2.6 we combine the two integrals in (5.2.26) to get

$$\begin{aligned} R_{\mathcal{C}}(\gamma,z) =& \frac{1}{2\pi i} L_{\mathcal{C}}(1,a/c) + \int_\infty^z r_{\mathcal{C}}(\gamma,z)dw + a_{\mathcal{C}}(0)\left(\frac{a+d}{c}\right) \\ =& \frac{1}{2\pi i} L_{\mathcal{C}}(1,a/c) + \varepsilon_{\mathcal{C}} \sum_{w'<-d/c<w} \log(z-w) - \log(z-w') + a_{\mathcal{C}}(0)\left(\frac{a+d}{c}\right). \end{aligned} \tag{5.2.27}$$

□

Since $a_{\mathcal{C}}(0) = \log\lambda$ is real, the following corollary easily follows from (5.2.25)

Corollary 5.2.7. *Let* $\Phi_{\mathcal{C}}(\gamma) = \frac{2}{\pi} \lim_{y\to\infty} \mathrm{Im}\, R_{\mathcal{C}}(\gamma, iy)$. *Then*

$$\Phi_{\mathcal{C}}(\gamma) = -\tfrac{1}{\pi^2} \mathrm{Re}\, L_{\mathcal{C}}(1,a/c).$$

In the rest of the section we will give two applications of Theorem 5.2.6 and the cocycle relation

$$R_{\mathcal{C}}(\sigma\gamma,z) = R_{\mathcal{C}}(\sigma,\gamma z) + R_{\mathcal{C}}(\gamma,z).$$

The first one is an analog of the Dedekind's reciprocity formula (5.1.5) for the Dirichlet series $L_{\mathcal{C}}(1,a/c)$. More precisely we have

Theorem 5.2.8. *Let* $(a,c)=1$ *and* $ac \neq 0$. *Then*

$$\frac{1}{i\pi}\left[L_{\mathcal{C}}(1,a/c) - L_{\mathcal{C}}(1,-c/a)\right] = -2\left(\tfrac{a^2+c^2+1}{ac}\right)\log\lambda - \nu_{\mathcal{C}}(a/c)) \tag{5.2.28}$$

where

$$\nu_{\mathcal{C}}(x) = \varepsilon_{\mathcal{C}} \sum_{w'<0<w} \left[\log\left(\frac{x-w}{x-w'}\right) - \log\left(\frac{1+xw}{1+xw'}\right)\right].$$

Here we interpret the imaginary part of the logarithm of a negative real number to be π.

Proof. Let $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ and $S = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. From (5.2.27) we have

$$R_{\mathcal{C}}(\gamma, z) = \frac{1}{2\pi i} L_{\mathcal{C}}(1, a/c) + \int_{i\infty}^{z} r_{\mathcal{C}}(\gamma, w)dw + a_{\mathcal{C}}(0)\left(\frac{a+d}{c}\right)$$

Since $R_{\mathcal{C}}(\gamma, z)$ is a cocycle it satisfies

$$R_{\mathcal{C}}(S\gamma, z) = R_{\mathcal{C}}(S, \gamma z) + R_{\mathcal{C}}(\gamma, z). \tag{5.2.29}$$

Hence

$$\begin{aligned} R_{\mathcal{C}}(S\gamma, z) =& \frac{1}{2\pi i} L_{\mathcal{C}}(1, -c/a) + \int_{i\infty}^{z} r_{\mathcal{C}}(S\gamma, w)dw + a_{\mathcal{C}}(0)\left(\frac{b-c}{a}\right) \qquad (5.2.30) \\ =& \frac{1}{2\pi i} L_{\mathcal{C}}(1, 0) + \int_{i\infty}^{\gamma z} r_{\mathcal{C}}(S, w)dw \\ &+ \frac{1}{2\pi i} L_{\mathcal{C}}(1, a/c) + \int_{i\infty}^{z} r_{\mathcal{C}}(\gamma, w)dw + a_{\mathcal{C}}(0)\left(\frac{a+d}{c}\right) \end{aligned}$$

We let $z \to i\infty$ to get

$$\begin{aligned} \frac{1}{2\pi i}\left[L_{\mathcal{C}}(1, -c/a) - L_{\mathcal{C}}(1, a/c)\right] =& a_{\mathcal{C}}(0)\left(\frac{a^2+c^2+1}{ac}\right) \qquad (5.2.31) \\ &+ \frac{1}{2\pi i} L_{\mathcal{C}}(1, 0) + \int_{i\infty}^{a/c} r_{\mathcal{C}}(S, w)dw \qquad (5.2.32) \end{aligned}$$

Hence

$$\begin{aligned} &\frac{1}{2\pi i}\left[L_{\mathcal{C}}(1, -c/a) - L_{\mathcal{C}}(1, a/c)\right] \\ &\qquad = a_{\mathcal{C}}(0)\left(\frac{a^2+c^2+1}{ac}\right) + \frac{1}{2\pi i} L_{\mathcal{C}}(1, 0) \\ &\qquad\quad + \varepsilon_{\mathcal{C}} \sum_{w'<0<w} \log\left(\frac{a}{c} - w\right) - \log\left(\frac{a}{c} - w'\right) \qquad (5.2.33) \end{aligned}$$

Now replacing the roles of $-c$ with a and a with c gives

$$\begin{aligned} &\frac{1}{2\pi i}\left[L_{\mathcal{C}}(1, a/c) - L_{\mathcal{C}}(1, -a/c)\right] \\ &\qquad = -a_{\mathcal{C}}(0)\left(\frac{a^2+c^2+1}{ac}\right) + \frac{1}{2\pi i} L_{\mathcal{C}}(1, 0) \\ &\qquad\quad + \varepsilon_{\mathcal{C}} \sum_{w'<0<w} \log\left(\frac{-c}{a} - w\right) - \log\left(\frac{-c}{a} - w'\right) \qquad (5.2.34) \end{aligned}$$

Finally noting that $a_{\mathcal{C}}(0) = \log \lambda$ and taking the difference of the last two equations prove (5.2.28). □

As a second application we have the following geometric interpretation of the special value of the Dirichlet series $L_{\mathcal{C}}(s, a/c)$.

Theorem 5.2.9. *Let $L_{\mathcal{C}}(s, a/c)$ be the Dirichlet series associated to $F_{\mathcal{C}}(z)$. Then*

$$\frac{1}{2\pi i}\left[L_{\mathcal{C}}(1, a/c) + L_{\mathcal{C}}(1, -d/c)\right] \tag{5.2.35}$$

$$= -\varepsilon_{\mathcal{C}} \sum_{w' < \frac{-d}{c} < w} \log(\tfrac{-d}{c} - w) - \log(\tfrac{-d}{c} - w') \tag{5.2.36}$$

$$= -\,\varepsilon_{\mathcal{C}}\left(2\log\left|\prod_{w' < \frac{-d}{c} < w} \tan\left(\frac{\theta_w}{2}\right)\right| + i\pi \sum_{w' < \frac{-d}{c} < w} 1\right) \tag{5.2.37}$$

where the sum and the product runs over elements $(w', w) \in \mathcal{W}_{\mathcal{C}}$ that are separated by $\frac{-d}{c}$. θ_w is the angle of intersection of the vertical line $(-d/c, -d/c + i\infty)$ with the semicircle with end points w' and w. Here θ_w is the angle containing the line segment connecting this intersection to w'.

Proof. Let $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Using the cocycle relation $0 = R_{\mathcal{C}}(\gamma, \gamma^{-1}z) + R_{\mathcal{C}}(\gamma^{-1}, z)$, the formula (5.2.25) and taking the limit as $z \to i\infty$ leads to the first equality (5.2.36). Since $-d/c, w, w'$ all lie on the real axis, the argument of each logarithm term in the sum in (5.2.36) is π. Here we interpret the imaginary part of the logarithm of a negative real number to be π. This proves that the imaginary part of (5.2.36) is indeed given by $\pi \sum_{w' < \frac{-d}{c} < w} 1$.

The fact that the real part (5.2.36) is given as in (5.2.37) follows easily using elementary geometry. (See also [9] p.116.)

□

5.2.4 Intersection numbers

In this section we restrict ourselves to the imaginary part of $R_{\mathcal{C}}(\gamma, z)$. Recall from (5.2.7) that

$$\Phi_{\mathcal{C}}(\gamma) = \tfrac{2}{\pi} \lim_{y \to \infty} \operatorname{Im} R_{\mathcal{C}}(\gamma, iy) = -\tfrac{1}{\pi^2} \operatorname{Re} L_{\mathcal{C}}(1, a/c).$$

Our first goal is to prove that $\Phi_{\mathcal{C}}(\gamma)$ is an intersection number, hence an integer.

We start by noting that Theorem 5.2.9 gives

$$\Phi_{\mathcal{C}}(\gamma) + \Phi_{\mathcal{C}}(\gamma^{-1}) = 2 \sum_{w' < \frac{-d}{c} < w} 1$$

and hence as a simple corollary we have

Proposition 5.2.10. *Let $\mathcal{C}$ be the conjugacy class of a hyperbolic element σ, $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ another hyperbolic element in Γ and $I_{\mathcal{C}}(\gamma^{-1}(i\infty), i\infty) = I_{\mathcal{C}}(-d/c, i\infty)$ be as defined in (5.2.12). Then*

$$\Phi_{\mathcal{C}}(\gamma) + \Phi_{\mathcal{C}}(\gamma^{-1}) = -2|I_{\mathcal{C}}(-d/c, i\infty)|$$

The next result shows that $\Phi_{\mathcal{C}}(\gamma)$ is already an integer.

Theorem 5.2.11. *Let $\gamma \in \Gamma$ be a hyperbolic element. Then $\Phi_{\mathcal{C}}(\gamma) = -|I_{\mathcal{C}}(-d/c, i\infty)|$ and hence $\Phi_{\mathcal{C}}(\gamma) \in \mathbb{Z}$.*

Proof. For $\gamma_1 = \left[\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right]$, $\gamma_2 = \left[\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right]$, two not necessarily hyperbolic elements of Γ, let

$$\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = \Phi_{\mathcal{C}}(\gamma_1\gamma_2) - \Phi_{\mathcal{C}}(\gamma_1) - \Phi_{\mathcal{C}}(\gamma_2).$$

Note that $I_{\mathcal{C}}(-d_1/c_1, i\infty) = I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty)$. We will show that

$$\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = |I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty)| + |I_{\mathcal{C}}(\gamma_2^{-1}i\infty, i\infty)| - |I_{\mathcal{C}}((\gamma_1\gamma_2)^{-1}i\infty, i\infty)|. \tag{5.2.38}$$

This will prove the theorem since this then $\gamma \mapsto \Phi_{\mathcal{C}}(\gamma) + |I_{\mathcal{C}}(\gamma^{-1}i\infty, i\infty)|$ is a homomorphism of Γ into $\mathbb{C}$ and so is identically 0.

First note that if either γ_1 or γ_2 is T^n for some $n \in \mathbb{Z}$ then $\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = 0$ and the identity holds trivially. So we assume that γ_1, γ_2 are not parabolic.

To prove (5.2.38) note that from definition (5.1.12) of $\Phi_C(\gamma)$ and the cocycle property we have

$$\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = \frac{2\varepsilon_{\mathcal{C}}}{\pi} \lim_{y\to\infty} \operatorname{Im}(R_{\mathcal{C}}(\gamma_1, \gamma_2 iy) - R_{\mathcal{C}}(\gamma_1, iy))$$

which by (5.2.27) equals

$$\frac{2\varepsilon_{\mathcal{C}}}{\pi} \lim_{y\to\infty} \left[\sum \arg\left(\frac{\gamma_2 iy - w}{\gamma_2 iy - w'}\right) - \sum \arg\left(\frac{iy - w}{iy - w'}\right)\right]$$

the sums are over $(w', w) \in \mathcal{W}_{\mathcal{C}}$, $w' < -d_1/c_1 < w$. The second sum in the limit clearly goes to zero. Since $\gamma_2 iy \to a_2/c_2$ when $y \to \infty$

$$\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = 2\varepsilon_{\mathcal{C}} n(\gamma_1^{-1}, \gamma_2) \tag{5.2.39}$$

where $n(\gamma_1^{-1}, \gamma_2)$ is the number of $(w', w) \in \mathcal{W}_{\mathcal{C}}$, for which $w' < -d_1/c_1, a_2/c_2 < w$. By the definition (5.2.12) we have

$$\varepsilon_{\mathcal{C}} n(\gamma_1^{-1}, \gamma_2) = |I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty) \cap I_{\mathcal{C}}(\gamma_2 i\infty, i\infty)|$$

Any geodesic that does not go through the vertices of an ideal hyperbolic triangle intersects exactly two sides of the triangle if it intersects the triangle at all. Applying this fact to the ideal hyperbolic triangle with vertices $i\infty, a_2/c_2 = \gamma_2 i\infty$ and $-d_1/c_1 = \gamma_1^{-1}i\infty$ shows that the sets

$$\begin{gathered} I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty) \cap I_{\mathcal{C}}(\gamma_2 i\infty, i\infty), \\ I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty) \cap I_{\mathcal{C}}(\gamma_2 i\infty, \gamma_1^{-1}i\infty) \text{ and} \\ I_{\mathcal{C}}(\gamma_2 i\infty, \gamma_1^{-1}i\infty) \cap I_{\mathcal{C}}(\gamma_2 i\infty, i\infty) \end{gathered}$$

are mutually disjoint. A standard inclusion exclusion argument then gives

$$\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = |I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty)| + |I_{\mathcal{C}}(\gamma_2 i\infty, i\infty)| - |I_{\mathcal{C}}(\gamma_1^{-1}i\infty, \gamma_2 i\infty)|$$

Finally we use that $|I_{\mathcal{C}}(z_2, z_1)| = |I_{\mathcal{C}}(z_1, z_2)| = |I_{\mathcal{C}}(\gamma z_1, \gamma z_2)|$ for all $\gamma \in \Gamma$ to establish that

$$\delta_{\mathcal{C}}(\gamma_1, \gamma_2) = |I_{\mathcal{C}}(\gamma_1^{-1}i\infty, i\infty)| + |I_{\mathcal{C}}(\gamma_2^{-1}i\infty, i\infty)| - |I_{\mathcal{C}}(\gamma_2^{-1}\gamma_1^{-1}i\infty, i\infty)|. \tag{5.2.40}$$

$\square$

Note that formula (5.2.39) for the co-boundary $\delta_{\mathcal{C}}$ of $\Phi_{\mathcal{C}}$ allows one to calculate $\Phi_{\mathcal{C}}(\gamma)$ successively by writing γ in terms of some set of generators of the group Γ. We give an alternative approach for establishing that $\Phi_{\mathcal{C}}$ takes integer values. This method does not identify $\Phi_{\mathcal{C}}$ geometrically but also gives a fast algorithm to compute it.

Note that since $L_{\mathcal{C}}(1, a/c)$ depends only on $a/c \mod 1$, so does $\Phi_{\mathcal{C}}(\gamma)$ and hence for $c \neq 0$ we can write $\Phi_{\mathcal{C}}(a/c) = \Phi_{\mathcal{C}}(\gamma)$. The following is a simple corollary of Theorem 5.2.9 and Corollary 5.2.7.

Lemma 5.2.12. *Let* $S = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. *Then*

$$\Phi_{\mathcal{C}}(0) = \Phi_{\mathcal{C}}(S) = -\varepsilon_{\mathcal{C}} \sum_{w'<0<w} 1$$

The following theorem is an analogue of Dedekind's reciprocity formula. It allows for, via Euclid's algorithm, a quick computation of $\Phi_{\mathcal{C}}(\gamma)$.

