The arithmetic and geometry of some hyperbolic three manifolds

by

P. SARNAK[1]

Courant Institute
New York, N.Y., U.S.A.

Contents

1. Introduction

In section 304 of his "Disquestiones Arithmeticae" Gauss observed numerically that the asymptotic averages of class numbers of indefinite binary quadratic forms (over **Z**) when ordered by their discriminants are rather erratic. Today the behavior of the class number for large discriminant still remains a major unsolved problem. In [17] we showed that if we form the averages of these class numbers when ordered by the sizes of the corresponding fundamental units (or regulators) then there is an asymptotic law. The main result of this paper is to derive similar such asymptotic expressions for averages of binary quadratic forms over the integers of an imaginary quadratic number field. As will be seen there are some interesting differences. We believe that our results have an appropriate extension to an arbitrary number field.[2] To develop these results we will need to analyze in some detail various aspects of the geometry of certain hyperbolic three manifolds. Along the way various auxiliary theorems, which are of interest in their own right, concerning these manifolds, will be proven. These are stated in the following outline of the various sections.

[1] Research supported in part by NSF Grant NSF MCS 7900813.
[2] For the appropriate explicit trace formulas for totally real number fields see Efrat [4].

Let $D>0$ be a square free rational integer, and let $K_D=Q(\sqrt{-D})$. We denote by $\mathcal{O}_D$ the ring of integers of K_D. The discrete subgroups $SL_2(\mathcal{O}_D)$ of $SL_2(\mathbf{C})$ will be denoted by Γ_D. Occasionally we will also consider congruence subgroups of these. These groups act discontinuously on $H^3=\{(y,x_1,x_2):y>0\}$ equipped with the line element $ds^2=(dx_1^2+dx_2^2+dy^2)/y^2$.

If

$$\tau=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\in SL_2(\mathbf{C})$$

then τ acts on H^3 as an isometry by

$$(y,z)\mapsto\left(\frac{y}{|\gamma z+\delta|^2+|\gamma|^2y^2},\frac{(\alpha z+\beta)(\overline{\gamma z+\delta})+\alpha\bar{\gamma}y^2}{|\gamma z+\delta|^2+|\gamma|^2y^2}\right) \tag{1.1}$$

where $z=x_1+ix_2$.

Denote by M_D the quotient H/Γ_D. The geometric quantities of M_D which will be of interest to us are, volumes of M_D, lengths of closed geodesics on M_D, the spectrum of the Laplace Beltrami operator for M_D. As may be expected from the definitions of the groups Γ_D each of these has a number theoretic interpretation. The quotients M_D are of finite volume, and their volumes were computed by Humbert [10] to be

$$\frac{|d|^{3/2}\zeta_k(2)}{4\pi^2}\quad\text{where } d=\operatorname{disc}(K_D), \tag{1.2}$$

and ζ_K is the Dedekind zeta function of the field $K=K_D$.

In section 2 we begin by computing the Eisenstein–Maass series for these groups, explicitly in terms of L functions of K_D.

These are needed because of their close connection with the spectrum of the Laplacian (we will denote the last by Δ). The Eisenstein series are useful in many other respects as well—for example we use them to compute the volumes (1.2) above.

Besides the Eisenstein series we also introduce Poincaré type series, which are generalizations of those introduced by Selberg [20] (see section 2). A computation of the inner product of two such series, turns out to be a zeta type function whose coefficients are Kloosterman sums. Thus if A is a non zero ideal of O and ψ_1,ψ_2 are additive characters of O/A then the sums in question are

$$S(\psi_1,\psi_2,A)=\sum_{\delta\in(O/A)^*}\psi_1(\delta)\,\psi_2(\delta^{-1}).$$

The poles of this "Selberg–Kloosterman" type zeta function (see definition 3.3) are closely related to the spectrum of the Laplacian on M_D. This relationship is used in two natural ways. The first, which is the one used in this paper, is to derive lower bounds on the discrete spectrum of the Laplacian on M_D. We will prove in section 3 the following:

THEOREM 3.1. *If λ_1 is the first discrete eigenvalue of $-\Delta$ on $L^2(M_D)$ (thus $\lambda_0=0$ corresponds to the constant function and the rest of the spectrum is positive) then* $\lambda_1 \geq 3/4$.

The other way of exploiting the above relationship is to use it to study the associated Selberg–Kloosterman zeta function, and in particular to study sums of Kloosterman sums in imaginary quadratic number fields. This has been done in joint work with D. Goldfeld [7].

Theorem 1 has also been proven by completely different and sophisticated methods by Jacquet and Gelbart [28]. Our proof is direct and is in the spirit of Selberg [20], who proved a similar such bound $\lambda_1 \geq 3/16$, for congruence subgroups of the classical modular group.

For the rest of the paper, viz. sections 4–7 we will consider for the sake of simplicity only those D for which K_D is of class number one (i.e. $D=1, 2, 3, 7, 11, 19, 43, 67, 163$). The results go over for any D as long as one uses the notion of "primitive discriminants" as in Speizer [29]. The closed geodesics on M_D are identified through the theory of quadratic forms over $\mathcal{O}_D$ (we call these Dirichlet forms). We introduce the elements of this theory in section 4. Denote by $\mathcal{D}$ the set of discriminants of binary quadratic forms

$$\mathcal{D} = \{m \in \mathcal{O}_D \colon m \equiv x^2 \pmod 4 \text{ for some } x \in \mathcal{O}_D \text{ and } m \text{ is not a perfect square}\}.$$

The last condition in the definition of $\mathcal{D}$ ensures that a form of discriminant $d \in \mathcal{D}$ will not factor over $\mathcal{O}_D$.

Equivalence of forms is defined in the narrow sense i.e. if $Q=ax^2+bxy+cy^2$, $Q'=a'(x')^2+b'(x'y')+c'(y')^2$ then $Q \sim Q'$ if

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with } \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL_2(\mathcal{O}_D).$$

For $d \in \mathcal{D}$ let $h(d)$ be the number of classes of primitive quadratic forms of discriminant d. We discuss the solutions to the diophantine equation

$$t^2 - du^2 = 4, \quad d \in \mathcal{D},\ t, u \in \mathcal{O}.$$

Except for finitely many d, the solutions to this equation are generated by a fundamental "unit"

$$\varepsilon_d = \frac{t_0 + \sqrt{d}\, u_0}{2} \in Q(\sqrt{-D}, \sqrt{d}), \quad |\varepsilon_d| > 1.$$

The lengths of the closed geodesics as well as some other quantities related to closed geodesics (Poincaré maps) on M_D, are described in terms of $h(d)$ and ε_d, $d \in \mathcal{D}$ (see Theorem 4.1). Though of great interest and importance in number theory the numbers $h(d)$ and ε_d are not well understood even for forms over **Z**.

In section 5 we prove what we call the prime geodesic theorem. This gives the asymptotic distribution of the lengths of closed geodesics on the manifolds M_D. Let $\pi(x)$ be the number of primitive closed geodesics whose length is not greater than x, then we prove in Theorem 5.1:

$$\pi(x) = \mathrm{Li}\,(e^{2x}) + O(e^{\gamma x}) \quad \text{for any } \gamma > \tfrac{5}{3}$$

where

$$\mathrm{Li}\,(u) = \int_2^u \frac{1}{\log t}\, dt.$$

For more on the existing literature on such asymptotics see remarks 5.2.

Let $\mathcal{D}_x = \{d \in \mathcal{D} : |\varepsilon_d| \leq x\}$. As $x \to \infty$ these sets form an increasing exhausting family of sets. In section 6 by use of more standard methods of additive number theory we find the asymptotics of $|\mathcal{D}_x|$ = cardinality of $\mathcal{D}_x$, as $x \to \infty$.

Finally in section 7 the various results are put together to prove the results mentioned at the beginning of the introduction. We state the asymptotic results here:

THEOREM 7.2. *There is a constant $c_k > 0$ depending on K_D such that*

$$\frac{1}{|\mathcal{D}_x|} \sum_{d \in \mathcal{D}_x} h(d) = \frac{\mathrm{Li}\,(x^4)}{c_k x^2} + O(x^\gamma) \quad \textit{for any } \gamma > \tfrac{4}{3} \textit{ as } x \to \infty.$$

The constant c_k may be computed exactly, see section 6, and turns out to be a rational number times

$$\frac{\pi}{\sqrt{D}} \frac{(\zeta_k(2))^2}{\zeta_k(4)}.$$

Theorem 7.2 tells us that on the average, in this ordering h is about the square of the unit. In the case of forms over $\mathbf{Z}$, see [17], h turned out to be about the size of the unit. This difference is not too surprising in view of the fact that the class numbers are related to those of biquadratic extensions in the case of forms over $\mathcal{O}_D$, and the discriminants of these fields increased by a square from what the corresponding ones over $\mathbf{Z}$ are. On the other hand the units are the same size, so by the Siegel-Brauer theorem, see Lang [13] we expect h should increase accordingly.

2. Maass–Eisenstein and Poincaré series

Our aim in this section is to introduce Eisenstein and Poincaré series attached to the cusp at infinity of Γ_D. These are needed for the spectral analysis of Δ on M_D. The Eisenstein series furnish the continuous spectrum, and will be needed later on when we need the Selberg trace formula. The Poincaré series are used to make a non-trivial estimation of the discrete spectrum. The use of the Eisenstein series in volume computations is also carried out in this section.

We will always use the co-ordinates (y, x_1, x_2) for H^3, as in the introduction. The Laplace–Beltrami operator is then given by:

$$\Delta = y^2\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial y^2}\right)-y\frac{\partial}{\partial y}. \tag{2.1}$$

D is fixed as in section 1 (not necessarily corresponding to class number one) and let

$$\Gamma_\infty = \{\gamma\in\Gamma_D\colon \gamma(\infty)=\infty\} \tag{2.2}$$

be the stabilizer of infinity. For simplicity we assume $D\neq 1$, or 3 so the only units of $\mathcal{O}$ are ± 1. It is then not difficult to see that

$$\Gamma_\infty = \left\{\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}\colon b\in\mathcal{O}\right\} \tag{2.3}$$

(recall we are working mod $\pm I$ i.e. in $PSL_2(\mathbf{C})$).

The Eisenstein series corresponding to the cusp at infinity are defined by the sum:

$$E(s,\omega) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (y(\gamma\omega))^s \tag{2.4}$$

where $\omega=(y, x_1, x_2)$.

The series converges absolutely for $\operatorname{Re}(s)>2$ and since y^s is an eigenfunction of Δ, so is $E(s,\omega)$, the latter coming from the former by averaging by isometries. Thus

$$\Delta E(s,\omega)+s(2-s)E(s,\omega)=0 \tag{2.5}$$

and due to the averaging, we have

$$E(s,\gamma\omega)=E(s,\omega) \quad \text{for } \gamma\in\Gamma_D.$$

So $E(s,\omega)$ gives us eigenfunctions of Δ on M_D, though not in L^2. The series $E(s,\omega)$ may, by the general theory of Selberg, be continued to a meromorphic function of s. Its analytic nature is controlled by the zeroth coefficients of $E(s,\omega)$ in its Fourier expansion in the cusp. More precisely, since

$$E(s,y,x)=E(s,y,x+b), \quad \forall b\in\mathcal{O}$$

we may write

$$E(s,\omega)=\sum_{l\in\mathcal{O}^*} A_l(y,s)\,e(\langle l,x\rangle) \tag{2.6}$$

where $\mathcal{O}^*$ is the dual lattice to $\mathcal{O}$ in $\mathbf{R}^2$, and $e(z)=e^{2\pi iz}$. By separation of variables and (2.5)

$$E(s,\omega)=y^s+\varphi_{11}(s)\,y^{2-s}+\text{nonzero coefficient} \tag{2.6$'$}$$

we need to know $\varphi_{11}(s)$ explicitly.

