Vol. 79, No. 1 DUKE MATHEMATICAL JOURNAL © July 1995

CLASS GROUP L-FUNCTIONS

W. DUKE, J. FRIEDLANDER, AND H. IWANIEC

To Professor Wolfgang Schmidt on the occasion of his sixtieth birthday.

1. Introduction. We consider the imaginary quadratic field $K = \mathbb{Q}(\sqrt{-D})$ of discriminant $-D$. Our main interest here is to study the L-functions of K which are attached to the characters of the class group $\mathscr{H}$. Letting $\chi \in \hat{\mathscr{H}}$ be such a character, we denote

$$L_K(s, \chi) = \sum_{\mathfrak{a}} \chi(\mathfrak{a})(N\mathfrak{a})^{-s}, \tag{1.1}$$

where $\mathfrak{a}$ ranges over nonzero integral ideals, and $N\mathfrak{a}$ is the norm. Thus we have $h = h(-D) = |\mathscr{H}|$ such L-functions. It is known that the "class number" h satisfies

$$D^{1/2-\varepsilon} \ll h \ll D^{1/2} \log D. \tag{1.2}$$

Here the lower bound (ineffective) is due to C. L. Siegel [Si], and the upper bound is elementary.

For the trivial character $\chi = 1$, the L-function is just the Dedekind zeta-function of K and can be expressed as $\zeta_K(s) = \zeta(s)L(s, \chi_D)$, where $\zeta(s)$ is the Riemann zeta-function and $L(s, \chi_D)$ is the Dirichlet L-function for the field character $\chi_D(n) = (-D/n)$. More generally, if χ is real we have Kronecker's factorization

$$L_K(s, \chi) = L(s, \chi_{D_1})L(s, \chi_{D_2}), \tag{1.3}$$

where $-D = D_1 D_2$ is some factorization into fundamental discriminants $-D_1$ and $-D_2$, see [Si2, pp. 81 and 91]. Of course the number $2^{\omega(D)-1}$ of such real characters is quite small in comparison to h when D is large.

The L-functions defined in (1.1) for $\operatorname{Re} s > 1$, where the series converges absolutely, possess an analytic continuation to the whole complex s-plane. They are entire functions except for $\zeta_K(s)$, which has a simple pole at $s = 1$ of residue

$$\operatorname*{res}_{s=1} \zeta_K(s) = L(1, \chi_D) = 2\pi h w^{-1} D^{-1/2}, \tag{1.4}$$

Received 17 June 1994.
Duke and Iwaniec supported by NSF grant DMS-9202022.
Friedlander and Iwaniec supported by NSERC grant A5123.

where w is the order of the unit group. This is the famous Dirichlet class number formula. Moreover every $L_K(s, \chi)$ satisfies the functional equation

$$\Phi(s, \chi) = \Phi(1 - s, \chi), \tag{1.5}$$

where

$$\Phi(s, \chi) = (2\pi)^{-s}\Gamma(s)D^{s/2}L_K(s, \chi). \tag{1.6}$$

All the above properties follow at once from the integral representation due to E. Hecke [He];

$$\Phi(s, \chi) + \delta_\chi hw^{-1}s^{-1}(1 - s)^{-1} = \int_1^\infty (u^{s-1} + u^{-s})S(u, \chi)\, du, \tag{1.7}$$

where $\delta_\chi = 1$ if $\chi = 1$ or else it vanishes, and

$$S(u, \chi) = \sum_{\mathfrak{a}} \chi(\mathfrak{a}) \exp(-2\pi u N\mathfrak{a}/\sqrt{D}). \tag{1.8}$$

Our primary purpose is to give bounds for $L_K(s, \chi)$ at points on the critical line $s = 1/2 + it$. We shall concentrate on the dependence of these bounds on D rather than s. By the trivial estimate $|L_K(s, \chi)| \leqslant \zeta(\sigma)^2 \leqslant \sigma^2(\sigma - 1)^{-2}$ on the vertical line $\operatorname{Re} s = \sigma = 1 + (\log D)^{-1}$, the functional equation (1.5), and the convexity principle of Phragmen-Lindelöf, one gets immediately that

$$L_K(s, \chi) \ll s(1 - s)D^{1/4}(\log D)^2 \tag{1.9}$$

on the critical line $s = 1/2 + it$. One can do a bit better with respect to D by using the integral representation (1.7). Indeed we get $|\Phi(s, \chi)| \leqslant 4hw^{-1} + \Phi(1/2, 1)$ if $\chi \neq 1$, whence

$$L_K(s, \chi) \ll (\operatorname{ch} \pi t)^{1/2}hD^{-1/4}. \tag{1.10}$$

For a real character χ the convexity bound (1.9) as well as (1.10) can be improved substantially using the celebrated estimate of D. Burgess

$$L(s, \chi_D) \ll |s|\,|D|^{3/16+\varepsilon}. \tag{1.11}$$

This yields by (1.3) that for χ real,

$$L_K(s, \chi) \ll s(1 - s)D^{3/16+\varepsilon}. \tag{1.12}$$

We wish to get a comparable improvement for $L_K(s, \chi)$ with a general class character. To do this we intend to apply the method of our previous works

[FI], [DFI1], [DFI2]. This requires us to study various mean values of these L-functions.

It is quite easy to evaluate $L_K(s, \chi)$ on average as χ varies in $\hat{\mathscr{H}}$.

THEOREM 1. *If* Re $s = 1/2$ *we have*

$$\frac{w}{h} \sum_{\chi \in \hat{\mathscr{H}}} L_K(s, \chi) = \zeta(2s) + \zeta(2-2s)\frac{\Gamma(1-s)}{\Gamma(s)}\left(\frac{\sqrt{D}}{2\pi}\right)^{1-2s} + O(e^{-\pi\sqrt{D}}). \quad (1.13)$$

In particular, at the special point $s = 1/2$ *this is* $\log(\sqrt{D}/8\pi) + \gamma + O(e^{-\pi\sqrt{D}})$.

With more work, but still using mostly elementary means, we shall establish an asymptotic formula for the second power-moment.

THEOREM 2. *If* Re $s = 1/2$ *we have*

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} |L_K(s, \chi)|^2 = l_D(s) + O(L(1, \chi_D)), \quad (1.14)$$

where the main term $l_D(s)$ *is given as the sum of residues of* $w^{-1}D^{v-1/2}L(v, \chi_D)H(v)$ *at* $v = 1, 2s, 2-2s$. *Here* $H(v)$ *is a meromorphic function on the whole complex plane determined in Section 6 which does not depend on D.*

A complete analytic expression for $H(v)$ may be obtained by inserting (6.5) and (6.10) into (6.20). To compute the residues, we can use the Laurent series expansions (6.13) and (6.14). From these we find that in the half-plane Re $v > 1/2$, the poles are at $v = 1, 2s, 2-2s$ (all on the line Re $v = 1$). The pole at $v = 1$ has order 4 if $s = 1/2$, and it has order 2 if $s \neq 1/2$. In the latter case, the additional poles at $v = 2s, 2-2s$ are simple. Therefore, if $s = 1/2$ we have

$$l_D(1/2) = \textstyle\sum_{j+k \leqslant 3} c_{jk} L^{(j)}(1, \chi_D)(\log D)^k, \quad (1.15)$$

and if $s \neq 1/2$ we have

$$\begin{aligned} l_D(s) = {} & c_{01}(s)L(1, \chi_D)\log D + c_{00}(s)L(1, \chi_D) + c_{10}(s)L'(s, \chi_D) \\ & + c(s)L(2s, \chi_D)D^{s-1/2} + c(1-s)L(2-2s, \chi_D)D^{1/2-s}. \end{aligned} \quad (1.16)$$

Here the highest coefficients are quite simple, namely,

$$c_{03} = (w\zeta(2))^{-1} \quad (1.17)$$

and

$$c_{01}(s) = (12w\zeta(2))^{-1}\zeta(2s)\zeta(2-2s). \quad (1.18)$$

For $c(s)$ see (6.26). The other coefficients look more complicated. We certainly expect that the leading component of $l_D(s)$ is the one which has the highest power of $\log D$, but we cannot prove it yet. One can see this using the Riemann hypothesis for $L(s, \chi_D)$, which yields

$$(\log\log D)^{-A} \ll L^{(j)}(2s, \chi_D) \ll (\log\log D)^{A}. \tag{1.19}$$

Hence, if $s = 1/2$,

$$l_D(1/2) \sim c_{03} L(1, \chi_D)(\log D)^3, \tag{1.20}$$

and, if $s \neq 1/2$,

$$l_D(s) \sim c_{01}(s) L(1, \chi_D) \log D. \tag{1.21}$$

In both cases the asymptotic formula (1.14) is meaningful for large D since the error term is smaller (albeit slightly) than the main term.

At the central point $s = 1/2$, we can estimate the main term $l_D(1/2)$ unconditionally; in fact, our lower bound has the right order of magnitude. First we shall show that

$$l_D(1/2) \asymp L(\log D)^2, \tag{1.22}$$

where

$$L = \sum_{\mathfrak{a}} (N\mathfrak{a})^{-1} \exp(-2\pi N\mathfrak{a}/\sqrt{D}). \tag{1.23}$$

It follows by (1.11) that

$$L(1, \chi_D) \log D \ll L \ll L(1, \chi_D) \log D + L'(1, \chi_D). \tag{1.24}$$

Since $L'(1, \chi_D) \ll (\log D)^2$, we obtain unconditionally that

$$L(1, \chi_D)(\log D)^3 \ll l_D(1/2) \ll (\log D)^4. \tag{1.25}$$

Assuming there is no Siegel zero for the character χ_D, we have

$$\frac{L'}{L}(1, \chi_D) \ll \log D, \tag{1.26}$$

and hence we can improve (1.25) to get

$$l_D(1/2) \asymp L(1, \chi_D)(\log D)^3. \tag{1.27}$$

Notice that Theorem 2 yields the convexity bound for every individual $L_K(s, \chi)$ (apart from the factor in s) by ignoring the remaining L-functions and using positivity.

With a lot more work we shall establish a mean-value theorem in which the error term saves a positive power of D.

THEOREM 3. *If* Re $s = 1/2$ *we have*

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} |L_K(s, \chi)|^2 = l_D(s) + l_D^+(s) + O(D^{-1/28+\varepsilon}) \tag{1.28}$$

where $l_D(s)$ is as in Theorem 2 and $l_D^+(s) = w^{-1}c^+(s)L(1, \chi_D)$ is given by (11.17).

The first part $l_D(s)$ comes from the contribution of the relevant diagonal terms. The second part $l_D^+(s)$ looks simple but is not easy to obtain; it emerges from the off-diagonal terms which we shall treat in more detail than in Theorem 2 using the spectral theorem for the modular group. More precisely, $l_D^+(s)$ comes from the projection of a certain Poincaré series onto the constant eigenfunction (of eigenvalue $\lambda_0 = 0$). The problem of estimating the error term in (1.28) is even harder. It reduces eventually to an equidistribution property of Heegner points (2.4) with respect to the Poincaré measure on the hyperbolic plane which was established in [Du].

The improvement (1.28) over (1.14) does not however lead to better bounds for the individual L-functions. In order to effect such improvements, we follow ideas from our earlier papers [FI], [DFI1], [DFI2] and give estimates for the second power-moment weighted by a general character sum of type

$$\mathscr{A}(c, \chi) = \sum_{q \leqslant Q} c_q \gamma_q(\chi), \tag{1.29}$$

where

$$\gamma_q(\chi) = 2^{-\omega(q)} \sum_{\mathfrak{q}\bar{\mathfrak{q}}=q,\,(\mathfrak{q},\bar{\mathfrak{q}})=1} \chi(\mathfrak{q}). \tag{1.30}$$

Notice that $\gamma_q(\chi)$ is real, and the sum is void unless every prime factor of q has degree 1 and is unramified, i.e., $\chi_D(p) = 1$ for every $p|q$. In this case

$$\gamma_q(\chi) = \prod_{p^\alpha \| q} \operatorname{Re} \chi(\mathfrak{p}^\alpha), \tag{1.31}$$

where $\mathfrak{p}\bar{\mathfrak{p}} = p$. In particular, for the trivial character we have $\gamma_q(1) = 1$.

THEOREM 4. *Let $q \geqslant 1$ and* Re $s = 1/2$. *We have*

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} \gamma_q(\chi)|L_K(s, \chi)|^2 \ll (q^{-1/2} + q^6 D^{-1/28})D^\varepsilon \tag{1.32}$$

where the implied constant depends on ε and s.

With a little extra work one should be able to establish an asymptotic formula in place of the estimate (1.32) as in the case of $q = 1$ in (1.28), but we do not need such a precise result. For the applications in this paper, the key issue is the uniformity of (1.32) in both D and q. Such an asymptotic formula would reveal that the term $q^{-1/2}$ represents the right order of magnitude for the twisted mean value in (1.32) (assuming that q has only prime factors of degree 1, is coprime with D, and relatively small). However, the power q^6 in the second term could certainly be reduced if we were more economical in some arguments.

Theorem 4 shows that if q has the size of a small positive power of D then, as χ runs over the group $\hat{\mathscr{H}}$, the sign change (or vanishing) of $\gamma_q(\chi)$ causes a considerable cancellation of the terms $\gamma_q(\chi)|L_K(s, \chi)|^2$. Now, as the bound for such a twisted mean value is smaller than for the straight one, one may attempt to derive a sharper bound for the individual $L_K(s, \chi)$ by ignoring the contribution from the remaining ones. Yet, at this point we cannot do so because the positivity is lost. In order to restore the positivity, we average over q. From Theorem 4 one derives easily the following theorem.

THEOREM 5. *Let c_q be any complex numbers for $q \leqslant Q$. If* Re $s = 1/2$ *we have*

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} |\mathscr{A}(\boldsymbol{c}, \chi) L_K(s, \chi)|^2 \ll \|\boldsymbol{c}\|^2 (1 + Q^7 D^{-1/28}) D^{\varepsilon}, \tag{1.33}$$

where $\|\boldsymbol{c}\|$ is the l_2-norm of $\boldsymbol{c} = \{c_q\}$.

The sum $\mathscr{A}(\boldsymbol{c}, \chi)$ is often chosen with the object of smoothing out or “mollifying” the behavior of the accompanying L-functions. In our case we deduce a sharp bound for the individual $L_K(s, \chi)$ by choosing $\mathscr{A}(\boldsymbol{c}, \chi)$ to “amplify” the contribution of the particular character χ in which we are interested.

A natural choice of the coefficients is $c_q = \gamma_q(\chi)$ giving

$$\mathscr{A}_\chi = \sum_{q \leqslant Q} \gamma_q(\chi)^2. \tag{1.34}$$

Then we drop by positivity every term on the left-hand side of (1.33) other than the one corresponding to χ and divide by $h^{-1}\mathscr{A}_\chi^2$, getting

$$L_K^2(s, \chi) \ll h \mathscr{A}_\chi^{-1} (1 + Q^7 D^{-1/28}) D^{\varepsilon}. \tag{1.35}$$

Hence it is evident that if $\mathscr{A}_\chi$ can be taken to be a power of Q when Q is a small power of D, then we can save a power of D in comparison to the convexity bound (1.9). To make $\mathscr{A}_\chi$ that large requires the existence of many primes of degree 1 having small norm. It is well known that finding such primes is a serious problem, and, indeed, interesting on its own. The corresponding obstacle in our earlier works was sometimes trivial, other times not so, but always manageable.

Later, in Section 15, we shall establish a number of results about primes of degree 1 in $K = \mathbb{Q}(\sqrt{-D})$ under various natural conditions. Under these conditions we get a lower bound

$$\mathscr{A}_\chi \gg Q^{1/3-\varepsilon}. \tag{1.36}$$

We need this for $Q = D^{6\alpha+\varepsilon}$, say, with very small α. If $\alpha \leqslant 1/1156$ is available we conclude from (1.35) that

$$L_K(s, \chi) \ll D^{1/4-\alpha+\varepsilon}. \tag{1.37}$$

One of the conditions which is sufficient to yield (1.36) for every α is that the class number is large enough, or more precisely, that

$$h = h(-D) \gg \frac{\sqrt{D}}{\log D}(\log\log D)^3. \tag{1.38}$$

This bound is probably always true since, by the Riemann hypothesis for the Dirichlet L-function $L(s, \chi_D)$, much more follows, namely,

$$\sqrt{D}(\log\log D)^{-1} \ll h(-D) \ll \sqrt{D}\log\log D. \tag{1.39}$$

Combining (1.37), (1.38), and (1.11) we derive a slight but unconditional improvement over the trivial bound for $L_K(s, \chi)$, namely, that

$$L_K(s, \chi) \ll D^{1/4}(\log D)^{-1}(\log\log D)^3. \tag{1.40}$$

Another condition which suffices for us to establish (1.36) is a nontrivial bound for the field character sums. We require that

$$\sum_{m \leqslant M} \chi_D(m) \ll MD^{-\eta} \tag{1.41}$$

holds true if $M > D^\alpha$, where $\eta = \eta(\alpha) > 0$. Note that the Burgess estimate [Bu] proves this condition for any $\alpha > 1/4$, and the Lindelöf Hypothesis implies it for any $\alpha > 0$. If D has all prime factors $< D^{\alpha^2}$, then (1.41) follows from Theorem 5 of Graham and Ringrose [GrRi].

Finally we shall examine the set of exceptional discriminants $-D$ for which small primes, say $p \leqslant D^\alpha$, of degree 1 occur rarely; see (15.16). For those D we cannot claim (1.36). However, we shall show that the number of exceptional D's in any interval of type $X < D < X^2$ is bounded by a constant depending on α alone; see (15.17). Therefore, given α, the number of exceptional discriminants $-D$ with $0 < D < Y$ is bounded by $O(\log\log Y)$.

THEOREM 6. *Let χ be a character of the class group of the field $K = \mathbb{Q}(\sqrt{-D})$. Then for* Re $s = 1/2$ *we have*

$$L_K(s, \chi) \ll D^{1/4-\alpha+\varepsilon}, \tag{1.42}$$

with $\alpha = 1/1156$ subject to any one of the following conditions:
(i) *the class number satisfies (1.38);*
(ii) *the field character sums satisfy (1.41) for all $M > D^{\alpha}$;*
(iii) *D is not α-exceptional.*
The implied constant in (1.42) depends on ε and s.

Note that the estimate (1.42) is proved unconditionally for any D having all prime factors $< D^{\alpha^2}$, with $\alpha = 1/1156$, due to the result of Graham and Ringrose. Also, the corollaries to Theorem 7 below are unconditional for such discriminants.