Theorem 5.2.13. *Let* $\mathcal{C}$ *be a hyperbolic conjugacy class and* $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. *For* $ac \neq 0$ *we have*

$$\Phi_{\mathcal{C}}(a/c) = \Phi_{\mathcal{C}}(-c/a) - \frac{1}{\pi} \operatorname{Im} \nu_{\mathcal{C}}(a/c) \tag{5.2.41}$$

Proof. The formula follows from Theorem 5.2.8 and Corollary 5.2.7. Note that our definition of the argument gives $\operatorname{Im} \log x = 0$ for a positive real number x, and $\operatorname{Im} \log x = \pi$ for a negative real number x. $\square$

Remark 5.2.14. Note that

$$\frac{1}{\pi} \operatorname{Im} \nu_{\mathcal{C}}(a/c) = -\epsilon_{\mathcal{C}} \sum_{w'<0<w} \operatorname{sign}\left(\frac{-\frac{c}{a} - w}{-\frac{c}{a} - w'}\right) - \operatorname{sign}\left(\frac{\frac{a}{c} - w}{\frac{a}{c} - w'}\right)$$

and

$$\operatorname{sign}\left(\frac{-\frac{c}{a} - w}{-\frac{c}{a} - w'}\right) - \operatorname{sign}\left(\frac{\frac{a}{c} - w}{\frac{a}{c} - w'}\right)$$

is non zero only for those $w' < 0 < w$ for which exactly one of $\{\frac{a}{c}, -\frac{c}{a}\}$ is in the open interval $(-w', w)$. Therefore once all the conjugates of $\sigma \in \mathcal{C}$ whose fixed points are separated by 0 are listed (an easy task, see Remark 5.2.3) the right hand side in the above theorem is an easily computable elementary function of $\frac{a}{c}$. This in turn allows a fast calculation of $\Phi_{\mathcal{C}}(a/c)$ in view of $\Phi_{\mathcal{C}}(\frac{a}{c}) = \Phi_{\mathcal{C}}(\frac{a+nc}{c})$ for any $n \in \mathbb{Z}$. Since $\Phi_{\mathcal{C}}(0)$ is an integer, it also establishes that $\Phi_{\mathcal{C}}(\frac{a}{c})$ is an integer.

We finish this section by collecting some results about the hyperbolic geometry that will be needed to prove Theorem 5.1.5 from the introduction, In particular it will be important for us to compare $|I_{\mathcal{C}_\sigma}(\gamma^{-1}z_0, z_0)|$ and $|I_{\mathcal{C}_\sigma}(\gamma^{-1}i\infty, i\infty)|$. We start with a simple lemma about hyperbolic quadrangles. Recall that for $z_1, z_2 \in \overline{\mathcal{H}}$ the geodesic segment connecting z_1 and z_2 is denoted by ℓ_{z_1,z_2}.

Lemma 5.2.15. *Let* $z_1, z_2 \in \mathcal{H}$ *and* $x_1, x_2 \in \partial\mathcal{H}$. *If* ℓ *is a geodesic that intersects neither the geodesic half line* ℓ_{z_1,x_1} *nor the geodesic half line* ℓ_{z_2,x_2} *then* ℓ *intersects either both* ℓ_{x_1,x_2} *and* ℓ_{z_1,z_2} *or it intersects neither of them.*

Proof. By applying a hyperbolic isometry we may assume that $\ell = \ell_{0,i\infty}$. The geodesic arc from z_1 to x_1 does not intersect $\ell = (0, i\infty)$, so x_1 and $\mathrm{Re}(z_1)$ have the same sign. Similarly the geodesic arc from z_2 to x_2 does not intersect $(0, i\infty)$, so x_2 and $\mathrm{Re}(z_2)$ have the same sign. Finally the arc from z_1 to z_2 intersects $(0, i\infty)$ if and only if their real parts have opposite signs. This proves that ℓ either intersects both the arc from z_1 to z_2 and the geodesic from x_1 to x_2 or that intersects neither of them.

□

Proposition 5.2.16. *Let σ, γ be hyperbolic elements, and fix a point $z_0 \in S_\gamma$. Then*

$$\big|\,|I_{\mathcal{C}_\sigma}(\gamma^{-1}z_0, z_0)| - |I_{\mathcal{C}_\sigma}(\gamma^{-1}i\infty, i\infty)|\,\big| \le 2|I_{\mathcal{C}_\sigma}(z_0, i\infty)|. \tag{5.2.42}$$

Note that we do not assume γ to be primitive.

Proof. Let $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Consider the geodesic circular arc L_1 connecting $\gamma^{-1}z_0$ to $\gamma^{-1}i\infty = -d/c$ and the half-line L_2 connecting z_0 to $i\infty$. Assume that αS_σ intersects neither L_1 nor L_2. Then it follows from Lemma 5.2.15 that either αS_σ intersects both the arc from z_0 to $\gamma^{-1}z_0$ and the line from $-d/c$ to $i\infty$ or αS_σ intersects neither of them.

Hence we have shown that the symmetric difference of the sets $I_{\mathcal{C}}(-d/c, i\infty)$ and $I_{\mathcal{C}}(z_0, \gamma^{-1}z_0)$ is a subset of $I_{\mathcal{C}}(z_0, i\infty) \cup I_{\mathcal{C}}(-d/c, \gamma^{-1}z_0)$;

$$I_{\mathcal{C}}(-d/c, i\infty) \triangle I_{\mathcal{C}}(z_0, \gamma^{-1}z_0) \subset I_{\mathcal{C}}(z_0, i\infty) \cup I_{\mathcal{C}}(-d/c, \gamma^{-1}z_0,)$$

Since

$$\big|\,|I_{\mathcal{C}}(z_0, \gamma^{-1}z_0)| - |I_{\mathcal{C}}(-d/c, i\infty)|\,\big| \le |I_{\mathcal{C}}(-d/c, i\infty) \triangle I_{\mathcal{C}}(z_0, \gamma^{-1}z_0)|$$

and $I_{\mathcal{C}}(z_0, i\infty)$ and $I_{\mathcal{C}}(-d/c, \gamma^{-1}z_0)$ have the same cardinality $|I_{\mathcal{C}_\sigma}(z_0, i\infty)|$ this proves the proposition.

□

5.2.5 Linking numbers in $\Gamma\backslash SL_2(\mathbb{R})$

In this section we prove Theorem 5.1.5. This is based on results of the previous section and a theorem of Birkhoff [10].

If γ is a primitive hyperbolic element such that $\operatorname{tr}\gamma > 2$ there is an associated closed periodic orbit of the geodesic flow whose linking number with the trefoil is given by the Rademacher symbol (see [6], [7],[52]).

$$\Psi(\gamma) = \Phi(\gamma) - 3\,\mathrm{sign}\,c(a+d) = \lim_{n\to\infty} \frac{\Phi(\gamma^n)}{n}.$$

For the convenience of the reader we sketch Ghys' argument for the identification of the Rademacher symbol with the linking number with the trefoil in the Appendix.

Our goal in this section is to provide the background for a similar interpretation for the homogenization of $\Phi_{\mathcal{C}}(\gamma)$ of Theorem 5.1.5,

$$\Psi_{\mathcal{C}}(\gamma) := \lim_{n\to\infty} \frac{\Phi_{\mathcal{C}}(\gamma^n)}{n}$$

as a linking number.

As alluded above this is based on Theorem 5.2.19, originally due to Birkhoff, (cf. [10]) which relates this linking number to the geometry of the net $\mathcal{N}_\sigma$ of a primitive hyperbolic element $\sigma \in \mathcal{C}$. Birkhoff's theorem [10, Section 27] which proves the existence of a certain surface bounding symmetric curves which is a surface of section of the geodesic flow, is more general than what is needed for us. This theorem was popularized by Fried [49] who named them Birkhoff sections. The theorem holds in even more generality as shown in [1, 2, 58, 70]. As is clear from this rich history there are a number of proofs of this theorem especially for compact hyperbolic surfaces (see e.g. [22] and the references therein, also [32] and esp. section 3 of [33]). For the convenience of the reader we also give one which is self contained and very elementary; it is based on a simple computation of the sign of the triple product of three vectors in the Lie-algebra $\mathfrak{sl}_2(\mathbb{R})$, (Proposition 5.2.18). The relation to the invariant $\Psi_{\mathcal{C}}(\gamma)$ follows from a careful book-keeping of potential multiplicities (Lemmas 5.2.21, 5.2.22, and 5.2.23).

To make this explicit note that if $\gamma \in \Gamma$ has $\operatorname{tr}\gamma > 2$ and fixed points $w' < w$ then both γ and γ^{-1} are diagonalized by $M = \frac{1}{\sqrt{w-w'}}\left[\begin{smallmatrix} w & w' \\ 1 & 1\end{smallmatrix}\right]$. By replacing γ with γ^{-1} we may assume that

$$\gamma M = M \begin{bmatrix} \varepsilon & 0 \\ 0 & 1/\varepsilon \end{bmatrix}$$

where $\varepsilon > 1$. When $a + d > 2$ this is equivalent to $\operatorname{sign} c > 0$. Both

$$\tilde{\gamma}_+(t) = M\phi(t) \text{ and } \tilde{\gamma}_-(t) = MS\phi(t)$$

are periodic orbits of the geodesic flow $g \mapsto g\phi(t)$ on $\Gamma\backslash SL_2(\mathbb{R})$. Here $\phi(t) = \left[\begin{smallmatrix} e^t & 0 \\ 0 & e^{-t}\end{smallmatrix}\right]$.

We now move on to interpret linking numbers combinatorially as intersection numbers. Let $[\tilde{\gamma}_+]$ and $[\tilde{\gamma}_-]$ be the homology class of the curves $t \mapsto M\phi(t)$, $t \in [0, \log\varepsilon]$ and $t \mapsto MS\phi(t)$, $t \in [0, \log\varepsilon]$, respectively. Note that $\tilde{\gamma}_+(t)i$, $t \in [0 \log\varepsilon]$ maps into a geodesic arc in $\mathcal{H}$ connecting Mi to γMi on the semicircle with endpoints w and w'. On the quotient space $\Gamma\backslash\mathcal{H}$ this is a closed geodesic, and $\tilde{\gamma}_-(t)i$ simply travels this closed geodesic backwards. The natural Seifert surface bounding $[\tilde{\gamma}_+]$ and $[\tilde{\gamma}_-]$ is just formed by the collection of unit tangent vectors rotating counterclockwise continuously through 180 degrees from the one orientation of the circle to the other. This is the geometric content of the following

Lemma 5.2.17. $[\tilde{\gamma}_+] + [\tilde{\gamma}_-]$ *is null-homologous in* $\Gamma\backslash SL_2(\mathbb{R})$.

Proof. In fact we even have that $M\phi(t)$ and $MS\phi(-t)$ are homotopic via

$$\begin{aligned} h:\ & [0,\log\varepsilon]\times[0,\pi/2] && \to G \\ & (t,\theta) && \mapsto M\phi(t)k(\theta) \end{aligned}$$

where as usual

$$k(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

□

Note the image of h is an immersed sub-manifold X_γ in the quotient space $\Gamma\backslash SL_2(\mathbb{R})$. This follows readily from the fact that $\phi(t_1)k(\theta_1) = \phi(t_2)k(\theta_2)$, for $\theta_i \in [0, \pi/2]$ implies $t_1 = t_2$, $\theta_1 = \theta_2$ and so the image of h when viewed in $SL_2(\mathbb{R})$ is an embedded submanifold.

Now assume that $\mathcal{C}_\sigma$ and $\mathcal{C}_\gamma$ are two (different) primitive conjugacy classes. The above construction of the null-homologous chains associated to σ, γ have a well-defined linking number [54], [90] which we denote by $Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma)$. (This is well defined as the chains themselves depend only on the conjugacy class.) A geometric interpretation of this linking number between the trivial homology class $[\tilde{\gamma}_+] + [\tilde{\gamma}_-]$ and $[\tilde{\sigma}_+] + [\tilde{\sigma}_-]$ is given as the number of signed intersections of X_γ (the surface defined above by the homotopy map h) and $\tilde{\sigma}_+(s)$ and $\tilde{\sigma}_-(s)$, $s \in [0, \log \lambda]$, the closed orbits associated to σ. The geodesic flow has the interesting property that all intersections of X_γ and $\tilde{\sigma}_+$ have the same sign.

We fix the sign by fixing an orientation as follows. We think of $SL_2(\mathbb{R})$ as a subspace of the space of real 2×2 matrices. The tangent space at the identity is the set of 2×2 real matrices with trace 0 where we fix the basis (see [68] pg 27)

$$\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \text{ and } \mathbf{h} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and we say the orientation of three tangent vectors tangent to $SL_2(\mathbb{R})$ at g is positive, i.e. three matrices v_1, v_2, v_3 are positively oriented if $g^{-1}v_1, g^{-1}v_2, g^{-1}v_3$, are positively oriented at the identity. We then have the following proposition.

Proposition 5.2.18. *Let* $N = \frac{1}{\sqrt{w_\sigma - w'_\sigma}} \left[\begin{smallmatrix} w_\sigma & w'_\sigma \\ 1 & 1 \end{smallmatrix} \right]$, *where* w_σ, w'_σ *are the two fixed points of* σ. *Assume that the trajectory* $N\phi(s)$ *is disjoint from* $[\tilde{\gamma}_+] + [\tilde{\gamma}_-]$ *and intersects* $X_{\tilde{\gamma}}$ *at a point* g. *Then the sign of the intersection is negative.*

Proof. Let

$$g = M\phi(t)k(\theta) = N\phi(s).$$

To compute the sign of the intersection we have to compute the determinant of the coefficient matrix of the tangent vectors

$$g^{-1}M\phi'(t)k(\theta), \quad g^{-1}M\phi(t)\kappa'(\theta) \text{ and } g^{-1}N\phi'(s).$$

Since $\phi'(t) = \phi(t)\mathbf{h}$ and $\kappa'(\theta) = \kappa(\theta)(\mathbf{y} - \mathbf{x})$ we have

$$g^{-1}M\phi'(t)k(\theta) = k(-\theta)\mathbf{h}k(\theta),$$

$$g^{-1}M\phi(t)k'(\theta) = (\mathbf{y} - \mathbf{x}),$$

and

$$g^{-1}N\phi'(s) = \mathbf{h}.$$

Since $k(-\theta) \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix} \right] k(\theta) = \left[\begin{smallmatrix} \cos 2\theta & -\sin 2\theta \\ -\sin 2\theta & -\cos 2\theta \end{smallmatrix} \right] = -\sin 2\theta \mathbf{x} - \sin 2\theta \mathbf{y} + \cos 2\theta \mathbf{h}$ the value of the determinant we need to compute is $-2\sin 2\theta$, always negative since $\theta \in (0, \pi/2)$.

□

An immediate consequence of Proposition 5.2.18 is the following theorem.

Theorem 5.2.19. *Let* $M = \frac{1}{\sqrt{w_\gamma - w'_\gamma}} \left[\begin{smallmatrix} w_\gamma & w'_\gamma \\ 1 & 1 \end{smallmatrix} \right]$, $N = \frac{1}{\sqrt{w_\sigma - w'_\sigma}} \left[\begin{smallmatrix} w_\sigma & w'_\sigma \\ 1 & 1 \end{smallmatrix} \right]$, *with* $\{w_\gamma, w'_\gamma\}$ *and* $\{w_\sigma, w'_\sigma\}$, *the fixed points of* γ *and* σ *respectively so that*

$$\gamma M = M \begin{bmatrix} \varepsilon & 0 \\ 0 & 1/\varepsilon \end{bmatrix}, \quad \sigma N = N \begin{bmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{bmatrix}$$

and let

$$A = \{(s,t,\theta) \in [0,\log\lambda) \times [0,\log\varepsilon) \times [0,\pi/2) : \exists\alpha \in \Gamma, M\phi(t)k(\theta) = \alpha N\phi(s)\} \quad (5.2.43)$$

and

$$B = \{(s,t,\theta) \in [0,\log\lambda) \times [0,\log\varepsilon) \times [0,\pi/2) : \exists\alpha \in \Gamma, M\phi(t)k(\theta) = \alpha NS\phi(s)\}. \quad (5.2.44)$$

For the linking number we have

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = -|A| - |B|.$$

Proof. By definition each point in the set A corresponds to an intersection of the surface X_γ with the curve $[\tilde{\sigma}^+]$ and similarly points in B correspond to intersections of X_γ with the curve $[\tilde{\sigma}^-]$. Hence for the linking number, using Proposition 5.2.18, we have

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = -|A| - |B|$$

which proves Theorem 5.2.19. □

Note that it is natural to interpret (see for example [32]) the elements of A as values $\{(s,t) \in [0,\log\lambda) \times [0,\log\varepsilon) : M\phi(t)i = N\phi(s)i \in \Gamma\backslash\mathcal{H}\}$, i.e. the number of intersections of the closed geodesics in $\Gamma\backslash\mathcal{H}$ associated to γ, σ, and similarly for B, since each time the underlying path of σ in H crosses the underlying curve of γ, precisely one of its two lifts will intersect the Seifert surface. The proper interpretation of this geometric idea requires care due to both multiplicities arising from self-intersections and the presence of elliptic elements in $\Gamma = SL_2(\mathbb{Z})$. To avoid these complications we go directly to $|I_\mathcal{C}(z_0, \gamma z_0)|$ which counts the group elements in $I_\mathcal{C}(z_0, \gamma z_0)$. In this notation Birkhoff's theorem takes the following form:

Theorem 5.2.20 (Birkhoff)**.** *If we let $z_0 = Mi \in S_\gamma$ then*

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = -|I_{\mathcal{C}_\sigma}(z_0, \gamma z_0)|$$

The theorem will follow from a series lemmas relating $|A| + |B|$ to $|I_{\mathcal{C}_\sigma}(z_0, \gamma z_0)|$.