(2.7) *Explicit calculation of the zeroth coefficient.*

When we think of $\mathcal{O}$ as a lattice in $\mathbf{R}^2$ we denote it by L. A basis for L as a Z module is

$$1,\omega' \quad \text{where } \omega'=\begin{cases}\dfrac{1+i\sqrt{D}}{2} & \text{if } D\equiv 3(4)\\ i\sqrt{D} & \text{if } D\not\equiv 3(4)\end{cases}.$$

A coset representation for $\Gamma_\infty\backslash\Gamma$ is

$$\left\{\begin{pmatrix}* & *\\ c & d\end{pmatrix} : (c,d)=(1) \text{ and } (*,*) \text{ is any choice of two numbers in } \mathcal{O} \text{ so that } \begin{pmatrix}* & *\\ c & d\end{pmatrix}\in\Gamma_D\right\}.$$

Therefore,

$$E(s,\omega)=\sum_{\substack{(c,d)=(1)\\ \text{mod } \pm I}} \frac{y^s}{(|cz+d|^2+|c|^2y^2)^s}, \quad z=x_1+ix_2$$

$$= y^s + \frac{1}{2} \sum_{\substack{(c,d)=1 \\ c \neq 0}} \frac{y^s}{(|cz+d|^2 + |c|^2 y^2)^s}$$

$$= y^s + \frac{1}{2} \sum_{c \neq 0} \sum_{\substack{d \bmod c \\ (d,c)=1}} \sum_m \frac{y^s}{(|cz+d+mc|^2 + y^2|c|^2)^s}$$

$$= y^s + \frac{1}{2} \sum_{c \neq 0} \frac{y^s}{|c|^{2s}} \sum_{\substack{d \bmod c \\ (d,c)=1}} \sum_m \frac{1}{\left(\left|z + \frac{d}{c} + m\right|^2 + y^2\right)^s}.$$

Letting F_L be a fundamental domain for the lattice L in $\mathbf{R}^2$, we obtain the zeroth coefficient by integrating over F_L.

$$\int_{F_L} E(s, y, z)\, dx_1\, dx_2 = V(F_L)\, y^s + \frac{1}{2} \sum_{c \neq 0} \frac{y^s}{|c|^{2s}} \int_{F_L} \sum_{\substack{d \bmod c \\ (d,c)=1}} \sum_m \frac{dx_1\, dx_2}{\left(\left|z + m + \frac{d}{c}\right|^2 + y^2\right)^s}. \tag{2.8}$$

Using

$$\int_{F_L} \sum_{\gamma \in L} f(z+\gamma)\, dx_1\, dx_2 = \int\int_{\mathbf{R}^2} f(x_1, x_2)\, dx_1\, dx_2 \quad \text{with } f(z) = \frac{1}{\left(\left|z + \frac{d}{c}\right|^2 + y^2\right)^s}$$

in (2.8) and also making the change of variable $z' = z - d/c$ we have

$$\int_{F_L} E(s, \omega)\, dx_1\, dx_2 = V(F_L)\, y^s + \frac{1}{2} \sum_{c \neq 0} \sum_{\substack{d \bmod c \\ (d,c)=1}} \frac{y^s}{|c|^{2s}} \int\int_{\mathbf{R}^2} \frac{dx_1\, dx_2}{(x_1^2 + x_2^2 + y^2)^s}$$

$$= V(F_L)\, y^s + \frac{1}{2} \sum_{c \neq 0} \sum_{\substack{d \bmod c \\ (d,c)=1}} \frac{y^s}{|c|^{2s}} \int_0^\infty \int_0^{2\pi} \frac{r\, d\theta\, dr}{(r^2 + y^2)^s}$$

$$= V(F_L)\, y^s + \frac{\pi}{2} \frac{y^{2-s}}{(s-1)} \sum_{c \neq 0} \sum_{\substack{d \bmod c \\ (d,c)=1}} \frac{1}{|c|^{2s}}.$$

Thus

$$E(s, \omega) = y^s + \left(\frac{\pi}{V(F_L)\, 2(s-1)} \sum_{c \neq 0} \frac{\Phi(c)}{|c|^{2s}} \right) y^{2-s} + \text{nonzero Fourier coefficient},$$

where Φ is the Euler function

$$\Phi(x) = \#\{b \bmod x : (b, x) = 1\},$$

a more generally for any ideal A of O

$$\Phi(A) = \#\{\text{residue classes mod} A \text{ prime to } A\}.$$

Thus in (2.6)′

$$\varphi_{11}(s) = \left(\sum_{\substack{A \text{ principal} \\ A \neq 0}} \frac{\Phi(A)}{N(A)^s} \right) \frac{\pi}{V(F_2)(s-1)} \tag{2.9}$$

with $N(A)$=norm of the ideal A.

We express (2.9) in a more convenient form. Let I be the ideal class group of $\mathcal{O}$, with $|I|=h$ the class number of K_D. Let $\psi_1, \psi_2, \ldots, \psi_h$ be the character group to I, with ψ_1=trivial character. ψ_j may be extended to be defined on the ideals of $\mathcal{O}$ in the obvious way. Let L_j be the corresponding L function

$$L_j(s) = \sum_{A \neq 0} \psi_j(A) N(A)^{-s}$$

clearly then

$$\tilde{K}_j(s) = \sum_{A \neq 0} \Phi(A) \psi_j(A) N(A)^{-s} = \frac{L_j(s-1)}{L_j(s)}.$$

Noting that $L_1(s)$ is none other than the Dedekind zeta function of K, $\zeta_k(s)$. From (2.9)

$$\sum_{A \text{ principal}} \frac{\Phi(A)}{N(A)^s} = \frac{1}{h} \sum_{j=1}^{h} \tilde{K}_j(s) = \frac{1}{h} \sum_{j=1}^{h} \frac{L_j(s-1)}{L_j(s)}. \tag{2.10}$$

This gives us an expression for $\varphi_{11}(s)$ in terms of L functions of K. Notice that if $h=1$ then

$$\varphi_{11}(s) = \varphi(s) = \frac{\pi}{V(F_L)(s-1)} \frac{\zeta_k(s-1)}{\zeta_k(s)} \tag{2.11}$$

which will be used later.

Thus the zeroth coefficients of $E(s, \omega)$ is

$$y^s + \left(\frac{\pi}{V(F_L)(s-1)h} \sum_{j=1}^{h} \frac{L_j(s-1)}{L_j(s)} \right) y^{2-s}.$$

Our main interest in $E(s,\omega)$ is in their spectral properties, however they (the Eisenstein series) are useful in other respects. We digress to compute some volumes.

(2.12) *Volumes of* H^3/Γ_D.

LEMMA 2.13.

$$\operatorname{Res}(\varphi_{11}(s), s=2) = \frac{2\pi^2}{V(F_2)\sqrt{|d|}\,\zeta_k(2)\,w}$$

where $d = \operatorname{disc}(K_D)$, *and* $w = \#$ *of roots of* 1 *in* $\mathcal{O}$.

Proof. It is clear by summation by parts that for $\operatorname{Re}(s)>0$, $L_j(s)$ is analytic if $j\neq 1$. Thus the pole at $s=2$ of $\varphi_{11}(s)$ comes from $\tilde{K}_1(s)$ alone. Now $L_1=\zeta_k$ has a simple pole at $s=1$ with residue $h2\pi/W\sqrt{|d|}\,\zeta_k(2)$. (Dirichlet class number formula.) The lemma is now obvious.

Thus

$$\operatorname{Res}[E(s,\omega), s=2] = \frac{2\pi^2}{V(F_L)\,W\sqrt{|d|}\,\zeta_k(2)}. \tag{2.14}$$

Let F_D be a fundamental domain for Γ_D in H^3, and let $V(F_D)$ be its volume. Determining such a fundamental domain in terms of hemispheres that bound it, is not easy. An algorithm for doing so is given in Swan [21] but computations, even for small D, become very complicated. Luckily we do not need to know the fundamental domain explicitly in order to determine its volume. The general shape of a domain F_D looks like Figure 1.

LEMMA 2.15. $\operatorname{Res}(E(s,\omega), s=2) = V(F_L)/V(F_D)$.

Proof. The residue of $E(s,\omega)$ at $s=2$ is a constant function, since it is an eigenfunction of Δ on M_D with eigenvalue $s(2-s)|_{s=2}=0$. Let this constant be c. Thus

$$E(s,\omega) = y^s + \varphi_{11}(s)\,y^{2-s} + H(s,\omega)$$

where $H(s,\omega)$ is analytic in $\operatorname{Re}(s)\geq 2$ and is also uniformly square integrable over F_D for such s. This last claim is a consequence of the fact that the non zero Fourier coefficients in expansion (2.6) decay rapidly as functions of y, into the cusp. (See [12].) Therefore the function

$$b(s) = \int_{F_D} (E(s,\omega) - y^s)\,\frac{dx_1\,dx_2\,dy}{y^3}$$

has residue $cV(F_D)$ at $s=2$.

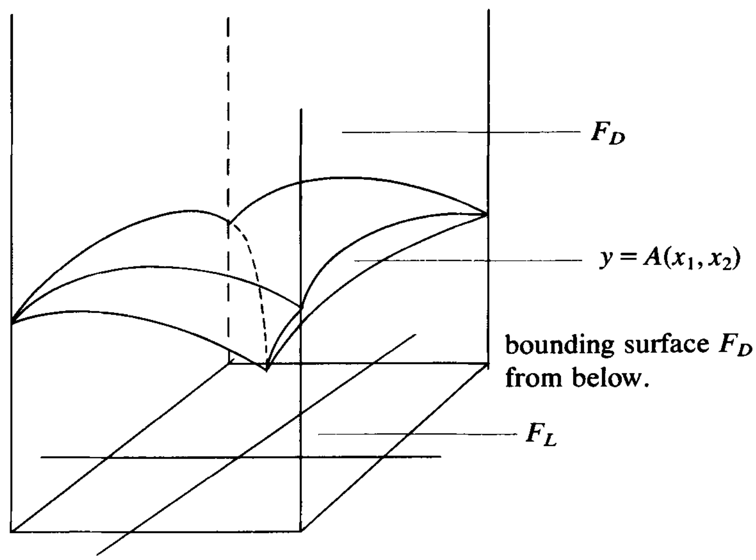

Fig. 1

On the other hand

$$E(s, \omega) - y^s = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma}' y(\gamma\omega)^s$$

the prime denoting, omit γ=identity. So that

$$\begin{aligned} b(s) &= \int_{F_D} (E(s, \omega) - y^s) \frac{dx_1\, dx_2\, dy}{y^3} \\ &= \int_{F_L} \int_0^{A(x)} y^s \frac{dy}{y^3}\, dx_1\, dx_2 \\ &= \int_{F_L} \frac{A(x)^{s-2}}{s-2}\, dx_1\, dx_2. \end{aligned}$$

From which

$$\operatorname{Res}(b(s), s = 2) = V(F_L)$$

$$c = V(F_L)/V(F_D).$$

If we put (2.14) and Lemma 2.15 together with the following elementary facts

$$V(F_L) = \begin{cases} \sqrt{D}/2 & \text{if } D \equiv 3 \pmod 4 \\ \sqrt{D} & \text{if } D \not\equiv 3 \pmod 4 \end{cases}$$

$$d=\begin{cases}-D & \text{if } D\equiv 3 \pmod 4\\ -4D & \text{if } D\not\equiv 3 \pmod 4\end{cases}$$

we obtain the following formula, which was proven by Humbert using other methods (see Humbert [10], or Thurston [22], chapter 7).

PROPOSITION 2.1.

$$V(F_D)=\frac{|d|^{3/2}\zeta_k(2)}{4\pi^2}.$$

Notes. (1) Though our proof does not apply directly to the cases of $Q(\sqrt{-1})$ and $Q(\sqrt{-3})$, the above formula is nevertheless valid since W and a factor of $V(F_L)$ cancel.

(2) In the case $h>1$, there are more cusps besides the one at infinity that we have been considering. In fact it is easy to see that the number of such inequivalent cusps is the class number h. It is then natural and necessary for many purposes to form Eisenstein series in each cusp. For this paper however they will not be needed.

(2.16) *Poincaré series.*

We now introduce a Poincaré type series associated to the cusp at infinity. These generalize the series introduced by Selberg in dimension 2 [20].

Definition 2.17. Let $m\in L^*$, $m\neq 0$ we define

$$P_m(s,\omega)=\sum_{\gamma\in\Gamma_\infty/\Gamma} y(\gamma\omega)^s e^{-2\pi|m|y+2\pi i\langle x,m\rangle}$$

where $\langle x,m\rangle$ is the $\mathbf{R}^2$ standard inner product. We observe that for $\operatorname{Re}(s)>2$ the series converges absolutely, since it is certainly dominated by the Eisenstein series. Since the function $e^{-2\pi|m|y}y^s e^{2\pi i\langle x,m\rangle}$ is Γ_∞ invariant it is clear that $P_m(s,\omega)$ is Γ_D invariant. An advantage that these have over the Eisenstein series is that they are in $L^2(M_D)$ for $\operatorname{Re}(s)>2$. In fact it is clear that for $\varepsilon>0$, $\exists C_\varepsilon$, s.t.

$$\|P_m(s,\cdot)\|_2\leq C_\varepsilon \tag{2.17'}$$

for $\operatorname{Re}(s)>2+\varepsilon$. A simple computation shows that

$$(\Delta+s(2-s))\{y^s e^{-2\pi|m|y}e^{2\pi i\langle x,m\rangle}\}=2\pi|m|(1-2s)\{y^{s+1}e^{-2\pi|m|y}e^{2\pi i\langle x,m\rangle}\}$$

from which

$$(\Delta+s(2-s))P_m(s,\omega)=2\pi|m|(1-2s)P_m(s+1,\omega) \tag{2.18}$$

follows.

We write this in a more convenient form

$$P_m(s,\omega)=2\pi|m|(1-2s)R_{s(2-s)}(P_m(s+1,\omega)) \tag{2.19}$$

R_λ being the resolvent of Δ on $L^2(M_D)$ at λ.