Since we do not display the dependence of our bounds in s, it may limit some applications. But it is easy to see, by tracking the arguments, that the implied constants in (1.32), (1.33), and (1.42) are $\ll |\Gamma(s)|^{-2} < e^{\pi|s|}$. This is a rather poor estimate in the s-aspect, yet sometimes it can be useful, for example, to derive a nontrivial bound in the D-aspect for a quite short but sufficiently well-smoothed character sum. In fact, one does not need to employ the class group L-functions to get such a result since it is already implicit in Proposition 1 (see Section 4). We infer from Proposition 1 the following theorem.

THEOREM 7. *Let $u \geqslant 1$. We have*

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}}^{*} \gamma_q(\chi) |S(u, \chi)|^2 \ll (q^{-1/2} + q^6 D^{-1/28}) u^{-1} D^{1/2+\varepsilon}, \tag{1.43}$$

where Σ^* means that the trivial character is omitted. Moreover, assuming that any one of the three conditions from Theorem 6 holds with $\alpha = 1/1156$, we have

$$S(u, \chi) \ll u^{-1/2} D^{1/2-\alpha+\varepsilon}. \tag{1.44}$$

The implied constant in the above estimates depends on ε only.

If the trivial character were present then (1.43) would be false since

$$S(u) = S(u, 1) = hw^{-1}u^{-1} + O(u^{-1/2} D^{7/16+\varepsilon}). \tag{1.45}$$

This follows from (1.4) and (1.13) by contour integration. Comparing (1.44) and (1.45), one sees instantly that a nontrivial class group character changes values on primitive ideals of norm $\ll D^{144/289+\varepsilon}$. The following corollaries are proved in Section 17 below.

COROLLARY 1. *Suppose that any one of the conditions (i), (ii), (iii) holds true. Then any coset of a subgroup of $\mathscr{H}$ contains nontrivial primitive ideals of norm $\ll k^2 D^{1/2-1/578+\varepsilon}$, where k is the index.*

Corollary 1 is related to the work of Baker and Schinzel [BS], who considered the subgroup of the principal genus (squares of classes). In this case only real characters are needed. Their result is unconditional, and the bounds for norms are sharper since they used only Burgess's estimate for character sums (an improvement was given in [H-B]). On the other hand, the genus cosets are quite large, while Corollary 1 yields a nontrivial bound for smaller cosets.

COROLLARY 2. *Suppose that any one of conditions (i), (ii), (iii) holds true. Then any cyclic subgroup of* $\mathscr{H}$ *may be generated by an ideal of norm* $\ll k^2 D^{1/2-1/578+\varepsilon}$, *where k is the index.*

After Gauss [Ga, Art. 306], a discriminant $-D$ is said to be regular if the principal genus of $\mathscr{H}$ is cyclic. Since in this case $k \ll D^{\varepsilon}$, we deduce from Corollary 2, under one of our assumptions, that *the principal genus of a regular discriminant* $-D$ *may be generated by an ideal of norm* $\ll D^{144/289+\varepsilon}$. This is somewhat analogous to the (unconditional) bound $p^{1/4+\varepsilon}$ of Burgess [Bu] for the smallest primitive root modulo p. Of course, the structure of the class group $\mathscr{H}$ is much more mysterious than that of the multiplicative group of integers modulo p. Thus, even the existence of infinitely many regular discriminants is not known. However, the numerical evidence supports the existence of a large positive proportion of regular discriminants [Bue]. In fact, it has been conjectured in [Ge] that the proportion of regular to all negative discriminants is $(\zeta(6)\prod_{n\geqslant 4}\zeta(n))^{-1} \approx .8469$.

Remarks. In this work our arguments stay within the quadratic field $K = \mathbb{Q}(\sqrt{-D})$ as much as we can afford in order to prove a number of results about the class group in addition to the main Theorem 6. One could enter the theory of automorphic forms at the beginning since, for a complex class group character, $L_K(s, \chi)$ is just the L-function of the cusp form $f(z) = S(-i\sqrt{D}z, \chi)$ of weight 1 for the group $\Gamma_0(D)$ and character χ_D (f is a newform with $2\gamma_p(\chi)$ as the Hecke eigenvalues, and corresponds to a dihedral representation; cf. [Se]). This, however, would not permit us to obtain another main result, Theorem 3. On the other hand, we can see a larger objective, that of breaking the convexity estimates for general automorphic L-functions of any integral weight and any Dirichlet character. This more general setting opens the possibility of exploiting the full spectrum of automorphic forms and may bring not only new results but also refine these for the class group L-functions. We intend to investigate this possibility on another occasion.

Acknowledgment. We thank Peter Sarnak for numerous discussions during the course of this work. W.D and H.I. are grateful to the Mathematics Department of the University of Toronto for providing hospitality and financial support.

2. Ideals in K. We recall some basic facts about ideal theory in the quadratic field $K = \mathbb{Q}(\sqrt{-D})$. Every rational prime p either factors as a product $p = \mathfrak{p}\bar{\mathfrak{p}}$ of two primes (not necessarily distinct) each of which has norm p and degree 1 or remains prime in K having norm p^2 and degree 2. In the first case $\mathfrak{p} \neq \bar{\mathfrak{p}}$ except for $p|D$ (the ramified primes). The decomposition of p in K as above is characterized by values of the field character; $\chi_D(p) = 1, 0, -1$, respectively. We say an ideal is of degree 1 (respectively 2) if every prime factor has degree 1 (respectively 2). Every ideal $\neq 0$ factors uniquely as the product of one of each of these.

We shall often factor an integral ideal $\neq 0$ uniquely as $(l)\mathfrak{a}$ where l is a positive integer and $\mathfrak{a}$ is a primitive ideal, i.e., $\mathfrak{a}$ has no rational integer factors other than ± 1. Note that the primitive ideals are characterized by the condition

$$(\mathfrak{a}, \bar{\mathfrak{a}}) | \sqrt{D}. \tag{2.1}$$

The ring of integers $\mathcal{O} \subset K$ is a free $\mathbb{Z}$-module of rank 2 generated by

$$\mathcal{O} = \left[1, \frac{D + i\sqrt{D}}{2}\right].$$

Every ideal $0 \neq \mathfrak{a} \subset \mathcal{O}$ is also a free $\mathbb{Z}$-module. If $\mathfrak{a}$ is primitive, then it is generated by

$$\mathfrak{a} = \left[a, \frac{b + i\sqrt{D}}{2}\right], \tag{2.2}$$

where $a = N\mathfrak{a}$, and b solves the congruence

$$b^2 + D \equiv 0 \pmod{4a} \tag{2.3}$$

and is determined modulo $2a$. Conversely, to every $b \pmod{2a}$ satisfying (2.3) corresponds the ideal $\mathfrak{a}$ generated by (2.2), and this is primitive. Therefore, there exists a one-to-one correspondence between the primitive ideals and the points

$$z_{\mathfrak{a}} = \frac{b + i\sqrt{D}}{2a} \tag{2.4}$$

in the upper half-plane $H = \{z = x + iy : x \in \mathbb{R}, y \in \mathbb{R}^+\}$ determined modulo 1. These will be called the Heegner points. The inverse $\mathfrak{a}^{-1}$ as a fractional ideal is a free $\mathbb{Z}$-module generated by

$$\mathfrak{a}^{-1} = [1, \bar{z}_{\mathfrak{a}}]. \tag{2.5}$$

The primitive ideals correspond in one-to-one fashion to the quadratic forms $[a, b, c] = aX^2 + bXY + cY^2$, with $a, c > 0$, $(a, b, c) = 1$, $4ac - b^2 = D$, and the Heegner point is just the one of the two roots which lies in H.

The modular group $\Gamma = SL_2(\mathbb{Z})$ acts on H by the linear fractional transformations

$$\sigma z = \frac{\alpha z + \beta}{\gamma z + \delta} \quad \text{if} \quad \sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma.$$

The set of Heegner points is mapped to itself. Precisely, we have

$$\sigma \frac{b + i\sqrt{D}}{2a} = \frac{b^\sigma + i\sqrt{D}}{2a^\sigma},$$

where

$$\begin{aligned} a^\sigma &= a\delta^2 + b\gamma\delta + c\gamma^2 \\ b^\sigma &= b + 2a\beta\delta + 2b\beta\gamma + 2c\alpha\gamma \\ D &= 4ac - b^2. \end{aligned} \tag{2.6}$$

Given a fundamental domain of Γ, say $\mathscr{D}$, we put

$$\Lambda_D = \{z_\mathfrak{a} \in \mathscr{D} \colon \mathfrak{a} \text{ primitive}\}. \tag{2.7}$$

The set Λ_D is finite; its cardinality is just the class number $|\Lambda_D| = h(-D)$. This gives a one-to-one correspondence between the primitive ideals and the points of the orbits

$$\{\sigma z \colon \sigma \in \Gamma_\infty \backslash \Gamma / \Gamma_z\} \qquad \text{for } z \in \Lambda_D,$$

where Γ_z is the stability group of z.

The above analysis can be generalized so as to replace the ring $\mathcal{O}$ by an ideal $\mathfrak{q} \subset \mathcal{O}$ such that

$$(\mathfrak{q}, \bar{\mathfrak{q}}) = 1. \tag{2.8}$$

Consider the integral ideals $\mathfrak{a} \subset \mathfrak{q}$ (i.e., divisible by $\mathfrak{q}$) which are primitive. The inclusion $\mathfrak{a} \subset \mathfrak{q}$ is expressed in terms of the Heegner point $z_\mathfrak{a}$ by the following congruence conditions

$$a \equiv 0 \pmod{q}, \qquad b + i\sqrt{D} \equiv 0 \pmod{2\mathfrak{q}} \tag{2.9}$$

where $q = N\mathfrak{q}$. From (2.6) it is clear that the group $\Gamma = \Gamma_0(q)$ acts on points $z_\mathfrak{a}$ for $\mathfrak{a} \subset \mathfrak{q}$ because it preserves the conditions (2.9). Given a fundamental domain of Γ, say $\mathscr{D}$, we put

$$\Lambda_D(\mathfrak{q}) = \{z_\mathfrak{a} \in \mathscr{D} \colon \mathfrak{a} \subset \mathfrak{q}, \mathfrak{a} \text{ primitive}\}. \tag{2.10}$$

This establishes a one-to-one correspondence between the primitive ideals $\mathfrak{a} \subset \mathfrak{q}$ and the points of the orbits

$$\{\sigma z \colon \sigma \in \Gamma_\infty \backslash \Gamma / \Gamma_z\} \qquad \text{for } z \in \Lambda_D(\mathfrak{q}).$$

When dealing with various sums over the Heegner points $z_{\mathfrak{a}}$, it is convenient for technical reasons to remove the second congruence condition of (2.9). This condition will disappear after summing over all factorizations of $q = \mathfrak{q}\bar{\mathfrak{q}}$ such that $(\mathfrak{q}, \bar{\mathfrak{q}}) = 1$. If q admits such a factorization, then every prime factor has degree 1 and is unramified; the number of such factorizations is then given by $2^{\omega(q)}$. In other words, we shall take all the primitive ideals $\mathfrak{a} \subset \mathcal{O}$ of norm $N\mathfrak{a} = a \equiv 0 \pmod{q}$ rather than $\mathfrak{a} \subset \mathfrak{q}$ only. Accordingly we introduce the set

$$\Lambda_D(q) = \{z_{\mathfrak{a}} \in \mathscr{D}: a \equiv 0 \ (\mathrm{mod}\ q),\ \mathfrak{a} \text{ primitive}\}. \tag{2.11}$$

3. The average value of $L_K(s, \chi)$. In this section we prove Theorem 1. We have the following identity

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} \chi(\mathfrak{a}) L_K(s, \chi) = w^{-1} \left(\frac{\sqrt{D}}{2}\right)^{-s} \zeta(2s) E(z_{\mathfrak{a}}, s), \tag{3.1}$$

where $\mathfrak{a}$ is any primitive ideal, $z_{\mathfrak{a}}$ is the Heegner point, and $E(z, s)$ is the Eisenstein series for the modular group. To see this, recall that

$$\zeta(2s) E(z, s) = y^s \sum\sum_{(m,n) \neq (0,0)} |m + nz|^{-2s}. \tag{3.2}$$

On the other hand, the left side of (3.1) is given by

$$\sum_{\mathfrak{b} \sim \mathfrak{a}} (N\mathfrak{b})^{-s} = w^{-1} (N\mathfrak{a})^{-s} \sum_{0 \neq \alpha \in \mathfrak{a}} |\alpha|^{-2s} = w^{-1} a^{-s} \sum\sum_{(m,n) \neq (0,0)} |m + n z_{\mathfrak{a}}|^{-2s}.$$

Combining these we get (3.1).

The Eisenstein series at the Heegner point $z_{\mathfrak{a}} = (b + i\sqrt{D})/2a$ can be estimated by using the Fourier expansion (see [Iw])

$$\Theta(s) E(z, s) = \Theta(s) y^s + \Theta(1 - s) y^{1-s} + 4 y^{1/2} \sum_{k=1}^{\infty} \left(\sum_{mn=k} \left(\frac{m}{n}\right)^{it} \right) K_{it}(2\pi k y) \cos(2\pi k x),$$

where $\Theta(s) = \pi^{-s}\Gamma(s)\zeta(2s)$. The contribution from $k \geqslant 1$ is bounded trivially by $O(e^{-2\pi y})$ for z in the standard fundamental domain. In particular, for given $\mathfrak{a}$ we choose $z = z_{\mathfrak{a}}$, the Heegner point in this fundamental domain, getting from (3.1) that

$$\left(\frac{\sqrt{D}}{2\pi}\right)^s \Gamma(s) \frac{w}{h} \sum_{\chi \in \hat{\mathscr{H}}} \chi(\mathfrak{a}) L_K(s, \chi) = \Theta(s) \left(\frac{\sqrt{D}}{2a}\right)^s + \Theta(1 - s) \left(\frac{\sqrt{D}}{2a}\right)^{1-s} + O(e^{-\pi\sqrt{D}/a}).$$

This formula is very precise if $a = N\mathfrak{a}$ is small compared to $\sqrt{D}$. Taking $\mathfrak{a} = (1)$ we obtain Theorem 1.

We are now ready to proceed to the second power-moment. By Fourier inversion we get from (3.1) that

$$L_K(s, \chi) = w^{-1}\left(\frac{\sqrt{D}}{2}\right)^{-s} \zeta(2s) \sum_{z_\mathfrak{a} \in \Lambda_D} \bar{\chi}(\mathfrak{a}) E(z_\mathfrak{a}, s). \tag{3.3}$$

By Plancherel's theorem we obtain

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} |L_K(s, \chi)|^2 = 2w^{-2} D^{-1/2} |\zeta(2s)|^2 \sum_{z_\mathfrak{a} \in \Lambda_D} |E(z_\mathfrak{a}, s)|^2. \tag{3.4}$$

Therefore, Theorems 2 and 3 are statements about the distribution of the values of this Eisenstein series at the Heegner points. Since the Heegner points themselves are equidistributed by a theorem of Duke [Du], one might hope to approximate the right-hand side by the corresponding integral over the fundamental domain. This leads to two problems, however. In the first place, the fact that the Eisenstein series is not square-integrable causes technical difficulties. In the second place, this method does not seem amenable to the twisted sums occurring in Theorem 4 and needed for the main applications. Therefore we shall use an alternative approach.

4. Preliminary transformations. Denote by $\mathscr{L}$ the left-hand side of (1.32):

$$\mathscr{L} = \frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} \gamma_q(\chi) |L_K(s, \chi)|^2. \tag{4.1}$$

The problem of evaluating $\mathscr{L}$ reduces to that of

$$\mathscr{M} = \frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} \gamma_q(\chi) |\Phi(s, \chi) + \delta_\chi h w^{-1} s^{-1} (1-s)^{-1}|^2. \tag{4.2}$$

Indeed, we have

$$\mathscr{M} = (2\pi)^{-1} |\Gamma(s)|^2 D^{1/2} \mathscr{L} + h(ws(1-s))^{-2} + 2\Phi(s)(ws(1-s))^{-1}. \tag{4.3}$$

where, in the last term,

$$\Phi(s) = \Phi(s, 1) = (2\pi)^{-s} \Gamma(s) D^{s/2} \zeta_K(s). \tag{4.4}$$

This is quite small; namely, by (1.11) we have

$$\Phi(s) \ll s(1-s) |\Gamma(s)| D^{3/16+\varepsilon}. \tag{4.5}$$

Next, using the integral representation (1.7), we transform $\mathcal{M}$ into

$$\mathcal{M} = \int_1^\infty \int_1^\infty (u_1^{s-1} + u_1^{-s})(u_2^{s-1} + u_2^{-s})\mathcal{M}(u_1, u_2)\, du_1\, du_2, \tag{4.6}$$

where

$$\mathcal{M}(u_1, u_2) = \frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} \gamma_q(\chi) S(u_1, \chi) S(u_2, \chi) \tag{4.7}$$

and $S(u, \chi)$ are defined by (1.8). We shall prove the following (see the conclusion of Section 12).

PROPOSITION 1. *For any $u_1, u_2 \geqslant 1$ we have*

$$\mathcal{M}(u_1, u_2) = \frac{h}{w^2 u_1 u_2} + O\left((q^{-1/2} + q^6 D^{-1/28})\left(\frac{D}{u_1 u_2}\right)^{1/2+\varepsilon}\right), \tag{4.8}$$

where the implied constant depends on ε only.

The contribution to $\mathcal{M}(u_1, u_2)$ from the trivial character is $h^{-1}S(u_1)S(u_2)$ where $S(u)$ satisfies (1.45). Therefore, the trivial character contributes

$$h^{-1}S(u_1)S(u_2) = \frac{h}{w^2 u_1 u_2} + O((u_1 u_2)^{-1/2} D^{7/16+\varepsilon}). \tag{4.9}$$

The main terms in (4.8) and (4.9) match. Subtracting, we infer that

$$\mathcal{M}^*(u_1, u_2) \ll (q^{-1/2} + q^6 D^{-1/28})(u_1 u_2)^{-1/2} D^{1/2+\varepsilon}. \tag{4.10}$$

Hence, taking $u_1 = u_2 = u$ we obtain (1.43). The derivation of the individual bound (1.44) goes by constructing an amplifier in the same fashion as (1.37) is derived from Theorem 5. To this end, one needs the lower bound (1.36) which will be established in Section 16. Applying this bound, we complete the proof of Theorem 7.

Inserting (4.10) into (4.6) we get

$$\mathcal{M} = h(ws(1-s))^{-2} + O((q^{-1/2} + q^6 D^{-1/28}) D^{1/2+\varepsilon}). \tag{4.11}$$

Then, comparing this with (4.3), we complete the proof of Theorem 4.