Lemma 5.2.21. *For A, B as in* (5.2.43), (5.2.44) *we have $A \cap B = \emptyset$.*

Proof. Recall that each point in A, (resp in B) corresponds to an intersection of X_γ with the curve $\tilde{\sigma}_+$ (resp $\tilde{\sigma}_-$).

Assume that $(s,t,\theta) \in A \cap B$ with $M\phi(t)k(\theta) = \alpha N\phi(s)$ and $M\phi(t)k(\theta) = \beta NS\phi(s)$ for some $\alpha, \beta \in \Gamma$. It follows that $\beta^{-1}\alpha = NSN^{-1}$. Recall that $N = \frac{1}{\sqrt{w_\sigma - w'_\sigma}} \left[\begin{smallmatrix} w_\sigma & w'_\sigma \\ 1 & 1 \end{smallmatrix}\right]$, where w_σ, w'_σ are the two fixed points of σ. Now a simple matrix multiplication shows that the matrix NSN^{-1} cannot have integer entries, contradicting $\beta^{-1}\alpha \in \mathrm{SL}(2,\mathbb{Z})$. Hence $A \cap B = \emptyset$. □

dc_1553_18

Lemma 5.2.22. *There is a bijection between B in* (5.2.44) *and*

$$B' = \{(s,t,\theta) \in [0, \log\lambda) \times [0, \log\varepsilon) \times [\pi/2, \pi) : \exists\alpha \in \Gamma, M\phi(t)k(\theta) = \alpha N\phi(s)\}$$

given for $s \neq 0$ by

$$(s,t,\theta) \mapsto (\log\lambda - s, t, \theta + \pi/2).$$

and for $s = 0$ by

$$(0,t,\theta) \mapsto (0,t,\theta + \pi/2)$$

Proof. Assume $(s,t,\theta) \in B$. The case $s = 0$ is trivial and otherwise $\exists\alpha \in \Gamma$ such that

$$M\phi(t)k(\theta) = \alpha NS\phi(s).$$

Since $\sigma N = N\phi(\log\lambda)$

$$M\phi(t)k(\theta) = \alpha\sigma^{-1}N\phi(\log\lambda - s)S.$$

This gives the claim since $S^{-1} = -k(\pi/2)$. □

Lemma 5.2.23. *There is a bijection between the set $A \cup B'$ and $I_{\mathcal{C}_\sigma}(z_0, \gamma z_0)$ and hence $|A \cup B'| = |I_{\mathcal{C}}(z_0, \gamma z_0)|$.*

Proof. We define a map

$$f : A \cup B' \to \Gamma/\Gamma_\sigma \tag{5.2.45}$$

$$(s,t,\theta) \mapsto \alpha\Gamma_\sigma. \tag{5.2.46}$$

Here α is the unique element in Γ given by

$$M\phi(t)k(\theta)\phi(-s)N^{-1} = \alpha. \tag{5.2.47}$$

To see that f is injective let $f(s,t,\theta) = f(s',t',\theta')$ with $M\phi(t)k(\theta)\phi(-s)N^{-1} = \alpha$ and $M\phi(t')k(\theta')\phi(-s')N^{-1} = \beta$. Then $\alpha\sigma^k = \beta$ for some $k \in \mathbb{Z}$. Hence

$$\phi(t)k(\theta)\phi(-s)N^{-1}\sigma^k N = \phi(t')k(\theta')\phi(-s').$$

Since $N^{-1}\sigma^k N = \phi(k\log\lambda)$ we have

$$\phi(t - t')k(\theta)\phi(k\log\lambda - s + s') = k(\theta').$$

Now a simple matrix multiplication shows that this equality holds only if $(s,t,\theta) = (s',t',\theta')$, proving the injectivity of f.

To show that $f(s,t,\theta) \in I_{\mathcal{C}}(z_0, \gamma z_0)$, let $(s,t,\theta), \alpha$ be such that

$$M\phi(t)k(\theta) = \alpha N\phi(s).$$

Now $M\phi(t)i$ is in $\mathcal{A}_\gamma$, the geodesic arc connecting $z_0 = Mi$ and γz_0 where as $N\phi(s)i$ is in S_σ and hence $\alpha\Gamma_\sigma \in I_{\mathcal{C}}(z_0, \gamma z_0)$.

Finally to see that this map is onto $I_{\mathcal{C}_\sigma}(z_0, \gamma z_0)$, let α be such that $\alpha\Gamma_\sigma \in I_{\mathcal{C}_\sigma}(z_0, \gamma z_0)$ so that there is $z \in S_\sigma$ for which $\alpha z \in A_\gamma$, and so $\alpha z = M\phi(t)i$ for some $t \in [0, \log\varepsilon)$, and also $z = \sigma^k N\phi(s)i$ for some $s \in [0, \log\lambda)$. Since the stabilizer of i in $SL_2(R)$ is $SO(2)$, there exists $\theta \in [0, 2\pi)$, such that

$$M\phi(t)k(\theta) = \alpha\sigma^k N\phi(s).$$

Replacing α by $-\alpha$ if necessary we may assume that $\theta \in [0, \pi)$ proving surjectivity. □

Proof of the Theorem 5.2.20.

By Birkhoff's theorem for the linking number we have

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = -|A| - |B|.$$

By Lemma 5.2.21 , $A \cap B = \emptyset$ and we have $Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = -|A \cup B|$. Finally by Lemma 5.2.22 and Lemma 5.2.23, $|A \cup B| = |I_{\mathcal{C}_\sigma}(z_0, \gamma z_0)|$.

This finishes the proof of the Theorem 5.2.20. □

We are now ready to prove

Theorem 5.2.24. *Let $\mathcal{C}_\sigma$ and $\mathcal{C}_\gamma$ be different primitive conjugacy classes. Then*

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = \Psi_{\mathcal{C}_\sigma}(\gamma)$$

Proof. By Theorem 5.2.20 we have

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_{\gamma^n}) = -|I_{\mathcal{C}}(z_0, \gamma^n z_0)|$$

and by Theorem 5.1.7

$$\Phi_{\mathcal{C}}(\gamma^n) = -|I_{\mathcal{C}}(\gamma^{-n} i\infty, i\infty)|$$

Clearly $I_{\mathcal{C}}(z_0, \gamma^{-n} z_0) = I_{\mathcal{C}}(z_0, \gamma^n z_0)$ and hence

$$|nLk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) - \Phi_{\mathcal{C}_\sigma}(\gamma^n)| = |\,|I_{\mathcal{C}}(z_0, \gamma^{-n} z_0)| - |I_{\mathcal{C}}(\gamma^{-n} i\infty, i\infty)|\,|$$

Now using Proposition 5.2.16 we have

$$|Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) - \frac{\Phi_{\mathcal{C}_\sigma}(\gamma^n)}{n}| \le \frac{2|I_{\mathcal{C}}(z_0, i\infty)|}{n}$$

Since $|I_{\mathcal{C}}(z_0, i\infty)|$ is independent of n this proves

$$Lk(\mathcal{C}_\sigma, \mathcal{C}_\gamma) = \lim_{n\to\infty} \frac{\Phi_{\mathcal{C}_\sigma}(\gamma^n)}{n} = \Psi_{\mathcal{C}_\sigma}(\gamma).$$

□

Chapter 6

Equidistribution of roots of quadratic congruences to prime moduli

6.1 Background and statements of results

Let $F(x) \in \mathbb{Z}[x]$ be a primitive, irreducible polynomial. By Lagrange's theorem the congruence

$$F(x) \equiv 0 \bmod p$$

cannot have more than $\deg F$ solutions, when p is prime. A natural question is to investigate the distribution of the roots among the various congruence classes. When $\deg F$ is at least 2, and irreducible, one conjectures that for any $0 \leq a < b < 1$ the frequency of roots $\nu \bmod p$ satisfying $a \leq \nu/p < b$ approaches $b - a$ as p runs through all prime numbers.

The main result of the paper is the proof of this conjecture for quadratic polynomials. As a corollary we get a similar statement about the equidistribution of the angles of the Salié sum $S(m, n; p)$ (defined below) as p runs through primes.

Our main theorem and the corollary extends a result of Duke, Friedlander, and Iwaniec [37]. They obtained the same result under the assumption that the discriminant of F is negative (or $mn < 0$ in the corollary).

We now proceed to give more precise statements of the results and an overview of the history of the problem. A sequence $x_n \in [0, 1]$ is said to be uniformly distributed with respect to Lebesgue measure (or simply uniformly distributed) if

$$\lim_{N\to\infty} \frac{\#\{1 \leq n \leq N : a < x_n < b\}}{N} = b - a$$

An equivalent requirement is that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i k x_n} = 0$$

for all $k \neq 0$. This criterion, which is due to Weyl [15], is the most convenient to use in practice.

Since the equidistribution, or lack of it, depends on the sequence and not on the set of points $\{x_n : n \in N\} \subset [0, 1]$, we make the following natural choice of ordering of the set

$$X = \{(n, \nu) : n \leq x, F(\nu) \equiv 0 \bmod n, 0 \leq \nu < n\}.$$

Let $(n, \nu) < (m, \mu)$ if either $n < m$ or $n = m$, $\nu < \mu$. In view of the estimate

$$\#\{\nu : F(\nu) \equiv 0 \bmod n, 0 \leq \nu < n\} \ll n^\epsilon$$

the ordering of the roots for a fixed modulus n is unimportant. When we speak of the distribution properties of the roots of the congruence $F(x) \equiv 0$ we are actually interested in the distribution of the sequence of points $\{\nu/n \in [0, 1] : (n, \nu) \in X\}$ (with the above ordering). Similarly, we consider the subsequence $\{\nu/p : (p, \nu) \in X,\ p \text{ prime}\}$. We simply say that (these sequences of) the roots are ordered by their denominators.

For linear functions the roots are not uniformly distributed (see Section 6.2.2. When $\deg F$ is at least 2, Hooley [67] showed that at least if the moduli are not restricted to primes but are allowed to take any positive integer as a value, the sequence ν/n is uniformly distributed when ordered by denominators.

Hooley's result is established via Weyl's criterion through the estimation of certain exponential sums. For congruences of degree 2, Hooley's techniques lead to better estimates (see [67]) but still do not yield an easy extension to prime moduli. However, for quadratic polynomials of negative discriminant, Duke, Friedlander, and Iwaniec [37] have succeeded in proving that the sequence ν/p arising from the solutions of the quadratic congruence is uniformly distributed in $[0, 1]$.

Both of the above-mentioned results as well as this work are based on estimates of sums of Kloosterman sums defined below. As is usual in analytic number theory, we use the notation $e(x) = e^{2\pi i x}$.

Definition 6.1.1. The Kloosterman sum is

$$K(m, n; p) = \sum_{xy \equiv 1(c)} e\left(\frac{mx + ny}{c}\right)$$

By Weil's famous result [128] there is an angle $0 \leq \theta p \leq \pi$ so that

$$K(m, n; p) = \sqrt{p}(e^{i\theta_p} + e^{-i\theta_p})$$

and so $|K(m, n; p)| \leq 2p^{1/2}$. Some elementary transformation properties of $K(m, n; c)$ then lead to

$$K(m, n; c) \leq (m, n, c)^{1/2} c^{1/2} \tau(n)$$

where $\tau(n)$ is the number of positive divisors of n. Since (2) uses the Riemann hypothesis for curves over finite fields, it is somewhat surprising that, for the Salié sum defined by

$$S(m, n; c) = \sum_{xy \equiv 1(c)} \left(\frac{x}{c}\right) e\left(\frac{mx + ny}{c}\right)$$

an estimate of the same quality is elementary in view of the identity (see [124]).

Theorem 6.1.2 ([124])**.** *The Salie sum is*

$$S(m, n; p) = \epsilon_p \left(\frac{n}{p}\right) \sqrt{p} \sum_{y^2 \equiv 4mn(p)} e\left(\frac{y}{p}\right).$$

(Here $\epsilon_p = 1$ *or* i*, depending on whether* $p \equiv 1$ *or* $-1 \mod 4$ *and* $\left(\frac{n}{p}\right)$ *is the Legendre symbol.)*

Our main result is the following.

Theorem 6.1.3. *Let $P(x) = Ax^2 + Bx + C$ be such that $B^2 - 4AC$ is not a square. Let S be an arithmetic progression that contains infinitely many primes. Then, as p runs through the prime numbers in S, the roots of the congruence*

$$P(\nu) \equiv 0 \mod p$$

are uniformly distributed with respect to Lebesgue measure.

Corollary 6.1.4. *Fix m, n so that mn is not a square. As p runs through the set of prime numbers in S the angles of the Salié sum $S(m, n; p)$ are uniformly distributed with respect to Lebesgue measure.*

(This is again in contrast with Kloosterman sums that are conjectured to follow a Sato-Tate distribution.)

The results of this paper are based on sieve results of [3] that we summarize in Section 7.2.2. What needs to be done is essentially an estimation of certain exponential sums, which can be transformed into sums of Kloosterman sums in the spirit of Hooley's work. This is the content of Section 7.2.5.

Although the possibility of this approach was suggested already in [67], two obstacles arise for an F with positive discriminant. One is the presence of the infinite group of automorphs U. This is handled using a U-invariant partition of unity.

Another less conceptual difficulty is the appearance of certain exponential sums whose expression in terms of classical Kloosterman sums would require considerable effort. Since the basic idea is to use Lagrange's presentation of a root of $F(\nu) \equiv 0 \pmod{p}$ in terms of quadratic forms, we overcome both of these obstacles by working on the group $SL(2, \mathbb{Z})$ (and on certain congruence subgroups) instead of the homogeneous space of quadratic forms. Since our estimations would require transferring all the sums of Kloosterman sums into this form anyway, this approach bypasses a significant amount of simple but tedious calculations.

The Kloosterman sums on congruence subgroups are reviewed in Section 7.2.4, where the estimation of the sums of Kloosterman sums in question is achieved along the lines of [34]. The estimates are completed in last section.

6.2 Proofs

6.2.1 Salie's identity

There are a number of proofs for Theorem 6.1.2 see e.g. [107, 129]. We present one from [124] that is shortest. Recall that

$$S(m, n; p) = \sum_{x\overline{x} \equiv 1(p)} \left(\frac{x}{p}\right) e\left(\frac{mx + n\overline{x}}{p}\right)$$

We may assume $(p, mn) = 1$, otherwise the sum is trivial. Then $S(m, n; p) = \left(\frac{n}{p}\right) S(mn, 1; p)$. Consider now

$$\sum_{y^2 \equiv mn(p)} e\left(\frac{2y}{p}\right).$$

We may write this as

$$\sum_{y \bmod p} \frac{1}{p} \sum_{x \bmod p} e\left(\frac{2y}{p}\right) e\left(\frac{x(y^2-mn)}{p}\right).$$

Interchanging the two sums leads to 0 when $x \equiv 0 \bmod p$ and otherwise to a Gauss sum

$$\sum_{y \bmod p} e\left(\frac{xy^2+2y}{p}\right) = \sum_{y \bmod p} e\left(\frac{\overline{x}(x^2y^2+2xy)}{p}\right) = \left(\frac{x}{p}\right) e\left(\frac{-\overline{x}}{p}\right) \sum_{t \bmod p} e\left(\frac{t^2}{p}\right)$$

whose evaluation is well known and this immediately leads to

$$S(mn,1;p) = \epsilon_p \left(\frac{n}{p}\right) \sqrt{p} \sum_{y^2 \equiv 4mn(p)} e\left(\frac{y}{p}\right)$$

(Here $\epsilon_p = 1$ or i, depending on whether $p \equiv 1$ or $-1 \mod 4$ and $\left(\frac{n}{p}\right)$ is the Legendre symbol.)