The spectral theory of Δ on $L^2(M_D)$ has been carefully studied by Selberg [19]. In fact the theory has been fully developed for any three manifold of finite volume. The spectrum consists of a finite number of discrete eigenvalues in $[0,1)$, call these $0=\lambda_0<\lambda_1\leqslant\lambda_2\ldots\leqslant\lambda_k<1$. Here $\lambda_0=0$ corresponds to the eigenfunction $u_0(\omega)=\text{constant}$. We denote the corresponding eigenfunction by $u_j(\omega)$ $j=1,2\ldots k$, and call these λ_j, u_j if they exist, *exceptional* spectrum. So $\Delta u_j+\lambda_j u_j=0$, $u_j\in L^2(M_D)$, $0<\lambda_j<1$.

The interval $[1,\infty)$ comprises two kinds of spectrum: Firstly the continuous spectrum which is spanned by the Eisenstein series $E_j(1+it,\omega)$, $t\in\mathbf{R}$, where the j denotes the Eisenstein series in the jth cusp. Each of these is defined in an analogous way to the Eisenstein series in the infinite cusp. Then there is also the discrete spectrum u_j,λ_j within $[1,\infty)$

$$\Delta u_j+\lambda_j u_j=0,\quad \lambda_j\geqslant 1,\ u_j\in L^2.$$

From these remarks, it follows that $R_{s(2-s)}$ is meromorphic in $\mathrm{Re}(s)>1$, and is analytic except at those s_j's which correspond to λ_j's in the spectrum $\lambda_j\in(0,1)$, so

$$s_j(2-s_j)=\lambda_j.$$

The corresponding residue of $R_{s(2-s)}$ at s_j is simply the projection operator onto the spectrum at λ_j.

Using (2.17)′ and (2.19) it follows that $P_m(s,\omega)$ may be continued to $\mathrm{Re}(s)>1$ and is analytic except possibly at the s_j's. Furthermore the residue of $P_m(s,\omega)$ at s_j is the projection of $P_m(s+1,\omega)$ on the eigenspace at λ_j, multiplied by $2\pi|m|(1-2s_j)$. To determine this more exactly, we expand a typical $u_j(\omega)$ in the cusp. Being an eigenfunction it is easy to see its coefficients must be Bessel functions and we have the expansion

$$u_j(\omega)=\sum_{n\in L^*}c_n(j)\,yK_{s_j-1}(2\pi|n|y)\,e(\langle m,x\rangle). \tag{2.20}$$

Therefore

$$\begin{aligned}\int_{F_D} P_m(s+1,\omega)\,u_j(\omega)\frac{dx_1\,dx_2\,dy}{y^3} &= \int_0^\infty \int_{F_L} y^{s+1} e^{-2\pi|m|y} e^{+2\pi i\langle m,x\rangle} u_j(\omega)\frac{dx_1\,dx_2\,dy}{y^3} \\ &= \int_0^\infty y^s e^{-2\pi|m|y} c_m(j)\, V(F_L)\, K_{s_j-1}(2\pi|m|y)\frac{dy}{y} \\ &= c_m(j)\, V(F_2)\frac{\sqrt{\pi}}{(4\pi|m|)^s}\,\frac{\Gamma(s+s_j-1)\,\Gamma(s-s_j+1)}{\Gamma(s+1/2)}. \qquad (2.21)\end{aligned}$$

It is also clear from the above set of calculation that for $m \neq 0$

$$\int_F P_m(s,\omega)\frac{dx_1\,dx_2\,dy}{y^3} = 0 \qquad (2.22)$$

so that $P_m(s,\omega)$ has no pole at $s=2$ (i.e. corresponding to $\lambda=0$).

3. Kloosterman sums and lower bounds for the discrete spectrum

To obtain information on the discrete spectrum we consider the inner product of two Poincaré series (all relative to a fixed Γ_D). More precisely consider

$$\int_F P_m(s,\omega)\,\overline{P_n(\bar{s}+2,\omega)}\,\frac{dx_1\,dx_2\,dy}{y^3} \qquad (3.1)$$

as a function of s. From the previous section we know the poles of the function in (3.1) are related to the spectrum in question, on the other hand a lengthy computation will show that the function in (3.1) is essentially a type of Zeta function whose co-efficients are "Kloosterman sums". We first introduce these sums.

Let $K=Q(\sqrt{-D})$, and $m,n\in(\mathcal{O})^*=(\mathcal{O}_D)^*$ the dual lattice of $\mathcal{O}$, as before. For $\gamma\neq 0$, $\gamma\in\mathcal{O}$ define the Kloosterman sum

$$S(m,n,\gamma)=\sum_{\alpha\in(\mathcal{O}/(\gamma))^*} e\left(\mathrm{Tr}_{K/Q}\left[\frac{\alpha\bar{m}+\alpha^{-1}\bar{n}}{\gamma}\right]\right) \qquad (3.2)$$

where $e(z)=e^{2\pi i z}$, $\mathrm{Tr}_{K/Q}$ is the trace from K to Q, and R^* is the group of invertible elements of a ring R.

S is well defined (i.e. independent of the choice of generator γ of (γ)), also note that $S(m,n,\gamma)$ is real and $S(m,n,\bar{\gamma})=S(\bar{m},\bar{n},\gamma)$.

Following Selberg [20] we define what we call the Selberg–Kloosterman zeta function

$$Z(m, n, s) = \sum_{\gamma \neq 0} \frac{S(m, n, (\gamma))}{N(\gamma)^s}. \tag{3.3}$$

From the trivial estimate, $|S(m, n, \gamma)| \leq N(\gamma)$ it follows that the series (3.3) converges absolutely for $\mathrm{Re}(s) > 2$. However if $mn \neq 0$ we can use the well known estimates of A. Weil [24] on exponential sums in finite fields to prove

PROPOSITION 3.4. *If* $mn \neq 0$ *the series* (3.3) *converges absolutely for* $\mathrm{Re}(s) > 3/2$.

Proof. We first need a factorization property of these sums, which will allow us to concentrate on the Kloosterman sum when it has a prime ideal in its third argument.

We need slightly more general sums than those in (3.2). Let $A \neq 0$ be an ideal of $\mathcal{O}$ and let ψ_1, ψ_2 be characters of the additive group $\mathcal{O}/A$, define

$$s(A, \psi_1, \psi_2) = \sum_{\delta \in (\mathcal{O}/A)^*} \psi_1(\delta)\, \psi_2(\delta^{-1}). \tag{3.5}$$

Clearly (3.2) is of this form with

$$\psi_1(\alpha) = e\left(\mathrm{Tr}_{K/Q}\left(\frac{\alpha \bar{m}}{\gamma}\right)\right)$$

$$\psi_2(\alpha) = e\left(\mathrm{Tr}_{K/Q}\left(\frac{\alpha \bar{n}}{\gamma}\right)\right).$$

Suppose that $A = P_1^{e_1} P_2^{e_2} \dots P_r^{e_r}$ is the factorization of A into prime ideals. Since

$$\mathcal{O}/A \cong \mathcal{O}/P_1^{e_1} \times \mathcal{O}/P_2^{e_2} \dots \times \mathcal{O}/P_r^{e_r} \tag{3.6}$$

we may factor any characters φ and ψ of $\mathcal{O}/A$ in $\varphi = \varphi_1 \times \varphi_2 \dots \times \varphi_r$, $\psi = \psi_1 \times \psi_2 \dots \times \psi_r$ with

$$\psi_i, \varphi_i \text{ characters of } \mathcal{O}/P_i^{e_i}.$$

So that $\psi(x_1, \dots, x_r) = \psi_1(x_1)\, \psi_2(x_2), \dots, \psi_r(x_r)$ in the factorization (3.6). The following lemma is then obvious.

LEMMA 3.7.

$$S(\psi, \varphi, A) = S(\psi_1, \varphi_1, P_1^{e_1})\, S(\psi_2, \varphi_2, P_2^{e_2}) \dots S(\psi_r, \varphi_r, P_r^{e_r}).$$

Returning to the sum (3.2) if $(\gamma) = P_1^{e_1} P_2^{e_2} \dots P_r^{e_r}$ then

$$S(m,n,\gamma)=S(\psi_1,\varphi_1,P_1^{e_1})\,S(\psi_2,\varphi_2,P_2^{e_2})\ldots S(\psi_r,\varphi_r,P_r^{e_r}).$$

To prove Proposition 3.4 it is of importance for us to know when φ_i and ψ_i are not both trivial.

LEMMA 3.8. *If P is not one of a finite number of primes (depending on n), then the "induced" character ψ_1 on P^e obtained from the factorization*

$$(\gamma)=P^eR,\quad (P,R)=1$$

where $\varphi(x)=e(\mathrm{Tr}_{K/Q}(\bar{n}x/\gamma))=\psi_1\times\psi_2(x)$, is nontrivial.

Proof. Suppose for the sake of argument that $D\not\equiv 3(4)$ in which case

$$\mathcal{O}=\mathbf{Z}\oplus\sqrt{-D}\,Z$$
$$L^*=\mathbf{Z}\oplus\frac{i}{\sqrt{D}}Z.$$

Say $n=n_1+in_2/\sqrt{D}$, $n_1,n_2\in\mathbf{Z}$.

Assuming that ψ_1 is trivial, we have

$$\varphi(x)=1,\quad \forall x\in R\quad\Rightarrow\quad \mathrm{Re}\left(\frac{x\bar{n}}{\gamma}\right)\in\mathbf{Z},\quad \forall x\in R. \tag{*}$$

Now $xi\sqrt{D}\in R$, $\forall x\in R$ so

$$\mathrm{Re}\left(\frac{xi\sqrt{D}\,\bar{n}}{\gamma}\right)\in\mathbf{Z},\quad x\in R \tag{**}$$

or

$$\mathrm{Im}\left(\frac{x\sqrt{D}\,\bar{n}}{\gamma}\right)\in\mathbf{Z},\quad x\in R$$

or

$$\mathrm{Im}\left(\frac{D\bar{n}x}{\gamma}\right)\in\sqrt{D}\,\mathbf{Z},\quad x\in R.$$

From (*) and (**)

$$\frac{x(\bar{n}D)}{\gamma}\in\mathcal{O}$$
$$\Rightarrow x(\bar{n}D)\in(\gamma)$$
$$\Rightarrow(\bar{n}D)\,C(\gamma)$$
$$\Rightarrow P|\bar{n}D.$$

Thus if P is not one of the finite number of prime divisors of $\bar{n}D$ then ψ_1 is not trivial.

Returning to the proof of Proposition 3.4, if P is a prime ideal and ψ and φ are not both trivial then

$$S(\psi,\varphi,P)=\sum_{x\neq 0}\psi(x)\,\varphi(x^{-1})$$

the sum being over the finite field $\mathcal{O}/P$. This sum can be written as

$$\sum_{x\neq 0}\chi(ax+bx^{-1})$$

with χ nontrivial, $ab\neq 0$. The estimates of Weil [24] now imply in the case that ψ and φ are not both trivial that

$$|S(\psi,\varphi,P)|\leqslant 2N(P)^{1/2}. \tag{3.9}$$

For our fixed n, m in Proposition 3.4 (say $n\neq 0$) we define a multiplicative function on the ideals of $\mathcal{O}$ by

$$F(P)=\begin{cases}2N(P)^{1/2} & \text{if } P\nmid nD\\ N(P) & \text{if } P\mid nD\end{cases}$$

and $F(P^e)=N(P^e)$ if $e>0$.

It follows from our considerations that

$$|S(n,m,\gamma)|\leqslant F((\gamma))\sum\frac{|S(n,m,\gamma)|}{N(\gamma)^\sigma}\leqslant\sum\frac{F((\gamma))}{N(\gamma)^\sigma}$$

$$=\prod_{P\mid nD}(1+N(P)\,N(P)^{-\sigma}+N(P)^2N(P)^{-2\sigma}\dots)$$

$$\times\prod_{P\nmid nD}(1+2N(P)^{1/2}N(P)^{-\sigma}+N(P^2)\,N(P)^{-2\sigma}\dots)$$

$$=\prod_{P\mid nD}(1-N(P)^{1-\sigma})^{-1}\prod_{P\nmid nD}\left(1+N(P)^{1/2-\sigma}+\frac{N(P)^{2-2\sigma}}{1-N(P)^{1-\sigma}}\right)$$

which converges absolutely if

$$\sum N(P)^{1/2-\sigma}\quad\text{and}\quad\sum N(P)^{2-2\sigma}$$

both converge, i.e. if $\sigma>3/2$. This completes the proof of Proposition 3.4.

The connection between (3.1) and $Z(m,n,s)$ is the following:

PROPOSITION 3.10.

$$\int_F P_m(s,\omega)\,\overline{P_n(\bar{s}+2,\omega)}\,\frac{dx_1\,dx_2\,dy}{y^3} = \frac{\pi^{-3/2}|n|^{-2}4^{-s-1}\Gamma(s)}{\Gamma(s)\,\Gamma(s+3/2)}\,Z(m,n,s)+\tilde{R}(s)$$

where $\tilde{R}(s)$ *is analytic in* $\mathrm{Re}(s)>1$.