Most of the problems from Section 1 are now transferred to the evaluation of $\mathcal{M}(u_1, u_2)$. Using the orthogonality of characters, we obtain

$$\mathcal{M}(u_1, u_2) = 2^{-\omega(q)} \sum_{\substack{\mathfrak{q}\bar{\mathfrak{q}} = q \\ (\mathfrak{q}, \bar{\mathfrak{q}})=1}} \sum_{\substack{\mathfrak{a}_1 \ \mathfrak{a}_2 \\ \mathfrak{a}_1 \mathfrak{q} \sim \mathfrak{a}_2}} \sum \exp\left(\frac{-2\pi}{\sqrt{D}}(u_1 a_1 + u_2 a_2)\right), \tag{4.12}$$

where a_1, a_2 are the norms of $\mathfrak{a}_1$, $\mathfrak{a}_2$, respectively, and $\sim$ denotes the ideal equivalence. To proceed further we appeal to the theory of ideals of the field $K = \mathbb{Q}(\sqrt{-D})$ described in Section 2. We write uniquely $\mathfrak{a}_1\mathfrak{q} = (l)\mathfrak{a}$, where l is a rational integer and $\mathfrak{a}$ is a primitive integral ideal. Note that as l ranges over positive integers the principal ideal (l) is encountered $w/2$ times. Moreover, the equivalence relation becomes $\mathfrak{a} \sim \mathfrak{a}_2$, and the congruence condition $(l)\mathfrak{a} \equiv 0 \pmod{\mathfrak{q}}$ becomes $\mathfrak{a} \equiv 0 \pmod{\mathfrak{q}_l}$, where $\mathfrak{q}_l = \mathfrak{q}/(l, \mathfrak{q})$. Hence, the norm satisfies

$$a = N\mathfrak{a} \equiv 0 \pmod{q_l}, \tag{4.13}$$

where $q_l = q/(l, q)$ because $(\mathfrak{q}, \bar{\mathfrak{q}}) = 1$. Furthermore, since $\mathfrak{a}$ is primitive, the condition $\mathfrak{a} \equiv 0 \pmod{\mathfrak{q}_l}$ as $\mathfrak{q}$ varies fixes the choice between $\mathfrak{p}$ and $\bar{\mathfrak{p}}$ for all prime factors $p = \mathfrak{p}\bar{\mathfrak{p}}$ of q_l. For each remaining prime factor of q, we have two choices. Therefore, given $\mathfrak{a}$ satisfying (4.13), we have

$$\#\{\mathfrak{q}\colon \mathfrak{q}\bar{\mathfrak{q}} = q, (\mathfrak{q}, \bar{\mathfrak{q}}) = 1, \mathfrak{q}_l \supset \mathfrak{a}\} = 2^{\omega(q)-\omega(q_l)}.$$

Hence,

$$\mathscr{M}(u_1, u_2) = \frac{2}{w}\sum_{l=1}^{\infty} 2^{-\omega(q_l)} \sideset{}{'}\sum_{\substack{\mathfrak{a}\sim\mathfrak{a}_2 \\ a\equiv 0(q_l)}}\sum \exp\left(-\frac{2\pi}{\sqrt{D}}(u_1 q^{-1}l^2 a + u_2 a_2)\right),$$

where $\sum'$ restricts the summation to primitive integral ideals. The equivalence of ideals says that $\mathfrak{a}_2 = (\alpha)\mathfrak{a}$ with $\alpha \in \mathfrak{a}^{-1}$. Since a principal ideal determines its generator up to a unit, we obtain

$$\mathscr{M}(u_1, u_2) = \frac{2}{w^2}\sum_{l=1}^{\infty} 2^{-\omega(q_l)} \sideset{}{'}\sum_{\substack{\mathfrak{a} \\ a\equiv 0(q_l)}}\sum_{\alpha\in\mathfrak{a}^{-1}} \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 q^{-1}l^2 + u_2|\alpha|^2)\right).$$

Every $\mathfrak{a}$ corresponds to a Heegner point $z_\mathfrak{a} \in H$ determined modulo 1, and $\mathfrak{a}^{-1}$ is a free $\mathbb{Z}$-module generated by $[1, \bar{z}_\mathfrak{a}]$. Therefore $\alpha = m + nz_\mathfrak{a}$, where m, n run freely over integers not both zero, giving

$$\mathscr{M}(u_1, u_2) = \frac{4}{w^2}\sum_{l=1}^{\infty} 2^{-\omega(q_l)} \sum\sum_{(m,n)\neq(0,0)} \sideset{}{'}\sum_{a\equiv 0(q_l)} \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 q^{-1}l^2 + u_2|m + nz_\mathfrak{a}|^2)\right).$$

The contribution from terms with $n = 0$ is called the diagonal part of $\mathscr{M}(u_1, u_2)$; it is given by

$$\mathscr{M}^0(u_1, u_2) = \frac{2}{w^2}\sum_{l=1}^{\infty} 2^{-\omega(q_l)} \sum_{m=1}^{\infty} \sideset{}{'}\sum_{a\equiv 0(q_l)} \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 q^{-1}l^2 + u_2 m^2)\right). \tag{4.14}$$

In other words, the diagonal part is the contribution from pairs of ideals whose primitive kernels (the maximal rational-free factors) are equal.

The contribution from the remaining terms (the off-diagonal pairs) is partitioned into

$$\mathscr{M}^+(u_1, u_2) = \frac{4}{w^2} \sum_{l=1}^{\infty} 2^{-\omega(q_l)} \sum_{n=1}^{\infty} M_{ln}(u_1, u_2), \tag{4.15}$$

where

$$\mathscr{M}_{ln}(u_1, u_2) = \sum_{m \in \mathbb{Z}} \sum_{\mathfrak{a} \equiv 0(q_l)}' \exp\left(\frac{-\pi}{\operatorname{Im} z_{\mathfrak{a}}}(u_1 q^{-1} l^2 + u_2 |m + n z_{\mathfrak{a}}|^2)\right). \tag{4.16}$$

Integrating over u_1, u_2 as in (4.6) we define the corresponding parts of $\mathscr{M}$;

$$\mathscr{M}^0 = \int_1^{\infty} \int_1^{\infty} (u_1^{s-1} + u_1^{-s})(u_2^{s-1} + u_2^{-s}) \mathscr{M}(u_1, u_2)\, du_1\, du_2 \tag{4.17}$$

$$\mathscr{M}^+ = \int_1^{\infty} \int_1^{\infty} (u_1^{s-1} + u_1^{-s})(u_2^{s-1} + u_2^{-s}) \mathscr{M}^+(u_1, u_2)\, du_1\, du_2 \tag{4.18}$$

$$\mathscr{M}_{ln} = \int_1^{\infty} \int_1^{\infty} (u_1^{s-1} + u_1^{-s})(u_2^{s-1} + u_2^{-s}) \mathscr{M}_{ln}(u_1\, u_2)\, du_1\, du_2 . \tag{4.19}$$

Therefore we have $\mathscr{M} = \mathscr{M}^0 + \mathscr{M}^+$ and

$$\mathscr{M}^+ = \frac{4}{w^2} \sum_{l=1}^{\infty} 2^{-\omega(q_l)} \sum_{n=1}^{\infty} M_{ln} . \tag{4.20}$$

In view of (4.3) we define the diagonal part of $\mathscr{L}$ by

$$\mathscr{L}^0 = 2\pi |\Gamma(s)|^{-2} D^{-1/2} \mathscr{M}^0 . \tag{4.21}$$

The complementary part $\mathscr{L}^+ = \mathscr{L} - \mathscr{L}^0$ is given by

$$\mathscr{L}^+ = 2\pi |\Gamma(s)|^{-2} D^{-1/2} \{\mathscr{M}^+ - h(ws(1-s))^{-2} - 2\Phi(s)(ws(1-s))^{-1}\} . \tag{4.22}$$

5. Estimation of the diagonal part. We apply here mostly elementary means to give simple, yet sharp estimates for the diagonal parts $\mathscr{M}^0(u_1, u_2)$ and $\mathscr{M}^0$.

To get an upper bound we proceed as follows:

$$\begin{aligned} \mathscr{M}^0(u_1, u_2) &\ll \sum_{l=1}^{\infty} \sum_{m=1}^{\infty} \sum_{\mathfrak{a}} \exp\left(\frac{-2\pi a}{\sqrt{D}(l, q)}(u_1 l^2 + u_2 q m^2)\right) \\ &\leqslant \sum_{rs=q} \sum_{l=1}^{\infty} \sum_{m=1}^{\infty} \sum_{\mathfrak{a}} \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 r l^2 + u_2 s m^2)\right) \\ &\ll \tau(q) \left(\frac{D}{q u_1 u_2}\right)^{1/2} \exp\left(\frac{-\pi}{\sqrt{D}}(u_1 + u_2)\right) L . \end{aligned} \tag{5.1}$$

For the last step in (5.1) we have applied the general inequality

$$x^{-1/2}e^{-x} \ll \sum_{k=1}^{\infty} e^{-xk^2} \ll x^{-1/2}e^{-x/2},$$

followed by $a(u_1 r + u_2 s) \geqslant a(u_1 + u_2) \geqslant 2a + u_1 + u_2 - 2$. Recall that L is given by (1.23). Integrating over u_1, u_2, we get

$$\mathcal{M}^0 \ll \tau(q)q^{-1/2}D^{1/2}(\log D)^2 L. \tag{5.2}$$

Next we give a lower bound but only in the case of $q = 1$ and $s = 1/2$. In this case

$$\mathcal{M}^0(u_1, u_2) = \frac{4}{w^2}\sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\sum_{\mathfrak{a}}{}' \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 l^2 + u_2 m^2)\right). \tag{5.3}$$

Hence, as before, we infer that

$$\mathcal{M}^0(u_1, u_2) \gg \left(\frac{D}{u_1 u_2}\right)^{1/2} \sum_{\mathfrak{a}}{}' a^{-1} \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 + u_2)\right).$$

Integrating over u_1, u_2, we obtain

$$\mathcal{M}^0 \gg D^{1/2} \sum_{\mathfrak{a}}{}' a^{-1}\left(\log^+ \frac{\sqrt{D}}{2\pi a}\right)^2. \tag{5.4}$$

Then, using the asymptotic formula

$$\sum_{D^\alpha < N\mathfrak{a} < D^\beta} (N\mathfrak{a})^{-1} \sim (\beta - \alpha)L(1, \chi_D) \log D, \tag{5.5}$$

valid for any $\beta > \alpha > 3/8$ (this can be proved by a standard contour integration using (1.2), (1.4), (1.11)), we deduce that

$$\begin{aligned}
\sum_{\mathfrak{a}}{}' a^{-1}\left(\log^+ \frac{\sqrt{D}}{2a}\right)^2 &\gg (\log D)^2 \sum_{N\mathfrak{a} < D^{1/2-\gamma}}{}' (N\mathfrak{a})^{-1} \\
&\gg (\log D)^2 \sum_{N\mathfrak{a} < D^{1/2-\gamma}} (N\mathfrak{a})^{-1} \\
&\gg (\log D)^2 \sum_{N\mathfrak{a} < D^{1/2+\gamma}} (N\mathfrak{a})^{-1} \gg (\log D)^2 L,
\end{aligned}$$

where γ is a small positive constant. Hence, we conclude by (5.4) that

$$\mathcal{M}^0 \gg D^{1/2}(\log D)^2 L. \tag{5.6}$$

Finally we estimate L. By (5.5) we get the lower bound in (1.24). Moreover, we have

$$L = \frac{1}{2\pi i}\int_{(1/2)} (2\pi)^{-v}\Gamma(v)D^{v/2}\zeta_K(v+1)\,dv$$

$$= L(1,\chi_D)\log\frac{\sqrt{D}}{2\pi} + L'(1,\chi_D) + O(D^{-1/16+\varepsilon}), \tag{5.7}$$

by moving to the line $\mathrm{Re}\, v = -1/2$ and using (1.11). Notice that the above error term is negligible because $L \geqslant 1$ from the first term $\mathfrak{a} = 1$ in L. By (5.7) we get the upper bound in (1.24).

6. Asymptotic evaluation of the diagonal part. In this section we examine $\mathscr{M}^0$ for $q = 1$ and any s on the critical line $\mathrm{Re}\, s = 1/2$ with greater detail than given in Section 5. First, we pull out the greatest common factor of l and m in (5.3) and attach it to the ideal $\mathfrak{a}$, getting

$$\mathscr{M}^0(u_1, u_2) = \frac{2}{w}\sum_{\mathfrak{a}}\sum_{(l,m)=1}\sum \exp\left(\frac{-2\pi a}{\sqrt{D}}(u_1 l^2 + u_2 m^2)\right),$$

where the outer summation ranges over all integral ideals. Next, using the integral

$$e^{-x} = \frac{1}{2\pi i}\int_{(\sigma)} \Gamma(v)x^{-v}\,dv, \tag{6.1}$$

with $\sigma > 1$, we get by (5.3) that

$$\mathscr{M}^0(u_1, u_2) = \frac{2}{w}\frac{1}{2\pi i}\int_{(\sigma)} \Phi(v)G(u_1, u_2; v)\,dv, \tag{6.2}$$

where

$$G(u_1, u_2; v) = \sum_{l=1}^{\infty}\sum_{\substack{m=1\\(l,m)=1}}^{\infty} (u_1 l^2 + u_2 m^2)^{-v}. \tag{6.3}$$

This can be expressed by means of the Eisenstein series for the modular group whose properties are well known. Nevertheless, we choose a direct (really faster) approach.

We begin by decoupling the variables u_1, u_2 and l, m in (6.3) by means of the following formula:

$$\Gamma(v)(x+y)^{-v} = \frac{1}{2\pi i}\int_{(\eta)} \Gamma(z)\Gamma(v-z)x^{-z}y^{z-v}\,dz,$$

with $1/2 < \eta < \sigma - 1/2$. We get

$$G(u_1, u_2; v) = \frac{\pi^v}{\Gamma(v)\zeta(2v)} \frac{1}{2\pi i} \int_{(\eta)} \Theta(z)\Theta(v - z) u_1^{-z} u_2^{v-z}\, dz, \tag{6.4}$$

where

$$\Theta(z) = \pi^{-z}\Gamma(z)\zeta(2z). \tag{6.5}$$

The function $\Theta(z)$ is meromorphic on the whole complex plane, and it satisfies the following functional equation (inherited from that for the Riemann zeta-function):

$$\Theta(z) = \Theta(1/2 - z). \tag{6.6}$$

Moreover $\Theta(z)$ has only simple poles at $z = 1/2$ and $z = 0$ with the Laurent series at $z = 1/2$ being

$$\Theta(z) = \frac{1}{2}(z - 1/2)^{-1} + \alpha_0 + \alpha_1(z - 1/2) + \alpha_2(z - 1/2)^2 + \cdots, \tag{6.7}$$

where $\alpha_0 = (1/2)(\gamma - \log 4\pi)$ and $\alpha_1, \alpha_2, \ldots$ are more involved constants.

Now we can integrate (6.4) over u_1, u_2 explicitly getting

$$\begin{aligned} G(v) &= \int_1^\infty \int_1^\infty (u_1^{s-1} + u_1^{-s})(u_2^{s-1} + u_2^{-s}) G(u_1, u_2; v)\, du_1\, du_2 \\ &= \pi^v \Gamma(v)^{-1} \zeta(2v)^{-1} R(v), \end{aligned}$$

where

$$R(v) = \frac{1}{2\pi i} \int_{(\eta)} \Theta(z)\Theta(v - z) \left(\frac{1}{z - s} + \frac{1}{z - 1 + s}\right)\left(\frac{1}{v - z - s} + \frac{1}{v - z - 1 + s}\right) dz. \tag{6.8}$$

Inserting this into (6.2), after integration over u_1, u_2 we arrive at

$$\mathscr{M}^0 = \frac{2}{w} \frac{1}{2\pi i} \int_{(\sigma)} \left(\frac{\sqrt{D}}{2}\right)^v \frac{\zeta(v)}{\zeta(2v)} L(v, \chi_D) R(v)\, dv. \tag{6.9}$$

Hence, we shall evaluate $\mathscr{M}^0$ by moving the integration to the vertical line $\operatorname{Re} v = 1/2 + \varepsilon$. To this end we need an analytic continuation of $R(v)$ to the half-plane $\operatorname{Re} v > 1/2$.

Clearly (6.8) shows that $R(v)$ is holomorphic in $\operatorname{Re} v > 1$ as η can be chosen arbitrarily close to 1/2. Since $\Theta(z)$ decays exponentially on vertical lines, we can move the contour in (6.8) from $\operatorname{Re} z = \eta$ to any line $\operatorname{Re} z = \gamma$ with $0 < \gamma < 1/2$. The poles may occur at $z = 1/2, s, 1 - s$. If $s \neq 1/2$ all poles are simple. In fact, $z = 1/2$ is a removable singularity. At $z = s$ the residue is

$$R_s(v) = \left(\frac{1}{v-1} + \frac{1}{v-2s}\right)\Theta(s)\Theta(v-s), \tag{6.10}$$

and at $z = 1 - s$ the residue is given by the above expression with s changed into $1 - s$. If $s = 1/2$ the only pole at $z = 1/2$ has order 2, and the residue is $2R_{1/2}(v)$, where

$$R_{1/2}(v) = \frac{1}{v-1}\left(\frac{1}{v-1} + 2\alpha_0\right)\Theta(v - 1/2) - \frac{1}{v-1}\Theta'(v - 1/2). \tag{6.11}$$

Gathering the above results, we obtain

$$R(v) = R_s(v) + R_{1-s}(v) + \tilde{R}(v), \tag{6.12}$$

where $\tilde{R}(v)$ is given by the integral (6.8) on the line $\operatorname{Re} z = \gamma$ with $0 < \gamma < 1/2$.

