

Weyl Sums for Quadratic Roots

W. Duke, J. B. Friedlander and H. Iwaniec

1 Introduction

The topic of exponential sums entered firmly into the mainstream of number theory with the work [We] of H. Weyl. Although today very general sums

$$\sum_x e(F(x)) , \quad e(z) = \exp(2\pi iz) ,$$

are studied, Weyl was concerned with the case of F being a polynomial and, whether x runs over a segment of integers, elements of a finite field, or some other set of points, such sums are still referred to as “Weyl sums”.

An important instance is provided by the Weyl sum over roots of quadratic congruences:

$$(1.1) \quad W_h(D; c) = \sum_{\substack{b \pmod{c} \\ b^2 \equiv D \pmod{c}}} e\left(\frac{hb}{c}\right) .$$

The most powerful methods for handling these sums exploit the modern theory of automorphic forms; see [DFI1] for spectral aspects and [DIT] for more arithmetical connections.

The sum (1.1) has only a few terms, bounded by the divisor function, so there is not much room for cancellation, but for applications there is a lot of interest in bounds for sums of these as the modulus c varies, say

$$(1.2) \quad \mathcal{W}_h(D) = \sum_{c \equiv 0 \pmod{q}} f(c) W_h(D; c) ,$$

where f is a nice test function.

The best results to date [By], [DFI1], are extremely strong when the discriminant D is fixed, or at least small relative to the modulus. Our main goal is to obtain bounds which demonstrate such cancellation, even if relatively little, but for D as large as we can. Such a saving, even extremely modest in size, can be crucial for applications. Throughout the paper we assume that D is a positive fundamental discriminant (apart from some remarks and statements concerning negative discriminants in the final section). For simplicity, we only give details of proof for D odd, a condition that may be removed by suitable modifications of the arguments. Extensions to non-fundamental discriminants are also possible. We prove:

Theorem 1.1. *Let $h \geq 1$, $q \geq 1$ and D a positive (odd) fundamental discriminant. Let $f(y)$ be a smooth function supported on $Y \leq y \leq 2Y$ with $Y \geq 1$, such that*

$$(1.3) \quad |f(y)| \leq 1, \quad y^2 |f''(y)| \leq 1 .$$

Then

$$(1.4) \quad \mathcal{W}_h(D) \ll h^{\frac{1}{4}} (Y + h\sqrt{D})^{\frac{3}{4}} D^{\frac{1}{8} - \frac{1}{1331}} ,$$

where the implied constant is absolute.

Here, the parameter h (the frequency) is not very important to us, nevertheless we manage to give an explicit dependence. A similar remark applies to the divisor q . What is important is the relation between the size Y of the moduli and that of the discriminant D for which the result exhibits cancellation; namely, for h and q fixed,

$$\mathcal{W}_h(D) \ll Y^{1-1/1331 \cdot 665} \quad \text{if } Y > D^{\frac{1}{2} - \frac{1}{1332}} .$$

Note that the result retains its significance even when Y is somewhat smaller than \sqrt{D} . Since the class number is frequently as large as \sqrt{D} one cannot hope to achieve a result covering this wide a range by using any method which involves decomposing the sum into contributions from individual classes.

The spectral theory enters our problem by way of Kloosterman sums (see Section 2). Let $K(m, n; c)$ be the Kloosterman sum

$$(1.5) \quad K(m, n; c) = \sum_{d \pmod{c}} \varepsilon_d \left(\frac{c}{d} \right) e \left(\frac{m\bar{d} + nd}{c} \right)$$

for $c > 0$, $c \equiv 0 \pmod{4}$, where $\varepsilon_d = 1$, i according as $d \equiv 1 \pmod{4}$ or $d \equiv 3 \pmod{4}$ and $\left(\frac{c}{d}\right)$ denotes the extended Jacobi symbol as defined in Section 6. These are the Kloosterman sums which appear in the theory of half-integral weight modular forms with respect to the classical theta multiplier, see Section 2.

In order to prove Theorem 1.1 we require a bound for a sum of these, namely for

$$(1.6) \quad \mathcal{K}_g(m, n) = \sum_{c \equiv 0 \pmod{q}} c^{-1} K(m, n; c) g\left(\frac{4\pi\sqrt{mn}}{c}\right),$$

where g is a nice test function. Specifically, we prove

Theorem 1.2. *Let $m \geq 1$, $n \geq 1$, $q \geq 4$, $4 \mid q$, n squarefree. Let $g(x)$ be a smooth function, supported on $X \leq x \leq 2X$ with $X > 0$, such that*

$$(1.7) \quad |g(x)| \leq 1, \quad x^2 |g''(x)| \leq 1.$$

Then we have

$$(1.8) \quad \mathcal{K}_g(m, n) \ll (X^{-\frac{1}{4}} + X^{\frac{1}{2}+\varepsilon}) (mn)^{\frac{1}{4}} n^{-1/1330} \tau(m) \log 2m,$$

where the implied constant depends only on ε .

We remark that, in the sum (1.6) one can easily introduce the extra restriction $(c, \delta) = 1$ by means of the Möbius function. Then, the upper bound (1.8) remains true if we introduce an extra factor $\tau(\delta)$.

We give four arithmetic applications of Theorem 1.1. Weyl sums were originally architected by him for the study of uniform distribution of sequences modulo one. Hence, it is natural to give a result in this spirit as our first application.

Theorem 1.3. *Let $D > 1$, $D \equiv 1 \pmod{4}$, D squarefree, and $\chi_D(n) = \left(\frac{D}{n}\right)$. Let q be a positive integer, $q \equiv 0 \pmod{4}$, such that $D \equiv 1 \pmod{8}$ if $8 \mid q$ and*

$$\chi_D(p) = 1 \text{ if } p \mid q, \quad p \neq 2.$$

For $0 \leq \alpha < \beta \leq 1$ and $x \geq 1$, define $N(x; \alpha, \beta)$ to be the number of fractions b/c in the segment $[\alpha, \beta]$ with $1 \leq c \leq x$, $c \equiv 0 \pmod{q}$ and $b \pmod{c}$, $b^2 \equiv D \pmod{c}$. Then we have

$$(1.9) \quad N(x; \alpha, \beta) = (\beta - \alpha) \frac{3}{\pi^2} R(q) x L(1, \chi_D) + O\left((x + \sqrt{D}) x^{-\frac{1}{7987}}\right),$$

where

$$R(q) = [2 + (2, q/4)] \frac{2^{\omega(q)}}{q} \prod_{p|q} \left(1 + \frac{1}{p}\right)^{-1}$$

and the implied constant is absolute. In particular, $R(4) = 1$.