Actually more than Proposition 3.10 can be said, especially about the growth properties of the above as functions of t, where $s=\sigma+it$. For a more detailed analysis of this type and especially the behavior of $Z(m, n, s)$ and of sums of Kloosterman sums see Goldfeld and Sarnak [7], from where some of our computations here are borrowed.

Proof. We first consider

$$\int_F P_m(s_1,\omega)\,e^{-2\pi i\langle n,x\rangle}\,dx_1\,dx_2. \tag{3.11}$$

We use the notations $x=(x_1,x_2)$ or $z=x_1+ix_2$ interchangeably. Notice that $\langle n,z\rangle=n,\ n_1x_1+n_2x_2=\mathrm{Re}(\bar{n}z)$. Also if

$$\tau=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$$

then

$$y(\tau\omega)=\frac{y}{|\gamma z+\delta|^2+|\gamma|^2y^2},\quad x(\tau\omega)=z(\tau\omega)=\frac{\alpha}{\gamma}-\frac{1}{\gamma}\left[\frac{\overline{\gamma z+\delta}}{|\gamma z+\delta|^2+|\gamma|^2y^2}\right].$$

Finally let $\tilde{e}(z)=e^{2\pi i\,\mathrm{Re}(z)}$. Now (3.1) may be written as

$$\int_F y^{s_1}e^{-2\pi|m|y}\,e^{2\pi i\langle m,x\rangle-2\pi i\langle n,x\rangle}\,dx_1\,dx_2$$

$$+\sum_{\substack{\tau=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\in\Gamma_\infty\backslash\Gamma\\ \gamma\neq 0}}\int_F\frac{y^{s_1}e^{(-2\pi|m|y)/|\gamma z+\delta|^2+|\gamma|^2y^2}}{(|\gamma z+\delta|^2+|\gamma|^2y^2)^{s_1}}\,\tilde{e}\left(\frac{\alpha\bar{m}}{\gamma}-\frac{\bar{m}}{\gamma}\,\frac{\overline{\gamma z+\delta}}{(|\gamma z+\delta|^2+|\gamma|^2y^2)}-\bar{n}z\right)dx_1\,dx_2$$

$$=V(F)\,y^{s_1}e^{-2\pi|m|y}\,\delta_{m,n}$$

$$+\sum_{\substack{\gamma\neq 0\\ \tau\in\Gamma_\infty\backslash\Gamma}}\frac{y^{s_1}\tilde{e}(\alpha\bar{m}/\gamma)}{|\gamma|^{2s_1}}\int_F\frac{\exp(-2\pi|m|\,y/(|z+\delta/\gamma|^2+y^2)\,|\gamma|^2)}{(|z+\delta/\gamma|^2+y^2)^{s_1}}$$

$$\times\tilde{e}\left(-\frac{\bar{m}(\overline{z+\delta/\gamma})}{y^2(|z+\delta/\gamma|^2+y^2)}-\bar{n}z\right)dx_1\,dx_2$$

$$= V(F)\, y^{s_1} e^{-2\pi|m|y}\, \delta_{m,n}$$

$$+ \sum_{\substack{\gamma \neq 0,\ \gamma \bmod \pm 1 \\ \delta \bmod \gamma \\ \alpha\delta \equiv 1 \pmod{\gamma}}} \frac{y^{s_1} \tilde{e}(\alpha \bar{m}/\gamma)}{|\gamma|^{2s_1}} \sum_{k \in \mathcal{O}} \int_F \frac{\exp(-2\pi|m|\, y/(|z+k+\delta/\gamma|^2+y^2)\,|\gamma|^2)}{(|z+k+\delta/\gamma|^2+y^2)^{s_1}}$$

$$\times \tilde{e}\left(\frac{-\bar{m}(\overline{z+k+\delta/\gamma})}{y^2(|z+k+\delta/\gamma|^2+y^2)} - \bar{n}z \right) dx_1\, dx_2$$

$$= V(F)\, y^{s_1} e^{-2\pi|m|y}\, \delta_{m,n}$$

$$+ \sum_{\substack{\gamma \neq 0 \\ \delta \bmod \gamma \\ \alpha\delta \equiv 1 \pmod{\gamma}}} \frac{y^{s_1} \tilde{e}(\alpha \bar{m}/\gamma)}{|\gamma|^{2s_1}} \int_{\mathbf{R}^2} \frac{\exp(-2\pi|m|\, y/(|z+\delta/\gamma|^2+y^2)\,|\gamma|^2)}{(|z+\delta/\gamma|^2+y^2)^{s_1}}$$

$$\times \tilde{e}\left(\frac{-(\overline{z+\delta/\gamma})\, \bar{m}}{y^2(|z+\delta/\gamma|^2+y^2)} - \bar{n}z \right) dx_1\, dx_2.$$

Letting $z' = z+\delta/\gamma$ and then $z = z'/y$ we get, above

$$= V(F)\, y^{s_1} e^{-2\pi|m|y}\, \delta_{m,n}$$

$$+ \sum_{\gamma \neq 0} \frac{S(m,n,\gamma)}{|\gamma|^{2s_1}}\, y^{2-s_1} \int_{\mathbf{R}^2} \frac{\exp(-2\pi m/(|z|^2+1)\, y|\gamma|^2)}{(|z|^2+1)^{s_1}}\, \tilde{e}\left(\frac{-z\bar{m}}{\gamma^2 y(|z|^2+1)} - y\bar{n}z \right) dx_1\, dx_2. \tag{3.12}$$

It follows that

$$\int_F P_m(s_1,\omega)\, \overline{P_n(s_2,\omega)}\, \frac{dx_1\, dx_2\, dy}{y^3} = \int_0^\infty \int_F P_m(s_1,\omega)\, y^{\bar{s}_2} e^{-2\pi|n|y}\, e^{2\pi i\langle n,x\rangle}\, \frac{dx_1\, dx_2\, dy}{y^3}$$

which by (3.12) gives

$$\frac{\delta_{m,n} V(F)}{(4\pi|b|)^{s_1+\bar{s}_2-2}}\, \Gamma(s_1+\bar{s}_2-2) + \sum_{\substack{\gamma \neq 0 \\ \bmod \pm 1}} \frac{S(m,n,\gamma)}{|\gamma|^{2s_1}} \int_0^\infty \int_{\mathbf{R}^2} \frac{y^{\bar{s}_2-s_1} e^{-2\pi|n|y} \exp(-2\pi|m|/(|z|^2+1)\, y|\gamma|^2)}{(|z|^2+1)^{s_1}}$$

$$\times \tilde{e}\left(\frac{-z\bar{m}}{\gamma^2 y(|z|^2+1)} - y\bar{n}z \right) \frac{dx_1\, dx_2\, dy}{y^3}.$$

Now let $\bar{s}_2 = s_1+2$ and $s_1 = s$, then we have

$$\int_F P_m(s,\omega)\, \overline{P_n(\bar{s}+2,\omega)}\, \frac{dx_1\, dx_2\, dy}{y^3}$$

$$= \frac{\delta_{m,n} V(F)\,\Gamma(2s)}{(4\pi|n|)^{2s}} + \sum_{\substack{\gamma \bmod \pm 1 \\ \gamma \neq 0}} \frac{S(m,n,\gamma)}{|\gamma|^{2s}} \int_0^\infty \int_{\mathbf{R}^2} \frac{y^2 e^{-2\pi|n|y}\,\bar{e}(-y\bar{n}z)}{(|z|^2+1)^s} \frac{dx_1\,dx_2\,dy}{y}$$

$$+ \sum_{\gamma} \frac{S(m,n,\gamma)}{|\gamma|^{2s}} R_{m,n}(s,\gamma) \tag{3.13}$$

where

$$R_{m,n}(s,\gamma)$$
$$= \int_0^\infty \int_{\mathbf{R}^2} \frac{y\, e^{-2\pi|n|y}\,\bar{e}(-y\bar{n}z)}{(|z|^2+1)^s} \left[\exp\left(\frac{-2\pi|m|}{(|z|^2+1)\,y|\gamma|^2}\right) \bar{e}\left(\frac{-z\bar{m}}{\gamma^2 y(|z|^2+1)}\right) - 1\right] dx_1\,dx_2\,dy.$$

It is clear that

$$R_{m,n}(s,\gamma) \ll \int_0^\infty \left[\int_0^{|\gamma|^{-2}} y\,dy + \int_{|\gamma|^{-2}}^\infty e^{-2\pi|n|y} \frac{y}{|\gamma|^2}\,dy\right] \frac{r\,dr}{(r^2+1)^\sigma} \ll \frac{|\gamma|^{-2}}{\sigma-1}$$

and hence that

$$\sum_{\gamma} \frac{S(m,n,\gamma)}{|\gamma|^{2s}} R_{m,n}(s,\gamma)$$

is holomorphic in $\operatorname{Re}(s)>1$. Returning to (3.13), we must evaluate

$$I = \int_0^\infty \int_{\mathbf{R}^2} \frac{y^2 e^{-2\pi|m|y}\,\bar{e}(-y\bar{n}z)}{(|z|^2+1)^s} \frac{dx_1\,dx_2\,dy}{y}.$$

Now

$$\int_{\mathbf{R}^2} \frac{e^{+2\pi i(y(n_1x_1+n_2x_2))}}{(x_1^2+x_2^2+1)^s}\,dx_1\,dx_2 = \int_0^\infty \int_0^{2\pi} \frac{e^{-2\pi i yr\sin(\theta+\alpha)|n|}}{(r^2+1)^s}\, r\,d\theta\,dr$$
$$= \int_0^\infty \frac{r}{(r^2+1)^2} \int_0^{2\pi} e^{-2\pi i yr|n|\sin\theta}\,d\theta\,dr = \int_0^\infty \frac{J_0(2\pi ry|n|)\,r}{(r^2+1)^s}\,dr$$
$$= \frac{K_{-s+1}(2\pi|n|y)\,(2\pi|n|y)^{s-1}}{\Gamma(s)\,2^{s-1}} \quad \text{(see [8, p. 686])}.$$

$$I = \frac{1}{\Gamma(s)\,2^{s-1}} \int_0^\infty (2\pi|n|)^{s-1} y^{s+1} e^{-2\pi|n|y} K_{1-s}(2\pi|n|y) \frac{dy}{y}$$

$$= \frac{(2\pi|n|)^{s-1}\sqrt{\pi}\,\Gamma(2s)}{\Gamma(s)\,2^{s-1}(4\pi|n|)^{s+1}\Gamma(s+3/2)} \quad \text{(see [8, p. 712])}$$
$$= \frac{1}{\pi^{3/2}4^{s+1}|n|^2}\,\frac{\Gamma(2s)}{\Gamma(s)\,\Gamma(s+3/2)}$$

which completes the proof of Proposition 3.10.

We are now in the position to prove an important estimate, concerning λ_1, (as introduced following (2.19) and in the introduction).

THEOREM 3.1. $\lambda_1 \geq 3/4$.

Proof. In view of Proposition 3.4 we see that for $mn \neq 0$, the right hand side of Proposition 3.10 is analytic in $\mathrm{Re}(s) > 3/2$. However by the analysis of section 2, $P_m(s,\omega)$ and hence the left hand side of Proposition 3.4 may have a pole at an s_j in (3/2, 1) if any such exist. We can calculate the corresponding residue: if $s_j(2-s_j)$ is an eigenvalue with corresponding eigenfunctions $u_1(\omega), \ldots, u_r(\omega)$ (a basis for this space) then the projection of the residue of $P_m(s,\omega)$ on u_j is by (2.21)

$$2\pi|m|(1-2s_j)\frac{c_m(j)\,V(F_L)\sqrt{\pi}}{(4\pi|m|)^{s_j}}\,\frac{\Gamma(2s_j-1)}{\Gamma(s_j+1/2)}\,u_j.$$

So the residue of the left hand side in Proposition 3.4 is

$$2\pi|m|(1-2s_j)\sum_{k=1}^{r}\frac{C_m(k)\,V(F_L)\sqrt{\pi}\,\Gamma(2s_k-1)}{(4\pi|m|)^{s_k}\Gamma(s_k+1/2)}\int_F u_k\,\overline{P_n(s_k+2,\omega)}\,\frac{dx_1\,dx_2\,dy}{y^3}.$$

Now just as in the calculation (2.21) this becomes

$$2\pi|m|(1-2s_j)\sum_{k=1}^{r}\frac{V(F_L)^2\pi\Gamma(2s_k-1)\,\Gamma(2s_k)\,c_m(k)\,c_n(k)}{(4\pi|m|)^{s_k}(4\pi|n|)^{s_k+1}\Gamma(s_k+1/2)\,\Gamma(s_k+3/2)}.$$

If, as we may, we choose the co-efficients of u_j to be real, and choose $m=n$ so that $c_m(1) \neq 0$, we see that the residue at the ficticious s_j is not zero. Our previous remarks then imply that no $s_j \in (3/2, 2)$ exists. This means that $\lambda_1 > 3/4$.