Since γ can be chosen arbitrarily small, it proves that $\tilde{R}(v)$ is holomorphic in $\operatorname{Re} v > 1/2$ and that $\tilde{R}(v)$ decays exponentially on vertical lines. The other functions $R_s(v)$ defined by (6.10) if $s \neq 1/2$ and (6.11) if $s = 1/2$ are holomorphic in $\operatorname{Re} v > 1/2$ except for a finite number of poles on the line $\operatorname{Re} v = 1$, and they also decay exponentially. If $s \neq 1/2$ then $R_s(v)$ has simple poles at $v = 1$ and $v = 2s$ (but not at $v = s + 1/2$) with the Laurent series

$$R_s(v) = (v-1)^{-1}\Theta(s)\Theta(1-s) + \Theta(s)\left(\frac{\Theta(1-s)}{1-2s} + \Theta'(1-s)\right) + \cdots \tag{6.13}$$

and

$$R_s(v) = (v-2s)^{-1}\Theta(s)^2 + \Theta(s)\left(\frac{\Theta(s)}{2s-1} + \Theta'(s)\right) + \cdots, \tag{6.14}$$

respectively. If $s = 1/2$ then $R_{1/2}(v)$ has only a pole at $v = 1$ of order 3. By (6.11) and (6.7), we find that the Laurent series is

$$R_{1/2}(v) = (v-1)^{-3} + 2\alpha_0(v-1)^{-2} + 2\alpha_0^2(v-1)^{-1} + 2\alpha_0\alpha_1 - \alpha_2 + \cdots. \tag{6.15}$$

Finally, by (6.12) we obtain the desired continuation of $R(v)$ to $\operatorname{Re} v > 1/2$ with poles at $v = 1, 2s, 2 - 2s$ as described above.

Having established analytic properties of $R(v)$ we can now evaluate $\mathcal{M}^0$ by moving the integration in (6.9) to the line $\operatorname{Re} v = 1/2 + \varepsilon$. From the poles on the line $\operatorname{Re} v = 1$, we get

$$P = \frac{2}{w} \sum_{\text{poles}} \operatorname{res} \left(\frac{\sqrt{D}}{2}\right)^v \frac{\zeta(v)}{\zeta(2v)} L(v, \chi_{\mathrm{D}})[R_s(v) + R_{1-s}(v) + \tilde{R}(1)], \tag{6.16}$$

and the resulting integral on $\operatorname{Re} v = 1/2 + \varepsilon$ is estimated by using Burgess inequality for $L(v, \chi_{\mathrm{D}})$. We obtain

$$\mathcal{M}^0 = P + O(D^{7/16+\varepsilon}). \tag{6.17}$$

This yields the corresponding asymptotic for the diagonal part of $\mathcal{L}$ (see (4.21)),

$$\mathcal{L}^0 = l_D(s) + O(D^{-1/16+\varepsilon}). \tag{6.18}$$

Here the main term is given by

$$l_D(s) = 2\pi |\Gamma(s)|^{-2} D^{-1/2} P; \tag{6.19}$$

thus $l_D(s)$ is the sum of residues of $w^{-1} D^{v-1/2} L(v, \chi_{\mathrm{D}}) H(v)$ at $v = 1,\ 2s,\ 2 - 2s$, where

$$H(v) = (2\pi)^{-1} |\Gamma(s)|^{-2} \zeta(v) \zeta(2v)^{-1} [R_s(v) + R_{1-s}(v) + \tilde{R}(1)]. \tag{6.20}$$

Recall that by (6.8)

$$\tilde{R}(1) = \frac{-1}{2\pi i} \int_{(\gamma)} \frac{\zeta(2z)\zeta(2-2z)}{\sin \pi z} \left(\frac{2z-1}{(z-s)(z-1+s)}\right)^2 dz,$$

where $0 < \gamma < 1/2$. We complete this section by computing quite explicitly the residues in P.

If $s = 1/2$ there is only one pole at $v = 1$ of order 4, and the residue gives

$$P = \frac{2}{w} \operatorname*{res}_{v=1} \left(\frac{\sqrt{D}}{2}\right)^v \frac{2\zeta(v)}{\zeta(2v)} L(v, \chi_{\mathrm{D}}) R_{1/2}(v) + \frac{6\sqrt{D}}{\pi^2 w} L(1, \chi_{\mathrm{D}}) \tilde{R}(1). \tag{6.21}$$

Hence, by (6.15),

$$P = \frac{\sqrt{D}}{2} \sum_{j+k \leqslant 3} c_{jk} L^{(j)}(1, \chi_{\mathrm{D}}) (\log D)^k \tag{6.22}$$

for some constants c_{jk}. Here the highest coefficients is equal to $c_{03} = 6\pi^{-2} w^{-1}$.

If $s \neq 1/2$ there are three different poles at $v = 1,\ 2s,\ 2 - 2s$. Acoordingly P splits into

$$P = P_1 + P_{2s} + P_{2-2s}, \tag{6.23}$$

say. The pole at $v = 1$ has order 2. By (6.13), the residue gives

$$\begin{aligned} P_1 &= \frac{2}{w} \operatorname*{res}_{v=1} \left(\frac{\sqrt{D}}{2}\right)^v \frac{\zeta(v)}{\zeta(2v)} L(v, \chi_D) \\ &\quad \times \{2(v-1)^{-1}\Theta(s)\Theta(1-s) + \Theta'(s)\Theta(1-s) + \Theta'(1-s)\Theta(s)\} \\ &\quad + \frac{6\sqrt{D}}{\pi^2 w} L(1, \chi_D)\tilde{R}(1) \\ &= (2\pi)^{-1}\Gamma(s)\Gamma(1-s)\sqrt{D} \sum_{j+k\leqslant 1} c_{jk}(s)L^{(j)}(1, \chi_D)(\log D)^k \end{aligned} \tag{6.24}$$

for some coefficients $c_{jk}(s)$ depending on s. The highest coefficient is equal to $c_{01}(s) = (2\pi^2 w)^{-1}\zeta(2s)\zeta(2-2s)$. The pole at $v = 2s$ is simple. By (6.14) the residue gives

$$P_{2s} = \frac{2}{w}\left(\frac{\sqrt{D}}{2}\right)^{2s} \frac{\zeta(2s)}{\zeta(4s)} L(2s, \chi_D)\Theta(s)^2 = (2\pi)^{-1}\Gamma(s)\Gamma(1-s)D^s c(s) L(2s, \chi_D), \tag{6.25}$$

where

$$c(s) = (2\pi^2 w)^{-1}(2\pi)^{-2s}\Gamma(s)\Gamma(1-s)^{-1}\zeta(2s)^3\zeta(4s)^{-1}. \tag{6.26}$$

Finally, putting together (6.24) and (6.25) for both s and $1 - s$, we obtain from (6.23) the desired expression for P in the case of $s \neq 1/2$.

7. Estimation of the off-diagonal part. In this section we give an elementary estimate for the off-diagonal part $\mathscr{M}^+$. We shall do so only for $q = 1$ since this is what is needed to complete the proof of Theorem 2. In this case (4.15) and (4.16) become

$$\mathscr{M}^+(u_1, u_2) = \frac{4}{w^2} \sum_{l=1}^{\infty} \sum_{n=1}^{\infty} \mathscr{M}_{ln}(u_1, u_2) \tag{7.1}$$

and

$$M_{ln}(u_1, u_2) = \sum_{\mathfrak{a}}{}' \sum_{m \in \mathbb{Z}} \exp\left(\frac{-\pi}{\operatorname{Im} z_{\mathfrak{a}}}(u_1 l^2 + u_2|m + nz_{\mathfrak{a}}|^2)\right). \tag{7.2}$$

First, summing over m we get

$$\mathcal{M}_{ln}(u_1, u_2) \ll \sum_{\mathfrak{a}}{}' \left(\frac{\sqrt{D}}{2a}\right)^{1/2} \exp\left(-\pi u_1 l^2 \frac{2a}{\sqrt{D}} - \pi u_2 n^2 \frac{\sqrt{D}}{2a}\right)$$

$$\ll \sum_{\mathfrak{a}} \left(\frac{\sqrt{D}}{a}\right)^{1/2} \exp\left(-\pi \frac{2a}{\sqrt{D}} - \pi \frac{\sqrt{D}}{2a} - \pi ln\sqrt{u_1 u_2}\right)$$

$$\ll \sum_{\mathfrak{a}} \exp\left(\frac{-2\pi a}{\sqrt{D}} - \pi ln\sqrt{u_1 u_2}\right) \ll h \exp(-\pi ln\sqrt{u_1 u_2}). \tag{7.3}$$

Next, summing over l, n we get

$$\mathcal{M}^+(u_1, u_2) \ll h \exp(-\pi\sqrt{u_1 u_2}). \tag{7.4}$$

Finally, integrating over u_1, u_2 as in (4.18), we conclude that

$$\mathcal{M}^+ \ll h. \tag{7.5}$$

Inserting (7.5) into (4.22) and using (4.5), we obtain the corresponding estimate for the off-diagonal part of $\mathcal{L}$:

$$\mathcal{L}^+ \ll D^{-1/2} h \ll L(1, \chi_D). \tag{7.6}$$

Then, combining with (6.18), we arrive at

$$\mathcal{L} = l_D(s) + O(L(1, \chi_D)). \tag{7.7}$$

This completes the proof of Theorem 2.

For the proof of Theorem 3 we must refine the estimate (7.5). Our goal is an asymptotic formula for the off-diagonal part $\mathcal{M}^+$ in which the error term saves a fixed positive power of D. To get a result that strong we shall employ the spectral theory of automorphic functions.

8. Spectral expansion of the off-diagonal part. Recall that $\mathcal{M}^+$ is given in (4.18) in terms of $\mathcal{M}^+(u_1, u_2)$, defined in (4.15), which is in turn a sum over $\mathcal{M}_{ln}(u_1, u_2)$, defined by (4.16). In this section we establish a spectral expansion for $\mathcal{M}_{ln}(u_1, u_2)$.

By the correspondence described in Section 2, we arrange the summation over ideals in (4.16) into the orbits of Heegner points (2.4) in the set (2.11) with q replaced by $q_l = q/(l, q)$ with respect to the group $\Gamma = \Gamma_0(q_l)$. We obtain

$$\mathcal{M}_{ln}(u_1, u_2) = \sum_{z \in \Lambda_D(q_l)} |\Gamma_z|^{-1} P_{ln}(z), \tag{8.1}$$

where

$$P_{ln}(z) = \sum_{m \in \mathbb{Z}} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \exp\left(\frac{-\pi}{\operatorname{Im} \gamma z}(u_1 q^{-1} l^2 + u_2 |m + n\gamma z|^2)\right). \tag{8.2}$$

Clearly $P_{ln}(z)$ is an automorphic function bounded on H (in fact, it has exponential decay at all cusps), so it is square-integrable on the fundamental domain of Γ. Therefore, the spectral theorem [Iw] gives

$$P_{ln}(z) = \sum_j < P_{ln}, \quad u_j > u_j(z) + \sum_\kappa \frac{1}{4\pi} \int_{-\infty}^{\infty} < P_{ln}, \quad E_\kappa > E_\kappa\left(z, \frac{1}{2} + it\right) dt,$$

where $\{u_j(z)\}$, $j \geqslant 0$, is an orthonormal system of cusp forms together with the constant function $u_0(z) = V^{-1/2}$ for the eigenvalue $\lambda_0 = 0$, and $E_\kappa(z, v)$ is the Eisenstein series associated with cusp κ for the group $\Gamma = \Gamma_0(q_l)$.

Summing over the Heegner points, we get

$$\mathscr{M}_{ln}(u_1, u_2) = \sum_j < P_{ln}, \quad u_j > W_j + \sum_\kappa \frac{1}{4\pi} \int_{-\infty}^{\infty} < P_{ln}, \quad E_\kappa > W_\kappa\left(\frac{1}{2} + it\right) dt, \tag{8.3}$$

where

$$W_j = \sum_{z \in \Lambda_D(q_l)} |\Gamma_z|^{-1} u_j(z) \tag{8.4}$$

and

$$W_\kappa(v) = \sum_{z \in \Lambda_D(q_l)} |\Gamma_z|^{-1} E_\kappa(z, v). \tag{8.5}$$

The sums W_j and $W_\kappa(v)$ are called Weyl sums for the Heegner points of level q_l.

The main term in the asymptotic formula for $\mathscr{M}_{ln}(u_1, u_2)$ will come from the projection onto the constant eigenfunction. This indicates that the Maass forms are natural harmonics for the study of sums over ideals of a quadratic field. For further evidence of this, see the evaluation of the Weyl sums.

9. Evaluation of the inner products. Define the linear operator U_n, acting on automorphic functions by

$$U_n Q(z) = \sum_{\nu \pmod{n}} Q\left(z + \frac{\nu}{n}\right). \tag{9.1}$$

By (8.2) we may write

$$P_{ln}(z) = \sum_{\gamma \in \Gamma} \sum_{\nu \pmod{n}} \exp\left(-\left(\alpha + \beta\left|\frac{\nu}{n} + \gamma z\right|^2\right)\Big/\operatorname{Im} \gamma z\right),$$

with $\alpha = \pi u_1 q^{-1} l^2$ and $\beta = \pi u_2 n^2$. Now, suppose $Q(z)$ has at most polynomial growth. Then, by the unfolding method we get

$$\langle P_{ln}, Q \rangle = \int_H \exp(-(\alpha + \beta |z|^2) y^{-1}) U_n \bar{Q}(z) d\mu z. \tag{9.2}$$

Next, suppose $Q(z)$ is an eigenform with eigenvalue $\lambda = 1/4 - \nu^2$, so that it has a Fourier expansion of the type

$$Q(z) = y^{1/2}\left(Ay^\nu + By^{-\nu} + \sum_{k \neq 0} \rho(k) K_\nu(2\pi|k|y) e(kx)\right).$$

Then $n^{-1} U_n Q(z)$ has a Fourier expansion as above but supported on $k \equiv 0 \pmod n$. We insert the Fourier series for $U_n Q(z)$ into (9.2), interchange the summation with the integration, and compute the resulting integrals for each $k \equiv 0 \pmod n$ separately.

We shall use the following formulas (see [GR, 3.462.2, 3.471.9, 6.653.2]);

$$\int_{-\infty}^{\infty} \exp(-ax^2 + 2bx)\, dx = \sqrt{\pi/a}\, \exp(b^2/a), \tag{9.3}$$

$$\int_0^\infty \exp(-ay^{-1} - by) y^{\nu-1}\, dy = 2(a/b)^{\nu/2} K_\nu(2\sqrt{ab}), \tag{9.4}$$

$$\int_0^\infty \exp\left(-\frac{ab}{2y} - \frac{a^2+b^2}{2ab} y\right) K_\nu(y) y^{-1}\, dy = 2K_\nu(a) K_\nu(b). \tag{9.5}$$

First, we integrate in (9.2) horizontally, using (9.3), to get

$$\int_{-\infty}^{\infty} \exp(-\beta y^{-1} x^2) e(kx)\, dx = \left(\frac{\pi y}{\beta}\right)^{1/2} \exp\left(\frac{-\pi^2 k^2 y}{\beta}\right).$$

Then we integrate vertically to get

$$\int_0^\infty \exp\left(-\frac{\alpha}{y} - \beta y\right)(Ay^\nu + By^{-\nu}) y^{-1}\, dy = 2\left(A\left(\frac{\alpha}{\beta}\right)^{\nu/2} + B\left(\frac{\alpha}{\beta}\right)^{-\nu/2}\right) K_\nu(2\sqrt{\alpha\beta}),$$

by (9.4), and

$$\int_0^\infty \exp\left(-\frac{\alpha}{y} - \left(\beta + \frac{\pi^2 k^2}{\beta}\right) y\right) K_\nu(2\pi|k|y) y^{-1}\, dy = 2K_\nu(2\sqrt{\alpha\beta}) K_\nu\left(2\pi|k|\sqrt{\frac{\alpha}{\beta}}\right)$$

for $k \neq 0$, by (9.5). Collecting these results, we conclude that

$$\langle P_{ln}, Q\rangle = 2\left(\frac{\pi}{\sqrt{\alpha\beta}}\right)^{1/2} K_\nu(2\sqrt{\alpha\beta})U_n Q\left(i\sqrt{\frac{\alpha}{\beta}}\right), \tag{9.6}$$

where we recall that $\alpha = \pi u_1 q^{-1} l^2$ and $\beta = \pi u_2 n^2$.

For the constant function $Q(z) = u_0(z) = V^{-1/2}$ of zero eigenvalue we have $\nu = 1/2$ and

$$K_{1/2}(y) = \left(\frac{\pi}{2y}\right)^{1/2} e^{-y},$$

whence

$$\langle P_{ln}, u_0\rangle = l^{-1}\left(\frac{q}{u_1 u_2 V}\right)^{1/2} \exp\left(-2\pi ln\sqrt{\frac{u_1 u_2}{q}}\right). \tag{9.7}$$

The other spectral eigenforms (cusp forms and the Eisenstein series) with eigenvalue $\lambda = 1/4 - \nu^2 \geqslant 1/4$ satisfy

$$Q(z) \ll \lambda(y^{1/2} + y^{-1/2}), \tag{9.8}$$

where the implied constant is absolute (see Lemma 14.1). Hence

$$\langle P_{ln}, Q\rangle \ll \lambda n(\alpha^{-1/2} + \beta^{-1/2})|K_\nu(2\sqrt{\alpha\beta})|.$$

From the trivial bound

$$K_\nu(y) \ll \lambda^{-A} y^{-1} e^{-y/2},$$

valid for A any positive constant, we conclude that

$$\langle P_{ln}, Q\rangle \ll \lambda^{-A} q \exp\left(-\pi ln\sqrt{\frac{u_1 u_2}{q}}\right). \tag{9.9}$$

10. The Weyl sums for Eisenstein series. For notational simplicity, we shall carry out the computations for the level q since the case of $q_l = q/(l, q)$ is the same.

The Weyl sum $W_\kappa(v)$ of an Eisenstein series $E_\kappa(z, v)$ for the group $\Gamma = \Gamma_0(q)$ over the Heegner points $z \in \Lambda_D(q)$ will be expressed in terms of the L-functions of the field $K = \mathbb{Q}(\sqrt{-D})$,

$$L_K(s, \psi \circ N) = \sum_{\mathfrak{a}} \psi(N\mathfrak{a})(N\mathfrak{a})^{-v}, \tag{10.1}$$

where ψ is a Dirichlet character to modulus $d|q$.

First we consider, for $\operatorname{Re} v > 1$, the Eisenstein series given by

$$E_\kappa(z, v) = \sum_{\gamma \in \Gamma_\kappa \backslash \Gamma} (\operatorname{Im} \sigma_\kappa^{-1} \gamma z)^v. \tag{10.2}$$

Here Γ_κ is the stability group of the cusp κ, and σ_κ is a scaling matrix which has the following properties:

$$\sigma_\kappa \infty = \kappa, \qquad \sigma_\kappa^{-1} \Gamma_\kappa \sigma_\kappa = \Gamma_\infty. \tag{10.3}$$

The scaling matrix σ_κ is determined by the cusp κ up to a translation on the right side, and the Eisenstein series does not depend on the choice of σ_κ, nor on the choice of κ within its equivalence class. Every cusp for the group $\Gamma_0(q)$ is equivalent with exactly one rational number of the type

$$\kappa = \frac{t}{s} \quad \text{with } (t, s) = 1, \tag{10.4}$$

where $q = rs$ and t is determined modulo $d = (r, s)$. Thus, the total number of inequivalent cusps is given by

$$m(q) = \sum_{rs=q} \varphi(d), \quad \text{where } d = (r, s). \tag{10.5}$$

For κ of this type, we may take the scaling matrix

$$\sigma_\kappa = \begin{pmatrix} \kappa\sqrt{\gamma} & -1/\sqrt{\gamma} \\ \sqrt{\gamma} & 0 \end{pmatrix}, \tag{10.6}$$

where $\gamma = qsd^{-1}$, whence

$$\sigma_\kappa^{-1} = \begin{pmatrix} 0 & 1/\sqrt{\gamma} \\ -\sqrt{\gamma} & \kappa\sqrt{\gamma} \end{pmatrix}. \tag{10.7}$$

These facts can be found in Section 2 of [DI].