6.2.2 Criterion for the uniform distribution of roots of congruences

In order for the interested reader to develop some feel for the subject, we briefly mention the case of a linear function $F(x) = ax+b$. For an integer a, $\overline{a}$ stands for the multiplicative inverse of a to a modulus whose value should always be clear from the context. Now, assuming $(p \nmid a) = 1$, the only solution of $ax \equiv b \bmod p$ is $\overline{a}b \bmod p$, and so

$$\sum e\left(\frac{k\nu}{p}\right) = \sum e\left(\frac{k\overline{a}b}{p}\right) + O(1).$$

In view of

$$\frac{\overline{a}}{p} + \frac{\overline{p}}{a} - \frac{1}{ap} \in \mathbb{Z}$$

the second sum above can be transformed to

$$\sum_{p \le x} e\left(\frac{k\overline{a}b}{p}\right) = \sum_{p \le x} e\left(\frac{kb\overline{p}}{a}\right) + O(\log\log x).$$

Since the right-hand side depends only on the values of $p \bmod a$, the estimation of the above sums is equivalent to the prime number theorem for arithmetic progressions. However, this case is uncharacteristic; the roots tend to accumulate around the $\phi(a)$ points of $\{k/a \in [0,1] : gcd(k,a) = 1\}$, and this shows that the conditions of Theorem 1.2 are necessary.

Higher-order polynomials present some difficulties because the natural ordering of the roots of the polynomial congruences mod p cannot be enumerated using an explicit expression. However, for $S = \{n \equiv n_0 \bmod s\}$ and $S(x) = \{n \in S : 0 < n \le x\}$, we have the following asymptotic:

$$\#(p,\nu) : p \in S(x), F(\nu) \equiv 0 \bmod p\} \sim \frac{C_S x}{\log x}$$

where the constant C_S, whose value depends on the field extension $K = Q(\zeta, z)$, is never zero (ζ is a primitive s-th root of unity, and z is any root of $F(x)$). When $\deg F = 2$, the result is an easy corollary of quadratic reciprocity and the prime number theorem for arithmetic progressions. (In general one needs the Chebotarev density theorem for the normal closure of the field K.)

Weyl's criterion now takes the following form: the roots of the polynomial congruence $F(x) \equiv 0 \bmod p$ give rise to a sequence that is uniformly distributed with respect to Lebesgue measure if and only if, for every $k \neq 0$,

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sum_{p\in S(x)} \sum_{F(\nu)\equiv 0\,(n)} e\left(\frac{k\nu}{p}\right).$$

To simplify notation, we follow [67] and [37] and introduce

$$\rho_k(n) = \sum_{F(\nu)\equiv 0\,(n)} e\left(\frac{k\nu}{n}\right)$$

Thus, to verify Weyl's criterion, we need to estimate $\sum \chi(p)\rho_k(p)$ for all Dirichlet characters $\chi \mod s$. However, there does not seem to be a way to convert this problem into a result about the non-vanishing of a Dirichlet series, and the only way to proceed is through sieve methods.

In [37] a sieve powerful enough to give a simple criterion for any polynomial is developed, and we now recall this theory.

Theorem 6.2.1. *Let $\rho_k(n)$ be as above, and let $0 < \varepsilon < 1/3$ be arbitrary. Assume that when summing over $q < x^{1/2-\varepsilon}$ and $nq < x$, we have*

$$\sum_q \sum_n \lambda_q \chi(n)\rho_k(qn) \ll \frac{x}{(\log x)^2} \max\{|\lambda_d|\}, \tag{6.2.1}$$

and also assume that when summing over $p < x^{1/3-\varepsilon}$ and $\{(n,p) = 1 : n < x/p\}$, we have that

$$\sum_p \sum_n \alpha_n\beta_p\rho_k(pn) \ll \frac{x}{(\log x)^{2\deg F+1}} \max\{|\alpha_n|, |\beta_p|\} \tag{6.2.2}$$

holds in case β_p is supported on primes. (The character values are now absorbed in the α_n, β_p.) Then

$$\limsup_{x\to\infty} \frac{1}{\pi(x)} \left|\sum_{p\leq x} \chi(p)\rho_k(p)\right| \leq 10\varepsilon$$

It is also established in [37], and this is crucial, that the second condition (6.2.2) can be reduced to the first one (6.2.1) based on the following. Let $B(N,P) = \sum \alpha_n\beta_p\rho_k(np)$. Here $N \leq n \leq 2N, P \leq p \leq 2P$, p is restricted to primes and $\gcd(n,p) = 1$.

Theorem 6.2.2. *Assume that, for some $\varepsilon, \eta > 0$,*

$$\sum_{n\leq N} \rho_k(qn) \ll N^{1-\eta}$$

uniformly in the range $h \ll q \ll N^{1-\varepsilon}$. If $P^2 \leq N^{1-\varepsilon}$, then

$$B(N,P) \ll N^{1-\eta/2} P \max\{|\alpha_n\beta_p|\}.$$

Proof. □

. See [37, Section 5, pp. 432-433]. The lemma is not stated in this form explicitly, but the proof works without any modifications.

6.2.3 Kloosterman sums

In this section we derive the estimates for sums of Kloosterman sums. We briefly review the history of these sums and their relation to spectral properties of congruence subgroups, which culminates in the Kuznecov formulas and their various generalizations for Fuchsian groups of the first kind. For a nice introduction see [72]. Motohashi's book [96] is also useful, although he concentrates on the full modular group.

After setting up notation we illustrate the power of this machinery by deriving estimates in the style of Deshouillers and Iwaniec [34]. Our treatment concentrates on estimates with specific application to our problem. For a more complete treatment, see the above-mentioned book by Iwaniec and the references therein. Recall the definition of the classical Kloosterman sum

$$K(m,n;c) = \sum_{xy\equiv\ 1(c)} e\left(\frac{mx+ny}{c}\right).$$

What is needed in most applications is a good estimate for sums of Kloosterman sums, such as

$$\sum_{c\le x} \frac{1}{c} K(m,n;c) \tag{6.2.3}$$

For an individual term, Weil's bound $K(m,n;c) \le \tau(n)(m,n,c)^{1/2}c^{1/2}$ is the best possible, but Linnik [91] and Selberg [114] conjectured that (6.2.3) is majorized by any positive power of x whenever $\gcd(m,n)^{1/2} < x$.

We now define general Kloosterman sums for congruence subgroups. Let $\mathfrak{a}, \mathfrak{b}$ be two cusps for $\Gamma = \Gamma_0(q)$, and let $\Gamma_{\mathfrak{a}}, \Gamma_{\mathfrak{b}}$ be the corresponding stabilizer subgroups of these cusps in Γ. For any cusp $\mathfrak{c}$ we can choose $\sigma_{\mathfrak{c}} \in SL_2(\mathbb{R})$ such that

$$\sigma_{\mathfrak{c}}^{-1}\Gamma_{\mathfrak{c}}\sigma_{\mathfrak{c}} = B = \left\{\pm \left[\begin{smallmatrix} 1 & k \\ 0 & 1 \end{smallmatrix}\right] \ : k \in Z\right\}. \tag{6.2.4}$$

Definition 6.2.3. The Kloosterman sum $K_{\sigma_a,\sigma_{\mathfrak{b}}}(m,n)$ (corresponding to the above choices) is defined to be

$$K_{\sigma_{\mathfrak{a}},\sigma_{\mathfrak{b}}}(m,n;c) = \sum e\left(\frac{ma+nd}{c}\right),$$

where the sum is over $g = \left[\begin{smallmatrix} a & * \\ c & d \end{smallmatrix}\right] \in B\backslash\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}/B$ or, equivalently, over the set $\{g : \sigma_{\mathfrak{a}} g \sigma_{\mathfrak{b}}^{-1} \in \Gamma_{\mathfrak{a}}\backslash\Gamma/\Gamma_{\mathfrak{b}}\}$.

Some simple observations follow [72] . Let $\mathfrak{c}$ be a cusp of $\Gamma_0(q)$. Then $\mathfrak{c}$ is equivalent to some $\frac{u}{v}$ for which $v|q$. Moreover $\frac{u}{v}$, and $\frac{u'}{v'}$ give rise to the same cusp, if and only if $v = v'$, and $u \equiv u' \bmod (v, q/v)$. The width of the cusp $\frac{u}{v}$ is $\frac{q}{(v^2,q)}$.

Definition 6.2.4. If $\mathfrak{c} = \frac{u}{v}$ be a cusp for $\Gamma_0(q)$ we will denote $(v, q/v)$ by $C(\mathfrak{c})$.

Lemma 6.2.5. *Assume that*

$$g_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}g_{\mathfrak{a}} = \{(\pm\left[\begin{smallmatrix}1 & kA\\ 0 & 1\end{smallmatrix}\right] : k \in \mathbb{Z}\}$$

and

$$g_{\mathfrak{b}}^{-1}\Gamma_{\mathfrak{b}}g_{\mathfrak{b}} = \{(\pm\left[\begin{smallmatrix}1 & kB\\ 0 & 1\end{smallmatrix}\right] : k \in \mathbb{Z}\}$$

Let also $\sigma_{\mathfrak{a}} = g_a \begin{bmatrix}\sqrt{A} & 0\\ 0 & 1/\sqrt{A}\end{bmatrix}$ *and similarly* $\sigma_{\mathfrak{b}} = g_b \begin{bmatrix}\sqrt{B} & 0\\ 0 & 1/\sqrt{B}\end{bmatrix}$. *Then*

1. $\sigma_{\mathfrak{a}}, \sigma_{\mathfrak{b}}$ *satisfy (6.2.4) and*

$$K_{\sigma_{\mathfrak{a}},\sigma_{\mathfrak{b}}}(m,n;c\sqrt{AB}) = \sum e\left(\frac{ma}{Ac} + \frac{nd}{Bc}\right),$$

where the sum is over $\{g : g_{\mathfrak{a}}\left[\begin{smallmatrix}a & *\\ c & d\end{smallmatrix}\right] g_{\mathfrak{b}}^{-1} \in \Gamma_{\mathfrak{a}}\backslash\Gamma/\Gamma_{\mathfrak{b}}\}$;

2. *if* $g_{\mathfrak{a}}, g_{\mathfrak{b}} \in SL_2(\mathbb{Z})$, *then*

$$K_{\sigma_{\mathfrak{a}},\sigma_{\mathfrak{b}}}(m,n;c\sqrt{AB}) \le (m,n,ABc)^{1/2}(ABc)^{1/2}\tau(ABc)$$

Lemma 6.2.6. *Assume that* $\Gamma' = g^{-1}\Gamma g$. *Let* $\sigma_a', = g\sigma_{\mathfrak{a}}, \sigma_b' = g\sigma_{\mathfrak{b}}$. *Then we have*

$$K^{\Gamma}_{\sigma_{\mathfrak{a}},\sigma_{\mathfrak{b}}}(m,n,c) = K^{\Gamma'}_{\sigma_a'\sigma_b'}(m,n,c);$$

Remark 6.2.7. This lemma is particularly useful for us in the following situation. Let $\Gamma' = \Gamma_0(M\,[M,N])$, and let

$$\Gamma = \{\left[\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right] \in SL(2,\mathbb{Z}) : b \equiv 0 \bmod N, c \equiv 0 \bmod [M,N]\}.$$

Then we can use the lemma with $g = \begin{bmatrix}\sqrt{N} & 0\\ 0 & 1/\sqrt{N}\end{bmatrix}$.

Let $u_j(z)$ be an orthonormal basis in the space of Maass forms for $\Gamma_0(q)$. Then $\lambda_j = 1/4 + t_j^2 \ge 0$; Selberg conjectured that $\lambda_j \ge 1/4$ with the exception of $\lambda_0 = 0$, that is, when $u(z)$ is constant. It is well known that $u_j(z)$ have a Fourier expansion at each cusp of the form

$$u_j(\sigma_{\mathfrak{a}}z) = y^{1/2}\sum_{n\neq 0}\rho_{\mathfrak{a}j}(n)K_{it_j}(2\pi|n|y)e(nx)$$

We normalize the Fourier coefficients by taking

$$\nu_{\mathfrak{a}j} := \left(\frac{4\pi|n|}{\cosh \pi t_j}\right)^{1/2}\rho_{\mathfrak{a}j}(n).$$

From now on all Kloosterman sums are for $\Gamma_0(q)$ for some q (or sums that can be transformed into them by the remark above). Let f be of compact support; then Kuznetsov's formula (see [8], [2], and [1]) states that

$$\sum_c \frac{1}{c}K_{\mathfrak{ab}}(m,n;c)f\left(\frac{4\pi\sqrt{|mn|}}{c}\right) = \sum_{t_j}\hat{f}(t_j)\overline{\nu_{\mathfrak{a}j}(m)}\nu_{\mathfrak{b}j}(n) + \cdots,$$

where the contribution from the Eisenstein series and holomorphic modular forms is suppressed. The transform $\hat{f}$ is given by

$$\hat{f}(t) = \int_0^\infty N_{2it}(x) f(x) \frac{dx}{x}$$

Here $N_\nu(x)$ depends on the sign of mn as follows:

$$N_\nu(x) = \begin{cases} (2\sin(\pi\nu/2))^{-1}(J_{-\nu}(x) - J_\nu(x)) & \text{when } mn > 0; \\ K_\nu(x) & \text{when } mn < 0. \end{cases}$$

The Bessel transforms of f that arise can be estimated in terms of the following quantities (see [34]):

$$\Delta_0 = \int_0^\infty |f(x)| \frac{(1 + |\log(x)|)}{x} dx;$$

$$\Delta_i = \int_0^\infty \left| \left(x \frac{d}{dx} \right)^i f(x) \right| \frac{dx}{x}, \qquad i = 1, 2;$$

$$\Delta = \int_0^\infty |f(x)| x^{-3/2} dx.$$

By splitting the spectral sums at $T = 1 + \Delta_2/\Delta_1$ and using the large sieve inequalities for the Fourier coefficients such as

$$\sum_{|t_j| \le T} \left| \sum_{nN} a_n \nu_{aj}(n) \right|^2 \ll (T^2 + q^{-1} N \log N \|a_n\|_2^2),$$

one arrives at (see [72]) the following theorem.

Theorem 6.2.8. *Let a, b be cusps of $\Gamma_0(q)$, and assume that $C(\mathfrak{a}), C(\mathfrak{b}) \ll 1$ (Def.6.2.4). Then*

$$\sum_{nN} a_n \sum_c \frac{1}{c} K_{\mathfrak{ab}}(m, n; c) f\left(\frac{4\pi\sqrt{|mn|}}{c} \right)$$
$$\ll \left\{ (\Delta_0 + \Delta_1) \left(1 + \frac{m}{q} \right)^{1/2} \left(1 + \frac{N}{q} \right)^{1/2} + (\Delta_1\Delta_2)^{1/2} + E_{m,N}(q, f) \right\}$$
$$X(\log 2mN)\|a\|,$$

where $E_{m,N}(q, f)$ is the contribution from the exceptional spectrum, for which we have the following estimate:

$$E_{m,N}(q, f) \ll \left(\frac{\Delta}{q} (q + m)^{1/4} (q + N)^{1/4} (mN)^{1/4} \right) \times (\log 2mN)\|a\|.$$

It is conjectured by Selberg that exceptional eigenvalues do not exist for $\Gamma_0(q)$ and, therefore, $E_{m,N}(q, f) = 0$. At present the best result due to Kim and Sarnak [78] is that, for all eigenvalues, $\lambda \ge 975/4096 = 0.238...$. The above estimate for $E_{m,N}(q, f)$ uses density theorems for the exceptional spectrum to get around this problem.