Remarks 3.14. (1) From a completely geometric point of view one can derive lower bounds for λ_1 for a hyperbolic 3-manifold, however their bounds get worse as the volumes increase, and in fact go to zero as the volumes tend to infinity. For example Schoen (see Buser [1]) has shown that for a very small constant c

$$\lambda_1 \geq \frac{c}{V(M)^2}.$$

In the general case one can make λ_1 as small as one pleases for some suitably chosen 3-manifold.

(2) It seems a natural conjecture that $\lambda_1 \geq 1$ for our manifolds M_D. This conjecture could be stated in representation theoretic language. Also, a similar conjecture for congruence subgroups of the modular group in dimension 2 (that $\lambda_1 \geq 1/4$), exists in the literature (see Selberg [20]). For D very small, one may use methods analogous to those of Roelcke [16], i.e. estimating the Rayleigh quotient, to show that $\lambda_1 \geq 1$.

(3) Finally we remark that for certain number theoretically defined subgroups of Γ_3, Kubota (see Patterson [15]) has shown that $\lambda_1 = 8/9$. The corresponding eigenfunction is of great importance in Patterson and Heath-Brown's solution of the Kummer conjecture.

4. Diophantine equations, quadratic forms and closed geodesics

From this section on, we assume that K_D is a field of class number one. We will describe, in terms of standard number theoretic quantities, the closed geodesics on the manifolds M_D.

(4.1) *A diophantine equation.*

Consider the equation

$$t^2 - du^2 = 4 \tag{4.2}$$

which is to be solved for $(t, u) \in \mathcal{O} \times \mathcal{O}$, and $d \in \mathcal{O}$ is fixed (a "Pell" type of equation).

First, if d is a perfect square (in $\mathcal{O}$) then the left hand side of (4.2) factors over $\mathcal{O}$ and one can easily calculate all the solutions. So assume that d is not a perfect square

PROPOSITION 4.3. *$t^2 - du^2 = 4$ has infinitely many solutions.*

Proof. The easiest way of proving this is probably by diophantine approximation of $\sqrt{d}$, see for example [5].

The d's which are of interest to us are those which are discriminants of binary quadratic forms. Thus

$$d \equiv \beta^2 \pmod 4, \quad \text{for some } \beta \in \mathcal{O}. \tag{4.4}$$

For such d's one can introduce a convenient group structure on the solutions of (4.2).

LEMMA 4.5. *Let $x, y \in \mathcal{O}$ with*

$$(x-y)(x+y) \equiv 0 \pmod 4$$

then

$$x+y \equiv 0 \pmod 2.$$

Proof. If 2 is a prime in $\mathcal{O}$ then $2|x-y$ or $2|x+y$ and so $2|x\pm y$ since $x+y \equiv x-y \pmod 2$.

If $2=p^2$ where p is prime, then clearly $p|(x-y)(x+y) \Rightarrow p^2$ divides one of the factors and so $p^2=2$ divides both factors.

If $2=pq$ with p and q primes, then since $x-y \equiv x+y \pmod 2$ we see that $(x-y)^2 \equiv 0 \pmod 2$, and therefore p and q divide $x-y$ or $(x-y) \equiv 0 \pmod 2 \Rightarrow (x+y) \equiv 0 \pmod 2$.

Now suppose that d satisfies (4.4). To each solution (t, u) of (4.2) we associate

$$\varepsilon_{t,u} = \frac{t+\sqrt{d}\,u}{2} \tag{4.6}$$

(where argument $\sqrt{d}$ is chosen in $[0, \pi)$). Now if $\varepsilon_1, \varepsilon_2$ correspond to (t_1, u_1) and (t_2, u_2) respectively then

$$\varepsilon_1 \varepsilon_2 = \frac{a+\sqrt{d}\,b}{2}$$

where

$$a = \frac{t_1 t_2 + d u_1 u_2}{2}, \quad b = \frac{t_1 u_2 + t_2 u_1}{2}.$$

Thus clearly $a^2 - db^2 = 4$, and

$$a \equiv \beta^2 \pmod 4 \Rightarrow t_k^2 \equiv \beta^2 u_k^2 \pmod 4, \quad k = 1, 2.$$

Thus

$$(t_k - \beta u_k)(t_k + \beta u_k) \equiv 0 \pmod 4.$$

By Lemma 4.5 we have

$$t_k \equiv \pm \beta u_k \pmod 2$$

so that

$$t_1 t_2 \equiv \beta^2 u_1 u_2 \pmod 2$$

is

$$t_1, t_2 \equiv du_1 u_2 \pmod 2 \quad \Rightarrow a \in \mathcal{O}.$$

Also,

$$t_1 u_2 + t_2 u_1 \equiv 0 \pmod 2 \quad \Rightarrow b \in \mathcal{O}.$$

Thus, when $\varepsilon_1 \varepsilon_2$ is written in the form (4.6) we see the corresponding (t, u) is a solution of (4.2). The structure of this group, with the exception of some finitely many d's (which we easily identify), will be seen to be infinite cyclic.

There is always the trivial solution of (4.2), viz., $t=\pm 2$, $u=0$. Any solution with $u \neq 0$ will be called nontrivial. Corresponding to each nontrivial solution (t_0, u_0) one has three others $(\pm t_0, \pm u_0)$ and correspondingly $\pm \varepsilon^{\pm 1}$.

First, we want all solutions with $|\varepsilon|=1$. In this case $|\varepsilon^{-1}|=1$ as well, so that

$$|t_0+\sqrt{d}\, u_0| = |t_0-\sqrt{d}\, u_0| = 2 \quad \text{where } \varepsilon = \frac{t_0+\sqrt{d}\, u_0}{2}$$

therefore

$$|t_0+\sqrt{d}\, u_0|^2+|t_0-\sqrt{d}\, u_0|^2 = 8$$

so that

$$|t_0|^2+|d|\,|u_0|^2 = 4. \tag{4.7}$$

It follows that if there is a nontrivial solution with $|\varepsilon|=1$ then

$$|d| \leq 4. \tag{4.8}$$

Thus with the exception of a finite number of d's there are no nontrivial solutions with $|\varepsilon|=1$. For these we choose a solution (t_0, u_0) for which

$$|\varepsilon_0|^2+|\varepsilon_0^{-2}|$$

is as small as possible (of course it is larger than 2). It is clear that

$$\pm \varepsilon_0^{\pm n}, \quad n \geq 0,\ n \in \mathbf{Z} \tag{4.8$'$}$$

will yield all the solutions of (4.2). This solution which by convention we choose to be in absolute value greater than one, is called the fundamental solution. This shows the group of solutions to be infinite cyclic. The fundamental solution is denoted by ε_d.

Returning to (4.8) we find that for $|d|\leq 4$ it may happen that nontrivial solutions, with $|\varepsilon|=1$ exist. For example take K_D with $D\neq 1$ or 3 so that the units of $\mathcal{O}_D$ are ± 1.

By (4.7), $N(d)\leq 16$ and $|d|=\sqrt{N(d)}$ is in Q. Therefore

$$|d| = 1, 2, 3 \text{ or } 4.$$

(i) $|d| = 4 \Rightarrow t = 0,\ |u| = 1, \quad \Rightarrow u = \pm 1.$

Therefore by (4.2)

$$d = -4 \quad \text{and} \quad \varepsilon = \pm\frac{\sqrt{-4}}{2}.$$

(ii) $|d| = 3 \Rightarrow |u|^2 = 1 = |t|^2.$

Therefore

$$u = t = \pm 1 \quad \Rightarrow d = -3$$

$$= \frac{\pm 1 \pm \sqrt{-3}}{2}, \quad \text{a generator is } \frac{1+\sqrt{-3}}{2}.$$

(iii) $|d| = 2$

can only happen if $D=2$ and then $d=2$ (else d is a perfect square)

$$\Rightarrow t = 0, \quad u = \pm\sqrt{-2}$$

$$\varepsilon = \pm\frac{\sqrt{-2}\,\sqrt{2}}{2}$$

(iv) $|d| = 1$, so $d = -1$

(otherwise again d is a perfect square). However $d=-1$ is never a square modulo 4 for these fields.

Thus $d=-4$ has the group of solutions to (4.2) of the form $\mathbf{Z}\times\mathbf{Z}/(4\mathbf{Z})$.

$$d = -3 \text{ has } \mathbf{Z}\times\mathbf{Z}/(6\mathbf{Z})$$

and $d=+2$ only in the case of $Q(\sqrt{-2})$ has the group structure

$$\mathbf{Z}\times\mathbf{Z}/(4\mathbf{Z}). \tag{4.9}$$

All the others are infinitely cyclic.

Even in the noncyclic case it makes sense to talk of the modulus of a generator of the free part of the group of solutions to (4.2).

(4.10) *Quadratic forms.*

Let $Q(x, y)=ax^2+byy+cy^2$ be a binary quadratic form over $\mathcal{O}$. We call Q primitive if $(a, b, c)=(1)$. Dirichlet [3] was the first to develop Gauss's theory of binary forms to forms with coefficients in $\mathbf{Z}[\sqrt{-1}]$. We therefore will call these forms (for any $Q(\sqrt{-D})$) Dirichlet forms. As usual $SL_2(\mathcal{O})$ acts on these forms by linear substitutions. This gives us the notion of equivalent forms (in the narrow sense). The discriminant d of the form is of course invariant. The number of classes of primitive forms with a given discriminant is finite (Dirichlet [3]) and denoted $h(d)$. As far as computing $h(d)$ for various d we refer the reader to Bianchi [2], who computes fundamental domains for the action of $SL_2(\mathcal{O})$ on $\mathcal{H}^3$, as in section 2, and shows how there may be used to find a representative set for these classes.

By an automorph of Q is meant a transformation

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL_2(\mathcal{O})$$

which fixes Q. These clearly form a subgroup of $SL_2(\mathcal{O})$.

PROPOSITION 4.11. *The group of automorphs of a primitive form Q is given by the set*

$$\begin{pmatrix} \frac{t-bu}{2} & -cu \\ au & \frac{t+bu}{2} \end{pmatrix} \quad \textit{such that } t^2-du^2=4$$

where d is the discriminant of Q.

Proof. Very similar to the familiar case of forms over $\mathbf{Z}$, we give the proof anyhow. First note that

$$\frac{t-bu}{2}, \frac{t+bu}{2} \in \mathcal{O} \tag{4.12}$$

since

$$t^2 \equiv b^2u^2 \pmod 4 \quad (d \equiv b^2 \pmod 4)$$

so that

$$(t-bu)(t+bu) \equiv 0 \pmod 4$$

which by Lemma 4.5 implies (4.12).

It is simple algebra to verify that each transformation of the above form fixes Q.

Now we show that every automorph is of the above form. Let

$$\begin{pmatrix} r & s \\ m & n \end{pmatrix}$$

be such a transformation. It follows if $Q = ax^2 + bxy + cy^2$ then

$$a = ar^2 + brm + cm^2$$

$$b = 2ars + b(1+2sm) + 2cmn.$$

Therefore

$$0 = ars + bsm + cmn.$$

Eliminating b and c gives

$$as = -cm$$

$$a(n-r) = bm.$$

Since $(a, b, c) = (1)$ it follows from a/cm and a/bm that a/m; so let $m = au$; then $s = -cu$, $n - r = bu$, $(n+r)^2 = du^2 + 4$. So if $n + r = t$ then

$$r = \frac{t-bu}{2}, \quad n = \frac{t+bu}{2} \quad \text{as needed.}$$

It is not difficult to see that fixing Q with discriminant d, and mapping solutions of (4.2) to matrices

$$(t, u) \mapsto \begin{pmatrix} \frac{t-bu}{2} & -cu \\ au & \frac{t+bu}{2} \end{pmatrix}$$

is a group isomorphism.

From this we see that if

$$d \in \mathscr{D} = \{m \in \mathcal{O} \colon m \equiv \beta^2 \ (\mathrm{mod}\, 4), \quad \text{some } \beta \in \mathcal{O},\ m \text{ not a perfect square}\}$$

and d is not one of the finitely many exceptions described in (4.9), then the group of automorphs of Q is infinite cyclic and is generated by

$$\varepsilon_Q = \begin{pmatrix} \frac{t_0 - bu_0}{2} & -cu_0 \\ au_0 & \frac{t_0 + bu_0}{2} \end{pmatrix}$$

where (t_0, u_0) is a fundamental solution of (4.2).

(4.12) *Closed geodesics.*

Quite generally if $M_\Gamma = H^3/\Gamma$ where Γ is a discrete subgroup of $PSL_2(\mathbf{C})$ acting discontinuously on $\mathcal{H}^3$, we can describe the closed geodesics of M in terms of the transformations in Γ. We are being a little sloppy in that if $\gamma \in \Gamma$, $\gamma \neq$identity has fixed points in $\mathcal{H}^3$ then M is not really a manifold at such points; however this has no bearing on our considerations.