Now, by the correspondence described in Section 2 of this paper, we infer by unfolding that

$$W_\kappa(v) = \sum_{z \in \Lambda_D(q)} |\Gamma_z|^{-1} E_\kappa(z, v) = \sum_{\mathfrak{a} \equiv 0(q)}'' (\operatorname{Im} \sigma_\kappa^{-1} z_\mathfrak{a})^v. \tag{10.8}$$

In the last sum $\mathfrak{a}$ ranges over the primitive ideals of norm $N\mathfrak{a} = a$ divisible by q, and the double prime means that the points $\sigma_\kappa^{-1} z_\mathfrak{a}$ are identified if they differ by an integral translation.

Recall that

$$z_{\mathfrak{a}} = \frac{b + i\sqrt{D}}{2a}, \qquad b^2 + D = 4ac. \tag{10.9}$$

The linear fractional transformation (10.7) maps the Heegner point $z_{\mathfrak{a}}$ to another one of the form

$$\sigma_{\kappa}^{-1} z_{\mathfrak{a}} = \frac{B + i\sqrt{D}}{2Cq/d}, \qquad B^2 + D = 4AC \tag{10.10}$$

(in the denominator we have extracted the factor q/d to simplify the forthcoming notation), where

$$\begin{aligned} A &= as^{-1} \\ B &= 2ats^{-1} - b \\ C &= at^2 s^{-1} - bt + cs. \end{aligned} \tag{10.11}$$

Conversely, a point (10.10) corresponds to (10.9) with

$$\begin{aligned} a &= As \\ b &= 2At - B \\ c &= (At^2 - Bt + C)s^{-1}. \end{aligned} \tag{10.12}$$

The condition that a, b, c are integers with $a \equiv 0 \pmod{q}$ is equivalent to the condition that A, B, C are integers such that $A \equiv 0 \pmod{r}$ and

$$At^2 - Bt + C \equiv 0 \pmod{s}. \tag{10.13}$$

Substituting $A = (B^2 + D)/4C$, these conditions become

$$\begin{aligned} B^2 + D &\equiv 0 \pmod{4Cr} \\ (B - 2C\bar{t})^2 + D &\equiv 0 \pmod{4Cs}, \end{aligned} \tag{10.14}$$

where $\bar{t}$ is a fixed integer such that $t\bar{t} \equiv 1 \pmod{s}$. Since the points (10.10) are identified in the sum (10.8) by integral translations, it means that B ranges modulo $2Cq/d$. Therefore, inserting (10.10) into (10.8) we obtain

$$W_{\kappa}(v) = \left(\frac{d\sqrt{D}}{2q}\right)^{v} \sum_{C} \eta(C) C^{-v}, \tag{10.15}$$

where $\eta(C)$ denotes the number of solutions to the system of congruences (10.14) in $B \pmod{2Cq/d}$.

Next we evaluate $\eta(C)$. To achieve symmetry we change B into $B + C\bar{t}$, getting a new system

$$\begin{aligned} (B + C\bar{t})^2 + D &\equiv 0 \pmod{4Cr} \\ (B - C\bar{t})^2 + D &\equiv 0 \pmod{4Cs}, \end{aligned} \tag{10.16}$$

which has the same number of solutions as (10.14). Subtracting, we infer that $B \equiv 0 \pmod{d}$. Changing B into Bd, we get

$$\begin{aligned} (Bd + C\bar{t})^2 + D &\equiv 0 \pmod{4Cr} \\ (Bd - C\bar{t})^2 + D &\equiv 0 \pmod{4Cs}, \end{aligned} \tag{10.17}$$

where now B ranges modulo $2Cqd^{-2}$ (note that $d^2 | q$). If there is any solution in B, then C must satisfy the following condition

$$C^2 + Dt^2 \equiv 0 \pmod{d}, \tag{10.18}$$

which we henceforth assume to hold true. In particular, this condition implies that $(C, d) = 1$.

We continue to modify the system (10.17) with the intention of splitting it into independent congruences. To this end we make two variables $X \pmod{2C}$ and $Y \pmod{qd^{-2}}$ out of B by writing $B = X + 2CY$. In these variables the system (10.17) is equivalent to the following three congruences:

$$(Xd \pm C\bar{t})^2 + D \equiv 0 \pmod{4Cd} \tag{10.19}$$

$$CdY^2 + (Xd + C\bar{t})Y + A^+ \equiv 0 \pmod{r/d} \tag{10.20}$$

$$CdY^2 + (Xd - C\bar{t})Y + A^- \equiv 0 \pmod{s/d}, \tag{10.21}$$

where $A^\pm = [(Xd \pm C\bar{t})^2 + D]\,(4Cd)^{-1}$. Given any X satisfying (10.19), we first count the solutions in Y of (10.20)–(10.21). Notice that the moduli r/d and s/d are coprime so that we can count separately. In fact, by the Chinese remainder theorem, this reduces to counting the solutions to prime moduli p. If $p | Cd$, there is one solution, and if $p \nmid Cd$, there are two solutions because the congruence

$$x^2 + D \equiv 0 \pmod{p}$$

has two solutions (recall that $p | q$, so p has degree 1 in the field $K = \mathbb{Q}(\sqrt{-D})$). From this local consideration we infer that the number of $Y \pmod{qd^{-2}}$ satisfying (10.20)–(10.21) is equal to 2^n, where n is the number of prime factors of qd^{-2} which do not divide Cd, so

$$n = \omega(q) - \omega(d) - \omega((q, C)). \tag{10.22}$$

This number does not depend on X. Since $(C, d) = 1$ the number of X (mod $2C$) which satisfy (10.19) is equal to the number of solutions to

$$X^2 + D \equiv 0 \pmod{4C} \tag{10.23}$$

by a change of variables, and the latter is equal to $\nu(C)$, the number of primitive ideals of norm C. More precisely, every solution to (10.23) corresponds to the ideal

$$\mathfrak{c} = \left[C, \frac{X + i\sqrt{D}}{2}\right].$$

Therefore

$$\eta(C) = 2^{\omega(q)-\omega(d)-\omega((q,C))}\nu(C) \tag{10.24}$$

if C satisfies (10.18), or else $\nu(C)$ vanishes.

Inserting (10.24) into (10.25), we obtain

$$W_\kappa(v) = \left(\frac{d\sqrt{D}}{2q}\right)^v 2^{\omega(q)-\omega(d)} \sideset{}{'}\sum_{C^2+Dt^2\equiv 0(d)} 2^{-\omega((q,C))}C^{-v}, \tag{10.25}$$

where the summation ranges over the primitive ideals whose norm satisfies the congruence (10.18). Observe that if q is squarefree, then $d = (r, s) = 1$, so that all the Weyl sums are equal.

To remove the congruence (10.18) we employ the Dirichlet characters ψ (mod d). Put

$$\rho_D(\psi) = 2^{-\omega(d)} \sum_{\delta^2+D\equiv 0(d)} \psi(\delta). \tag{10.26}$$

Then

$$\frac{1}{\varphi(d)} \sum_{\psi (\mathrm{mod}\, d)} \rho_D(\bar\psi)\bar\psi(t)\psi(C) = 2^{-\omega(d)} \tag{10.27}$$

if C satisfies (10.18), or else the sum vanishes. Hence, by (10.25) we obtain

$$W_\kappa(v) = \left(\frac{d\sqrt{D}}{2q}\right)^v \frac{2^{\omega(q)}}{\varphi(d)} \sum_{\psi (\mathrm{mod}\, d)} \bar\psi(t)\rho_D(\bar\psi)G(v, \psi), \tag{10.28}$$

where

$$\begin{aligned}G(v, \psi) &= \sum_{c}{}' 2^{-\omega((c,q))}\psi(c)c^{-v} \\ &= \Big(\sum_{(c,q)=1}{}' \psi(c)c^{-v}\Big)\Big(\sum_{c|q^\infty} \psi(c)c^{-v}\Big) \\ &= \frac{w}{2}\frac{L_K(v, \psi \circ N)}{L(2v, \psi)}\Big(\sum_{c|q^\infty} 2^{\omega(c)}\psi(c)c^{-v}\Big)^{-1}\Big(\sum_{c|q^\infty} \psi(c)c^{-v}\Big) \\ &= \frac{w}{2}\frac{L_K(v, \psi \circ N)}{L(2v, \psi)}\prod_{p|q}(1 + 2\psi(p)p^{-v}(1 - p^{-v})^{-1})^{-1}(1 - \psi(p)p^{-v})^{-1} \\ &= \frac{w}{2}\frac{L_K(v, \psi \circ N)}{L(2v, \psi)}\prod_{p|q}(1 + \psi(p)p^{-v})^{-1}. \end{aligned} \tag{10.29}$$

Inserting (10.29) into (10.28) we arrive at the following formula:

$$W_\kappa(v) = \frac{w}{2}\left(\frac{d\sqrt{D}}{2q}\right)^v \frac{2^{\omega(q)}}{\varphi(d)} \sum_{\psi(\mathrm{mod}\, d)} \bar{\psi}(t)\rho_D(\bar{\psi})\frac{L_K(v, \psi \circ N)}{L(2v, \psi)}\prod_{p|q}(1 + \psi(p)p^{-v})^{-1}. \tag{10.30}$$

This formula was derived for $\mathrm{Re}\, v > 1$, but it is valid everywhere by analytic continuation.

For v on the line $\mathrm{Re}\, v = 1/2$ we obtain by Burgess's estimate that

$$W_\kappa(v) \ll v(1 - v)D^{7/16+\varepsilon}, \tag{10.31}$$

where the implied constant depends on ε only.

For the cusp $\kappa = \infty$ the formula (10.30) simplifies into

$$W_\infty(v) = \frac{w}{2}\left(\frac{\sqrt{D}}{2q}\right)^v \frac{\zeta_K(v)}{\zeta(2v)} 2^{\omega(q)} \prod_{p|q}(1 + p^{-v})^{-1}. \tag{10.32}$$

Hence, by (1.4),

$$\operatorname*{res}_{v=1} W_\infty(v) = 2^{\omega(q)}hV^{-1}, \tag{10.33}$$

where

$$V = \frac{\pi}{3} q \prod_{p|q}\left(1 + \frac{1}{p}\right) \tag{10.34}$$

is the volume of the fundamental domain of $\Gamma_0(q)$. In fact, the residue of the Weyl sums at $v = 1$ is the same for any cusp because it is true for the Eisenstein series; namely, for any cusp κ we have

$$\operatorname*{res}_{v=1} E_\kappa(z, v) = V^{-1}; \tag{10.35}$$

see Lemma 3.7 of [DI]. Hence, we compute the Weyl sum W_0 for the constant eigenfunction $u_0(z) = V^{-1/2}$, getting

$$W_0 = 2^{\omega(q)} h V^{-1/2}. \tag{10.36}$$

11. The projection on the constant eigenfunction. The additional contribution $l_D^+(s)$ to the main term in (1.28) will come from the off-diagonal part $\mathscr{L}^+$ given by (4.22). This part will get its main term from that of $\mathscr{M}^+$ after subtracting the quantity $h(ws(1-s))^{-2}$ which has emerged from the trivial character. Then, in turn, the main term of $\mathscr{M}^+$ will be derived from the projection on the constant eigenfunction $u_0(z) = V^{-1/2}$ in the spectral expansion (8.3). In this section we shall evaluate this projection asymptotically.

We denote

$$\mathscr{M}_{ln}^{\sharp}(u_1, u_2) = \langle P_{ln}, u_0 \rangle W_0 . \tag{11.1}$$

By (9.7) and (10.36), together with (10.34), with q upgraded to $q_l = q/(l, q)$, we get

$$\mathscr{M}_{ln}^{\sharp}(u_1, u_2) = \frac{3h}{\pi}(u_1 u_2 q)^{-1/2}\frac{(l, q)}{l} 2^{\omega(q_l)} \prod_{p|q_l}\left(1 + \frac{1}{p}\right)^{-1} \exp\left(-2\pi l n \sqrt{\frac{u_1 u_2}{q}}\right). \tag{11.2}$$

Inserting this into (4.15), we get

$$\mathscr{M}^{\sharp}(u_1, u_2) = \frac{12h}{\pi w^2}(u_1 u_2 q)^{-1/2} S_q\left(2\pi \sqrt{\frac{u_1 u_2}{q}}\right), \tag{11.3}$$

where

$$S_q(x) = \sum_{l=1}^{\infty}\sum_{n=1}^{\infty}\frac{(l, q)}{l}\prod_{p|q_l}\left(1 + \frac{1}{p}\right)^{-1} \exp(-lnx).$$

To evaluate $S_q(x)$, we use (6.1), getting

$$S_q(x) = \frac{1}{2\pi i}\int_{(\sigma)} x^{-v}\Gamma(v)\zeta(v)\lambda(v)\, dv \tag{11.4}$$

with $\sigma > 1$, where

$$\begin{aligned}\lambda(v) &= \sum_{l=1}^{\infty} (l, q) l^{-1-v} \prod_{p|q_l} \left(1 + \frac{1}{p}\right)^{-1} \\ &= \prod_{p|q} \left(1 + \frac{1}{p}\right)^{-1} \sum_{l=1}^{\infty} (l, q)^{-1-v} \prod_{p|(l,q),\, p \nmid q/(l,q)} (1 + p^{-1}) \\ &= \prod_{p|q} (1 + p^{-1})^{-1} \prod_{p \nmid q} (1 - p^{-1-v})^{-1} \\ &\quad \times \prod_{p^{\alpha} \| q} \left\{1 + p^{-v} + \cdots + p^{-(\alpha-1)v} + p^{\alpha}\left(1 + \frac{1}{p}\right) \sum_{\beta \geqslant \alpha} p^{-\beta(1+v)}\right\} \\ &= \zeta(1 + v) \prod_{p^{\alpha} \| q} \{(1 + p^{-v} + \cdots + p^{-(\alpha-1)v})(1 + p^{-1})^{-1}(1 - p^{-1-v}) + p^{-\alpha v}\} \\ &= \zeta(1 + v) T_q(v), \qquad (11.5)\end{aligned}$$

say, where $T_q(v)$ is the above finite product. Clearly $T_q(v)$ is an entire function such that

$$T_q(1) = 1 \qquad (11.6)$$

$$|T_q(v)| \leqslant \tau(q) \qquad \text{in } \operatorname{Re} v \geqslant 0. \qquad (11.7)$$

By (11.4) and (11.5) we obtain

$$S_q(x) = \frac{1}{2\pi i} \int_{(\sigma)} x^{-v} \Gamma(v) \zeta(v) \zeta(1 + v) T_q(v)\, dv. \qquad (11.8)$$

Moving the integration to the line $\operatorname{Re} v = \varepsilon$ with $0 < \varepsilon < 1$, we get by (11.6) that

$$S_q(x) = \zeta(2) x^{-1} + S_q^{\flat}(x), \qquad (11.9)$$

where $S_q^{\flat}(x)$ is given by the integral (11.8) on the line $\operatorname{Re} v = \varepsilon$. Accordingly, by (11.3) we split

$$\mathscr{M}^{\sharp}(u_1, u_2) = h w^{-2} (u_1 u_2)^{-1} + \mathscr{M}^{\flat}(u_1, u_2), \qquad (11.10)$$

say, where

$$\mathscr{M}^{\flat}(u_1, u_2) = \frac{12h}{\pi w^2} (u_1 u_2 q)^{-1/2} S_q^{\flat}\left(2\pi \sqrt{\frac{u_1 u_2}{q}}\right). \qquad (11.11)$$

Here, by the trivial bound $S_q^\flat(x) \ll x^{-\varepsilon}\tau(q)$ we have

$$\mathcal{M}^\flat(u_1, u_2) \ll hq^{-1/2+\varepsilon}(u_1u_2)^{-1/2-\varepsilon}, \tag{11.12}$$

where the implied constant depends on ε only.

Next, integrating (11.10) over u_1, u_2 as in (4.18), we obtain

$$\mathcal{M}^\sharp = h(ws(1-s))^{-2} + O(hq^{-1/2+\varepsilon}). \tag{11.13}$$

More precisely,

$$\mathcal{M}^\sharp = h(ws(1-s))^{-2} + \mathcal{M}^\flat, \tag{11.14}$$

say, where

$$\mathcal{M}^\flat = 12h\pi^{-1}w^{-2}q^{-1/2}I_q(s) \tag{11.15}$$

and

$$I_q(s) = \int_1^\infty \int_1^\infty \frac{(u_1^{s-1} + u_1^{-s})(u_2^{s-1} + u_2^{-s})}{(u_1u_2)^{1/2}} S_q^\flat\left(2\pi\sqrt{\frac{u_1u_2}{q}}\right) du_1\, du_2 .$$

By (11.8), with $\sigma = \varepsilon$, we integrate over u_1, u_2, explicitly getting

$$I_q(s) = \frac{4}{2\pi i}\int_{(\varepsilon)} \left(\frac{\sqrt{q}}{2\pi}\right)^v \frac{\Gamma(v)\zeta(v)\zeta(1+v)T_q(v)v^2}{(v+1-2s)^2(v-1+2s)^2}\, dv.$$

For $q = 1$, this simplifies a bit as follows:

$$I(s) = \frac{4}{2\pi i}\int_{(\varepsilon)} \frac{(2\pi)^{-v}\Gamma(v)\zeta(v)\zeta(1+v)v^2}{(v+1-2s)^2(v-1+2s)^2}\, dv. \tag{11.16}$$

Hence, by the duplication formula for the gamma function, we can also write

$$I(s) = \frac{2}{2\pi i}\int_{(\varepsilon)} \frac{\Theta(v/2)\Theta(-v/2)v^2}{(v+1-2s)^2(v-1+2s)^2}\, dv.$$

From (11.13)–(11.16) we find that the main term of $\mathcal{L}^+$ in (4.22) when $q = 1$ is

$$l_D^+(s) = \frac{12}{\pi w}\Gamma(s)\Gamma(1-s)I(s)L(1, \chi_D). \tag{11.17}$$

This is the additional contribution to the main term in (1.28).