We now proceed to derive some estimates that are used in the last section.

Theorem 6.2.9. *Let $N, C \geq 1$, $G(n,c)$ be supported in $[N, 2N] \times [C, 2C]$, and suppose that, for $i, j = 0, 1, 2$,*

$$\frac{\partial^{i+j}}{(\partial n)^i (\partial c)^j} G(n,c) \ll N^{-i} C^{-j}.$$

Let

$$A = \sum_c \frac{1}{c} \sum_n a_n G(n,c) K_{\mathfrak{ab}}(m,n;c).$$

Then

$$A \ll \|a\| \left\{ \left(1 + \frac{m}{q}\right)^{1/2} \left(1 + \frac{N}{q}\right)^{1/2} + \frac{C^{1/2}}{q} (q+m)^{1/4} (q+N)^{1/4} \right\} \log^2 mNCq.$$

Proof. We start by defining

$$F(t,x) = \int_{-\infty}^{\infty} G\left(u, \frac{4\pi\sqrt{mu}}{x}\right) e(tu) du.$$

Then

$$A = \int_{-\infty}^{\infty} \sum_{n,c} a_n \frac{1}{c} F\left(t, \frac{4\pi\sqrt{mn}}{c}\right)$$

We apply Theorem 3.3 for each individual t and then integrate over t. It is easy to establish that

$$\Delta_0 \ll N(1 + \log(mNCq))$$

and that

$$\Delta_i \ll N \text{ for } i = 1, 2.$$

This is used for small t.

For large values of t we use

$$F(t,x) = \int_{-\infty}^{\infty} \frac{d^2}{du^2} \left[G\left(u, \frac{4\pi\sqrt{mu}}{x}\right)\right] \frac{e(ut)}{-4\pi^2 t^2} dx$$

Then $|F(t,x)| \ll 1/Nt^2$, and so

$$\Delta_0(t) \ll \frac{1 + \log(mNCq)}{Nt^2}$$

and

$$\Delta_i(t) \ll \frac{1}{Nt^2}$$

for $i = 1, 2$

Combining these estimates,

$$\Delta_0(t) \ll \frac{N}{1 + N^2 t^2} (1 + \log(mNCq))$$

and

$$\Delta_i(t) \ll \frac{N}{1+N^2t^2}$$

for $i = 1, 2$.

Note that Δ is needed to estimate the contribution from the exceptional eigenvalues. Since $|F(t,x)| \ll N$ and $|F(t,x)| \ll 1/Nt^2$, we have

$$\Delta(t) \ll X^{-1/2}\frac{N}{1+N^2t^2},$$

where $[X, 2X]$ is the support of $F(t,\cdot)$, $X \simeq \sqrt{mN}/C$; that is,

$$\Delta(t) \ll \frac{C^{1/2}}{\sqrt[4]{mN}}\frac{N}{1+N^2t^2}.$$

The theorem follows after integrating Theorem 6.2.8 with respect to t. □

6.2.4 Reduction to sums of Kloosterman sums

Recall that with $\rho_k(n) = \sum_{P(\nu)\equiv 0(n)} e(k\nu/n)$ we want to estimate

$$\mathcal{L}_q(X) = \sum_{X\le n\le 2X} \rho_k(qn)e\left(\frac{jn}{s}\right) \tag{6.2.5}$$

or the more general sum

$$\mathcal{L}_q(f) = \sum_{n\equiv 0(q)} \rho_k(n)f(n), \tag{6.2.6}$$

where the function f has support in $[qX, 2qX]$. (The factor $e(nj/s)$ is absorbed in f.) At first any f will do, but later it will be convenient for us to use a smooth f instead of the characteristic function of the interval $[qX, 2qX]$. In the rest of the section we transform $\mathcal{L}_q(f)$ into sums of Kloosterman sums based on the theory of binary quadratic forms. In this transformation a small error term arise, and we must keep track of the dependence on k, q and X as this is crucial in applying the sieve argument, dependence on other fixed factors will be ignored.

Although this transformation could be done as in [[67], without ever mentioning the congruence subgroups $\Gamma_0(q)$, we follow a slightly different path. Apart from some technical simplifications, the main advantage is that we arrive at the generalized Kloosterman sums to which the results of Section 3 are directly applicable.

Let $P(x) = Ax^2 + Bx + C$, with discriminant $D = B^2 - 4AC$.

Lemma 6.2.10. *We have*

$$\rho_k(n) = e\left(\frac{-kB)}{2An}\right)\sum_{y\in Y_n} e\left(\frac{ky}{2An}\right)$$

where

$$Y_n = \{y \bmod 2An : y \equiv B \bmod 2A \textit{ and } y^2 \equiv D \bmod 4An\}. \tag{6.2.7}$$

dc_1553_18

Proof. Assume $P(\nu) \equiv 0 \bmod n$, and let $y = 2A\nu + B$. Then $y \bmod 2An$ is uniquely determined by $\nu \bmod n$, $y \equiv B \bmod 2A$ and $y^2 \equiv D \bmod 4An$. $\square$

To handle $\mathcal{L}_q(f)$ we use the classical correspondence between solutions of $y^2 \equiv D$ and representations of the modulus by quadratic forms as in [5].

Let $Q(X,Y) = MX^2 + 2RXY + NY^2$ be a quadratic form of discriminant $4D$ so that $R^2 - MN = D$. On occasion we will write $[M, R, N]$ for Q and also identify it with the matrix $\left[\begin{smallmatrix} M & R \\ R & N \end{smallmatrix}\right]$.

Recall that one of the many possible action of $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in SL_2(\mathbb{Z})$ on Q is

$$gQ = g \left[\begin{smallmatrix} M & R \\ R & N \end{smallmatrix}\right] g^t.$$

This is clearly a left-action, $(g_1 g_2)Q = g_1(g_2 Q)$. We will need the the explicit form of this action

$$g \cdot (MX^2 + 2RXY + NY^2) = M(g)X^2 + 2R(g)XY + N(g)Y^2.$$

Here for $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, $M(g) = Q(a,b)$, $N(g) = Q(c,d)$, and $R(g) = \left[\begin{smallmatrix} a & b \end{smallmatrix}\right] \left[\begin{smallmatrix} M & R \\ R & N \end{smallmatrix}\right] \left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$, so

$$M(g) = Ma^2 + Rab + Nb^2 \tag{6.2.8}$$
$$R(g) = Mac + R(ad + bc) + Nbd \tag{6.2.9}$$
$$N(g) = Mc^2 + Rcd + Nd^2. \tag{6.2.10}$$

Definition 6.2.11. Let
$$\Lambda = \left\{ \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : b, c \equiv 0 \bmod 2A \right\},$$

Lemma 6.2.12. *If*

$$V = \{(M, R, N) : R^2 - MN = D, R \equiv B \ (2A), M \equiv 0 \ (4A)\}$$

then V is invariant under the action of Λ and Y_n in Lemma 6.2.7 is in bijection with $\Lambda_\infty \backslash V$.

Proof. If y is in Y_n it gives rise to an element in V via

$$Q_y = \left[(y^2 - D)/n, y, n \right].$$

Note, that $y_1 \equiv y_2 \bmod 2An$ if and only if the corresponding quadratic forms transform into one another by an element of the form $\left[\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right] \in \Lambda_\infty$. $\square$

It is well known that $\Lambda \backslash V$ is finite, and we choose a finite set $\{Q_j(X,Y) = M_j X^2 + 2R_j XY + N_j Y^2 : j = 1, ..., h\}$, so that

$$V = \bigcup_{j=1}^{h} \Lambda \cdot (M_j, R_j, N_j).$$

Remark 6.2.13. By an old theorem of Weber we may even choose Q_j in such a way as to make sure that N_j is prime. This is unnecessary for the moment, and is only needed to simplify an elementary argument in Section 6.2.5.

We summarize the above in the next proposition.

Proposition 6.2.14.

$$\mathcal{L}_q(f) = \sum_{n\equiv 0(q)} f(n)\rho_k(n) = \sum_{j=1}^{h}\left(\sum\nolimits^{q} f(N_j(g))e\left(\frac{-kB)}{2An}\right)e\left(\frac{kR_j(g)}{2AN_j(g)}\right)\right),$$

where $\sum^q$ restricts the sum to those $g \in \Lambda_\infty\backslash\Lambda/\mathcal{U}_j$ for which $Q_j(c,d) \equiv 0 \mod q$. Since $N(\tau\cdot Q) = N(Q)$ for every Q and every $\tau \in \Lambda_\infty$ this property does not depend on the representative of the Λ_∞-coset. The group $\mathcal{U}_j$ is the group of automorphs of Q_j in Λ,

$$\mathcal{U}_j = \{g \in \Lambda : g\cdot Q_j = Q_j\}.$$

The functions $R_j(g), N_j(g)$ are those defined in (6.2.9) and (6.2.10) for Q_j.

Remark 6.2.15. The change of notation to U_j for the stabilizers is to further emphasize the fact that we are now using a different action of SL_2 on quadratic forms.

Definition 6.2.16. Let Q be a quadratic form, U its group of automorphs in Λ. A partition of unity for the group U is a function $\psi : SL_2(\mathbb{R}) \to [0,1]$ with the following properties $\psi(\tau g) = \psi(g)$ for $\tau \in \Lambda_\infty$ and

$$\sum_{u\in\mathcal{U}} \psi(gu) = 1$$

for all g for which $N(g) = Q(g_{21}, g_{22})$ is positive.

Of course in the case $D < 0$, $\#\mathcal{U}_j < \infty$, and one could simply use the constant function $\psi(g) = (\#\mathcal{U}_j)^{-1}$. For positive discriminants the existence of a partition of unity is established in the following

Proposition 6.2.17. *Let $U \in SL_2(\mathbb{R})$ be of hyperbolic type, with row eigenvectors $(1, w_1)$ and $(1, w_2)$, $w_1 < w_2$. Then there exists a smooth function ψ on $SL_2(\mathbb{R})$, such that for all $g = \left[\begin{smallmatrix} g_{11} & g_{12} \\ g_{21} & g_{22}\end{smallmatrix}\right] \in SL_2(\mathbb{R})$ for which $w_1 < g_{22}/g_{21} < w_2$:*

$$\sum_{k=-\infty}^{\infty} \psi(gU^k) = 1.$$

In addition we may assume that $\psi(g) = \phi(g_{22}/g_{21})$ for some smooth, compactly supported function ϕ whose support is contained in (w_1, w_2), and therefore $\psi(\left[\begin{smallmatrix} 1 & t \\ 0 & 1\end{smallmatrix}\right] g) = \psi(g)$, for any $t \in \mathbb{R}$.

Proof. Since the action of U on the second row (g_{21}, g_{22}) is linear, we can consider the induced fractional linear transformation on $t = g_{22}/g_{21}$:

$$(1 : t) \mapsto (1 : \tfrac{b+dt}{a+ct}).$$

We will denote this map $t \mapsto \frac{b+dt}{a+ct}$ by $\tilde{U}$. Since $\det U = 1$, $\tilde{U}(t) = \frac{d}{c} - \frac{1}{c^2x+ac}$, and so it is continuous and strictly increasing on both $(-\infty, -\frac{a}{c})$ and $(-\frac{a}{c}, \infty)$. Moreover if $t_1 < -\frac{a}{c} < t_2$, then $c^2t_1 + ac < 0 < c^2t_2 + ac$, and so $\tilde{U}(t_1) > \tilde{U}(t_2)$. It follows, that if $w_1 < w_2$ are the fixed points of $\tilde{U}$, then $-\frac{a}{c} \notin (w_1, w_2)$, and so $\tilde{U}$ takes this open interval to itself bijectively.

Consider now $s(t) = \frac{t-w_1}{w_2-t}$. Since $\left[\begin{smallmatrix} 1 & w_1 \\ 1 & w_2\end{smallmatrix}\right] U = \left[\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda\end{smallmatrix}\right]\left[\begin{smallmatrix} 1 & w_1 \\ 1 & w_2\end{smallmatrix}\right]$ we have that

$$s(\tilde{U}(t)) = \frac{1}{\lambda^2}s(t).$$

The rest of the argument now follows the standard proof of a partition of unity. By replacing U by U^{-1} if necessary, (which we can, since the proposition is about the cyclic group generated ny U), we may assume that $\lambda > 1$.

Let h_0 be smooth non-negative function on $(0, \infty)$, which is supported on $[1/\lambda^2, \lambda^2]$, such that it is strictly positive (say 1) on $[1/\lambda, \lambda]$. Then the function $h_1(s) = \sum_{k\in\mathbb{Z}} h_0(\lambda^{2k}s)$ is smooth since the sum is locally finite.

We also have that $h_0 \le h_1$, and that $h_1(\lambda^2 s) = h_1(s)$. It follows that h_1 is everywhere positive on $(0, \infty)$ and we may therefore define $h(s) = h_0(s)/h_1(s)$, which is smooth. Clearly $\sum_{k\in\mathbb{Z}} h(\lambda^{2k}s) = 1$ for all $s \in (0, \infty)$.

Finally let

$$\phi(t) = h\left(\tfrac{t-w_1}{w_2-t}\right) \qquad \text{and} \qquad \psi(g) = \phi(g_{22}/g_{21}). \tag{6.2.11}$$

The property $\sum_{k=-\infty}^{\infty} \psi(gU^k) = 1$ is valid by the above construction. Since the function ψ only depends on g_{22}/g_{21}, the property $\psi([\begin{smallmatrix}1 & t\\ 0 & 1\end{smallmatrix}]\, g) = \psi(g)$, holds trivially for any $t \in \mathbb{R}$. $\square$

Proposition 6.2.18. *Let $Q = (M, R, N)$ with $N < 0$, $R^2 - MN = D$, and stabilizer $\mathcal{U}$. Let ψ and ϕ be as in the proposition above for the group of automorphs of Q, and let*

$$F(c, d) = f(Q(c, d))\phi\left(\frac{d}{c}\right).$$

Assume also that $\operatorname{supp} f \subset [qX, 2qX]$. *With the notation (6.2.9) and (6.2.10) we have*

$$\sum_{g\in\Lambda_\infty\backslash\Lambda/\mathcal{U}}^{q} f(N(g))e\left(\frac{-kB)}{2AN(g)}\right) e\left(\frac{kR(g)}{2AN(g)}\right) =$$

$$\sum_{g\in\Lambda_\infty\backslash\Lambda}^{q} e\left(\frac{ka}{2Ac}\right) F(c, d) + O(k\log(qX)),$$

where on both sides $\sum^q$ restricts the sums to $g = [\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}]$ for which $Q(c, d) \equiv 0 \bmod q$. The implied constant depends on M, R, N, A, B and $\max|f|$, but does not depend on k, q or X.