A transformation γ in $PSL_2(\mathbf{C})$ is conjugate to one of the following.

(a) $$\begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}, \quad z \in \mathbf{C},\ z \neq 0$$

in which case, it is called parabolic.

(b) $$\begin{pmatrix} \eta & 0 \\ 0 & \eta^{-1} \end{pmatrix}, \quad |\eta| = 1,\ \eta \neq \pm 1$$

called elliptic.

(c) $$\begin{pmatrix} re^{i\theta} & 0 \\ 0 & r^{-1}e^{-i\theta} \end{pmatrix}, \quad r > 1,\ r \in \mathbf{R}$$

called hyperbolic.

In the latter case many authors call γ hyperbolic if $\theta = 0$, otherwise they call it loxodromic.

In case (c) let $N(\gamma) = r^2 e^{2i\theta}$. Also in this case the transformation γ has two distinct fixed points in the extended plane. The geodesic between these is called the axis of the transformation. γ has the property that it "closes" its axis (and no other geodesic), which is clear since the axis is moved into itself. Two conjugates (in Γ) of hyperbolic transformations will give rise to the same closed geodesic in M_Γ. A hyperbolic transformation is called primitive if it is not a power (not trivially) of any other transformation in Γ. It is not difficult to see therefore, that the primitive conjugacy classes of hyperbolic transformations give rise to the closed geodesics on M_Γ. Also from the canonical form (c) the length of the closed geodesics is $\log |N(\gamma)|$, and the Poincaré map about the closed geodesic is a rotation by

$$2 \arg (N(\gamma)). \tag{4.13}$$

With a primitive form $ax^2 + bxy + cy^2$ with $d \in \mathcal{D}$, we may associate the roots of

$a\theta^2+b\theta+c=0$ in $K(\sqrt{d})$. We will give the roots an order. As before, we choose $\arg\sqrt{d}\in[0,\pi)$, and then associate with $[a,b,c]$ the roots

$$(\theta_1,\theta_2)=\left(\frac{-b-\sqrt{d}}{2a},\frac{-b+\sqrt{d}}{2a}\right).$$

This then distinguishes between $[a,b,c]$ and $[-a,-b,-c]$. The group of automorphs of Q has θ_1,θ_2 as common fixed points. Conversely any $PSL_2(\mathcal{O})$ transformation fixing these points is an automorph of Q. In this way we associate to each form, a directed axis of its automorphs.

To make our convention in (4.8)′ about choice of (t_0,u_0) for (4.2) more unique, we always choose $\sqrt{d}$ and u_0 with arguments in $[0,\pi)$ and then choose t_0 so that $|\varepsilon_{t_0,u_0}|>1$. This defines ε_d uniquely (as well as t_d,u_d).

We now make the correspondence between primitive forms and primitive hyperbolic transformations precise.

To the primitive form $[a,b,c]$ associate of discriminant $d\in\mathfrak{D}$ associate the primitive hyperbolic transformation

$$\begin{pmatrix}\dfrac{t_d-bu_d}{2} & -cu_d\\ au_d & \dfrac{t_d+bu_d}{2}\end{pmatrix}. \tag{4.14}$$

A computation shows that equivalent forms are sent to conjugate transformations. Thus we have a mapping from classes of (primitive) forms and conjugacy classes of primitive hyperbolic transformations. This mapping is onto since if γ is primitive hyperbolic in $PSL_2(\mathcal{O}_D)$ then γ has an axis between fixed points θ_1,θ_2 say. If

$$\gamma=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

then θ_1,θ_2 are roots of

$$\frac{az+b}{cz+d}=z$$

or $cz^2+(d-a)z+b=0$. Thus if $\alpha_1z^2+\alpha_2z+\alpha_3=0$, $\alpha_i\in\mathcal{O}$ is the minimal equation of θ_1,θ_2 over $\mathcal{O}$ then γ is an automorph of $[\alpha_1,\alpha_2,\alpha_3]$ and since it is primitive, it is one of (4.14) or its inverse (which corresponds to $[-\alpha_1,-\alpha_2,-\alpha_3]$).

The preceding discussion proves the following

THEOREM 4.1. *By choosing a representative set for the classes of primitive* *quadratic forms of discriminant* $d\in\mathcal{D}$, *and associating to it its fundamental automorph as in* (4.14), *one gets a representative set for the conjugacy classes of primitive hyperbolic transformations in* Γ_D.

COROLLARY 4.1. *The norms of the conjugacy classes of primitive hyperbolic transformations are the numbers* ε_d^2, *with* $d\in\mathcal{D}$ *and with multiplicity* $h(d)$. *In particular the lengths of the closed geodesics (primitive) are the numbers* $2\log|\varepsilon_d|$ *with multiplicity* $h(d)$, $d\in\mathcal{D}$ *and the rotations of the Poincaré maps about these are* $2\arg\varepsilon_d$ *with* $d\in\mathcal{D}$ *and multiplicity* $h(d)$.

Proof. When

$$\begin{pmatrix} \frac{t_d-bu_d}{2} & -cu_d \\ au_d & \frac{t_d+bu_d}{2} \end{pmatrix}$$

is brought in canonical form

$$\begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix}.$$

We will have

$$z+z^{-1}=t_d$$

therefore

$$z^2-t_dz+1=0$$

or

$$z=\frac{t_d\pm\sqrt{t_d^2-4}}{2}=\frac{t_d\pm\sqrt{d}\,u_d}{2}=\varepsilon_d.$$

The important point is that the length and rotation of the closed geodesic associated to a primitive form $Q=[a,b,c]$ depends only on the discriminant of Q.

5. Prime geodesic theorems

By a prime geodesic theorem we mean an asymptotic formula for the lengths of the closed geodesics on a Riemannian manifold M. Let C denote the set of closed geodesics (we distinguish between the two orientations of such a geodesic) on M. For $\gamma\in C$ we let $\tau(\gamma)$ be its length. The counting function whose asymptotics are of interest to us is

$$\pi(x)=\#\{\gamma\in C:\tau(\gamma)\leqslant x\}. \tag{5.1}$$

It is remarkable, that quite generally for any compact manifold of negative sectional curvature Margulis [14] has proved an asymptotic formula for $\pi(x)$. In [18] various aspects of such theorems are discussed.

Our interest here is in M's which are quotients of $\mathcal{H}^3$ by discrete groups, and for which the quotient is of finite volume (i.e. cofinite quotients).

Let M be such a manifold and as before let λ_1 be the first discrete eigenvalue of $-\Delta$ on M.

THEOREM 5.1. *If* $\lambda_1 \geq 5/9$ *then*

$$\pi(x) = \mathrm{Li}\,(e^{2x}) + O(e^{\gamma n}) \quad as\ x \to \infty$$

where γ *is any number greater than* 5/3, *and where*

$$\mathrm{Li}\,(u) = \int_2^u \frac{1}{\log t}\, dt$$

as in the ordinary prime number theorem.

Remarks 5.2. (i) If $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_k < 5/9$ then the above formula will hold with the addition of k Li terms.

(ii) Theorem 5.1 corresponds to a similar asymptotic formula for lengths of closed geodesics on M's which are compact quotients of $\mathcal{H}^2$, which is due to Selberg and Huber and is analyzed in great detail in Hejhal [9]. In fact using the same methods, i.e. those of the Selberg zeta function, Gangolli and Warner [6] have proven a weaker form of (5.1).

$$\pi(x) \sim \frac{e^{2x}}{2x} \quad \text{as } x \to \infty.$$

They prove this in greater generality (viz. in the case of cofinite quotients of rank one symmetric spaces). As pointed out to the author by Selberg, the zeta function method may be used to prove (5.1). However, since at present the result is not in the literature we will prove it here by a different method which is quite short and direct. For more on this latter method in general rank one case (3), see Sarnak and Woo, to appear.

(iii) Of corse Theorem 3.1 of section 3, tells us that the λ_1 condition of (5.1) holds for the manifolds M_D of this paper (any D).

In order not to have to introduce any new notation we will prove (5.1) in the case of M_D with one cusp only. Indeed the general case with a finite number of cusps, can

(3) See also D. L. DeGeorge, Ecole Normale Superieure, Annales Sc. Ser. 4, 10 (1977), p. 133.

be handled similarly and is only notationally more involved. We will need the Selberg trace formula as it applies to such manifolds.

(5.3) Let $\varphi(s)$ be as in (2.19), the constant term of the Eisenstein series in the cusp at infinity, and let $r_j^2+1=\lambda_j$ as in (3.17). Let g be an even $C_0^\infty(\mathbf{R})$ test function and h its Fourier transform. Then the trace formula takes the following form (we are a little sloppy in not describing various constants precisely as they are of no consequence in what follows) see [27], or Gangolli–Warner [6].

$$\sum_j h(r_j)+c_1\int_{-\infty}^{\infty} h(t)\frac{\varphi'(1+it)}{\varphi(1+it)}\,dt = c_2\int_{-\infty}^{\infty} h(r)\,r^2\,dr$$
$$+\text{hyperbolic}+\text{elliptic}+c_3\,g(0)+c_4\,h(0)+c_5\int_{-\infty}^{\infty} h(t)\frac{\Gamma'}{\Gamma}(1+it)\,dt \tag{5.3}$$

The elliptic term is a sum over conjugacy classes of elliptic elements of Γ, while the hyperbolic is a sum over primitive hyperbolic conjugacy classes. The elliptic sum is a finite sum with a typical term of the form

$$\text{constant}\times\int_{-\infty}^{\infty} h(r)\,dr. \tag{5.4}$$

The hyperbolic sum, with the exception of a possible finite number of terms (those hyperbolic transformations whose axes are also fixed by some elliptic transformation) is

$$2\sum_{\substack{\{\gamma\}\text{ primitive}\\ \text{hyperbolic}}}\sum_{k=1}^{\infty}\frac{\log|N(\gamma)|}{|N(\gamma)^{k/2}-N(\gamma)^{-k/2}|^2}\,g(k\log|N(\gamma)|). \tag{5.5}$$

We will write this as

$$2\sum_{\{\gamma\}}\sum_{k=1}^{\infty} a_{k,\gamma}\,g(k\tau(\gamma)). \tag{5.6}$$

To prove (5.1) we will apply the trace formula with carefully chosen test functions. Define

$$K_T(x)=\begin{cases}\left(1-\frac{|x|}{T}\right)^2 & \text{for } |x|\le T\\ 0 & \text{for } |x|>T\end{cases}. \tag{5.7}$$

Thus its transform $\hat{K}_T(\xi)$ is

$$\hat{K}_T(\xi) = \frac{4}{T\xi^2} - \frac{4 \sin T\xi}{T^2\xi^3}. \tag{5.8}$$

fix $\alpha > 1$ and let

$$H_T(x) = \alpha K_{\alpha T}(x) - K_T(x).$$

Thus

$$\hat{H}_T(\xi) = \frac{4}{T^2\xi^3}\left[\sin T\xi - \frac{1}{\alpha}\sin(\alpha T\xi)\right]. \tag{5.9}$$

The following lemmas will allow us to eliminate most of the terms in the trace formula as far as certain asymptotics are concerned.

LEMMA 5.10. *Let μ be an even positive measure satisfying $\mu\{[0,x]\} = O(x^3)$ as $x \to \infty$ and let ψ be a fixed element of Scwartz class $\mathcal{S}$, with ψ even, $\psi \geq 0$, $\hat{\psi} \geq 0$, $\hat{\psi}(0) = 1$. Then the following estimate is valid.*

$$\int_{-\infty}^{\infty} |\hat{H}_T(\xi)|\,\hat{\psi}(\varepsilon\xi)\,d\mu(\xi) = O\left(T + \frac{1}{T^2}\log\frac{1}{\varepsilon}\right)$$

as $T \to \infty$, $\varepsilon \to 0$ (the implied constants depend on ψ and α).

Proof. Write the integral as $\int_0^1 + \int_1^\infty$ and make direct estimates on $\int_0^1$ and for the second case $|H| \leq 1/|\xi|^3$ and integrate by parts.