Note that the asymptotic formula (1.9) is meaningful if

$$x > q D^{\frac{1}{2} - \frac{1}{16000}}.$$

Of course, we believe that the study of such exponential sums with wide uniformity in the discriminant (positive or negative) is of interest on its own, as is the case with many other arithmetic quantities. But actually, the primary motivation by which we were originally drawn to the problem comes from three specific external sources, all of which are concerned with positive discriminant.

Let $D > 1$ be a fundamental discriminant, odd for simplicity, so $D \equiv 1 \pmod{4}$, D squarefree. We consider the indefinite integral binary quadratic forms with discriminant D .

$$(1.10) \quad Q(x, y) = ax^2 + bxy + cy^2, \quad b^2 - 4ac = D.$$

These are primitive, that is $(a, b, c) = 1$, because D is squarefree. Then, the roots of $Q(x, y)$

$$(1.11) \quad \theta_1 = \frac{-b - \sqrt{D}}{2a}, \quad \theta_2 = \frac{-b + \sqrt{D}}{2a}$$

are the endpoints of a geodesic in \mathbb{H} . This geodesic induces a unique, primitive positively-oriented closed geodesic on the modular surface $PSL_2(\mathbb{Z}) \backslash \mathbb{H}$, of length $\log \varepsilon_D$ (or $2 \log \varepsilon_D$ if Q is not equivalent to $-Q$), where

$$\varepsilon_D = \frac{1}{2}(u_D + v_D \sqrt{D})$$

and u_D, v_D is the smallest positive solution of $u^2 - Dv^2 = 4$.

Equivalent forms give rise to the same closed geodesic, the corresponding hyperbolic conjugacy class in $PSL_2(\mathbb{Z})$ being that of

$$(1.12) \quad g_Q = \pm \begin{pmatrix} \frac{1}{2}(u_D + bv_D) & cv_D \\ -av_D & \frac{1}{2}(u_D - bv_D) \end{pmatrix}.$$

For more information see the classical paper [Sa] of P. Sarnak,

The closed geodesics of length $\log \varepsilon_D$ (or $2 \log \varepsilon_D$) appear with multiplicity $h(D)$, the class number of the group of equivalence classes of the forms, which is also the class number of the field $K = \mathbb{Q}(\sqrt{D})$. Recall the celebrated formula of Dirichlet

$$(1.13) \quad h(D) \log \varepsilon_D = L(1, \chi_D) \sqrt{D}.$$

Hence, it is not difficult to establish an asymptotic formula for the sum

$$\sum_{0 < D \leq X} h(D) \log \varepsilon_D$$

(cf [IK]). A deeper problem occurs when the averaging is not ordered by the discriminant, but instead by the size of the fundamental unit. Here, the celebrated trace formula of A. Selberg provides an adequate tool leading to

$$\sum_{\log \varepsilon_D \leq X} h(D) \log \varepsilon_D = X + O(X^{\frac{3}{4}}).$$

To be precise, the above sum is extended to all positive discriminants, not only the fundamental ones (see [Sa]).

Our objective is still more difficult because we want to work with each term individually. One would like to analyze the distribution of the closed geodesics of a given length, as D tends to infinity. But, because $h(D)$ can be small (Gauss conjectured that $h(D) = 1$ indefinitely often) the task seems to be hopelessly difficult. Therefore we look at the problem from a somewhat larger perspective.

Every form of our discriminant D is equivalent (under the action of $SL_2(\mathbb{Z})$ as a unimodular transformation of the variables) to some form (1.10)

with

$$(1.14) \quad \begin{aligned} 0 < b < \sqrt{D} , \\ \sqrt{D} - b < 2|a| < \sqrt{D} + b . \end{aligned}$$

These conditions translate in terms of the roots into

$$(1.15) \quad |\theta_2| < 1 < |\theta_1| ,$$

and such a form is called “reduced”. In general such a reduction does not always determine a unique representative, because some reduced forms can still be equivalent to each other (cf [Bu]). Nevertheless, it is easy to see that the number of reduced forms of a given discriminant $D > 0$ is finite.

Having set the stage, we are now ready to go for our objective. The problem concerns the distribution of integral points (a, b, c) on the one-sheeted hyperboloid

$$(1.16) \quad b^2 - 4ac = D$$

reduced by (1.14). Without losing generality we also assume that

$$(1.17) \quad a > 0$$

because the reduction conditions (1.14) depend only on $|a|$. Let $\mathcal{H}(D)$ be the set of such integral points. By the way, every point in $\mathcal{H}(D)$ gives rise to an integral ideal of $\mathbb{Q}(\sqrt{D})$:

$$\mathfrak{a} = a\mathbb{Z} + \frac{1}{2}(b + \sqrt{D})\mathbb{Z} ,$$

of norm $N\mathfrak{a} = a$. First, we state the following asymptotic formula.

As $D \rightarrow \infty$ over odd fundamental discriminants we have

$$(1.18) \quad |\mathcal{H}(D)| \sim 6\pi^{-2}(\log 2)\sqrt{D} L(1, \chi_D) .$$

We want to know how the first root (1.11) varies with points in $\mathcal{H}(D)$. To this end we consider the subset $\mathcal{H}(D; \lambda)$ of points in $\mathcal{H}(D)$ (the negatives of the roots) which satisfy

$$(1.19) \quad 1 < \frac{b + \sqrt{D}}{2a} < \lambda$$

for any fixed $\lambda > 1$. One may say that we are asking how the generators of integral ideals in $\mathbb{Q}(\sqrt{D})$ are distributed. Here is the answer.