Proof. By the construction of ψ,

$$\sum_{g\in\Lambda_\infty\backslash\Lambda/\mathcal{U}}^{q} f(N(g))e\left(\frac{kR(g)}{2AN(g)}\right) = \sum_{g\in\Lambda_\infty\backslash\Lambda} f(N(g))\psi(g)e\left(\frac{kR(g)}{2AN(g)}\right) =$$

$$\sum_{g\in\Lambda_\infty\backslash\Lambda} f(Q(c, d))\phi\left(\frac{d}{c}\right) e\left(\frac{kR(g)}{2AN(g)}\right) = \sum_{g\in\Lambda_\infty\backslash\Lambda}^{q} F(c, d)e\left(\frac{kR(g)}{N(g)}\right).$$

The main idea going back to [67] is to use Bruhat decomposition which leads to the identity

$$\frac{Mac + R(ad + bc) + Nbd}{Mc^2 + 2Rcd + Nd^2} = \frac{a}{c} - \frac{Rc + Nd}{c(Mc^2 + 2Rcd + Nd^2)}.$$

Therefore we have

$$e\left(\frac{kR(g)}{N(g)}\right) = e\left(\frac{ka}{c}\right) e\left(\frac{k(Rc + Nd)}{cQ(c, d)}\right)$$

dc_1553_18

and here $\frac{k(Rc+Nd)}{cQ(c,d)} = \frac{k(R+Nd/c)}{Q(c,d)} \ll \frac{1}{qX}$ by our assumptions, and so

$$e\left(\frac{k(Rc+Nd)}{cQ(c,d)}\right) = 1 + O\left(\frac{k}{Q(c,d)}\right)$$

the implied constant depending on the form $[M,R,N]$ only. We therefore have

$$\sum_{g\in\Lambda_\infty\backslash\Lambda/\mathcal{U}}^{q} f(N(g))e\left(\frac{kR(g)}{2AN(g)}\right) = \sum_{g\in\Lambda_\infty\backslash\Lambda}^{q} e\left(\frac{ka}{2Ac}\right) + O(k)\sum_{g\in\Lambda_\infty\backslash\Lambda}^{q,f,\phi} \frac{1}{Q(c,d)}$$

where $\sum^{q,f,\phi}$ now indicates that the sum is over those (c,d), for which d/c is in the support of ϕ, $Q(c,d)$ is in the support of f and also $Q(c,d) \equiv 0 \bmod q$. The simplest way to estimate this sum is to compare it to

$$\iint_\Omega \frac{dxdy}{Q(x,y)}$$

where $\Omega = \{(x,y) \in \mathbb{R}^2 : y/x \in [t_1,t_2], Q(x,y) \in [qX, 2qX]\}$, where $\operatorname{supp}\phi \subset [t_1,t_2]$. This leaves the congruence condition out, but still gives a satisfactory answer. The integral is easily evaluated by the change of variables $t = x/y$, $s = \sqrt{Q(x,y)}$, and leads to the $\log(qX)$ error term. □

Alternatively, we may remove the $\log q$ term, by noticing that $\{(c,d) : d/c \in [t_1,t_2] : Q(c,d) = n\} \le C\tau(n)$, where $\tau(n)$ is the number of divisors of n. (In fact with the construction in Proposition 6.2.17 we may chose C to be 2.), Applying Dirichlet's theorem on the sum of $\tau(n)$ gives the improvement, which plays no role here, and therefore the details are omitted.

Let $\Gamma = \Gamma(q,A) = \Gamma_0(q) \cap \Lambda$.

Remark 6.2.19. We will not show the dependence on q, which is fixed for the identity that we are about to derive. It would also make the sub-index indicating the stabilizer of a cusp very inconvenient to show. Note, that

$$\Gamma_\infty = \{g \in \Gamma : g\infty = \infty\} = \{\left[\begin{smallmatrix}1 & kA\\ 0 & 1\end{smallmatrix}\right] = \Lambda_\infty.$$

Recall from Definition 3.1 that, for the congruence subgroup Γ_q and cusps $\mathfrak{a},\mathfrak{b}$ of Γ_q, the generalized Kloosterman sum is

$$K_{\mathfrak{ab}}(m,n) = \sum e\left(\frac{ma+nd}{c}\right),$$

the sum being over $\{\sigma_\mathfrak{a} g\sigma_\mathfrak{b}^{-1} \in \Gamma_\mathfrak{a}\backslash\Gamma/\Gamma_\mathfrak{b}\}$. We are now ready to state the main proposition.

Theorem 6.2.20. *Assume the notation of Proposition 6.2.18.*

$$\sum_{g\in\Lambda_\infty\backslash\Lambda} F(c,d)e\left(\frac{ka}{c}\right) = \sum_{\mathfrak{c}}^{q}\sum_{m,c}\frac{1}{c\sqrt{2aw_c}}K_{\infty\mathfrak{c}}(m,k;c)G_\mathfrak{c}(m,c).$$

Here $K_{\infty\mathfrak{c}}(m,k;c)$ is a Kloosterman sum for the group Γ_q the sum is over certain cusps $\mathfrak{c}$ whose description is given in the proof, and where $m \in \mathbb{Z}$, and $c \in qw\mathbb{Z}$, w being the width of the cusps $\mathfrak{c}$. The function G is defined by

$$G_\mathfrak{c}(m,c) = \int_{-\infty}^{\infty} F(c,y)e\left(-\frac{my}{cw_\mathfrak{c}}\right)dy$$

Proof. First, if $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ we will overload the notation and use $F(g)$ for the function $g \mapsto F(c,d)$.

Recall that $\sum^q$ means that g is subject to the condition $Q(c,d) \equiv 0 \bmod q$. To handle this condition we split the sum as follows:

$$\sum_{g \in \Lambda_\infty \backslash \Lambda}^{q} F(g) e\left(\frac{ka}{2Ac}\right) = \sum_{g} \sum_{h}^{q} F(gh) e\left(\frac{ka}{2Ac}\right)$$

where $gh = \left[\begin{smallmatrix} a & * \\ c & * \end{smallmatrix}\right]$, and where g runs through $\Gamma_\infty \backslash \Gamma$, and the sum in h is over those $h \in \Gamma \backslash \Lambda$ for which $N(h \cdot Q) \equiv 0 \bmod q$. This is well defined since for all $g \in \Gamma_q$ and for any quadratic form Q:

$$N(g \cdot Q) \equiv N(Q) \bmod q.$$

Note that as g runs through $\Gamma_\infty \backslash \Gamma$, gh runs through $\Gamma_\infty \backslash \Gamma h$. Let w be the width of the cusp $\mathfrak{c} = h(\infty)$, and let

$$B_w = \left\{ \begin{pmatrix} 1 & mw \\ 0 & 1 \end{pmatrix} : m \in \mathbb{Z} \right\}.$$

Now let $g = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, then

$$\sum_{g \in \Gamma_\infty \backslash \Gamma h} F(g) e\left(\frac{ka}{2Ac}\right) = \sum_{g \in \Gamma_\infty \backslash \Gamma h / B_w} e\left(\frac{ka}{2Ac}\right) \sum_{m \in \mathbb{Z}} F\left(g \left[\begin{smallmatrix} 1 & mw \\ 0 & 1 \end{smallmatrix}\right]\right)$$

By Poisson summation the inner sum equals

$$\sum_{m \in \mathbb{Z}} \int_{-\infty}^{\infty} F\left(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\left[\begin{smallmatrix} 1 & xw \\ 0 & 1 \end{smallmatrix}\right]\right) e(-mx) dx.$$

After the change of variable $x = t - d/wc$, the whole sum becomes

$$\sum_{m \in \mathbb{Z}} \sum_{g \in \Gamma_\infty \backslash \Gamma h / B_w} e\left(\frac{ka}{2Ac} + \frac{md}{wc}\right) \int_{-\infty}^{\infty} F\left(\left[\begin{smallmatrix} * & * \\ c & cwt \end{smallmatrix}\right]\right) e(-mt) dt$$

Since $hB_w = \Gamma_{\mathfrak{c}} h$, we recognize the sum

$$\sum_{gh^{-1} \in \Gamma_\infty \backslash \Gamma / \Gamma_{\mathfrak{c}}} e\left(\frac{ma}{2Ac} + \frac{kd}{wc}\right)$$

as a Kloosterman sum $K_{\infty\mathfrak{c}}(m,k; c\sqrt{2Aw})$. (Use Lemma 6.2.5 with $\mathfrak{a} = \infty$, $g_{\mathfrak{a}} = I$, $\mathfrak{b} = \mathfrak{c}$, and $g_{\mathfrak{b}} = h$.)

Finally substituting $y = cwt$ we have

$$\int_{-\infty}^{\infty} F(c, cwt)\, e(-mt) dt = \tfrac{1}{cw} \int_{-\infty}^{\infty} F(c, y)\, e\left(-\tfrac{my}{cw}\right) dy = G_{\mathfrak{c}}(m, c)$$

□

.

Remark. By Lemma6.2.6 the above Kloosterman sums can be replaced by Kloosterman sums for the group $\Gamma_0(qA^2)$, and we make this identification in what follows.

6.2.5 Proof of the main theorem

For the estimation of the right-hand side of (13), the sum in m is split at some parameter to be defined later. It is best to denote this parameter qM for some $M > 1$ to ease the notation in the final estimates.

We will first bound the sum of Kloosterman sums that arose for our problem in the range when

$$|m| \geq qM,$$

where we simply use Weil's bound

$$K(m, k; c\sqrt{2Aw}) \ll \tau(c)(m, k, wc)^{1/2}(wc)^{1/2}$$

together with some elementary estimates.

First we need the following

Proposition 6.2.21. *The function*

$$G(m, c) = \int_{-\infty}^{\infty} F(c, cwt)\, e(-mt)dt$$

decays rapidly in m:

$$G(m, c) \ll w^{L-1}m^{-L} \tag{6.2.12}$$

for any L, the implied constant depending on L only.

Since

$$F(c, cwt) = f(Q(c, cwt))\phi(wt)$$

the proposition is a simple consequence of the following lemma.

Lemma 6.2.22. *Let $f(x)$ be supported in $[X, 2X]$. Let*

$$C_l = \max\{|X^j f^{(j)}(x)| : x \in [X, 2X], j \leq l\}.$$

Let $u(x) = Q(c, cwt)$ for some quadratic form Q. Assume $c \asymp X^{1/2}$. Let $D = d/dt$. Then

$$D^l(f(u(t))) \ll C_l w^l,$$

where the implied constant only depends on the degree of the derivation l and the coefficients of Q.

Proof. The derivative in question is a sum of monomials of the form (see e.g.[120, Chapter 5] for a more precise form)

$$f^{(j)}(u)(c^2u')^{2j-l}(c^2u'')^{l-j}$$

and this is bounded by $C_l X^{-j}(Xw)^{2j-l}(Xw^2)^{l-j}$. □

.

Proof of the Proposition. Now (6.2.12) is a standard property of Fourier integrals that follows from Lemma 5.1 after integrating by parts. □

Since $|\Gamma : \Lambda|$ is large (about q), we still need to establish that the number of $h \in \Gamma\backslash\Lambda$ that contribute to the sum is negligible. This is shown by first grouping the cosets according to what cusp arises as $h(\infty)$ and then counting for each cusp separately.

Definition 6.2.23. Let $Q = [M, R, N]$, and let

$$A_{\mathfrak{c}} = \sum_{h(\infty)=\mathfrak{c}}^{q} 1$$

where $h \in \Gamma\backslash\Lambda$, $N(h) \equiv 0 \bmod q$.

Lemma 6.2.24. *In Definition 6.2.23, choose $\mathfrak{c} = h(\infty) = \frac{a}{c}$ such that $c|q$. (This can always be done.)*

a) Let the cusp $\mathfrak{c} = h(\infty) = \frac{a}{c}$ be as above. If $c \not| N$ then $A_{\mathfrak{c}} = 0$.

b) Make the assumption that N is prime. (See Remark 6.2.13). Then, when $A_{\mathfrak{c}} \neq 0$, it is majorized by $\tau(q)$.

Proof. It is well known that the representative elements in $\Gamma\backslash\Lambda$ can be chosen so that $c|q$.

Now, assume $h = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \Gamma\backslash\Lambda$ has the property that $q|N(h) = Q(c, d)$. By our choice $c|q$, and this implies that $c|Nd^2$, and since $\gcd(c, d) = 1$, $c|N$. This restriction on h shows that only the above-specified cusps can arise as $h(\infty)$ with property $A_c \neq 0$.

From now on we will make the simplifying assumption that N is prime, see Remark 6.2.13.

First,t he number of cusps that arise in the estimates have either $c = 1$, with width $w = q$, or they have $c = N$, with width $q/(q, N^2)$.

We now bound the number of h for each of the two types. If $h(\infty) = h'(\infty)$, then there is $\tau \in \Lambda_\infty$ such that $h' = \tau h$. Suppose $\tau = \left[\begin{smallmatrix} 1 & m \\ 0 & 1 \end{smallmatrix}\right]$. The condition $N(h') \equiv 0 \bmod q$ leads to the following quadratic congruence for m : $(2Rc^2 + 2Ncd)m + Nc^2m^2 \equiv 0 \bmod q$,. Write $q = N^\alpha q_1$, with $(q_1, N) = 1$ to conclude that this has at most $O(\tau(q))$ solutions with an absolute implied constant. □

Corollary 6.2.25. *Let $\mathfrak{c}$ be a cusp of Γ such that $A_{\mathfrak{c}} \neq 0$. Let w be the width of $\mathfrak{c}$. Then $w \asymp q$.*

We are now ready to prove the following

Proposition 6.2.26. *For any positive integer L, we have*

$$\sum_{qM\leq|m|} \sum_{c} \frac{1}{c} G_c(m, c)K(m, k; c\sqrt{2Aw}) \ll q^{-1/4}X^{3/4}M^{-L}$$

where the implied constant depends on L only.

Proof. We will use the Weil-bound from Lemma 6.2.5

$$|K(m, k; 2c\sqrt{Aw})| \leq (m, k, 2Ac)^{1/2}(2Awc)^{1/2}\tau(2Awc)$$

Since

$$G(m, c) = \int_{-\infty}^{\infty} F\left(c, cwt\right) e(-mt)dt = \int_{-\infty}^{\infty} f(Q(c, cwt))\phi(wt)e(-mt)dt$$

in c its support is contained in $\{c : f(c^2Q(1, cwt)) \neq 0\}$. Now the construction of the partition of unity ϕ in Proposition 6.2.17 shows that for wt in the support of ϕ, $Q(1, wt)$ is positive, and so there are constants $0 < \mu_1, \mu_2$, such that when $\phi(wt) \neq 0$, we have $\mu_1 < Q(1, wt) < \mu_2$. (These depend on Q, and ϕ only.) Therefore $c \asymp \sqrt{qX}$. Now use Proposition 6.2.21 with $L+1$ instead of L and sum all these estimates get the Proposition. □

We now move to the treat the sum over the range $m \leq qM$

Lemma 6.2.27.

$$\sum_{|m|\leq qM} \sum_{c} \frac{1}{c} G_{\mathfrak{c}}(c, y) K_{\infty\mathfrak{c}}(m, k; c\sqrt{2Aw}) \ll q^{1/4} X^{3/4} M.$$

Proof. We interchange the sums and the integral that defines $G(m, c)$. For simplicity we make the substitution $y = cwt$. If we split the interval $[0, qM]$ into dyadic intervals, then we are in position to apply Theorem 6.2.9 for each individual y[1] Note that the conditions of Theorem 6.2.20 are satisfied with $C = C(y) \ll \sqrt{qX}$ and $N = qM$, giving

$$\sum_{c,m} \frac{1}{cw} F(c, y) K_{\infty\mathfrak{c}}(m, k; \sqrt{2Awc})$$

$$\ll \left\{ \left(1 + \frac{k}{q}\right)^{1/2} \left(1 + \frac{M}{q}\right) + \frac{(qX)^{1/4}}{q} (q+k)^{1/4} (q+M)^{1/4} \right\} (\log^2 kMX) \frac{\sqrt{M}}{\sqrt{w}}.$$

Under the assumptions $k \ll q$ we have by Corollary 5.4 that

$$\sum_{c,m} \frac{1}{c} F_c(c, y) K_{\infty c}(m, k; c\sqrt{2Ac}) \ll \left(\frac{X}{q}\right)^{1/4} M.$$

The integration is along an interval of length $\sqrt{qX}$, and the lemma follows. □

The estimates depend on L, that is, the bounds on the first L derivatives of f, and on M. For the unsmoothing below, we choose $M = q^{1/(2L)}$ for some integer L to get (for $q \ll X$)

$$\mathcal{L}_q(f) \ll q^{1/4} X^{3/4(1+1/(2L))}$$

(the implied constant depending on L).

The passage from the estimation of $\mathcal{L}_q(f)$ to that of $\mathcal{L}_q(X)$ gives rise to a loss in the quality of the estimates. However, the new estimates still suffice to prove the analogue of (6.2.1).

Proposition 6.2.28. *Let $\mathcal{L}_q(X)$ be as in (6.2.5). We have*

$$\mathcal{L}_q(X) \ll \left(\frac{q}{X}\right) X^{1+1/L^2}$$

where the implied constant depends on L, the coefficients A, B, C of the original quadratic equation, but independent of q or X.

[1] To be more precise, one needs a smooth version of this; the error that results can be estimated trivially using the above Proposition.