12. Evaluation of the off-diagonal part. In this section we evaluate $\mathscr{M}^+$ asymptotically with an error term which saves a power of D. This saving is due to nontrivial estimates for the Weyl sums, which is the main ingredient in this work. For those sums which are associated to the Eisenstein series, we have just established (10.31) and (10.36). We still need similar estimates for the sums W_j associated to the cusp forms $u_j(z)$, $j \geqslant 1$. The following estimate will be established in the next two sections:

$$W_j \ll \lambda_j^6 q^3 D^{13/28+\varepsilon}. \tag{12.1}$$

Here we use this result to complete the proofs of Theorem 3, Proposition 1 and Theorem 4. Multiplying by the estimate (9.9), we get

$$\langle P_{ln}, u_j \rangle W_j \ll \lambda_j^{-A} q^4 \exp\left(-\pi ln \sqrt{\frac{u_1 u_2}{q}}\right) D^{13/28+\varepsilon}, \tag{12.2}$$

where A is any positive constant. We choose $A = 2$ and take the bound

$$\sum_{j=1}^{\infty} \lambda_j^{-2} \ll V, \tag{12.3}$$

which follows by Weyl's law, where V is the volume given by (10.34). Next summing over l, n, we get

$$\sum_{l=1}^{\infty} 2^{-q_l} \sum_{n=1}^{\infty} \exp\left(-\pi ln \sqrt{\frac{u_1 u_2}{q}}\right) \ll \left(\frac{q}{u_1 u_2}\right)^{1/2} \exp\left(-\sqrt{\frac{u_1 u_2}{q}}\right) \log q. \tag{12.4}$$

Then, integrating over u_1, u_2 as in (4.18), we get

$$\int_1^{\infty} \int_1^{\infty} \exp\left(-\sqrt{\frac{u_1 u_2}{q}}\right) \frac{du_1 \, du_2}{u_1 u_2} \ll (\log q)^2. \tag{12.5}$$

Collecting (12.2)–(12.5), we infer that the contribution from the cuspidal spectrum in (8.3) to $\mathscr{M}^+$ is bounded by

$$\mathscr{M}^+(\text{cuspidal}) \ll q^6 D^{13/28+\varepsilon}. \tag{12.6}$$

We deal with the contribution from the continuous spectrum in the same manner. We have (10.31) in place of (12.1) and

$$\sum_{\kappa} \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{dv}{1+v^2} \ll m(q) \tag{12.7}$$

in place of (12.3), where $m(q)$ denotes the number of inequivalent cusps (the multiplicity of the continuous spectrum) given by (10.5). We obtain

$$\mathcal{M}^+(\text{continuous}) \ll qD^{7/16+\varepsilon}. \tag{12.8}$$

Finally, we recall that the contribution $\mathcal{M}^\#$ from the zero eigenvalue has been evaluated in Section 11. Combining these sections we complete the proof of Theorem 3 (here we need only the results for $q = 1$). Also, using (5.2), we complete the proof of Proposition 1 (for any q) and Theorem 4 as its consequence.

13. Weyl sums as Fourier coefficients. Next we will realize the needed Weyl sums for u_j in terms of the Fourier coefficients of an associated Maass cusp form f_j of weight 1/2, which is related to u_j by means of a theta correspondence. Applying a result from [Du], the proof of (12.1) is reduced to an estimate for the L_2-norm of f_j.

For fixed $q \in \mathbb{Z}^+$ and $m = (m_1, m_2, m_3)$, let $F(m) = m_2^2 - 4qm_1m_3$, and define for any $n \neq 0$ the hyperboloid

$$H_n = \{m \in \mathbb{R}^3; F(m) = n \quad \text{and} \quad m_1 > 0 \text{ if } n < 0\}.$$

H_n is acted on by $\tilde{G}$, the connected component of the identity of

$$\{\tilde{g} \in SL_3(\mathbb{R}); F(\tilde{g}m) = F(m)\}.$$

For $n < 0$ we have the bijection $H_n \to H$ defined by

$$m \mapsto z_m = \frac{m_2 + i\sqrt{|n|}}{2qm_1}$$

with inverse map

$$z \mapsto {}^t m_z = \frac{\sqrt{|n|}}{y}\left(\frac{1}{2q}, x, \frac{|z|^2}{2}\right).$$

This induces a homomorphism $G \to \tilde{G}$ given by

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = g \mapsto \tilde{g} = \begin{bmatrix} \delta^2 & \gamma\delta/q & \gamma^2/q \\ 2q\beta\delta & \alpha\delta + \beta\gamma & 2\alpha\gamma \\ q\beta^2 & \alpha\beta & \alpha^2 \end{bmatrix} \tag{13.1}$$

so that $gz_m = z_{\tilde{g}m}$ and $\tilde{g}m_z = m_{gz}$. Its kernel is $\{\pm 1\}$, and it is surjective since $G/\pm 1$ and $\tilde{G}$ are connected and have the same dimension. For $\Gamma = \Gamma_0(q)$, we have from (12.1) that

$$\tilde{\Gamma} \subset \tilde{G} \cap SL(3, \mathbb{Z}). \tag{13.2}$$

Let H^* be the set of all oriented geodesics on H, so

$$H^* = \{C(x_1, x_2): x_1, x_2 \in \mathbb{R} \cup \{\infty\}, x_1 \neq x_2\},$$

where $C(x_1, x_2)$ is the unique directed geodesic from x_1 to x_2. Now H^* is acted upon transitively by $G = SL(2, \mathbb{R})$ by linear fractional maps, and for $n > 0$ we have a bijection $H_n \to H^*$ given by

$$m \mapsto C_m = \begin{cases} C\left(\dfrac{m_2 - \sqrt{n}}{2qm_1}, \dfrac{m_2 + \sqrt{n}}{2qm_1}\right) & m_1 \neq 0 \\ C\left(\dfrac{m_3}{m_2}, \infty\right) & \text{if } m_1 = 0 \text{ and } m_2 > 0 \\ C\left(\infty, \dfrac{m_3}{m_2}\right) & \text{if } m_1 = 0 \text{ and } m_2 < 0, \end{cases}$$

which also respects the above isomorphism $G/\pm 1 \to \tilde{G}$. For $C \in H^*$ let $g_C \in G$ be such that $g_C C(0, \infty) = C$. For

$$A = \left\{\begin{pmatrix} a & \\ & a^{-1} \end{pmatrix}: a \in \mathbb{R}^+\right\},$$

we have $G_C = \{g \in G: gC = C\} = g_C(\pm A)g_C^{-1}$. Let $\Gamma_C = G_C \cap \Gamma$. Then, Γ_C is either $\{\pm 1\}$ or $\left\{\pm\begin{pmatrix} \varepsilon & \\ & \varepsilon^{-1} \end{pmatrix}^j : j \in \mathbb{Z}\right\}$ for some $\varepsilon > 1$. Define $\int_{\Gamma \backslash C} u_j(z)$ to be $\int_0^\infty u_j(g_C\, y^i)\, dy/y$ in the first case and $\int_1^{\varepsilon^2} u_j(g_C\, y^i)\, dy/y$ in the second, observing that they are independent of the choice of g_C.

We now recall Siegel's theta function for F. It follows from (13.5) that $({}^t g)^\sim = u^{-1}\, {}^t(\tilde{g})u$ where $u = \operatorname{diag}(2q^2, 1, 2)$ and, hence, that for $F_+(m) = 2q^2m_1^2 + m_2^2 + 2m_3^2$ and $K = SO(2)$ we have

$$\tilde{K} = \{\tilde{g} \in \tilde{G}: F_+(\tilde{g}m) = F_+(m)\}.$$

In this way we may identify H with the set of positive forms

$$\{F_+(\tilde{g}^{-1}m): g \in G\},$$

which is Siegel's representation space for $\tilde{G}$. Define now for $z = x + iy \in H$ and $g \in G$

$$\Theta(z, g) = \sum_{m \in \mathbb{Z}^3} e(xF(m) + iyF_+(\tilde{g}^{-1}m))$$

which, in view of (13.2) and the above, satisfies for $\sigma \in \Gamma$ and $k \in K$

$$\Theta(z, \sigma g k) = \Theta(z, g). \tag{13.3}$$

Let u_j be a weight-0 Maass cusp form for $\Gamma_0(q)$ made into a K-invariant function on G. By (13.3) we may define

$$f_j(z) = y^{3/4} \int_{\Gamma\backslash G} \overline{\Theta}(z, g) u_j(g)\, dg \tag{13.4}$$

for a bi-invariant Haar measure dg. It follows from Theorem 4 in [Du] that f_j is a Maass cusp from of weight 1/2 and discriminant -4 for $\Gamma_0(4q)$ with eigenvalue $(1/4) + (t_j/2)^2$ if $(1/4) + t_j^2$ is the eigenvalue of u_j (see [Du] for definitions). Thus $f_j(z)$ has a Fourier expansion at ∞ of the form

$$f_j(z) = \sum_{n \neq 0} \rho_j(n) W_{\kappa, v_j}(4\pi |n| y) e(nx), \tag{13.5}$$

where $\kappa = n/4|n|$, $v_j = it_j/2$ and $W_{\kappa,v}(y)$ is the standard Whittaker function. In the case of anisotropic ternary forms, Maass [Ma] determined these Fourier coefficients; for cusp forms u_j his result also holds for our F. A very clear treatment was given recently by Katok and Sarnak [KS] in the case $q = 1$. Their proof, with straightforward modifications to handle general q, yields the following result.

LEMMA 13.1. *For $n < 0$ we have*

$$\rho_j(n) = c_0 |n|^{-3/4} \sum_z |\Gamma_z|^{-1} u_j(z),$$

where $c_0 \neq 0$ is an absolute constant and z ranges over the points

$$z = \frac{b + i\sqrt{|n|}}{2a}, \quad b^2 - 4ac = n,\ a, c > 0,\ a \equiv 0 \pmod q$$

in a fundamental domain of $\Gamma_0(q)$. For $n > 0$ we have

$$\rho_j(n) = c_1 |n|^{-3/4} \sum_C \int_{\Gamma\backslash C} u_j(z),$$

where $c_1 \neq 0$ is an absolute constant and $C = C_{(a/q,b,c)}$ ranges over Γ-inequivalent solutions to $b^2 - 4ac = n$, $a \equiv 0 \pmod q$.

This lemma reduces the estimation of the Weyl sums to that of the Fourier coefficient ρ_j. To bound the latter, we appeal to Theorem 5 in [Du] which, together with Lemma 13.1, proves the following proposition.

PROPOSITION 2. *Let $f_j(z)$ be defined in (13.4). Then for any $D > 0$ such that $-D$ is a fundamental discriminant we have*

$$W_j \ll (1 + |t_j|^5) \operatorname{ch}\left(\frac{\pi}{4} t_j\right) \|f_j\| D^{13/28+\varepsilon},$$

where $\|f_j\|$ is the L_2-norm of f_j and the implied constant depends on ε only.

This bound, in turn, reduces the problem to bounding the L_2-norm of f_j. It is remarked that, in case $q = 1$, Lemma 13.1 has been substantially refined in [KS] to give the exact relation with the Shimura lift. In this case we could use their result to complete the estimation of W_j.

14. Estimates for cusp forms. Our objective is to prove the following proposition which, together with Proposition 2, gives (12.1).

PROPOSITION 3. *For the cusp form f_j as in Proposition 2, we have*

$$\mathrm{ch}\left(\frac{\pi}{4}t_j\right)\|f_j\| \ll (\lambda_j q)^{3+\varepsilon}.$$

By definition, f_j is given in (13.4) as an inner product of u_j against the theta-function $\Theta(z, g)$. However, this formula appears to be very difficult to use directly for our purpose. Instead, we shall employ Lemma 13.1 which gives the Fourier coefficients of f_j as the Weyl sums for u_j. This will lead us to two other problems.

The first problem is to estimate $u_j(z)$ on the upper half-plane uniformly with respect to the eigenvalue λ_j and the level q. There are many results of this type in the literature. By (8.3) of [Iw], we get the following estimate.

LEMMA 14.1. *Let $u_j(z)$ be a cusp form for $\Gamma_0(q)$ of weight* 0 *and eigenvalue $\lambda_j = (1/4) + t_j^2$ normalized by $\|u_j\| = 1$. Then*

$$u_j(z) \ll \lambda_j^{1/4}(y + y^{-1})^{-1/2}, \tag{14.1}$$

where the implied constant is absolute.

We next apply this lemma together with Lemma 13.1 to estimate $\rho_j(n)$. In case $n < 0$, we estimate the number summands by ignoring the congruence $a \equiv 0 \pmod{q}$ and consider these points modulo the group $\Gamma_0(1)$. There are $h(n)$ of them in a fundamental domain of $\Gamma_0(1)$, where $h(n)$ is the class number of positive definite quadratic forms of discriminant n. Hence, the total number of points in a fundamental domain of $\Gamma_0(q)$ is bounded by $[\Gamma_0(1):\Gamma_0(q)]h(n)$. We have

$$\mu_q = [\Gamma_0(1):\Gamma_0(q)] = q\prod_{p|q}\left(1 + \frac{1}{p}\right) \ll q\tau(q) \tag{14.2}$$

and

$$h(n) \ll \tau(|n|)|n|^{1/2}\log(2|n|). \tag{14.3}$$

Using (14.1) and Lemma 13.1, we derive that

$$\rho_j(n) \ll \lambda_j\mu_q h(n)|n|^{-3/4}, \tag{14.4}$$

where the implied constant is absolute. The same bound can be derived for $n > 0$ in a similar fashion.

Next, having estimates for the Fourier coefficients of $f_j(z)$ in the cusp ∞, we wish to derive a bound for the L_2-norm. Clearly, the Fourier expansion in one cusp determines the form completely; nevertheless, it does not show immediately a fast decay in all cuspidal zones. In particular, if one is seeking good uniformity with respect to the group, there does not seem to be much in the literature. Some results are given in [IS] and [DI], but these do not cover our case of the half-integral weight cusp forms.

We shall consider the problem in some generality. Let $f(z)$ be a cusp form for $\Gamma = \Gamma_0(q)$ of weight $0 \leqslant k \leqslant 2$ and eigenvalue $\lambda = (1/4) - \nu^2$. Therefore $f(z)$ transforms by

$$f(gz) = \varepsilon(g)\left(\frac{cz + d}{|cz + d|}\right)^k f(z)$$

for any $g = \begin{pmatrix} ab \\ cd \end{pmatrix} \in \Gamma$, where $\varepsilon(g)$ is a multiplier with $|\varepsilon(g)| = 1$. Moreover, f satisfies the equation

$$(\Delta_k + \lambda)f = 0,$$

where

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

To simplify a bit, we assume that the multiplier is singular in every cusp. This implies that the Fourier expansion of f in any cusp runs over the nonzero integral frequencies. In particular, the Fourier expansion in the cusp ∞ is of the type

$$f(z) = \sum_{n \neq 0} \rho(n) W(4\pi |n| y) e(nx), \tag{14.5}$$

where $W(y) = W_{kn/2|n|, \nu}(y)$ is the Whittaker function. Since the spectrum of Δ_k satisfies $\lambda \geqslant (k/2)(1 - (k/2))$, we have $0 \leqslant \operatorname{Re} \nu \leqslant |(k - 1)/2| \leqslant 1/2$.

The Whittaker function $W_{\kappa,\nu}(y)$ is the solution to the equation

$$W'' + \left(\frac{\lambda}{y^2} + \frac{\kappa}{y} - \frac{1}{4}\right)W = 0,$$

which has exponential decay as $y \to \infty$; namely we have

$$W_{\kappa,\nu}(y) = y^\kappa e^{-y/2}\{1 + O((\kappa^2 + |\nu|^2 + 1)y^{-1})\}, \tag{14.6}$$

where the implied constant is absolute (see [MO]). Notice that $W_{\kappa,\nu}(y) = W_{\kappa,-\nu}(y)$ is real. By (14.6) we infer that

$$\int_Y^\infty W(y)^2 y^{-2}\,dy \leqslant 2\int_Y^{\Delta+Y} W(y)^2 y^{-2}\,dy \tag{14.7}$$

for any $Y > 0$, where $\Delta \asymp \lambda$. We shall also need the bound

$$\int_0^\infty W(y)^2 y^{\sigma-1}\,dy \ll \lambda^{1+\sigma} e^{-\pi|\nu|}, \tag{14.8}$$

which is valid for any $\sigma \geqslant 1/2$, the implied constant depending on σ. One can prove (14.8) by using the Fourier integral representation (cf. [GR, p. 321])

$$W_{\kappa,\nu}(4\pi y) = \pi^{-1}(\pi y)^{1/2-\nu}\Gamma\left(\nu + \frac{\kappa+1}{2}\right)\int_{-\infty}^{\infty}(1-ix)^{-\kappa}(1+x^2)^{(\kappa-1)/2-\nu}e(-xy)\,dx.$$

By the Plancherel theorem this gives

$$\int_0^\infty W(y)^2 y^{\sigma-1}\,dy = (4\pi)^{-3/2}4^{\sigma}\frac{\Gamma\left(\sigma+\frac{1}{2}\right)}{\Gamma(\sigma+1)}\left|\Gamma\left(\nu+\frac{\kappa+1}{2}\right)\right|^2, \tag{14.9}$$

where $\sigma = 2\,\mathrm{Re}\,\nu$. Hence, applying Stirling's formula and (14.6), one derives (14.8) for any $\sigma \geqslant 1/2$.

Having collected the above information about f, we are now ready to estimate the L_2-norm of f in terms of the Fourier coefficients $\rho(n)$.

LEMMA 14.2. *Let $f(z)$ be a cusp form for $\Gamma_0(q)$ of weight $0 \leqslant k \leqslant 2$ and eigenvalue $\lambda = (1/4) - \nu^2$ whose Fourier expansion in the cusp ∞ is given by (14.5). Then the L_2-norm of f is bounded by*

$$\|f\| \ll \tau(q)(\lambda q)^{\sigma+3/2}e^{-(\pi/2)|\nu|}\left(\sum_{n\neq 0}|\rho(n)|^2|n|^{-\sigma}\right)^{1/2},$$

where σ is any number $\geqslant 1/2$ such that the series converges. The implied constant depends on σ only.