Theorem 1.4. *Let $\lambda > 1$. Then, as $D \rightarrow \infty$ over odd fundamental discriminants we have*

$$(1.20) \quad \frac{|\mathcal{H}(D; \lambda)|}{|\mathcal{H}(D)|} \sim \frac{1}{\log 2} \log \frac{2\lambda}{\lambda + 1} .$$

In other words, the first root (1.11) is equidistributed with respect to the measure

$$(1.21) \quad d\sigma(x) = \frac{1}{\log 2} \frac{dx}{x(x+1)} .$$

The problem has a long, fascinating history. First, partially successful attempts were made by B. F. Skubenko [Sk] and Yu. V. Linnik (see Chapter VI of [Li] and [TV]) by ideas from ergodic theory. These ideas have been recently revised and developed much further by P. Michel and A. Venkatesh (see their survey article [MV] in the ICM, Madrid 2006). In the Rutgers Ph.D. thesis by R. Chelluri [Ch], the Skubenko–Linnik problem was given a rather different treatment from that here, based on the arguments in [D1].

Our next application is to the solution of a problem posed in [DIT] concerning the asymptotic behavior of cycle integrals of the classical modular function $j(z)$. Recall that $j(z)$ is holomorphic on \mathbb{H} , satisfies $j(\gamma z) = j(z)$ for all $\gamma \in SL_2(\mathbb{Z})$, and that $j(z)$ has a simple pole at $i\infty$ with integral Fourier coefficients

$$j(z) = e(-z) + 744 + 196884 e(z) + \dots .$$

More generally, for each integer $m \geq 0$ there is a unique modular function $j_m(z)$ that is holomorphic on \mathbb{H} and whose Fourier expansion begins

$$j_m(z) = e(-mz) + O(e(z)) .$$

Thus,

$$j_0 = 1, j_1 = j - 744, j_2 = j^2 - 1489j + 159768, \dots .$$

The cycle integrals are defined by

$$Tr_D(j_m) = \frac{1}{2\pi} \sum_Q \int_{C_Q} j_m(z) \frac{dz}{Q(z)} ,$$

where C_Q is any smooth curve from any $z \in \mathbb{H}$ to $g_Q(z)$ where g_Q was defined in (1.12) and where the sum is over $SL_2(\mathbb{Z})$ -inequivalent forms Q of discriminant D . It is easily shown (see [DIT]) that $Tr_D(j_m)$ is well-defined. Our third application of Theorem 1.1 is to give an asymptotic formula for $Tr_D(j_m)$ for fixed m and large D .

Theorem 1.5. *We have*

$$Tr_D(j_m) \sim -\frac{12}{\pi} \sigma_1(m) L(1, \chi_D)$$

as $D \rightarrow \infty$ through (odd) fundamental discriminants.

Here $\sigma_1(m)$ is the usual divisor function $\sum_{d|m} d$. Actually, we get a power saving in D ; see Theorem 14.1. Since j_m has exponential growth in the cusp this theorem shows that there is a remarkable amount of cancellation in the sum of the integrals defining $Tr_D(j_m)$. A version of this result for negative discriminants was given in [D2] which does have an exponential main term. Another proof of that result can be given using the methods of this paper. We refer to [DIT] for further information about $Tr_D(j_m)$ and for its connection with the Fourier coefficients of mock modular forms.

Some weeks after the June 2010 Beijing lectures on our proofs by one of us (H. I.), we received a preprint [Ms] from R. Masri, devoted to the [DIT] problem about the cycle integrals and based on the results of [D1].

Our final application is the resolution of a problem, postponed in Section 14.8 of [FI], concerning prime points on the sphere. One knows [D1] that, as $N \rightarrow \infty$ through suitable residue classes modulo eight, the integer points on the sphere

$$(1.22) \quad a^2 + b^2 + c^2 = N$$

are equidistributed.

It would be wonderful to be able to make a similar statement (subject to the appropriate congruence conditions) for points having prime co-ordinates (a, b, c) . Such a result seems hopelessly far off at the current state of affairs so an easier question was proposed in [FI], to estimate the number of solutions

with at least one prime co-ordinate. Even this is too much to handle unconditionally but, subject to some well-known conjectures, the problem becomes accessible.

The semi-linear sieve allows one to approach this question in the form of a slightly modified counting. Instead of recording the number of solutions of (1.22) with $c = p$ a prime we count only (so without multiplicity) the number of primes which occur. Thus we are interested in $S(N) = \sum_{p < \sqrt{N}} B(N - p^2)$ where $B(n)$ is the characteristic function of integers representable as a sum of two squares.

Theorem 1.6. *Let $N \equiv 2 \pmod{4}$, $N \not\equiv 1 \pmod{3}$. Assume the Riemann Hypothesis holds for the Dirichlet L -functions attached to the character χ_N and also to the same character twisted by the non-principal character of modulus four. Assume that the Elliott-Halberstam conjecture holds. Then we have*

$$S(N) \asymp \mathfrak{S}(N)N(\log N)^{-\frac{3}{2}} .$$

Here, $\mathfrak{S}(N)$ is positive arithmetic function of modest size (under RH it is bounded above and below by different powers of $\log \log N$). For the precise definition, and for information about the semi-linear sieve and the Elliott-Halberstam conjecture concerning the distribution of primes in arithmetic progressions, see [FI].

Theorem 1.6 is stated in [FI] as Theorem 14.15. The upper bound is relatively standard and holds unconditionally but the lower bound was proved there subject to a Proposition 14.16 on Weyl sums, the demonstration of which was postponed. Here, we finish that job.

To conclude this section we describe the plan of the paper. We begin in Section 2 with a survey of the automorphic theory required for our proofs. Although standard for a few readers, the material is widely scattered and not easily accessible, especially as it concerns delicate aspects to the extent it applies to multiplier systems.