Let $\psi \in \mathcal{S}$ as in the last lemma. For $\varepsilon > 0$ let

$$\psi_\varepsilon(x) = \frac{1}{\varepsilon}\psi\left(\frac{x}{\varepsilon}\right) \tag{5.11}$$

so

$$\hat{\psi}_\varepsilon(\xi) = \hat{\psi}(\varepsilon\xi).$$

Finally let

$$g_{T,\varepsilon}(x) = (H_T * \psi_\varepsilon)(x) \quad \text{so that } h_{T,\varepsilon}(\xi) = \hat{H}_T(\xi)\,\hat{\psi}(\varepsilon\xi). \tag{5.12}$$

Now it follows from Weyl's law (and it may be easily deduced from the trace formula, or in our case one may also use the form of $\varphi(s)$) that

$$\int_{-x}^{x}\left|\frac{\varphi'}{\varphi}(1+it)\right|dt + \sum_{|r_j| \leq x} 1 = O(x^3). \tag{5.13}$$

Now using Lemma 5.10 in the trace formula (5.3) together with (5.13) and standard facts about Γ'/Γ, etc., that all the terms except $\Sigma_{r_j} h(r_j), r_j$ imaginary and the hyperbolic term, drop out to give

$$2\sum_{\gamma}\sum_{k} \mathscr{A}_{\gamma,k}\, g_{T,\varepsilon}(k\tau(\gamma)) = \sum_{r_j \notin \mathbf{R}} h_{T,\varepsilon}(r_j) + O\left(T + \frac{1}{T^2}\log\frac{1}{\varepsilon}\right). \tag{5.14}$$

By (5.12) we see that the leading term on the right hand side is

$$\sum_{r_j \notin \mathbf{R}} \hat{H}_T(r_j)\,\hat{\psi}(\varepsilon r_j) = \sum_{r_j \notin \mathbf{R}} \hat{H}_T(r_j) + O(\varepsilon\, e^{aT}). \tag{5.15}$$

The above r_j's are of the form $r_j = it_j$ with $t_j \in \mathbf{R}$ and with $0 < t_j \leq 1$, $t_0 = 1$. (5.15) becomes

$$\frac{4}{T^2}\left(\sum_{t_j} \frac{e^{aTt_j}}{at_j^3} - \sum_{t_j} \frac{e^{Tt_j}}{t_j^3}\right) + O(\varepsilon\, e^{aT}). \tag{5.16}$$

On the other hand on the left hand side of (5.14) we would like to replace $g_{T,\varepsilon}$ by H_T.

Now

$$|G_{T,\varepsilon}(x) - H_T(u)| = O(\varepsilon)$$

and for $|x| > T + \varepsilon$, $g_{T,\varepsilon}(x) \equiv H_T(x) \equiv 0$.

A trivial estimate shows that

$$\sum_{\tau(\gamma) \leq T} \mathscr{A}_\gamma = O(e^{\mu T})$$

for some constant μ, therefore

$$\left|\sum_{\gamma}\sum_{k} \mathscr{A}_{\gamma,k}\, g_{T,\varepsilon}(k\tau(\gamma)) - \sum_{\gamma}\sum_{k} \mathscr{A}_{\gamma,k}\, H_T(k\tau(\gamma))\right| = O(e^{Ta\mu}\varepsilon).$$

Thus we may write the left hand side of (5.14) as

$$2\sum_{\gamma}\sum_{k} \mathscr{A}_{\gamma,k}\, H_T(k\tau(\gamma)) + O(\varepsilon\, e^{Ta\mu}). \tag{5.17}$$

Putting (5.16), (5.17) and (5.14) together we have

$$2\sum_{\gamma}\sum_{k} \mathscr{A}_{\gamma,k}\, H_T(k\tau(\gamma)) = \frac{4}{T^2}\left(\sum_{t_j} \frac{e^{aTt_j}}{at_j^3} - \frac{e^{Tt_j}}{t_j^3}\right) + O\left(\varepsilon\, e^{aT} + \varepsilon\, e^{Ta\mu} + T + \frac{1}{T^2}\log\frac{1}{\varepsilon}\right)$$

which on choosing $\varepsilon = e^{-\alpha T\mu}$ yields

$$\sum_{\gamma}\sum_{k}\mathscr{A}_{\gamma,k}H_T(k\tau(\gamma)) = \frac{2}{T^2}\left(\sum_{t_j}\left(\frac{e^{\alpha Tt_j}}{\alpha t_j^3}-\frac{e^{Tt_j}}{t_j^3}\right)\right)+O(T) \tag{5.18}$$

(now all O signs are with respect to $T\to\infty$).

We prefer to rewrite this slightly: instead of summing over γ primitive and then over k, we sum over all conjugacy classes of hyperbolic transformations, and then the coefficient $\mathscr{A}_\gamma$ depends on the primitive element covered by γ. Actually in our asymptotics the nonprimitive terms are easily seen to be smaller than the O term. So we write the above as

$$\sum_{\gamma}\mathscr{A}_\gamma H_T(\tau(\gamma)) = \frac{2}{T^2}\left(\sum_{t_j}\left(\frac{e^{\alpha Tt_j}}{\alpha t_j^3}-\frac{e^{Tt_j}}{t_j^3}\right)\right)+O(T). \tag{5.19}$$

Let

$$F(T) = \sum_{\gamma,\,\tau(\gamma)\leq T}\mathscr{A}_\gamma\left(1-\frac{|\tau(\gamma)|}{T}\right)^2 \tag{5.19$'$}$$

so that (5.18) can be written as

$$\alpha F(\alpha T)-F(T) = \frac{2}{T^2}\left(\sum_{t_j}\left(\frac{e^{\alpha Tt_j}}{\alpha t_j^3}-\frac{e^{Tt_j}}{t_j^3}\right)\right)+O(T). \tag{5.20}$$

Now define

$$\pi_1(x) = \sum_{\tau(\gamma)\leq x}\mathscr{A}_\gamma. \tag{5.21}$$

Since $\mathscr{A}_\gamma\geq 0$, π_1 is increasing. It is clear that

$$\left.\begin{aligned} F(s) &= \frac{2}{s}\int_0^s\left(1-\frac{u}{s}\right)\pi_1(u)\,du \\ \text{and if } G(s) &= \int_0^s\int_0^v \pi_1(u)\,du\,dv \\ \text{then } G(s) &= \frac{s^2}{2}F(s), \quad \text{and} \quad G'' = \pi_1. \end{aligned}\right\} \tag{5.22}$$

Then from (5.20)

$$\alpha\frac{2}{(\alpha T)^2}G(\alpha T)-\frac{2}{T^2}G(T) = \frac{2}{T^2}\left(\sum_{t_j}\left(\frac{e^{\alpha Tt_j}}{\alpha t_j^3}-\frac{e^{Tt_j}}{t_j^3}\right)\right)+O(T)$$

or

$$\frac{G(\alpha T)}{\alpha} - G(T) = \sum_{t_j} \left(\frac{e^{\alpha T t_j}}{\alpha t_j^3} - \frac{e^{T t_j}}{t_j^3} \right) + O(T^3). \tag{5.23}$$

Since G is increasing and $\alpha > 1$ the last allows us to conclude that

$$G(s) = O(e^s)$$

and setting $\alpha T = s$ in (5.23) yields

$$G(s) = \sum_{t_j} \frac{e^{s t_j}}{(t_j)^3} + O(e^{s/\alpha}), \quad \text{as } s \to \infty;$$

since α is arbitrary we have

$$G(s) = \sum_{t_j} \frac{e^{s t_j}}{(t_j)^3} + O(e^{\varepsilon s}), \quad \text{as } \forall \varepsilon > 0. \tag{5.24}$$

Since π_1 is increasing (though not continuous), we have by (5.22) that for any $h > 0$

$$\frac{G(x-2h)+G(x)-2G(x-h)}{h^2} \leq \pi_1(x) \leq \frac{G(x+2h)+G(x)-2G(x+h)}{h^2} \tag{5.25}$$

and by (5.24) this gives

$$\pi_1(x) = \sum_{t_j} \frac{e^{t_j x}}{t_j} + O\left(h e^x + \frac{e^{\varepsilon x}}{h^2} \right)$$

choosing $h = e^{-x/3}$

$$\pi_1(x) = \sum_{t_j} \frac{e^{t_j x}}{t_j} + O(e^{(2/3+\varepsilon)x}). \tag{5.26}$$

It is easy to see that only the primitive γ's come into the above asymptotics also by assumption of the theorem i.e. that $\lambda_1 \geq 5/9$ means that, $t_1 < 2/3$, so we have

$$\sum_{\substack{\gamma \text{ prim} \\ \log|N(\gamma)| \leq x}} \frac{\log|N(\gamma)|}{|N(\gamma)^{1/2} - N(\gamma)^{-1/2}|^2} = e^x + O(e^{(2/3+\varepsilon)x}).$$

A simple calculation from here shows that

$$\sum_{\log|N(\gamma)| \leq x} \log|N(\gamma)| = \frac{x^2}{2} + O(x^{5/3+\varepsilon})$$

from which (5.1) follows.

6. Unit asymptotics

In order to obtain our main result on the class number asymptotics we need to know the asymptotics of the cardinality of the sets

$$\mathcal{D}_x = \{d \in \mathcal{D}: |\varepsilon_d| \leq x\}$$

as $x \to \infty$ (the notation and setting begin as in section 4). This section is devoted to this question.

THEOREM 6.1. *There is a constant* $c_k > 0$ *depending on the field* $K = Q(\sqrt{-D})$, *such that*

$$|\mathcal{D}_x| = c_k x^2 + O(x^\beta) \quad \textit{as } x \to \infty$$

where β *is any number greater than* 4/3.

The constant c_k may be computed and we will say more about its value later in this section.

The argument is similar to that in Sarnak [17] and so we will only outline the proof of Theorem 6.1.

Consider the sets

$$E_x = \{(d, k): d \in \mathcal{D},\ k \geq 1,\ k \in \mathbf{Z} \quad \text{and} \quad |\varepsilon_d^k| \leq x\}$$

and let $\psi_1(x) = |E_x|$, and $\psi(x) = |\mathcal{D}_x|$. It is clear that

$$\psi_1(x) = \psi(x) + \psi(x^{1/2}) + \psi(x^{1/3})$$

so that since we will show that $\psi_1(x) = O(x^2)$, we see that

$$\psi_1(x) = \psi(x) + O(x). \tag{6.1}$$

As we saw in section 4, ε_d^k $d \in \mathcal{D}$, $k \geq 1$ yield solutions to $t^2 - du^2 = 4$. Also every solution of $t^2 - du^2 = 4$ for which $|\varepsilon_{t,m}| > 1$ is one of ε_d^k or $-\varepsilon_d^k$ (we are assuming, as we may in these asymptotics, that d is not one of the finite number of exceptions for which the group of solutions is not cyclic).

Thus we are counting

$$\tfrac{1}{2}\#\{(t, u, d): t^2 - du^2 = 4, \quad d \in \mathcal{D} \quad \text{and} \quad 1 < |\varepsilon_{t,u}| \leq x\}.$$

Now

$$\varepsilon_{t,u}=\frac{t+\sqrt{d}\,u}{2}=\frac{t+\sqrt{t^2-4}}{2}=t+O\left(\frac{1}{|t|}\right).$$

Thus we want to count

$$\psi_2(x)=\tfrac{1}{2}\#\{(t,u,d)\colon |t|\leq x,\ t^2-du^2=4,\ d\in\mathscr{D}\}.$$

Indeed

$$\psi(x)=\psi_1(x)+O(x)=\psi_2(x)+O(x).$$

Without increasing the error term so far, we may allow d to be a perfect square in the above, so that we are looking at

$$\psi_3(x)=\tfrac{1}{2}\#\{(t,u,d)\colon |t|\leq x,\ t^2-du^2=4,\ d\equiv\beta^2,\ d\not\equiv 0\ (\text{mod}\,4),\ \text{some}\ \beta\}.$$

So if we fix $\beta_1,\beta_2,\beta_3,\beta_4$ which are representatives of the squares modulo 4, with say $\beta_1=0$, $\beta_2=1$ then we will need

$$\begin{aligned}N_j(x)&=\#\{(t,u,l)\colon t^2-(4l+\beta_j)\,u^2=4,\quad |u|\leq x,\ |t|\leq x\}\\&=\tfrac{1}{2}\#\{(t,u)\colon t^2\equiv(\beta_j u^2+4)\ \text{mod}\,(4u^2),\quad |u|\leq x,\ |t|\leq x\}+O(x)\end{aligned}\tag{6.2}$$

so that

$$\psi(x)=\frac{1}{2}\sum_{j=1}^{4}N_j(x)+O(x).$$

To calculate the asymptotics of (6.2) in (6.2) let

$$S_j(x,u)=\#\{|t|\leq x\colon t^2\equiv(\beta u^2+4)\ (\text{mod}\,4u^2)\}$$

and

$$S_j^*(u)=\#\,\text{residue class solutions of}\ t^2\equiv(\beta_j u^2+4)\ (\text{mod}\,4u^2)$$

then

$$N_j(x)=\sum_{|u|\leq x}S_j(u,x).\tag{6.3}$$

LEMMA 6.4.

$$S(x,u)=\frac{\pi x^2 S^*(u)}{V(\mathscr{F}_L)\,N(4u^2)}+O\left(\frac{S^*(u)\,x^{2/3}}{N(4u^2)^{1/3}}+S^*(u)\right)$$

LEMMA 6.5. *Let L be a fixed lattice in $\mathbf{R}^2$ and p be any point in $\mathbf{R}^2$, then*

$$\#\{\lambda \in L: |\lambda - p| \leqslant R\} = \frac{\pi R^2}{V(L)} + O(R^{2/3}) + O(1).$$

The $O(1)$ is put there so as to include $R \to 0$ as well, also the implied constant is independent of p.