Proof. Let $\mathcal{D}$ be the standard fundamental polygon for the modular group,

$$\mathcal{D} = \left\{z = x+iy : |x| \leqslant \frac{1}{2}, |z| \geqslant 1\right\}. \tag{14.10}$$

Then the union of $\sigma\mathcal{D}$, where σ runs over coset representatives of $\Gamma_0(q) \subset \Gamma_0(1)$, is a fundamental domain for $\Gamma_0(q)$. Accordingly, the L_2-norm of f splits into

$$\|f\|^2 = \sum_{\sigma} \int_{\mathscr{D}} |f(\sigma z)|^2 \, d\mu \, z. \tag{14.11}$$

For each σ we consider the Fourier expansion in the cusp $\sigma\infty$ (some of the cusps $\sigma\infty$ are equivalent),

$$f(\sigma z) = \varepsilon_\sigma \sum_{n \neq 0} \rho_\sigma(n) W(4\pi |n| y/\omega_\sigma) e(nx/\omega_\sigma),$$

where $|\varepsilon_\sigma| = 1$ and ω_σ is the width of the cusp, $1 \leqslant \omega_\sigma \leqslant q$. Hence,

$$\begin{aligned}
\int_{\mathscr{D}} |f(\sigma z)|^2 \, d\mu \, z &\leqslant \int_{\sqrt{3}/2}^{\infty} \int_0^{\omega_\sigma} |f(\sigma z)|^2 \, d\mu \, z \\
&= \omega_\sigma \int_{\sqrt{3}/2}^{\infty} \sum_{n \neq 0} |\rho_\sigma(n) W(4\pi |n| y/\omega_\sigma)|^2 y^{-2} \, dy \\
&\leqslant 2\omega_\sigma \int_{\sqrt{3}/2}^{\Delta_\sigma} \sum_{n \neq 0} |\rho_\sigma(n) W(4\pi |n| y/\omega_\sigma)|^2 y^{-2} \, dy \\
&= 2 \int_{\sqrt{3}/2}^{\Delta_\sigma} \int_{-\omega_\sigma/2}^{\omega_\sigma/2} |f(\sigma z)|^2 \, d\mu \, z,
\end{aligned} \tag{14.12}$$

where $\Delta_\sigma = \sqrt{3}/2 + \omega_\sigma \Delta/4\pi$. We can choose the coset representatives to satisfy

$$\sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \gamma | q, \quad |\delta| \leqslant \gamma/2. \tag{14.13}$$

For σ of this form, the cusp $\sigma\infty = \alpha/\gamma$ has width $\omega_\sigma = q/(\gamma^2, q) \leqslant q/\gamma$, so $\Delta_\sigma \leqslant \Delta q/\gamma$. If z is in the range of (14.11), then

$$\operatorname{Im} \sigma z = y((\gamma x + \delta)^2 + \gamma^2 y^2)^{-1} \geqslant 4y((q + \gamma)^2 + \gamma^2 y^2)^{-1} \geqslant \Delta^{-1} q^{-2}$$

and

$$\left| \operatorname{Re} \sigma z - \frac{\alpha}{\gamma} \right| = \gamma^{-2} |x + \delta/\gamma| ((x + \delta/\gamma)^2 + y^2)^{-1} \leqslant \frac{1}{2\gamma^2 y} < 1.$$

Therefore, changing the variable z into $\sigma^{-1} z$ in (14.12), we get

$$\int_{\mathscr{D}} |f(\sigma z)|^2 \, d\mu \, z \leqslant 4 \int_{\Delta^{-1} q^{-2}}^{\infty} \int_0^1 |f(z)|^2 \, d\mu \, z. \tag{14.14}$$

Inserting (14.14) into (14.11), we obtain

$$\|f\|^2 \leqslant 4\mu_q \int_{\Delta^{-1}q^{-2}}^{\infty} \int_0^1 |f(z)|^2 \, d\mu \, z, \tag{14.15}$$

where μ_q is the index given by (14.2). Then, inserting the Fourier expansion (14.5) into (14.15), we get

$$\begin{aligned} \|f\|^2 &\leqslant 4\mu_q \sum_{n\neq 0} |\rho(n)|^2 \int_{\Delta^{-1}q^{-2}}^{\infty} W(4\pi|n|y)^2 y^{-2} \, dy \\ &\leqslant 4\mu_q(\Delta q^2)^{\sigma+1}\left(\sum_{n\neq 0} |\rho(n)|^2 |n|^{-\sigma}\right) \int_0^{\infty} W(4\pi y)^2 y^{\sigma-1} \, dy, \end{aligned}$$

where σ is any number $\geqslant -1$ such that the series and the integral above converge. Applying (14.9), we complete the proof of Lemma 14.2.

In the case of our cusp form $f_j(z)$ of weight $k = 1/2$, we infer from (14.3) and (14.4) that Lemma 14.2 is applicable with $\sigma = (1/2) + \varepsilon$ proving Proposition 3.

15. Split primes. In this section we look for the primes p which split in the field $K = \mathbb{Q}(\sqrt{-D})$, i.e., those with

$$\chi_D(p) = \left(\frac{-D}{p}\right) = 1. \tag{15.1}$$

We need a lot of these to be small. They will be the building blocks used in constructing the amplifier (1.34). Our agruments take ideas from sieve methods and are mostly elementary.

Let us put

$$\zeta_K(s) = \zeta(s)L(s, \chi_D) = \sum_1^{\infty} a_D(n)n^{-s},$$

i.e.,

$$a_D(n) = \sum_{m|n} \chi_D(m).$$

Note that $a_D(n)$ is a nonnegative multiplicative function such that for p prime

$$a_D(p) = 1 + \chi_D(p) = \begin{cases} 2 & \text{if } p \text{ splits} \\ 1 & \text{if } p \text{ is ramified} \\ 0 & \text{if } p \text{ is inert}. \end{cases} \tag{15.2}$$

Our main objective is to establish lower bounds for sums of the type

$$P_D(z, w) = \sum_{w \leqslant p < z} a_D(p)p^{-1}, \tag{15.3}$$

where w and z are quite small relative to the discriminant. From these results one can infer, by a combinatorial argument, lower bounds for the sums

$$N_D(x) = \sum_{n \leqslant x} a_D(n). \tag{15.4}$$

The key issue is to control the uniformity in D, i.e., to obtain bounds which are valid for $x > D^\alpha$ with α as small as possible.

If $\alpha > 1/4$, then, by the Burgess estimate for (1.41), one infers that

$$N_D(x) = \frac{2\pi h}{w\sqrt{D}}x + O(x^{1-\delta}) = L(1, \chi_D)x + O(x^{1-\delta}). \tag{15.5}$$

Hence,

$$xD^{-\varepsilon} \ll N_D(x) \ll x \log D \tag{15.6}$$

for all $x > D^{1/4+\varepsilon}$, where the lower bound is ineffective since it is based on the Siegel estimate

$$L(1, \chi_D) \gg D^{-\varepsilon}. \tag{15.7}$$

In fact, the upper bound of (15.6) holds true for all $x \geqslant 1$ by the trivial estimate

$$a_D(n) \leqslant \tau(n). \tag{15.8}$$

Throughout, for notational simplicity, we shall drop the subscript D whenever it is obvious from the context.

Our first result requires a sharp bound for the class number.

THEOREM 8. *Suppose the class number of $K = \mathbb{Q}(\sqrt{-D})$ satisfies (1.38), or equivalently,*

$$L(1, \chi_D) \gg (\log D)^{-1}(\log\log D)^3. \tag{15.9}$$

Then, for any fixed α, $a > 0$, we have

$$P_D(D^\alpha, \log^a D) > \log\log\log D + O(1), \tag{15.10}$$

the implied constant depending on α, a. Moreover, for any $\varepsilon > 0$, we have

$$N_D(x) \gg xD^{-\varepsilon} \tag{15.11}$$

for all $x \geqslant 1$, the implied constant depending on ε (effectively).

Proof. By (15.5) and (15.9), we get

$$\sum_{n<D} a(n)n^{-1} \gg L(1, \chi) \log D \gg (\log \log D)^3 .$$

On the other hand, this sum is bounded from above by the product

$$\prod_{p<D} \left(1 + \frac{a(p)}{p}\right) \ll \exp\left(\sum_{p<D} \frac{a(p)}{p}\right).$$

Comparing these estimates, we get by taking logarithms that

$$\sum_{p<D} \frac{a(p)}{p} > 3 \log \log \log D + O(1).$$

Subtracting the contribution from small and large primes, and by using the trivial bound $0 \leqslant a(p) \leqslant 2$, we obtain (15.10).

For the proof of (15.11), we may assume that D is large and $D^\varepsilon < x < D$, since otherwise the assertion is either trivial or follows from (15.6). We take (15.10) with $\alpha = \varepsilon/3$ and $a = 3/\varepsilon$, getting

$$\sum_{w\leqslant p<z}' p^{-1} > 1,$$

where $w = (\log D)^{3/\varepsilon}$, $z = D^{\varepsilon/3}$, and the summation in $\sum'$ ranges over primes p of degree 1 in $K = \mathbb{Q}(\sqrt{-D})$. Hence, there exists $w \leqslant y < z$ such that

$$\sum_{p<y}' 1 > y^{1-\varepsilon/3}.$$

Raising this to the exponent $k = [\log x/\log y]$, we get

$$N_D(x) \geqslant \frac{1}{k!}\left(\sum_{p<y}' 1\right)^k > \frac{y^{k(1-\varepsilon/3)}}{k!} > \frac{xD^{-\varepsilon/3}}{yk!}.$$

Here, we have

$$yk! < zk^k \leqslant zx^{\log \log x/\log w} < D^{2\varepsilon/3}.$$

This completes the proof of Theorem 8.

With some extra work, one could weaken the hypothesis (15.9) by reducing the factor $(\log\log D)^3$ a bit. In the next theorem, we do not require this hypothesis. Instead, (15.7) will be used, and we need a nontrivial estimate for short character sums.

THEOREM 9. *Suppose that for any $M \geqslant D^\alpha$ we have*

$$\sum_{m \leqslant M} \chi_D(m) \ll MD^{-\eta}, \tag{15.12}$$

where the exponents α, η are fixed positive numbers. Then we have

$$P_D(D^{\alpha+10\varepsilon}, D^\varepsilon) \gg L(1, \chi_D)(\log D)^{-3}, \tag{15.13}$$

for any $\varepsilon > 0$ provided D is sufficiently large in terms of α, η, and ε.

Proof. We begin by estimating the sifting function

$$S(x, w) = \sum_{\substack{n \leqslant x \\ (n, P(w))=1}} a(n), \tag{15.14}$$

where $P(w)$ denotes the product of all primes $p < w$. We shall apply a Brun-type sieve of 'dimension' $\leqslant 2$ (cf. [HR]). To this end, one needs asymptotics for

$$A_d(x) = \sum_{\substack{n \leqslant x \\ n \equiv 0(d)}} a(n),$$

with $d|P(w)$ as large as possible. We proceed as follows:

$$\begin{aligned} A_d(x) &= \sum_{\substack{lm \leqslant x \\ lm \equiv 0(d)}} \chi(m) \\ &= \sum_{m \leqslant M} \chi(m) \sum_{\substack{l \leqslant x/m \\ l \equiv 0(d/(d,m))}} 1 + \sum_{l \leqslant x/M} \sum_{\substack{M < m \leqslant x/l \\ m \equiv 0(d/(d,l))}} \chi(m) \\ &= \sum_{m \leqslant M} \chi(m)\left(x\frac{(d,m)}{dm} + O(1)\right) + O\left(\sum_{l \leqslant x/M}\left(D^\alpha + x\frac{(d,l)}{dl}D^{-\eta}\right)\right) \\ &= \frac{x}{d}\sum_{m \leqslant M} \chi(m)\frac{(d,m)}{m} + O\left(M + \frac{x}{M}D^\alpha + \frac{\tau(d)}{d}D^{-\eta}x\log x\right), \end{aligned}$$

where M will be chosen later. Furthermore, we have

$$\sum_{m \leqslant M} \chi(m) \frac{(d, m)}{m} = \sum_{ab|d} \chi(ab)\mu(b)b^{-1} \sum_{m \leqslant M/ab} \chi(m)m^{-1}$$

$$= \sum_{c|d} \chi(c) \frac{\varphi(c)}{c} \{L(1, \chi) + O(cM^{-1}D^{\alpha} + D^{-\eta})\}$$

$$= \omega(d)L(1, \chi) + O(dM^{-1}D^{\alpha} + \tau(d)D^{-\eta}),$$

where $\omega(d)$ is the multiplicative function given by $\omega(p) = 1 + \chi(p)(1 - 1/p)$. Combining these evaluations, we obtain, upon choosing $M = (xD^{\alpha})^{1/2}$,

$$A_d(x) = \frac{\omega(d)}{d} xL(1, \chi) + O(x^{1/2}D^{\alpha/2} + \tau(d)d^{-1}D^{-\eta}x \log x).$$

Applying a sieve of level Δ with $\Delta \geqslant w^4$ (to hit the sieving limit for dimension 2), we get

$$S(x, w) \asymp \prod_{p<w} \left(1 - \frac{\omega(p)}{p}\right) xL(1, \chi), \tag{15.15}$$

subject to the condition that the resulting total remainder term

$$R = \Delta x^{1/2} D^{\alpha/2} + D^{-\eta} x (\log x)^3$$

has a smaller order of magnitude than the main term. This condition is satisfied for $\Delta = x^{1/2}D^{-\alpha/2-\varepsilon}$. Therefore, the formula (15.15) is applicable for $x = z = D^{\alpha+10\varepsilon}$ and $w = D^{\varepsilon}$.

On the other hand, we derive the trivial upper bound

$$S(x, w) = 1 + \sum_{w \leqslant p \leqslant x} \sum_{\substack{n \leqslant x/p \\ (n, P(p))=1}} a(np) \ll 1 + \left(\sum_{w \leqslant p \leqslant x} \frac{a(p)}{p}\right) x \log x.$$

Comparing this with (15.15), we obtain (15.13).

The hypotheses made in Theorems 8 and 9 are very natural; nevertheless, they may not be verified in the near future. In the next theorem, we shall establish the desired estimates unconditionally for almost all fundamental discriminants.

Let $\alpha > 0$. We say that D is α-exceptional if

$$P_D(D^{\alpha}, D^{\alpha/4}) \leqslant \frac{1}{2}. \tag{15.16}$$

Of course, if D is large enough, (15.16) should not be true, i.e., the number of exceptional discriminants should be finite. Indeed, assuming the class number

$h(-D)$ satisfies (1.38), it follows from Theorem 8 that the exceptional discriminants are bounded by a constant depending on α. This fact also follows from Theorem 9 and the Burgess estimate [Bu] for the character sums (15.12), provided $\alpha > 1/4$.

THEOREM 10. *Let $0 < \alpha \leqslant 1/2$. The number of α-exceptional discriminants in any interval $X < D \leqslant X^2$ is bounded by a constant depending on α only. More precisely, letting $\mathscr{E}_\alpha(X)$ be the set of such discriminants, we have*

$$|\mathscr{E}_\alpha(X)| < (200/\alpha)^{8/\alpha} \tag{15.17}$$

if X is sufficiently large in terms of α.

Proof. If $D \in \mathscr{E}_\alpha(X)$, then by (15.16) we get

$$\sum_{z<p<z^2} a_D(p)p^{-1} \leqslant \frac{1}{2},$$

where $z = X^{\alpha/2}$. Hence, if X is sufficiently large we get

$$-6 \sum_{z<p<z^2} \chi_D(p)p^{-1} > 1$$

by $\log 2 - (1/2) > (1/6)$. Raising this inequality to the exponent k and then summing over $D \in \mathscr{E}_\alpha(X)$, we get

$$\begin{aligned} |\mathscr{E}| &\leqslant \sum_{D\in\mathscr{E}} \left(-6 \sum_p \chi_D(p)p^{-1}\right)^k \\ &= (-6)^k \sum_{p_1,\ldots,p_k} (p_1 \cdots p_k)^{-1} \sum_{D\in\mathscr{E}} \chi_D(p_1 \cdots p_k). \end{aligned}$$

Hence, by Cauchy's inequality and the bound

$$\sum_{z<p<z^2} p^{-1} < 1,$$

we get

$$|\mathscr{E}|^2 \leqslant 6^{2k} k! S, \tag{15.18}$$

where

$$S = \sum_{\substack{N<n\leqslant N^2 \\ (n,P(z))=1}} n^{-1} \left(\sum_{D\in\mathscr{E}} \chi_D(n)\right)^2$$

with $N = z^k = X^{\alpha k/2}$. We relax the condition $(n, P(z)) = 1$ by applying a linear upper-bound sieve $\Lambda = \{\lambda_d\}$ of level $\Delta = z$, getting

$$S \leqslant \sum_{\substack{d|P(z)\\ d<\Delta}} \lambda_d \sum_{\substack{N<n\leqslant N^2\\ n\equiv 0(d)1}} \mathrm{n}^{-1}\Big(\sum_{D\in\mathscr{E}} \chi_{\mathrm{D}}(n)\Big)^2$$

$$= \sum_d \lambda_d \sum_D \sum_{D'} \sum_{n\equiv 0(d)} n^{-1}\chi_{\mathrm{D}}(n)\chi_{D'}(n)$$

$$= \sum_d \lambda_d \sum_{(DD',d)=1}\sum \left\{\delta_{DD'}\frac{\varphi(D)}{dD}\log N + O\left(\frac{DD'}{N}\right)\right\}$$

$$= (\log N)\sum_D \frac{\varphi(D)}{D}\sum_{(d,D)=1}\lambda_d d^{-1} + O\left(\Delta N^{-1}\Big(\sum_D D\Big)^2\right),$$

where D runs over the set $\mathscr{E}$. Here we have

$$\sum_{(d,D)=1}\lambda_d d^{-1} \ll \prod_{p<z,p|D}(1-p^{-1}) \ll \frac{D}{\varphi(D)}(\log z)^{-1},$$

whence

$$S \ll \frac{\log N}{\log z}|\mathscr{E}| + \Delta N^{-1}X^4|\mathscr{E}|^2 = k|\mathscr{E}| + X^{4-\alpha/2(k-1)}|\mathscr{E}|^2.$$

Inserting this into (15.18), we conclude that $|\mathscr{E}| \ll 6^{2k}(k+1)!$, provided $k > 1 + 8\alpha^{-1}$. Taking $k = 3 + [8\alpha^{-1}]$, we get (15.17).

COROLLARY. *Let $\alpha > 0$. The number of α-exceptional discriminants with $0 < D < X$ is bounded by $c(\alpha)\log\log X$, where $c(\alpha)$ is a positive constant depending on α.*

Note that if D is not α-exceptional, then (15.13) holds true.