The third section exposes a very neat inequality involving a trio of sums of Kloosterman sums. It is physically reminiscent of the Cauchy inequality, which indeed is one ingredient in its proof. However, the complete demonstration is highly non-trivial, requiring the full spectral resolution of the space

of automorphic forms. This result provides a decisive step in our arguments. Once this Cauchy inequality is granted, no further use of the spectral theory is needed. What has been accomplished is the reduction of Kloosterman sums with parameters m, n to those of m, m and of n, n . This has the effect, at the Weyl sum level, of producing quadratic polynomials which are reducible. It is no surprise that, in such a case, we can explicitly parametrize the solutions.

Sections 4 and 5 are more analytic in nature. They deal with the construction of various test functions and their integral transforms. Both the construction and the proofs are quite delicate because we need to come up with functions which simultaneously satisfy a number of positivity constraints.

All of the sums we deal with, whether Weyl sums or Kloosterman sums, are related to classical Salié sums by the reciprocity law. Section 6 provides the relevant transformations and estimates. The whole project is powered by estimates for sums of Salié sums, which we take from our earlier paper [DF12], albeit in slightly modified form, as presented in Section 7.

As the Cauchy inequality of Section 3 reduces sums $\mathcal{K}(m, n)$ to the product of sums $\mathcal{K}(m, m)$ and $\mathcal{K}(n, n)$ it is sufficient to make a crucial saving in just one of these factors. Hence we can afford to give two different treatments, a direct method in Section 8 which uses only bounds for the individual Kloosterman sums and, in Section 9, the more sophisticated bound which applies the results of Section 7.

In Section 10 we input the bounds of the two previous sections into the Cauchy inequality to obtain Theorem 1.2. In Section 11 we move from Kloosterman sums to Weyl sums to deduce Theorem 1.1 from Theorem 1.2.

We next turn to applications of Theorem 1.1. In Section 12 we use it to deduce Theorem 1.3. At the end of the section we provide some L -function computations which are needed in this proof and as well in the following Section 13 where we treat the Linnik-Skubenko problem. Then, in Section 14 we still continue to apply Theorem 1.1, now to the equidistribution problem for the cycle integrals. In Section 15 we make some generalizations of the previous arguments (involving Fourier expansions about various cusps) and of Theorem 1.1, resulting in a version which fully covers the needs of the application to the problem of prime points on the sphere. In the final Section

16 we make some remarks and statements about the analogues to Theorems 1.1 and 1.2 for negative discriminants D .

2 Spectral Decomposition of Kloosterman Sums

In this section we give an overview of the spectral theory of automorphic forms with applications to Kloosterman sums. Most of the facts are standard and can be found in the original sources [Ma], [Ro], [Se]. We also recommend [He], [DI], [I2], [DFI3].

Although we only need the theory of forms on the group $\Gamma_0(q)$ with $q \equiv 0 \pmod{4}$ with respect to the theta multiplier of weight $k = \frac{1}{2}$, our presentation looks more transparent in a more general setting.

The group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{H} = \{z = x + iy; x \in \mathbb{R}, y \in \mathbb{R}^+\}$ by

$$gz = \frac{az + b}{cz + d}, \quad \text{if } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$$

For any $g \in \mathrm{SL}_2(\mathbb{R})$ we define the function $j_g : \mathbb{H} \rightarrow \mathbb{C}^*$ by

$$j_g(z) = \frac{cz + d}{|cz + d|} = e^{i \arg(cz+d)}.$$

The powers z^s are defined for every $z \in \mathbb{C}^*$ and $s \in \mathbb{C}$ by

$$z^s = e^{s(\log |z| + i \arg z)}$$

with $\log |z| \in \mathbb{R}$ and $-\pi < \arg z \leq \pi$. In particular

$$(j_g(z))^k = \left(\frac{cz + d}{|cz + d|} \right)^k = e^{ik \arg(cz+d)}.$$

We are interested only in k real. For every $g, h \in \mathrm{SL}_2(\mathbb{R})$ we set

$$\sigma_k(g, h) = j_{gh}(z)^{-k} j_g(hz)^k j_h(z)^k.$$

This product does not depend on $z \in \mathbb{H}$. Note that

$$\sigma_k(g, h) = 1 \quad \text{if } g = \begin{pmatrix} * & * \\ 0 & 1 \end{pmatrix}, \quad h = \begin{pmatrix} * & * \\ 0 & 1 \end{pmatrix},$$

$$\sigma_k(g, h) = e^{2\pi ik} \quad \text{if} \quad g = \begin{pmatrix} * & * \\ 0 & -1 \end{pmatrix}, \quad h = \begin{pmatrix} * & * \\ 0 & -1 \end{pmatrix} .$$

Let Γ be a subgroup of $\text{SL}_2(\mathbb{R})$ which is discrete and has fundamental domain of finite volume with respect to the invariant measure $d\mu z = y^{-2} dx dy$. For simplicity let us assume that

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \Gamma .$$

Furthermore let us assume that ∞ is a cusp for Γ and its stability group consists of translations $z \rightarrow z + b$ by the integers b , that is

$$\Gamma_\infty = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix}; \quad b \in \mathbb{Z} \right\} .$$

A multiplier of weight k on Γ is any function $\vartheta : \Gamma \rightarrow \mathbb{C}^*$ satisfying the following conditions:

$$\begin{aligned} |\vartheta(\gamma)| &= 1 \\ \vartheta(\gamma\delta) &= \sigma_k(\gamma, \delta)\vartheta(\gamma)\vartheta(\delta) \end{aligned}$$

for every γ, δ in Γ , and

$$\vartheta \left(\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right) = e^{-\pi ik} .$$

This last one is called the consistency condition. Note that

$$\vartheta \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = 1 .$$

We see that the complex conjugate $\bar{\vartheta}$ is a multiplier of weight $-k$. If ϑ_1, ϑ_2 are multipliers of weight k_1, k_2 respectively, then $\vartheta_1\vartheta_2$ is a multiplier of weight $k_1 + k_2$. Finally, if ϑ is a multiplier of weight k , then it is also a multiplier of every weight $\ell \equiv k \pmod{2}$. For this reason the Kloosterman sum (our main object of interest) depends only on $k \pmod{2}$.