Proof of Lemma 6.5. This is a standard lattice point result—the proof shows that the error term is independent of p, see [25].

Proof of Lemma 6.4.

$$S_j(x, u) = \#\{t: |t| \leqslant x,\ t^2 \equiv u^2\beta_j \pmod{4u^2}\}.$$

Let $t_1, \ldots, t_k$ where $k = S_j^*(u)$ be a representative set for the residue classes of the congruence defining the set in S_j. Now

$$\begin{aligned}\#\{|t| \leqslant x: t \equiv t_j \bmod (4u^2)\} &= \#\{\lambda \in \mathcal{O}: |t_j + 4\lambda u^2| \leqslant x\} \\ &= \#\left\{\lambda \in \mathcal{O}: \left|\lambda - \left(\frac{-t_j}{4u^2}\right)\right| \leqslant \frac{x}{|4u^2|}\right\}\end{aligned}$$

which by Lemma 6.5

$$= \frac{\pi x^2}{V(\mathcal{F}_L)\, N(4u^2)} + O\left(\frac{x^{2/3}}{|4u^2|^{2/3}}\right) + O(1)$$

from which the result follows by adding over the classes.

$$N_j(x) = \sum_{|u| \leqslant x}^{*} S_j(u, x)$$

We will split this sum as follows (j fixed)

$$\left(\sum_{|u| \leqslant x^{1/2}} + \sum_{x^{1/2} < |u| \leqslant x^{2/3}} + \sum_{x^{2/3} < |u| \leqslant x}\right) S(u, x) = \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

$$\mathrm{I} = \frac{\pi x^2}{V(\mathcal{F}_L)} \sum_{|u| \leqslant x^{1/2}}^{*} \frac{S^*(U)}{N(4u^2)} + O\left(x^{2/3} + \sum_{|u| \leqslant x^{1/2}} \frac{1}{N(4u^2)^{1/3}}\right) + O(x), \quad \forall \varepsilon > 0$$

which follows from Lemma 6.4 and the estimate

$$|S^*(u)| = O(|u|^{\varepsilon}), \quad \forall \varepsilon > 0.$$

Therefore

$$\mathrm{I} = \frac{\pi x^2}{V(\mathscr{F}_L)} \sum_{u \neq 0} \frac{S^*(u)}{N(4u^2)} + O(x^{1+\varepsilon}).$$

$$\mathrm{II} = \sum_{x^{1/2} < |u| \leq x^{2/3}} S(x, u).$$

For $x^{1/2} < |u| \leq x^{2/3}$

$$\frac{x^2}{N(u^2)} \leq 1$$

therefore

$$|S(x, u)| = O\left\{ \sum_{|u| \leq x^{2/3}} S^*(u) \right\} = x^{4/3+\varepsilon}, \quad \forall \varepsilon > 0.$$

Finally

$$\mathrm{III} = \sum_{x^{2/3} < |u| \leq x} S(u) \leq \# \{(t, n, u) \colon |t| \leq x,\ t^2 - 4u^2 = 4,\ x^{2/3} < |u| \leq x\}.$$

The condition implies

$$ku^2 = (t-2)(t+2).$$

Now if for example $(t-2, t+2) = (1)$ (the other possibilities are similar) we have

$$u = yz \quad \text{with } y^2 | t-2,\ z^2 | t+2$$

$$\Rightarrow t \equiv \pm 2 \,(\operatorname{mod} v^2),\ |t| \leq x$$

where

$$x^{1/3} < |v| < x^{1/2}.$$

For each t the number of choices for (k, u) is $O(|t|^{\varepsilon})$, $\forall \varepsilon > 0$. Therefore

$$\begin{aligned} \mathrm{III} &\leq \text{const.} \sum_{x^{1/3} < |v| \leq x^{1/2}} \left(\frac{x^{2+\varepsilon}}{|v|^4} + \frac{x^{2/3+\varepsilon}}{|v|^{4/3}} \right) \\ &= O\left(x^{2+\varepsilon} \int_{x^{1/3}}^{\infty} \frac{1}{r^4}\, r\, dv + x^{2/3+\varepsilon} \int_{x^{1/3}}^{x^{1/2}} \frac{r}{r^{4/3}}\, dr \right) \\ &= O(x^{4/3+\varepsilon}). \end{aligned}$$

Therefore combining estimates for I, II and III

$$N_j(x) = \frac{\pi x^2}{V(\mathscr{F}_L)} \sum_{u \neq 0} \frac{S_j^*(u)}{N(4u^2)} + O(x^{4/3+\varepsilon}), \quad \forall \varepsilon > 0;$$

and so

$$\psi(x) = \frac{1}{2} \sum_{j=1}^{4} N_j(x) = c_k x^2 + O(x^{4/3+\varepsilon}), \quad \forall \varepsilon > 0$$

where

$$c_k = \frac{1}{2} \frac{\pi}{V(\mathscr{F}_L)} \sum_{u \neq 0} \left(\sum_{j=1}^{4} S_j^*(u) \right) N(4u^2)^{-1}.$$

This completes the proof of Theorem 6.1. One can "evaluate" the c_k even further. For example if $D \equiv 3 \pmod{16}$ then using Euler products ($\Sigma_{j=1}^4 S_j^*(u)$ being closely related to a multiplicative function) we can show that

$$c_k = \frac{\pi}{2\sqrt{D}} \frac{(\zeta_k(2))^2}{\zeta_k(4)} \times (\text{rational}).$$

Since not much is known about rationality properties of $\zeta_k(2)$, $\zeta_k(4)$, we do not give the details of this computation. The computation is tedious but straightforward.

7. Class number asymptotics and the Selberg zeta functions

By Corollary 4.1 and Theorem 5.1 and (3.1) we have

$$\sum_{d \in \mathscr{D}_{(e^{x/2})}} h(d) = \operatorname{Li}(e^{2x}) + O(e^{\gamma x}), \quad \text{where } \gamma > \tfrac{5}{3}.$$

THEOREM 7.1.

$$\sum_{d \in \mathscr{D}_x} h(d) = \operatorname{Li}(x^4) + O(x^\gamma), \quad \gamma > \tfrac{10}{3}.$$

THEOREM 7.2.

$$\frac{1}{|\mathscr{D}_x|} \sum_{d \in \mathscr{D}_x} h(d) = \frac{1}{c_k} \frac{\operatorname{Li}(x^4)}{x^2} + O(x^\gamma), \quad \textit{where } \gamma > \tfrac{4}{3}.$$

Proof. Use Theorems 7.1 and 6.1.

Remark 7.3. Using congruence subgroups of Γ_D of level P a prime ideal of $\mathcal{O}_D$, one can in a fashion similar to Sarnak [17], obtain average asymptotic for

$$\frac{1}{|\mathcal{D}_{P,x}|}\sum_{d\in D_{P,x}} h(d)$$

where

$$D_{P,x}=\left\{d\in\mathcal{D}_x:\varepsilon_d=\frac{x+\sqrt{\alpha}\,y}{2}\quad\text{satisfies } y\in P\right\}.$$

(7.4) *Selberg zeta function.*

We now show how the Selberg zeta function for $\mathcal{H}/\Gamma_D$ may be viewed as a completely arithmetically defined function associated to K. By section 4 we have identified the lengths and Poincaré maps of the closed geodesics. Using this and Selberg's definition of a zeta function, see Gangolli–Warner [6] we define, for $\mathrm{Re}\,(s)>2$,

$$Z_K(s)=\prod_{d\in\mathcal{D}}\prod_m\prod_n(1-\varepsilon_d^{-m}\bar{\varepsilon}_d^{-n}|\varepsilon_d|^{-2s})^{h(d)}. \tag{7.5}$$

The point is that the logarithmic derivative of $Z_K(s)$, is the hyperbolic term in the trace formula (5.3) with

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

From this, as is by now standard in such situations, meromorphic continuation, functional equation, and zeros and poles of $Z_K(s)$ may be read off. The functional equation is about the line $\mathrm{Re}\,(s)=1$. $Z_K(s)$ has zeros at $1\pm ir_j$, and these are the only ones in $\mathrm{Re}\,(s)\geq 1$. Thus there is a zero at $s=2$ and possibly a finite number in the interval $(1,2)$. So the analogue of the Riemann hypothesis is essentially true, and by the main theorem of section 3 besides the zero at 2 (which is the analogue of the pole in the more familiar zeta function cases) there are no zeros in $\mathrm{Re}\,(s)>3/2$.

By Remark 3.14, we can show that for $K=Q(\sqrt{-1})$ there are no zeros in $(1,2)$. Besides these zeros one also has zeros coming from the φ'/φ term in the trace formula and from (2.19) these are at the point ϱ where ϱ is a nontrivial zero of $\zeta_K(s)$. That the so simply defined zeta functions of (7.5) should have such nice and well understood analytic properties, is to us rather remarkable.

Acknowledgements. I would like to thank Professors P. Garrett, D. Goldfeld, and V. S. Varadarajan for interesting discussions.

References

[1] BUSER, P., Private communication (1981).
[2] BIANCHI, L., Geometrische Darstellung der Gruppen linearer Substitutionen mit ganzen complexen Coefficienten nebst Anwendungen auf die Zahlenteorie. *Math. Ann.*, 38 (1891), 313–333.
[3] DIRICHLET, G. L., *Werke* I, pp. 565–596.
[4] EFRAT, I., *Selberg trace formula, rigidity and cusp forms*. Ph.D. thesis, New York University, 1982.
[5] FJELLSTEDT, L., On a class of Diophantine equations of the second degree in imaginary quadratic fields. *Ark. Mat.*, 2 (1954), 435–461.
[6] GANGOLLI, R. & WARNER, G., Zeta functions of Selberg's type for some noncompact quotients of symmetric spaces of rank one. *Nagoya Math. J.*, 78 (1980), 1–44.
[7] GOLDFELD, D. & SARNAK, P., Kloosterman sums in imaginary quadratic fields. To appear.
[8] GRADSHTEYN, I. & RYZHIK, I., *Tables of integrals, series and products*. Academic Press (1980).
[9] HEJHAL, D., *The Selberg trace formula for* $PSL(2, \mathbf{R})$, vol. I. Lecture notes in mathematics, 548 (1976). Springer Verlag.
[10] HUMBERT, G., Sur la mesure de classes d'Hermite de discriminant donne dans un corp quadratique imaginaire. *C.R. Acad. Sci. Paris*, 169 (1919), 448–454.
[11] KLOOSTERMAN, H. D., *Hamb. Abh.*, 5 (1927), 337–352.
[12] KUBOTA, T., *Elementary theory of Eisenstein series*. Halsted Press (1973).
[13] LANG, S., *Algebraic number theory* (chapter XVI). Addison-Wesley (1970).
[14] MARGULIS, G., Applications of ergodic theory to the investigations of manifolds of negative curvature. *Functional Anal. Appl.*, 3 (1969), 335–336.
[15] PATTERSON, S. J., A cubic analogue of the theta series. *J. Reine Angew. Math.*, 296 (1977), 125–161.
[16] ROELCKE, W., *Über die Wellengleichung bei Grenzenkreisgruppen erster Art*, 49–51. Springer Verlag, Heidelberg, 1956.
[17] SARNAK, P., Class numbers of indefinite binary quadratic forms. To appear in *J. Number Theory*.
[18] — *Prime geodesic theorems*. Thesis, Stanford, 1980.
[19] SELBERG, A., Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Indian Math. Soc.*, 20 (1956), 47–87.
[20] — On the estimation of Fourier coefficients of modular forms, in *Proceedings of symposia in pure mathematics*, 8 (1965).
[21] SWAN, R., Generators and relations for certain special linear groups. *Bull. Amer. Math. Soc.*, 74 (1968), 576–581.
[22] THURSTON, W., *The geometry and topology of three-manifolds*. Princeton University, 1981.
[23] — Three dimensional manifolds, Kleinian groups and hyperbolic geometry. Preprint.
[24] WEIL, A., On some exponential sums. *Proc. Nat. Acad. Sci.*, 34 (1948), 204–207.
[25] WALFISZ, A., *Gitterpunkte in mehrdimensionalen Kugeln*. Warszawa, 1957.
[26] WHITTAKER, *E.* & WATSON, G., *A course of modern analysis*. Cambridge University, 4th ed.
[27] TANIGAWA, Y., Selberg trace formula for Picard groups, in *Proc. Int. Symp. Algebraic Number Theory, Tokyo*, 1977, pp. 229–242.

[28] JACQUET, H. & GELBART, S., A relation between automorphic representations of $GL(2)$ and $GL(3)$. *Ann. Sci. École Norm. Sup.*, 11 (1978), 471–542.
[29] SPEIZER, A., *The theory of binary quadratic forms with coefficients in an arbitrary number field.* Dissertation, University of Göttingen, 1909.

Received March 3, 1982
Received in revised form October 30, 1982