Proof. Fix a positive integer L and a function $g(\mu,t)$ such that $g(\mu,t)=0$ when $t\le 1$ or $2\le t$, such that $g(\mu,t)=\mu^L$ when $1+\mu\le t\le 2-\mu$. Note that for $0<\mu<1/4$ the first L derivatives $(\partial/\partial t)^j(\mu,t)\ll 1$ with a constant that depends on our choice of g (that is L) but not the other parameters q and X.

To apply our estimates for primes in arithmetic progressions to some modulus s we choose our f in $\mathcal{L}_q(f)$ to be

$$f(x)=g\left(\mu,\tfrac{x}{qX}\right)e\left(\tfrac{jx}{s}\right).$$

The exponential factor is harmless, and will have no effect in the estimates that follow. First,

$$\mathcal{L}_q(X)=\mu^{-L}\mathcal{L}_q(f)+O(\mu X)\log X.$$

Here in the intervals $[X,X+\mu X]$, $[2X-\mu X,2X]$ we trivially estimate using $|\rho_k(n)|\ll\tau(n)$. We choose $\mu=[(q/X)^{1/4}X^{1+(1/L)}]^{1/L}$. By (15),

$$\mathcal{L}_q(f)\ll\left(\frac{q}{X}\right)^{1/4}X^{1+1/(2L)},$$

where the implied constant depends on L only. Therefore,

$$\mathcal{L}_q(X)\ll\left(\frac{q}{X}\right)^{1/(4L)}X^{1+1/L^2}$$

as claimed. □

Finally we have the following

Theorem 6.2.29. *Let $0<\varepsilon<1/3$, and assume that $q\le x^{1/2-\varepsilon}$. Then there exists an $\eta>0$ such that*

$$\sum_{q<x^{1/2-\varepsilon}}\lambda_d\sum_{qn\le x}\rho_h(qn)\ll x^{1-\eta}\max\{|\lambda_q|\}$$

holds with a constant that depends on ε alone.

Proof. Choose $L\ge 2/\varepsilon$ in (16). Then

$$\sum_{q<x^{1/2-\varepsilon}}\lambda_q\sum_{qn\le x}\rho_k(qn)\ll\max\{|\lambda_q|\}x^{1-\eta}$$

with $\eta=(1/200)\varepsilon^2$.

□

In view of the criteria of Theorems 2.1 and 2.2, this concludes the proof of Theorem 1.2.

Appendix A

Special functions

A.1 Whittaker functions

A standard reference for the theory of Whittaker functions is [126, Chap. 16]. Another good reference is [93]. For the convenience of the reader we will record here some of the properties of these special functions that we need.

For fixed μ, ν with $\mathrm{Re}(\nu \pm \mu + 1/2) > 0$ the Whittaker functions may be defined for $y > 0$ by [93, pp. 311, 313]

$$M_{\mu,\nu}(y) = y^{\nu+\frac{1}{2}} e^{\frac{y}{2}} \tfrac{\Gamma(1+2\nu)}{\Gamma(\nu+\mu+\frac{1}{2})\Gamma(\nu-\mu+\frac{1}{2})} \int_0^1 t^{\nu+\mu-\frac{1}{2}}(1-t)^{\nu-\mu-\frac{1}{2}} e^{-yt}\, dt \quad \text{and} \tag{A.1.1}$$

$$W_{\mu,\nu}(y) = y^{\nu+\frac{1}{2}} e^{\frac{y}{2}} \tfrac{1}{\Gamma(\nu-\mu+\frac{1}{2})} \int_1^\infty t^{\nu+\mu-\frac{1}{2}}(t-1)^{\nu-\mu-\frac{1}{2}} e^{-yt}\, dt. \tag{A.1.2}$$

Both $M_{\mu,\nu}(y)$ and $\mathcal{W}_{\mu,\nu}(y)$ satisfy the second order linear differential equation

$$\frac{d^2w}{dy^2} + \Big(-\tfrac{1}{4} + \mu y^{-1} + (\tfrac{1}{4} - \nu^2) y^{-2} \Big) w = 0. \tag{A.1.3}$$

Their asymptotic behavior as $y \to \infty$ for fixed μ, ν is easily found from (A.1.1) and (A.1.2) by changing variable $t \mapsto t/y$:

$$M_{\mu,\nu}(y) \sim \tfrac{\Gamma(1+2\nu)}{\Gamma(\nu-\mu+\frac{1}{2})} y^{-\mu} e^{y/2} \quad \text{and} \quad W_{\mu,\nu}(y) \sim y^{\mu} e^{-y/2}. \tag{A.1.4}$$

In particular, they are linearly independent. For small y we get directly from (A.1.1) that

$$M_{\mu,\nu}(y) = y^{\nu+\frac{1}{2}} \Big(1 + \emptyset_{\mu,\nu}(y) \Big). \tag{A.1.5}$$

It is also apparent from (A.1.1) and (A.1.2) that when $\nu - \mu = 1/2$ we have

$$M_{\mu,\nu}(y) + (2\mu+1) W_{\mu,\nu}(y) = \Gamma(2\mu+2) y^{-\mu} e^{y/2}, \tag{A.1.6}$$

while when $\nu + \mu = 1/2$ we have from (A.1.2) that

$$W_{\mu,\nu}(y) = y^{\mu} e^{-y/2}. \tag{A.1.7}$$

The I-Bessel and K-Bessel functions are special Whittaker functions [93]:

$$I_\nu(y) = 2^{-2\nu-\frac{1}{2}} \Gamma(\nu+1)^{-1} y^{-\frac{1}{2}} M_{0,\nu}(2y) \qquad \text{and} \qquad K_\nu(y) = \sqrt{\tfrac{\pi}{2y}}\, W_{0,\nu}(2y).$$

Their asymptotic properties for large y thus follow from (A.1.4).

A.2 An integral

We give the proof of the following evaluation. For $\mathrm{Re}(s) > 0$ we have

$$\int_0^\pi \cos(t\cos\theta) I_{s-\frac{1}{2}}(t\sin\theta) \frac{d\theta}{(\sin\theta)^{1/2}} = 2^{s-1}\frac{\Gamma(\frac{s}{2})^2}{\Gamma(s)} J_{s-1/2}(t).$$

Proof. Denote the left hand side by $L_s(t)$. We use the definition of $I_{s-\frac{1}{2}}$ in (4.2.16) to get

$$L_s(t) = \sum_{k=0}^{\infty} \frac{(t/2)^{s+2k-1/2}}{k!\Gamma(s+k+\frac{1}{2})} \int_0^\pi \cos(t\cos\theta)(\sin\theta)^{s+2k-1} d\theta.$$

Lommel's integral representation [125, p. 47] gives for $\mathrm{Re}\, v > -1/2$ that

$$J_\nu(y) = \frac{(y/2)^\nu}{\Gamma(\nu+\frac{1}{2})\Gamma(\frac{1}{2})} \int_0^\pi \cos(y\cos\theta)(\sin\theta)^{2\nu} d\theta.$$

Thus for $\mathrm{Re}(s) > 0$ we have that

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

This Neumann series can be evaluated (see [125, p.143,eq.1]) giving for $\mathrm{Re}(s) > 0$

$$L_s(t) = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{s}{2})}{\Gamma(\frac{s}{2}+\frac{1}{2})} J_{s-1/2}(t).$$

The result follows by the duplication formula for $\Gamma(s)$. □

A.3 Another integral

In this appendix we give a proof of the following integral formula which was given in Lemma 2.2.11. For $\mu \in \mathbb{C}$, $t > 0$ and $Re(s) > 0$

$$\int_0^\pi e^{\pm i(t cos\theta+\mu\theta)} M_{\mu,s-1/2}(2t\sin\theta) \frac{d\theta}{\sin\theta} = G(s,\mu) t^{1/2} J_{s-1/2}(t) \tag{A.3.1}$$

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.3.1) are positive since the formula with negative signs follows by complex conjugation. To prove the Lemma we will prove that both sides of (A.3.1) 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

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

Hence we want to show that the left hand side of (A.3.1) also satisfies

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

dc_1553_18

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

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

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

leads to

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

Using this last equation gives for the integral in (A.3.1)

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

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

$$2i\int\limits_0^\pi e^{i(tcos\theta+\mu\theta)} M'_{\mu,\lambda}(2t\sin\theta) 2\cos\theta d\theta = -\frac{2i}{t}\int\limits_0^\pi \frac{d}{d\theta}\left(e^{i(tcos\theta+\mu\theta)}\right) M_{\mu,\lambda}(2t\sin\theta)d\theta$$

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

$$-\frac{2i}{t}\int\limits_0^\pi \frac{d}{d\theta}\left(e^{i(tcos\theta+\mu\theta)}\right) M_{\mu,\lambda}(2t\sin\theta)d\theta = \int_0^\pi \left(\frac{2\mu}{t} - 2\sin\theta\right) h(t)e^{i\mu\theta}d\theta$$

we have

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

This proves that both sides of (A.3.1) 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

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

Using the integral formula (see [105, 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

$$\int_0^\pi e^{i(tcos\theta+\mu\theta)} M_{\mu,s-1/2}(2t\sin\theta)\frac{d\theta}{\sin\theta}$$
$$= (2\pi)e(\mu/4)\frac{\Gamma(2s)}{\Gamma(s-\mu)}\sum_{\ell=0}^\infty \sum_{m+n=\ell} \frac{(-1)^m\Gamma(s-\mu+n)\Gamma(s+n)}{m!n!\Gamma(2s+n)\Gamma(\frac{n+s+m+\mu+1}{2})\Gamma(\frac{n+s-m-\mu+1}{2})}t^{s+\ell} \quad \text{(A.3.2)}$$

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.3.1)

$$G(s,\mu)t^{1/2}J_{s-1/2}(t) = \frac{(\pi)^{3/2}e(\mu/4)2^{2-2s}\Gamma(2s)}{\Gamma(\frac{s+1+\mu}{2})\Gamma(\frac{s+1-\mu}{2})}\sum_{r=0}^\infty \frac{(-1)^r 2^{-2r}}{r!\Gamma(s+1/2+r)}t^{s+2r} \quad \text{(A.3.3)}$$

A straightforward calculation shows that the coefficients of t^s, t^{s+1} and t^{s+2} in (A.3.2) and (A.3.3) match, which is more than what is needed to finish the proof of the Lemma. □

...

Bibliography

[1] A'Campo, N. Le groupe de monodromie du déploiement des singsingular isolées de courbes planes II. Actes du Congrés de Vancouver (1974).

[2] A'Campo, Norbert. "Generic immersions of curves, knots, monodry and Gordian number." Publications Mathématiques de l'IHÉS 88 (1998): 151-169.

[3] Andrews, George E. Mock theta functions. Theta functions 1987, Part 2 , 283-298, Proc. Sympos. Pure Math., 49, Part 2, Amer. Math. Soc., Providence, RI, 1989.

[4] Asai, T.; Kaneko, M.; Ninomiya, H., Zeros of certain modular functions and an application. Comment. Math. Univ. St. Paul. 46 (1997), no. 1, 93–101.

[5] Ash, Avner. Parabolic cohomology of arithmetic subgroups of SL (2, Z) with coefficients in the field of rational functions on the Riemann sphere. American Journal of Mathematics (1989): 35–51.

[6] Atiyah, Michael, The logarithm of the Dedekind η-function. Math. Ann. 278 (1987), no. 1-4, 33–380.

[7] Barge, J. ; Ghys, É. Cocycles d'Euler et de Maslov. (French) [Euler and Maslov cocycles] Math. Ann. 294 (1992), no. 2, 235–265.

[8] Baruch, Ehud Moshe; Mao, Zhengyu A generalized Kohnen-Zagier formula for Maass forms. J. Lond. Math. Soc. (2) 82 (2010), no. 1, 1–16.

[9] Beardon, Alan F. The geometry of discrete groups. Graduate Texts in Mathematics, 91. Springer-Verlag, New York, 1983.

[10] Birkhoff, George D. "Dynamical systems with two degrees of freedom." Proceedings of the National Academy of Sciences of the United States of America 3.4 (1917): 314.

[11] Biró, A. Cycle integrals of Maass forms of weight 0 and Fourier coefficients of Maass forms of weight 1/2. Acta Arith. 94 (2000), no. 2, 103–152.

[12] Blomer, V.; Harcos, G., Hybrid bounds for twisted L-functions. J. Reine Angew. Math. 621 (2008), 53–79.

[13] Blomer, V., Harcos, G., Michel, P.: A Burgess-like subconvex bound for twisted L-functions. Forum Math. 19, 61–106

[14] A. Booker, A. Strömbergsson, A. Venkatesh, Effective Computation of Maass Cusp Forms, IMRN 2006, Article ID 71281, 34 pages.

[15] Borcherds, R. E., Automorphic forms on $O_{s+2,2}(R)$ and infinite products. Invent. Math. 120 (1995), no. 1, 161–213.

[16] Borthwick, David Spectral theory of infinite-area hyperbolic surfaces. Progress in Mathematics, 256. Birkhauser Boston, Inc., Boston, MA, 2007. xii+355 pp.

[17] Bringmann, K. ; Ono, K., Arithmetic properties of coefficients of half-integral weight Maass-Poincaré series. Math. Ann. 337 (2007), no. 3, 591–612.

[18] Bruinier, J.H., Borcherds products and Chern classes of Hirzebruch-Zagier divisors. Invent. Math. 138 (1999), no. 1, 51–83.

[19] Bruinier, J. H.; Funke, J., Traces of CM values of modular functions. J. Reine Angew. Math. 594 (2006), 1–33.

[20] Bruinier, J. H.; Funke, J., On two geometric theta lifts. Duke Math. J. 125 (2004), no. 1, 45–90.

[21] Bruinier and Ono, K., Heegner divisors, L-functions and harmonic weak Maass forms, to appear.

[22] Brunella, Marco. "On the discrete Godbillon-Vey invariant and Dehn surgery on geodesic flows." Annales de la Faculte des sciences de Toulouse: Mathématiques. Vol. 3. No. 3. 1994.

[23] Burgess, D. A. On character sums and L-series. Proc. London Math. Soc. (3) 12 1962 193-206.

[24] Calegari, D. ; Louwsma, J. Immersed surfaces in the modular orbifold, Proc. Amer. Math. Soc. 139 (2011), 2295–2308.

[25] Choie, Yj.; Zagier, D. Rational period functions for $PSL(2, \mathbb{Z})$. A tribute to Emil Grosswald: number theory and related analysis, 89–108, Contemp. Math., 143, Amer. Math. Soc., Providence, RI, 1993.

[26] Chiera, F. L. On Petersson products of not necessarily cuspidal modular forms. J. Number Theory 122 (2007), no. 1, 13–24.

[27] Cohen, H. Sums involving the values at negative integers of L functions of quadratic characters. Séminaire de Théorie des Nombres, 1974-1975 (Univ. Bordeaux I, Talence), Exp. No. 3, 21 pp.

[28] Cohen, H.; Lenstra, H. W., Jr. Heuristics on class groups. Number theory (New York, 1982), 26-36, Lecture Notes in Math., 1052, Springer, Berlin, 1984.

[29] Conrey, J. B.; Iwaniec, H. The cubic moment of central values of automorphic L-functions. Ann. of Math. (2) 151 (2000), no. 3, 1175–1216.

[30] Davenport, Harold Multiplicative number theory. Third edition. Revised and with a preface by Hugh L. Montgomery. Graduate Texts in Mathematics, 74. Springer-Verlag, New York, 2000.

[31] Dedekind, R., Gesammelte mathematische Werke. Chelsea Publishing Co., New York 1968 Vol. I. 159–173.

[32] Dehornoy, Pierre, Enlacement entre géodésiques sur une orbifold. (French. English, French summary) [Linking between geodesics on an orbifold] C. R. Math. Acad. Sci. Paris 350 (2012), no. 1-2, 77–80.

[33] Dehornoy, Pierre, Intersection norms on surfaces and Birkhoff surfaces for geodesic flows, eprint arXiv:1604.06688

[34] Deshouillers, J.M. and Iwaniec, H.,Kloosterman sums and Fourier coefficients of cusp forms. Inventiones mathematicae, 70(2), pp.219-288, 1982.