A function $f : \mathbb{H} \rightarrow \mathbb{C}$ is called automorphic on Γ of weight k with respect to the multiplier ϑ if for every $\gamma \in \Gamma$ it satisfies the equation

$$j_\gamma(z)^{-k} f(\gamma z) = \vartheta(\gamma) f(z) .$$

Let $\mathcal{A}_k(\Gamma, \vartheta)$ denote the linear space of such automorphic functions. Let $\mathcal{L}_k(\Gamma, \vartheta)$ be the subspace of those which are square integrable on the fundamental domain $\Gamma \setminus \mathbb{H}$ with respect to the measure

$$d\mu z = y^{-2} dx dy.$$

Therefore, $\mathcal{L}_k(\Gamma, \vartheta)$ is a Hilbert space with the inner product

$$\langle f, g \rangle = \int_{\Gamma \setminus \mathbb{H}} f(z) \bar{g}(z) d\mu z.$$

Note that this integral does not depend on the choice of the fundamental domain.

The Laplace operator

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

acts on smooth automorphic functions, that is, if $f \in \mathcal{A}_k(\Gamma, \vartheta)$ is smooth, then $\Delta_k f \in \mathcal{A}_k(\Gamma, \vartheta)$. If $f \in \mathcal{A}_k(\Gamma, \vartheta)$ is an eigenfunction of Δ_k , say

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

then such an f is called an automorphic form of eigenvalue λ . Here, λ can be any complex number. We write

$$\lambda = s(1-s), \quad s = \frac{1}{2} + it, \quad \text{so} \quad \lambda = \frac{1}{4} + t^2.$$

Here, t can be any complex number (not necessarily real). Note that t and $-t$ give rise to the same eigenvalue λ .

From now on we assume that

$$0 \leq k < 2.$$

Let $\mathcal{B}_k(\Gamma, \vartheta)$ be the subspace of $\mathcal{A}_k(\Gamma, \vartheta)$ which consists of smooth functions f such that both f and $\Delta_k f$ are bounded. Clearly $\mathcal{B}_k(\Gamma, \vartheta) \subset \mathcal{L}_k(\Gamma, \vartheta)$ and one can show that $\mathcal{B}_k(\Gamma, \vartheta)$ is dense in $\mathcal{L}_k(\Gamma, \vartheta)$. Moreover, for $f, g \in \mathcal{B}_k(\Gamma, \vartheta)$ we have

$$\langle \Delta_k f, g \rangle = \langle f, \Delta_k g \rangle,$$

which means Δ_k is a symmetric operator on $\mathcal{B}_k(\Gamma, \vartheta)$. For any $f \in \mathcal{B}_k(\Gamma, \vartheta)$, we have

$$\langle f, -\Delta_k f \rangle \geq \frac{k}{2} \left(1 - \frac{k}{2}\right) \langle f, f \rangle,$$

which means $-\Delta_k$ is bounded from below by $\frac{k}{2} \left(1 - \frac{k}{2}\right) \geq 0$. Therefore, by a theorem of Friedrichs, the operator $-\Delta_k$ admits a self-adjoint extension (which we also denote by $-\Delta_k$), and by a theorem of von Neumann the space $\mathcal{L}_k(\Gamma, \vartheta)$ has a complete spectral resolution with respect to $-\Delta_k$.

There is a point spectrum (of finite multiplicity)

$$\frac{k}{2} \left(1 - \frac{k}{2}\right) \leq \lambda_0 \leq \lambda_1 \leq \dots$$

and a continuous spectrum covering $\left[\frac{1}{4}, \infty\right)$ with constant multiplicity. The basic eigenpackets of the continuous spectrum are provided by the Eisenstein series $E_{\mathfrak{a}}(z, s)$ on the critical line $\operatorname{Re} s = \frac{1}{2}$ which are defined for every singular cusp \mathfrak{a} .

Let \mathfrak{a} be a cusp of Γ and $\Gamma_{\mathfrak{a}} = \{\gamma \in \Gamma; \gamma\mathfrak{a} = \mathfrak{a}\}$ its stability group. There exists $\sigma_{\mathfrak{a}} \in \operatorname{SL}_2(\mathbb{R})$ such that

$$\sigma_{\mathfrak{a}} \infty = \mathfrak{a}, \quad \sigma_{\mathfrak{a}} \Gamma_{\infty} \sigma_{\mathfrak{a}}^{-1} = \Gamma_{\mathfrak{a}}.$$

The matrix $\sigma_{\mathfrak{a}}$ (we call it a scaling matrix) is unique up to a translation $\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$ from the right side. While Γ_{∞} is generated by $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $-\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, the group $\Gamma_{\mathfrak{a}}$ is generated by $\gamma_{\mathfrak{a}}$ and $-\gamma_{\mathfrak{a}}$ with

$$\gamma_{\mathfrak{a}} = \sigma_{\mathfrak{a}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \sigma_{\mathfrak{a}}^{-1}.$$

We say that the cusp \mathfrak{a} is singular with respect to the multiplier ϑ if

$$\vartheta(\gamma_{\mathfrak{a}}) = 1.$$

Recall that, by our assumption, the cusp ∞ is singular.

For every singular cusp \mathfrak{a} (and only for these) the Eisenstein series is defined by

$$E_{\mathfrak{a}}(z, s) = \sum_{\gamma \in \Gamma_{\mathfrak{a}} \backslash \Gamma} \overline{\vartheta(\gamma) \sigma_k(\sigma_{\mathfrak{a}}^{-1}, \gamma)} j_{\sigma_{\mathfrak{a}}^{-1} \gamma}(z)^{-k} (\operatorname{Im} \sigma_{\mathfrak{a}}^{-1} \gamma z)^s.$$

The series converges absolutely in $\operatorname{Re} s > 1$ and it has meromorphic continuation to the whole complex s -plane. In the half-plane $\operatorname{Re} s \geq \frac{1}{2}$ the Eisenstein series is holomorphic except for a finite number of simple poles at $s = s_j$ say, with

$$\frac{1}{2} < s_j \leq 1 .$$

The points s_j do not depend on the variable $z \in \mathbb{H}$, but the residues may depend on z . These residues are automorphic forms of eigenvalue $\lambda_j = s_j(1 - s_j) < \frac{1}{4}$. They may not be bounded functions, yet they are square-integrable, so they occur in the discrete spectrum. Since $\lambda_j \geq \frac{k}{2}(1 - \frac{k}{2})$ we have $s_j = \frac{1}{2} + it_j$ with $0 < it_j \leq \frac{1}{2}|k - 1|$.