16. Estimating the amplifier. Having established various estimates for the prime ideals in $K = \mathbb{Q}(\sqrt{-D})$ of degree 1 and relatively small norms as in Theorems 8, 9, and 10, we can now prove the lower bound (1.36) for

$$\mathscr{A}_\chi = \sum_{q\leqslant Q}\gamma_q(\chi)^2$$

under the relevant conditions, and consequently complete the proofs of Theorems 6 and 7.

Suppose $\chi_D(p) = 1$, so $p = \mathfrak{p}\bar{\mathfrak{p}}$ with $\mathfrak{p} \neq \bar{\mathfrak{p}}$. Then $\gamma_p(\chi) = (1/2)(\chi(\mathfrak{p}) + \bar{\chi}(\mathfrak{p}))$. Unfortunately, not every such p gives a significant contribution to $\mathscr{A}_\chi$ because $\gamma_p(\chi)$

is small if $\chi(\mathfrak{p})$ is close to $\pm i$. But

$$(\chi(\mathfrak{p}) + \bar{\chi}(\mathfrak{p}))^2 - (\chi(\mathfrak{p}^2) + \bar{\chi}(\mathfrak{p}^2)) = 2,$$

so we always have either

$$|\chi(\mathfrak{p}) + \bar{\chi}(\mathfrak{p})| \geqslant 1$$

or

$$|\chi(\mathfrak{p}^2) + \bar{\chi}(\mathfrak{p}^2)| \geqslant 1.$$

Define $\tilde{p} = p$ or p^2 according to which case above appears first. Denote by $\mathscr{B}$ the set of integers composed of distinct numbers of type $\tilde{p}$. Thus if $q \in \mathscr{B}$,

$$|\gamma_q(\chi)| \geqslant 2^{-\omega(q)}.$$

Hence,

$$\mathscr{A}_\chi \geqslant \sum_{q \in \mathscr{B}(Q)} 4^{-\omega(q)} \gg |\mathscr{B}(Q)| Q^{-\varepsilon}, \tag{16.1}$$

where $\mathscr{B}(Q)$ is the subset of numbers in $\mathscr{B}$ bounded by Q.

LEMMA 16.1. *Let $D^\varepsilon < z < D^{1/2}$. Suppose that*

$$P_D(z, D^\varepsilon) > 2D^{-\varepsilon^2} \log D. \tag{16.2}$$

Then

$$|\mathscr{B}(z^6)| \gg z^2 D^{-\varepsilon}. \tag{16.3}$$

Proof. By (16.2) it follows that

$$|\{p \leqslant y \colon \chi_D(p) = 1\}| > yD^{-\varepsilon^2} \tag{16.4}$$

for some $D^\varepsilon < y \leqslant z$. Let k be the integer such that $z^2 < y^k \leqslant yz^2$, so $3 \leqslant k < 1 + \varepsilon^{-1}$. Then each number of type $q = \tilde{p}_1 \dots \tilde{p}_k$ with $\chi_D(p_j) = 1$, $p_j \leqslant y$, all p_j distinct, is in $\mathscr{B}(z^6)$. By (16.4) we obtain

$$|\mathscr{B}(z^6)| \gg (yD^{-\varepsilon^2})^k > z^2 D^{-\varepsilon},$$

since $k > 2 \log z / \log y$, as claimed.

Finally, we verify the hypothesis (16.2) with $z = D^{\alpha + 10\varepsilon}$ by applying any one of Theorems 8, 9, or 10. Hence (16.3) holds true. This, together with (16.1), yields (1.36) for $Q = z^6 = D^{6\alpha + 60\varepsilon}$, except that ε needs to be upgraded.

17. Applications. In this final section we prove Corollaries 1 and 2 of Theorem 7 in the introduction. Recall that $\alpha = 1/1156$.

Proof of Corollary 1. Let $\mathscr{G} \subset \mathscr{H}$ be a subgroup of index k and let $\mathscr{C}\mathscr{G}$ be a coset. Then we get by (1.44) and (1.45) that

$$S_{\mathscr{C}\mathscr{G}}(u) = \sum_{\mathfrak{a} \in \mathscr{C}\mathscr{G}} \exp(-2\pi u N\mathfrak{a}/\sqrt{D}) = \frac{1}{k} \sum_{\chi(\mathscr{G})=1} \bar{\chi}(\mathscr{C}) S(u, \chi) = \frac{h}{kwu} + O(u^{-1/2} D^{(1/2)-\alpha+\varepsilon}).$$

Hence, by Möbius inversion we have

$$\begin{aligned} S'_{\mathscr{C}\mathscr{G}}(u) &= \sum_{\mathfrak{a} \in \mathscr{C}\mathscr{G}}{}' \exp(-2\pi u N\mathfrak{a}/\sqrt{D}) \\ &= \frac{w}{2} \sum_{d=1}^{\infty} \mu(d) S_{\mathscr{C}\mathscr{G}}(ud^2) \\ &= \frac{3h}{\pi^2 ku} + O(u^{-1/2} D^{(1/2)-\alpha+\varepsilon}). \end{aligned}$$

By a standard Tauberian argument, one derives the asymptotic formula

$$\#\{\mathfrak{a} \in \mathscr{C}\mathscr{G}\colon \mathfrak{a} \text{ primitive with } N\mathfrak{a} \leqslant x\} \sim \frac{6h}{\pi k\sqrt{D}} x, \tag{17.1}$$

provided that $x > k^2 D^{(1/2)-2\alpha+\varepsilon}$, and Corollary 1 follows.

Proof of Corollary 2. Let $\mathscr{G} \subset \mathscr{H}$ be a cyclic subgroup of order g and index $k = h/g$. We shall first establish a formula for

$$\sigma_{\mathscr{G}}(\mathfrak{a}) = \begin{cases} 1 & \text{if } \mathfrak{a} \text{ generates } \mathscr{G}, \\ 0 & \text{otherwise}. \end{cases}$$

Choose a fixed generator $\mathfrak{g}$ for $\mathscr{G}$. Then $\mathfrak{a}$ generates $\mathscr{G}$ if and only if $\mathfrak{a} = \mathfrak{g}^l$, where $(l, g) = 1$. By orthogonality of characters we have

$$\frac{1}{h} \sum_{\chi \in \hat{\mathscr{H}}} \bar{\chi}(\mathfrak{a}) \chi(\mathfrak{g}^l) = \begin{cases} 1 & \text{if } \mathfrak{a} = \mathfrak{g}^l, \\ 0 & \text{otherwise}, \end{cases}$$

so

$$\sigma_{\mathscr{G}}(\mathfrak{a}) = \frac{1}{h} \sum_{\chi} \bar{\chi}(\mathfrak{a}) \sum_{\substack{l \pmod g \\ (l,g)=1}} \chi(\mathfrak{g}^l).$$

By Möbius inversion,

$$\sum_{(l,g)=1} \chi(\mathfrak{g}^l) = \sum_{d|g} \mu(d) \sum_{l (\mathrm{mod}\, g/d)} \chi(\mathfrak{g}^{dl}) = g \sum_{\substack{d|g \\ \chi(\mathfrak{g}^d)=1}} \frac{\mu(d)}{d},$$

and so we obtain the desired formula

$$\sigma_{\mathscr{G}}(\mathfrak{a}) = \frac{1}{k} \sum_{d|g} \frac{\mu(d)}{d} \sum_{\chi(\mathfrak{g})^d=1} \chi(\mathfrak{a}). \tag{17.2}$$

Writing

$$S_{\mathscr{G}}^*(u) = \sum_{\mathfrak{a}} \sigma_{\mathscr{G}}(\mathfrak{a}) \exp(-2\pi u N\mathfrak{a}/\sqrt{D}),$$

we get by (17.2), (1.44), and (1.45) that

$$S_{\mathscr{G}}^*(u) = \frac{1}{k} \sum_{d|g} \frac{\mu(d)}{d} \sum_{\chi^d(\mathscr{G})=1} S(u, \chi) = \frac{\varphi(g)}{wu} + O(\tau(g) u^{-1/2} D^{(1/2)-\alpha+\varepsilon}).$$

As above, this gives the asymptotic formula

$$\#\{\mathfrak{a}\colon \mathfrak{a} \text{ generates } \mathscr{G} \text{ and } N\mathfrak{a} \leqslant x\} \sim \frac{2\pi\varphi(g)}{w\sqrt{D}} x, \tag{17.3}$$

provided that $x > k^2 D^{(1/2)-2\alpha+\varepsilon}$ and, hence, finishes the proof of Corollary 2.

Appendix

In this appendix we give two additional results about class-group character sums. These were not needed in the main body of the work, but seem sufficiently basic to prove useful in other circumstances. The first of these is a general mean-value theorem.

Theorem A1. *For any complex numbers $\boldsymbol{c} = \{c_{\mathfrak{a}}\}$, we have*

$$\frac{1}{h} \sum_{\chi \in \mathscr{H}} \left| \sum_{N\mathfrak{a} \leqslant N} c_{\mathfrak{a}} \chi(\mathfrak{a}) \right|^2 = \{1 + O(D^{-1/2} N)\} \sum_{\mathfrak{a}}{}' \left| \sum_{(l)} c_{(l)\mathfrak{a}} \right|^2,$$

where $\sum'$ restricts the summation to primitive ideals.

The error term $O(D^{-1/2}N)$ can be suppressed if $2N < \sqrt{D}$.

For the proof of Theorem A1, we can assume that $\boldsymbol{c} = \{c_{\mathfrak{a}}\}$ is supported on primitive ideals. By the orthogonality of characters, the left-hand side is equal to

$$\sum_{\mathfrak{a} \sim \mathfrak{b}} \sum c_{\mathfrak{a}} \bar{c}_{\mathfrak{b}},$$

where $\sim$ denotes ideal equivalence. From the diagonal terms $\mathfrak{a} = \mathfrak{b}$, we get

$$\|c\|^2 = \sum |c_\mathfrak{a}|^2.$$

The contribution from the remaining off-diagonal terms is bounded by

$$\sum_\mathfrak{a} \#\{\mathfrak{b} \sim \mathfrak{a}\colon \mathfrak{b} \neq \mathfrak{a}, N\mathfrak{b} \leqslant N\} |c_\mathfrak{a}|^2.$$

We have $\mathfrak{b} = (\alpha)\mathfrak{a}$ with $\alpha \in \mathfrak{a}^{-1}$, $\alpha = m + n\bar{z}_\mathfrak{a}$, $1 \leqslant n \leqslant 2\sqrt{aN/D} \leqslant 2N/\sqrt{D}$, $|\alpha|^2 = N\mathfrak{b}/N\mathfrak{a} \leqslant N/a$. Hence, there are no off-diagonal terms if $2N < \sqrt{D}$. Given n as above, the number of m's does not exceed $1 + 2\sqrt{N/a}$. Hence,

$$\#\mathfrak{b} = \#(\alpha) \leqslant \#\{m, n\} \leqslant (1 + 2\sqrt{N/a})2\sqrt{aN/D} \leqslant 6N/\sqrt{D}.$$

This completes the proof of Theorem A1.

Our second result is a Poisson-type formula for these character sums.

Take a function $\varphi(s)$ which is holomorphic in the strip $-\varepsilon \leqslant \operatorname{Re} s \leqslant 1 + \varepsilon$ and such that $\varphi(s)\Gamma(s) \ll |s|^{-2}$ for $|s|$ large. Integrate $\varphi(s)\Phi(s, \chi)$ along the vertical line $\operatorname{Re} s = \sigma = 1 + \varepsilon$. Move to the line $\operatorname{Re} s = -\varepsilon$ passing through the pole at $s = 1$. When on the line $\operatorname{Re} s = -\varepsilon$ apply the functional equation (1.5) and change s into $1 - s$ getting

$$\varphi(1) \operatorname*{res}_{s=1} \Phi(s, \chi) = \frac{1}{2\pi i} \int_{(\sigma)} (\varphi(s) - \varphi(1 - s))\Phi(s, \chi)\, ds.$$

Writing $L_K(s, \chi)$ as a Dirichlet series and interchanging the order of summation with integration on the right-hand side, we get

$$\sum_\mathfrak{a} \chi(\mathfrak{a}) \frac{1}{2\pi i} \int_{(\sigma)} (\varphi(s) - \varphi(1 - s))\Gamma(s) \left(\frac{2\pi N\mathfrak{a}}{\sqrt{D}}\right)^{-s} ds.$$

Therefore,

$$\varphi(1) \operatorname*{res}_{s=1} \Phi(s, \chi) = S_f(\chi) - S_g(\chi), \tag{A.1}$$

say, where $S_f(\chi)$ and $S_g(\chi)$ are defined by

$$S_f(\chi) = \sum_\mathfrak{a} \chi(\mathfrak{a}) f(2\pi N\mathfrak{a}/\sqrt{D}). \tag{A.2}$$

Here f and g are the inverse Mellin transforms of $\varphi(s)\Gamma(s)$ and $\varphi(1 - s)\Gamma(s)$, respectively, i.e.,

$$f(x) = \frac{1}{2\pi i} \int_{(\sigma)} \varphi(s)\Gamma(s)x^{-s}\, ds \tag{A.3}$$

and

$$g(y) = \frac{1}{2\pi i} \int_{(\sigma)} \varphi(1-s)\Gamma(s)^{-s}\, ds. \tag{A.4}$$

The above relations between f, g can be made direct without passing through φ if we assume smoothness and proper decay conditions. Indeed, by Mellin inversion we have

$$\varphi(s)\Gamma(s) = \int_0^\infty f(x)x^{s-1}\, dx.$$

Inserting this into (A.3) we get

$$g(y) = \frac{1}{2\pi i} \int_{(\sigma)} \frac{\Gamma(s)}{\Gamma(1-s)} \left(\int_0^\infty f(x)(xy)^{-s}\, dx \right) ds.$$

In order to be able to interchange the integrations, first move to the line $\sigma = -\varepsilon$, getting

$$g(y) = \int_0^\infty f(x)\, dx + \int_0^\infty f(x) \left(\frac{1}{2\pi i} \int_{(-\varepsilon)} \frac{\Gamma(s)}{\Gamma(1-s)} (xy)^{-s} \right) dx.$$

Hence, by Barnes's formula,

$$\frac{1}{2\pi i} \int_{(-\varepsilon)} \frac{\Gamma(s)}{\Gamma(1-s)} z^{-s}\, ds = J_0(2\sqrt{z}) - 1$$

(cf. [G-R], (6.422.9)), we find that g is a Hankel-type transform of f. More precisely,

$$g(y) = \int_0^\infty f(x) J_0(2\sqrt{xy})\, dx. \tag{A.5}$$

THEOREM A2. *Suppose $f(x)$ is smooth on $\mathbb{R}^+$ such that $f^{(j)}(x) \ll (x + x^{-1})^{-2}$. Let $g(y)$ be given by (A.5). Then the class-group character sums (A.2) satisfy*

$$S_f(\chi) = S_g(\chi) + \delta_\chi h w^{-1} g(0), \tag{A.6}$$

where $\delta_\chi = 1$ if χ is the trivial character and $\delta_\chi = 0$ otherwise.

Proof. The decay condition on f guarantees every argument made before is valid. The function $\varphi(s)\Gamma(s)$ is holomorphic in the strip $-\varepsilon \leqslant \operatorname{Re} s \leqslant 1 + \varepsilon$ without poles. We also have

$$\varphi(1) = \int_0^\infty f(x)\,dx = g(0)$$

and

$$\operatorname*{res}_{s=1} \Phi(s, \chi) = \frac{\sqrt{D}}{2\pi} L(1, \chi_D) = hw^{-1}$$

if $\chi = 1$, or else $\Phi(s, \chi)$ has no poles. In any case, we get (A.6).

References

[BS] A. Baker and A. Schinzel, *On the least integers represented by the genera of binary quadratic forms*, Acta Arith. **28** (1971), 137–166.

[Bu] D. A. Burgess, *On character sums and primitive roots*, Proc. London Math. Soc. (3) **12** (1962), 179–192.

[Bue] D. Buell, *Class groups of quadratic fields*, Math. Comp. **30** (1976), 610–623.

[DI] J.-M. Deshouillers and H. Iwaniec, *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. Math. **70** (1982), 219–288.

[Du] W. Duke, *Hyperbolic distribution problems and half-integral weight Maass forms*, Invent. Math. **92** (1988), 73–90.

[DFI1] W. Duke, J. Friedlander, and H. Iwaniec, *Bounds for automorphic L-functions*, Invent. Math. **112** (1993), 1–8.

[DFI2] ———, *Bounds for automorphic L-functions II*, Invent. Math. **115** (1994), 219–239.

[FI] J. Friedlander and H. Iwaniec, *A mean-value theorem for character sums*, Michigan Math. J. **39** (1992), 153–159.

[Ga] C. F. Gauss, *Disquisitiones Arithmeticae*, English Edition, Springer-Verlag, New York, 1986.

[Ge] F. Gerth III, *Extension of conjectures of Cohen and Lenstra*, Exposition. Math. **5** (1987), 181–184.

[GR] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, Orlando, 1980.

[GrRi] S. W. Graham and C. J. Ringrose, "Lower bounds for least quadratic non-residue" in *Analytic Number Theory*, Progr. Math. **85**, Birkhäuser, Boston, 1990, 269–309.

[HR] H. Halberstam and H.-E. Richert, *Sieve Methods*, Academic Press, London, 1974.

[H-B] D. R. Heath-Brown, *On a paper of Baker and Schinzel*, Acta Arith. **35** (1979), 203–207.

[He] H. Hecke, *Eine neue Art von Zetafunktionen und ihre Benziehungen zur Verteilung der Primzahlen I*, Math. Z. **1** (1918), 357–376.

[Iw] H. Iwaniec, *Introduction to the spectral theory of automorphic forms*, Rev. Mat. Iberoamericana, to appear.

[IS] H. Iwaniec and P. Sarnak, *L^∞-norms of eigenfunctions of arithmetic surfaces*, Ann. of Math (2), to appear.

[KS] S. Katok and P. Sarnak, *Heegner points, cycles and Maass forms*, Israel J. Math. **84** (1993), 193–227.

[Ma] H. Maass, *Über die räumliche Verteilung der Punkte in Gittern mit indefiniter Metrik*, Math. Ann. **138** (1959), 287–315.

[MO] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics*, Chelsea, New York, 1949.

[Se] J-P. Serre, "Modular forms of weight one and Galois representation" in *Algebraic Number Fields*, ed. by A. Frölich, Academic Press, London, 1977, 193–268.

[Si] C. L. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arith. **1** (1935), 83–86.

[Si2] ———, *Lectures on Advanced Analytic Number Theory*, Tata Institute, Bombay, 1961.

Duke: Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903, USA

Iwaniec: Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903, USA

Friedlander: Department of Mathematics, University of Toronto, Toronto, Ontario M5S 1A1, Canada