[35] Duke, W. Hyperbolic distribution problems and half-integral weight Maass forms. Invent. Math. 92 (1988), no. 1, 73–90.

[36] Duke, W. "Almost a Century of Answering the Question: What Is a Mock Theta Function?." Notices of the AMS 61.11 (2014).

[37] Duke, W.; Friedlander, J. B.; Iwaniec, H. Equidistribution of roots of a quadratic congruence to prime moduli. Ann. of Math. (2) 141 (1995), no. 2, 423–441.

[38] Duke, W.; Friedlander, J.; Iwaniec, H. Bounds for automorphic L-functions. Invent. Math. 112 (1993), no. 1, 1–8.

[39] Duke, W.; Friedlander, J. B.; Iwaniec, H. Weyl sums for quadratic roots. Int. Math. Res. Not. IMRN 2012, no. 11, 2493–2549.

[40] Duke, W.; Imamoglu, Ö. and Tóth, A. Cycle integrals of the j-function and mock modular forms. Ann. of Math. (2) 173 (2011), no. 2, 947–981.

[41] Duke, W.; Imamoglu, Ö. and Tóth, A. Rational period functions and cycle integrals. Abh. Math. Semin. Univ. Hambg. 80 (2010), no. 2, 255–264.

[42] Duke, W.; Imamoglu, Ö. and Tóth, A. Modular cocycles and linking numbers. (Accepted at Duke Math Journal.)

[43] Duke, W., Imamoglu, Ö. and Tóth, Á., 2016. Geometric invariants for real quadratic fields. Annals of Mathematics, pp.949-990.

[44] Eichler, M. Grenzkreisgruppen und kettenbruchartige Algorithmen. (German) Acta Arith. 11 1965 169–180.

[45] Einsiedler, Manfred; Lindenstrauss, Elon; Michel, Philippe; Venkatesh, Akshay The distribution of closed geodesics on the modular surface, and Duke's theorem. Enseign. Math. (2) 58 (2012), no. 3-4, 249–313.

[46] Elstrodt, J; Die Resolvente zum Eigenwertproblem der automorphen Formen in der hyperbolischen Ebene. I Math. Ann. 203 (1973), 295–300, II Math. Z. 132 (1973), 99–134, III Math. Ann. 208 (1974), 99–132.

[47] Farb, Benson; Margalit, Dan A primer on mapping class groups. Princeton Mathematical Series, 49. Princeton University Press, Princeton, NJ, 2012. xiv+472 pp.

[48] Fay, John D. Fourier coefficients of the resolvent for a Fuchsian group. J. Reine Angew. Math. 293/294 (1977), 143–203.

[49] Fried, David. Transitive Anosov flows and pseudo-Anosov maps. Topology 22.3 (1983): 299-303.

[50] C. F. Gauss, Note dated 22 Jan. 1833, in Werke, Vol. V, ed. C. Sch afer Konigliche Gesellschaft der Wissenschaften zu Gottingen, Leipzig, Berlin, 1867, p. 605

[51] Goldfeld, Dorian; Hoffstein, Jeffrey Eisenstein series of 12-integral weight and the mean value of real Dirichlet L-series. Invent. Math. 80 (1985), no. 2, 185-208.

[52] Ghys,E., Knots and dynamics. International Congress of Mathematicians. Vol. I, 247–277, Eur. Math. Soc., Zürich, 2007.

[53] Ghys, E., Right-handed vector fields and the Lorenz attractor. Jpn. J. Math. 4 (2009), no. 1, 47–61.

[54] Geiges, Hansjörg. An introduction to contact topology. Cambridge Studies in Advanced Mathematics, 109. Cambridge University Press, Cambridge, 2008.

[55] Goldfeld, D.; Sarnak, P. Sums of Kloosterman sums. Invent. Math. 71 (1983), no. 2, 243–250.

[56] Gray, J.J., János Bolyai, non-euclidean geometry, and the nature of space (Vol. 1). Burndy Library MIT Press. 2004.

[57] Gross, B.; Kohnen, W.; Zagier, D., Heegner points and derivatives of L-series. II. Math. Ann. 278 (1987), no. 1-4, 497–562.

[58] Gusein-Zade, S. M. Matrices d'intersections pour certaines singularités de fonctions de 2 variables. Funkcional. Anal. i Prilozen 8.11-15 (1974).

[59] Hammond, William F. The modular groups of Hilbert and Siegel. Amer. J. Math. 88 1966 497–516.

[60] Hardy, Godfrey H., and Srinivasa Ramanujan. "Asymptotic formulae in combinatory analysis." Proceedings of the London Mathematical Society 2.1 (1918): 75-115.

[61] Hirzebruch, Friedrich E. P. Hilbert modular surfaces. Enseignement Math. (2) 19 (1973), 183–281.

[62] Hirzebruch, F.; Zagier, D., Intersection numbers of curves on Hilbert modular surfaces and modular forms of Nebentypus. Invent. Math. 36 (1976), 57–113.

[63] Hecke, Erich Mathematische Werke. 290–312

[64] Hecke, Erich Lectures on Dirichlet series, modular functions and quadratic forms. Edited by Bruno Schoeneberg. With the collaboration of Wilhelm Maak. Vandenhoeck & Ruprecht, Göttingen, 1983. 98 pp.

[65] Hejhal, Dennis A. The Selberg trace formula for PSL(2,R). Vol. 2. Lecture Notes in Mathematics, 1001. Springer-Verlag, Berlin, 1983. viii+806 pp.

[66] Hooley, C., An asymptotic formula in the theory of numbers. Proc. London Math. Soc. (3) 7 (1957), 396–413.

[67] Hooley, C., On the number of divisors of a quadratic polynomial. Acta Math. 110 1963 97–114.

[68] Humphreys, James E., Introduction to Lie algebras and representation theory. Second printing, revised. Graduate Texts in Mathematics, 9. Springer-Verlag, New York-Berlin, 1978.

[69] Ibukiyama, T.; Saito, H., On zeta functions associated to symmetric matrices. II. MPI Preprint.

[70] Ishikawa, Masaharu. Tangent circle bundles admit positive open book decompositions along arbitrary links. Topology 43.1 (2004): 215-232.

[71] Iwaniec, Henryk Fourier coefficients of modular forms of half-integral weight. Invent. Math. 87 (1987), no. 2, 385-401.

[72] Iwaniec, Henryk Spectral methods of automorphic forms. Second edition. Graduate Studies in Mathematics, 53. American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002. xii+220 pp.

[73] Iwaniec, Henryk. Topics in classical automorphic forms. Vol. 17. American Mathematical Soc., 1997.

[74] Iwaniec, Henryk; Kowalski, Emmanuel Analytic number theory. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004. xii+615 pp.

[75] Jenkins, P., Kloosterman sums and traces of singular moduli. J. Number Theory 117 (2006), no. 2, 301–314.

[76] Katok, Svetlana Coding of closed geodesics after Gauss and Morse. Geom. Dedicata 63 (1996), no. 2, 123–145.

[77] Katok, Svetlana; Sarnak, Peter Heegner points, cycles and Maass forms. Israel J. Math. 84 (1993), no. 1-2, 193–227.

[78] Kim, Henry, and Peter Sarnak. "Refined estimates towards the Ramanujan and Selberg conjectures." J. Amer. Math. Soc 16.1 (2003): 175-181.

[79] Kirby, R., Melvin, P., Dedekin sums, μ invariants and the signature cocycle, Math. Ann. 299 (1994), no. 2, 231–267.

[80] Koblitz, Neal I. Introduction to elliptic curves and modular forms. Vol. 97. Springer Science & Business Media, 2012.

[81] Kohnen, Beziehungen zwischen Modulformen halbganzen Gewichts und Modulformen ganzen Gewichts, Bonner mathematische Schrifte, 131, Bonn, 1981.

[82] Kohnen, W., Modular forms of half-integral weight on $\Gamma_0(4)$. Math. Ann. 248 (1980), no. 3, 249–266.

[83] Kohnen, Winfried Fourier coefficients of modular forms of half-integral weight. Math. Ann. 271 (1985), no. 2, 237–268.

[84] Kohnen, W.; Zagier, D.,Values of L-series of modular forms at the center of the critical strip. Invent. Math. 64 (1981), no. 2, 175–198.

[85] Knopp, Marvin I., Some new results on the Eichler cohomology of automorphic forms. Bull. Amer. Math. Soc. 80 (1974), 607–632.

[86] Knopp, Marvin I., Rational period functions of the modular group. With an appendix by Georges Grinstein. Duke Math. J. 45 (1978), no. 1, 47–62.

[87] Knopp, Marvin I., Rational period functions of the modular group. II. Glasgow Math. J. 22 (1981), no. 2, 185–197.

[88] Knopp, Marvin I., On the growth of entire automorphic integrals. Results Math. 8 (1985), no. 2, 146–152.

[89] Lebedev, N. N. ,Special functions and their applications. Revised edition, translated from the Russian and edited by Richard A. Silverman. Unabridged and corrected republication. Dover Publications, Inc., New York, 1972.

[90] Lickorish, WB Raymond. An introduction to knot theory. Vol. 175. Springer, 1997.

[91] Linnik, Y.V., Additive problems and eigenvalues of the modular operators. Proc. Intern. Cong. Math. Stockholm, 1962.

[92] Maass, H.: Über die räumliche Verteilung der Punkte in Gittern mit indefiniter Metrik. Math. Ann. 138, 287-315 (1959)

[93] Magnus, W., Oberhettinger, F., Soni, R. P., Formulas and theorems for the special functions of Mathematical physics, Springer-verlag, Berlin 1966

[94] Maskit, Bernard Kleinian groups. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 287. Springer-Verlag, Berlin, 1988. xiv+326 pp.

[95] Milnor, J., Introduction to algebraic K-theory. Annals of Mathematics Studies, No. 72. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1971.

[96] Motohashi, Y., 1997. Spectral theory of the Riemann zeta-function (Vol. 127). Cambridge University Press.

[97] Mozzochi, C. J., Linking numbers of modular geodesics. Israel J. Math. 195 (2013), no. 1, 71–95.

[98] H. Neunhöffer, Über die analytische Fortsetzung von Poincaréreihen, S.-B. Heidelberger Akad. Wiss. Math.-Natur. Kl. (1973), 33–90.

[99] D. Niebur, A class of nonanalytic automorphic functions, Nagoya Math. J. 52 (1973), 133–145.

[100] NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.19 of 2018-06-22. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, and B. V. Saunders, eds.

[101] Ono, K. , Unearthing the visions of a master: harmonic Maass forms and number theory. (English summary) Current developments in mathematics, 2008, 347–454, Int. Press, Somerville, MA, 2009.

[102] Parson, L. A., Modular integrals and indefinite binary quadratic forms. A tribute to Emil Grosswald: number theory and related analysis, 513–523, Contemp. Math., 143, Amer. Math. Soc., Providence, RI, 1993.

[103] Parson, L. A., Rational period functions and indefinite binary quadratic forms. III. A tribute to Emil Grosswald: number theory and related analysis, 109–116, Contemp. Math., 143, Amer. Math. Soc., Providence, RI, 1993.

[104] Poincaré, H. Théorie des groups fuchsiens, Acta Math. 1 (1882) 1–62.

[105] Rademacher, H.; Grosswald, E., "Dedekind Sums, The Carus Math." Monographs, MAA (1972).

[106] 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.

[107] Salie, H. Uber die Kloostermanschen Summen S(u,v; q), Math. Zeit. 34 (1931–32) pp. 91–109.

[108] Sarnak, Peter, Class numbers of indefinite binary quadratic forms. J. Number Theory 15 (1982), no. 2, 229–247.

[109] Sarnak, Peter Some applications of modular forms. Cambridge Tracts in Mathematics, 99. Cambridge University Press, Cambridge, 1990. x+111 pp.

[110] Sarnak, Peter Reciprocal geodesics. Analytic number theory, 217–237, Clay Math. Proc., 7, Amer. Math. Soc., Providence, RI, 2007.

[111] Sarnak, P., Letter to J. Mozzochi on linking numbers of modular geodesics,
http://publications.ias.edu/sarnak/paper/504

[112] Sawin, W.F., 2016. A Tannakian Category and a Horizontal Equidistribution Conjecture for Exponential Sums (Doctoral dissertation, Princeton University).

[113] Selberg, Atle, Uber die Mock-Theta funktionen siebenter Ordnung, Arch. Math. Naturvidenskab, 41 (1938), 3-15 in Collected Papers Vol 1

[114] Selberg, Atle, On the estimation of Fourier coefficients of modular forms. 1965 Proc. Sympos. Pure Math., Vol. VIII pp. 1–15 Amer. Math. Soc., Providence, R.I.

[115] Serre, Jean-Pierre. A course in arithmetic. Vol. 7. Springer Science & Business Media, 2012.

[116] Siegel, Carl Ludwig, Lectures on advanced analytic number theory. Tata Institute of Fundamental Research Lectures on Mathematics, No. 23 Tata Institute of Fundamental Research, Bombay 1980

[117] Siegel, Carl Ludwig, Die Functionalgleichungen einiger Dirichletscher Reihen, Math. Zeitschrift 63 (1956), 363–373

[118] Shimura, Goro. "Modular forms of half integral weight." Modular Functions of One Variable I. Springer Berlin Heidelberg, 1973. 57-74.

[119] Shintani, Takuro On construction of holomorphic cusp forms of half integral weight. Nagoya Math. J. 58 (1975), 83–126.

[120] Stanley, R.P., 1986. Enumerative combinatorics, Wadsworth Publ. Co., Belmont, CA.

[121] Stillwell, John. Sources of hyperbolic geometry. No. 10. American Mathematical Soc., 1996.

[122] Strömberg, Fredrik Computation of Maass waveforms with nontrivial multiplier systems. Math. Comp. 77 (2008), no. 264, 2375–2416.

[123] Tóth, Á., Roots of quadratic congruences. International Mathematics Research Notices,(14), pp.719-739, 2000.

[124] Tóth, Á., On the evaluation of Salié sums. Proc. Amer. Math. Soc. 133 , no. 3, 643–645, 2005.

[125] Waldspurger, J.-L. Sur les coefficients de Fourier des formes modulaires de poids demi-entier. (French) [On the Fourier coefficients of modular forms of half-integral weight] J. Math. Pures Appl. (9) 60 (1981), no. 4, 375–484.

[126] Whittaker, E. T.; Watson, G. N. 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 Mathematical Library. Cambridge University Press, Cambridge, 1996. vi+608 pp.

[127] Watson, G. N. The Final Problem : An Account of the Mock Theta Functions. J. London Math. Soc. 111 no. 1, 55. 1935

[128] Weil, A., On some exponential sums. Proceedings of the National Academy of Sciences, 34(5), pp.204-207, 1948.

[129] Williams, K. S. Note on the Kloosterman sum, Transactions of the American Mathematical Society 30(1), pages: 61–62, 1971.

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

[131] Zagier, D., Modular forms associated to real quadratic fields. Invent. Math. 30 (1975), no. 1, 1–46.

[132] Zagier, D, A Kronecker limit formula for real quadratic fields Math. Annalen 213 (1975) 153–184

[133] Zagier, D. Ramanaujan's mock theta functions and their applications [d'áprès Zwegers and Bringmann-Ono], Seminaire Bourbaki, 1986

[134] Zagier, D., Traces of singular moduli. Motives, polylogarithms and Hodge theory, Part I (Irvine, CA, 1998), 211–244, Int. Press Lect. Ser., 3, I, Int. Press, Somerville, MA, 2002.

[135] Zagier, D. B. Zetafunktionen und quadratische Körper. (German) [Zeta functions and quadratic fields] Eine Einführung in die höhere Zahlentheorie. [An introduction to higher number theory] Hochschultext. [University Text] Springer-Verlag, Berlin-New York, 1981.

[136] Zwegers, S. P., Mock Theta Functions, Utrecht PhD thesis, (2002) ISBN 90-393-3155-3

[137] Zwegers, S. P. Mock θ-functions and real analytic modular forms. q-series with applications to combinatorics, number theory, and physics (Urbana, IL, 2000), 269–277, Contemp. Math., 291, Amer. Math. Soc., Providence, RI, 2001.