The Eisenstein series $E_{\mathfrak{a}}(z, s)$ on the line $s = \frac{1}{2} + it$, with $t \in \mathbb{R}$ are not square-integrable and they form eigenpackets of the continuous spectrum with the spectral measure being $(4\pi)^{-1}dt$.

The remaining part of $\mathcal{L}_k(\Gamma, \vartheta)$, which is orthogonal to the Eisenstein parts, is spanned by cusp forms. Recall that an automorphic form $u_j(z)$ is a cusp form if

$$\int_0^1 u_j(\sigma_{\mathfrak{a}}z) dx = 0$$

for every singular cusp \mathfrak{a} . This means that the zero-th term in the Fourier expansion of $u_j(z)$ at singular cusps vanish identically. This property makes $u_j(z)$ bounded, so square-integrable, with its eigenvalue satisfying

$$\lambda_j = s_j(1 - s_j) \geq \frac{k}{2}(1 - \frac{k}{2}) .$$

Hence $s_j = \frac{1}{2} + it_j$ with either t_j real or purely imaginary with

$$0 < it_j \leq \frac{1}{2}|k - 1| .$$

The cusp forms with t_j purely imaginary are usually called exceptional, because one believes they do not exist for congruence groups Γ .

Now we are ready to state the following

Theorem 2.1. (Spectral Theorem) *Let $\{u_j(z)\}$ be an orthonormal basis of automorphic forms in the space of the discrete spectrum (including residues of Eisenstein series and cusp forms). Then, any $f \in \mathcal{B}_k(\Gamma, \vartheta)$ has the expansion*

$$f(z) = \sum_j \langle f, u_j \rangle u_j(z) + \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \langle f, E_{\mathfrak{a}}(*, \frac{1}{2} + it) \rangle E_{\mathfrak{a}}(z, \frac{1}{2} + it) dt ,$$

which converges absolutely.

Next we introduce the Fourier expansion of automorphic forms. This can be done at every cusp, but it looks nicer at the singular cusps in which case the frequencies are integers. Actually, we shall only consider here the Fourier expansion at the cusp ∞ (but see Section 15 for further generalizations). Throughout, $\{u_j(z)\}$ is an orthonormal basis in the space of the discrete spectrum. Suppose

$$(\Delta_k + \lambda_j)u_j(z) = 0 \quad \text{with} \quad \lambda_j = s_j(1 - s_j), \quad s_j = \frac{1}{2} + it_j.$$

Then,

$$u_j(z) = \rho_j(0)y^{1-s_j} + \sum_{n \neq 0} \rho_j(n)W_{\frac{kn}{2|n|}, it_j}(4\pi|n|y)e(nx)$$

where $\rho_j(n)$ are complex numbers, called the Fourier coefficients of $u_j(z)$. The zero-th coefficient $\rho_j(0)$ vanishes if $u_j(z)$ is a cusp form.

Similarly, every Eisenstein series has the Fourier expansion

$$E_{\mathbf{a}}(z, s) = \delta_{\mathbf{a}}y^s + \rho_{\mathbf{a}}(0, t)y^{1-s} + \sum_{n \neq 0} \rho_{\mathbf{a}}(n, t)W_{\frac{kn}{2|n|}, it}(4\pi|n|y)e(nx)$$

for $s = \frac{1}{2} + it$. Here $\delta_{\infty} = 1$ and $\delta_{\mathbf{a}} = 0$ if $\mathbf{a} \neq \infty$.

To get some feeling about the size of the Fourier coefficients $\rho_j(n)$, $\rho_{\mathbf{a}}(n, t)$ in the spectral terms and with respect to n , we need to understand the Whittaker function $W_{\kappa, \nu}(y)$. This can be introduced for $y > 0$ and any complex numbers κ, ν by the Mellin integral transform of the gamma functions:

$$W_{\kappa, \nu}(y) = \frac{e^{y/2}}{2\pi i} \int_{(\sigma)} \frac{\Gamma(u - \nu)\Gamma(u + \nu)}{\Gamma(u - \kappa + \frac{1}{2})} y^{\frac{1}{2}-u} du$$

where $\sigma > |\operatorname{Re} \nu|$. Notice that $W_{\kappa, \nu}(y)$ is holomorphic in both of the parameters κ, ν and it is even in ν . For $\operatorname{Re}(\nu - \kappa + \frac{1}{2}) > 0$ we have

$$W_{\kappa, \nu}(y) = \frac{y^{\kappa} e^{-y/2}}{\Gamma(\nu - \kappa + \frac{1}{2})} \int_0^{\infty} e^{-x} x^{\nu - \kappa + \frac{1}{2}} \left(1 + \frac{x}{y}\right)^{\nu + \kappa - \frac{1}{2}} dx.$$

For any complex numbers κ, ν we have the asymptotic formula

$$W_{\kappa, \nu}(y) \sim y^{\kappa} e^{-y/2}, \quad \text{as } y \rightarrow \infty,$$

and in special cases we have

$$\begin{aligned} W_{\kappa, \kappa - \frac{1}{2}}(y) &= y^\kappa e^{-y/2}, \\ W_{0, \nu}(y) &= \left(\frac{y}{\pi}\right)^{\frac{1}{2}} K_\nu\left(\frac{y}{2}\right). \end{aligned}$$

The following Fourier integral representation (see (3.384.9) of [GR]) is quite useful:

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

Using the spectral theory and estimates for the Whittaker function, one can show the following estimates:

$$\begin{aligned} \sum_{T-1 \leq t_j \leq T+1} \sum_{N \leq n \leq 2N} \frac{|\rho_j(n)|^2}{\text{ch}(\pi t_j)} &\ll T^{1-k}, \\ \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{T-1}^{T+1} \sum_{N \leq n \leq 2N} \frac{|\rho_{\mathfrak{a}}(n, t)|^2}{\text{ch}(\pi t)} &\ll T^{1-k}, \end{aligned}$$

for all $T \geq 1$ and $N \geq 1$. The proof is technically quite difficult so we do not provide it in this review (it is a by-product in the spectral development of Kloosterman sums). Our purpose in presenting the above estimates is to give some intuition as to how large are the Fourier coefficients of the basic automorphic forms in spectral terms. Since the number of eigenvalues $\lambda_j = \frac{1}{4} + t_j^2$ with $T-1 \leq t_j \leq T+1$ is about T the above estimate tells us that

$$n|\rho_j(n)|^2 \ll t_j^{-k} \text{ch}(\pi t_j)$$

on average (see also the exact formula for $|\rho_j(1)|^2$ in Proposition 19.1 of [DFI3], where k is a positive integer and the multiplier is given by a character).

Next we introduce the Kloosterman sums $K(m, n; c)$. In general Kloosterman sums are associated to pairs of cusps $\mathfrak{a}, \mathfrak{b}$, but we restrict our exposition only for $\mathfrak{a} = \mathfrak{b} = \infty$. Since the cusp ∞ is assumed to be singular the frequencies m, n are just integers. The modulus c is any positive left-lower entry in $\gamma = \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \Gamma$. Given $m, n \in \mathbb{Z}$ and $c > 0$ we set

$$K(m, n; c) = \sum_{\begin{pmatrix} a & * \\ c & d \end{pmatrix} \in \Gamma_\infty \backslash \Gamma / \Gamma_\infty} \bar{\vartheta} \left(\begin{pmatrix} a & * \\ c & d \end{pmatrix} \right) e\left(\frac{am + dn}{c}\right).$$

Estimates for sums of Kloosterman sums will be our main objective. In general we can only say that $|K(m, n; c)|$ is bounded by the number of terms of summation, getting

$$|K(m, n; c)| \leq c^2/c_1$$

and

$$\sum_{c \leq X} \frac{1}{c} |K(m, n; c)| \leq X/c_1 ,$$

where c_1 is the smallest positive modulus (see Proposition 2.8 of [I2]). By the way,

$$0 < c_1 < \text{vol}(\Gamma \backslash \mathbb{H}).$$

For example, if Γ is the group $\Gamma_0(q)$ the modulus runs over positive integers $c \equiv 0 \pmod{q}$, so $c_1 = q$. In this case, much stronger estimates for $K(m, n; c)$ are available.

We are interested in sums of Kloosterman sums:

$$(2.1) \quad \mathcal{K}_g(m, n) = \sum_c c^{-1} K(m, n; c) g\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

for fixed positive integers m, n and for various test functions $g(x)$ on \mathbb{R}^+ . Assuming

$$(2.2) \quad g(x) \ll x^{1+\varepsilon} ,$$

the series $\mathcal{K}_g(m, n)$ converges absolutely. The spectral theory yields a dual expression for $\mathcal{K}_g(m, n)$ as a sum (over the discrete spectrum) and integrals (over the continuous spectrum) of the Fourier coefficients of basic automorphic forms. These have the following shape:

$$(2.3) \quad \mathcal{L}_{\hat{g}}(m, n) = 4\pi\sqrt{mn} \sum_j \bar{\rho}_j(m) \rho_j(n) \frac{\hat{g}(t_j)}{\text{ch}(\pi t_j)} ,$$

$$(2.4) \quad \mathcal{M}_{\hat{g}}(m, n) = 4\pi\sqrt{mn} \sum_{\mathfrak{a}} \frac{1}{4\pi} \int_{-\infty}^{\infty} \bar{\rho}_{\mathfrak{a}}(m, t) \rho_{\mathfrak{a}}(n, t) \frac{\hat{g}(t)}{\text{ch}(\pi t)} dt ,$$

where $\hat{g}(t)$ is a holomorphic function in $|\text{Im } t| \leq \frac{1}{2}$. Note that the above sum and integral converge absolutely if

$$(2.5) \quad \hat{g}(t) \ll (|t| + 1)^{k-2-\varepsilon} .$$

In the relation between $\mathcal{K}_g(m, n)$ on one side and $\mathcal{L}_{\hat{g}}(m, n) + \mathcal{M}_{\hat{g}}(m, n)$ on the other side the test function $g(x)$ goes to $\hat{g}(t)$ through some transformations which are a combination of the integrals

$$(2.6) \quad \tilde{g}(s) = \int_0^\infty J_s(x)g(x)x^{-1} dx.$$

Here $J_s(x)$ is the Bessel function; it satisfies

$$J_s(x) \ll \min(x^{-\frac{1}{2}}, x^\sigma)$$

for $\sigma = \operatorname{Re} s \geq -\frac{1}{2}$. One should view $\tilde{g}(s)$ as the coefficient in the projection of $g(x)$ on $J_s(x)$ with respect to the measure $x^{-1} dx$. We recommend Appendix B in [I2] for further illuminating interpretations.

The coefficients $\tilde{g}(\ell - 1)$ give rise to the series

$$(2.7) \quad g^*(x) = \sum_{\substack{\ell \geq 2 \\ \ell \equiv k \pmod{2}}} 2(\ell - 1)\tilde{g}(\ell - 1)J_{\ell-1}(x),$$

which we call the Neumann series of $g(x)$. Suppose $g(x)$ is smooth on $[0, \infty)$ and it satisfies the following conditions:

$$(2.8) \quad g(0) = g'(0) = 0,$$

$$(2.9) \quad g(x), g'(x), g''(x) \ll (x + 1)^{-2-\varepsilon}.$$

Then, the Neumann series $g^*(x)$ converges absolutely and uniformly. We have

$$(2.10) \quad \int_0^\infty J_\mu(x)J_\nu(x)x^{-1}dx = \frac{2 \sin \frac{\pi}{2}(\mu - \nu)}{\pi(\mu - \nu)(\mu + \nu)}$$

if $\operatorname{Re}(\mu + \nu) > 0$. Hence, for $\mu, \nu > 0$, $\mu \equiv \nu \pmod{2}$ we get

$$(2.11) \quad 2\nu \int_0^\infty J_\mu(x)J_\nu(x)x^{-1}dx = \delta_{\mu, \nu}.$$

In view of this orthogonality of Bessel functions one can think of $g^*(x)$ as the projection of $g(x)$ on the subspace of $L^2(\mathbb{R}^2, x^{-1} dx)$ which is spanned by the

$J_{\ell-1}(x)$ with $\ell \equiv k \pmod{2}$. This subspace is not dense; actually it is rather thin, so for most nice functions $g(x) \neq g^*(x)$.

We still need to define the function $\hat{g}(t)$ in the spectral sums (2.3), (2.4). This is given in (84) of [Pr]. By the functional equation $\Gamma(z)\Gamma(1-z) \sin \pi z = \pi$ and by the formula

$$\operatorname{ch} 2\pi t + \cos \pi k = 2 \sin \pi \left(\frac{1-k}{2} + it \right) \sin \pi \left(\frac{1-k}{2} - it \right) ,$$

we find that

$$\Gamma \left(\frac{1-k}{2} + it \right) \Gamma \left(\frac{1-k}{2} - it \right) (\operatorname{ch} 2\pi t + \cos \pi k) = D(t)^{-1} ,$$

where

$$(2.12) \quad D(t) = \frac{1}{2} \pi^{-2} \Gamma \left(\frac{1+k}{2} + it \right) \Gamma \left(\frac{1+k}{2} - it \right) .$$

Hence, the function $\hat{g}(t)$ given in (84) of [Pr] can be written as

$$(2.13) \quad \hat{g}(t) = i \left[\tilde{g}(2it) \cos \pi \left(\frac{k}{2} + it \right) - \tilde{g}(-2it) \cos \pi \left(\frac{k}{2} - it \right) \right] \frac{D(t)}{\operatorname{sh} \pi t} .$$

Now we are ready to state the following result which is a generalization of Kuznetsov's formula due to Proskurin (see (83) of [Pr]).

Proposition 2.2. *Let $g(x)$ be a smooth function on $[0, \infty)$ which satisfies (2.8) and (2.9). Then for $m, n \geq 1$ we have*

$$(2.14) \quad \gamma_k \mathcal{K}_{g-g^*}(m, n) = \mathcal{L}_{\hat{g}}(m, n) + \mathcal{M}_{\hat{g}}(m, n) - \delta_{mn} I_k(g) ,$$

where

$$(2.15) \quad I_k(g) = \frac{1}{2\pi} \int_0^\infty J_k(x) g(x) dx$$

and

$$(2.16) \quad \gamma_k = e^{-\pi i k / 2} .$$

Remark: The factor γ_k belongs naturally to the Kloosterman sums $K(m, n; c)$; see Lemma 6.2.

To complete the spectral decomposition of $\mathcal{K}_g(m, n)$ we need to add to (2.14) the contribution of $\mathcal{K}_{g^*}(m, n)$. It turns out that $\mathcal{K}_{g^*}(m, n)$ also admits an expansion in terms of Fourier coefficients, but of classical holomorphic cusp forms of various weights rather than of Maass forms of fixed weight k .

Let $\ell \equiv k \pmod{2}$, so the multiplier ϑ of weight k is also a multiplier of weight ℓ . Denote by $S_\ell(\Gamma, \vartheta)$ the linear space of holomorphic cusp forms $f(z)$ on Γ of weight ℓ , that is satisfying the equation

$$(2.17) \quad (cz + d)^{-\ell} f(\gamma z) = \vartheta(\gamma) f(z)$$

for all $\gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \Gamma$. If you wish, putting $F(z) = y^{\ell/2} f(z)$ this equation takes the familiar shape

$$(2.18) \quad j_\gamma(z)^{-\ell} F(\gamma z) = \vartheta(\gamma) F(z).$$

The space of such cusp forms has finite dimension,

$$\dim S_\ell(\Gamma, \vartheta) \ll \ell.$$

Let $\mathcal{B}_\ell = \mathcal{B}_\ell(\Gamma, \vartheta)$ be an orthonormal basis of $S_\ell(\Gamma, \vartheta)$ with respect to the inner product

$$\int_{\Gamma \backslash \mathbb{H}} y^\ell f(z) \bar{g}(z) d\mu z.$$

For every $f \in \mathcal{B}_\ell(\Gamma, \vartheta)$ we have the Fourier expansion

$$(2.19) \quad f(z) = \sum_1^\infty \rho_f(n) (4\pi n)^{\frac{\ell-1}{2}} e(nz).$$

We have the following formula of Petersson which was a prototype of that in Kuznetsov's work [Ku] for Maass forms.

Proposition 2.3. *Let $\ell \geq 2$, $\ell \equiv k \pmod{2}$. For $m, n \geq 1$ we have*

$$(2.20) \quad \begin{aligned} & \Gamma(\ell - 1) \sum_{f \in \mathcal{B}_\ell} \bar{\rho}_f(m) \rho_f(n) \\ &= \delta_{m,n} + 2\pi e^{-\pi i \ell/2} \sum_c c^{-1} K(m, n; c) J_{\ell-1} \left(\frac{4\pi\sqrt{mn}}{c} \right). \end{aligned}$$

Note that, because $\ell \geq 2$, the series of Kloosterman sums converges absolutely. Taking $m = n$ and estimating the right side trivially, we deduce that

$$(2.21) \quad \Gamma(\ell - 1) \sum_{f \in \mathcal{B}_\ell} |\rho_f(n)|^2 \ll 1 ,$$

where the implied constant depends only on n and the group Γ . This bound tells us that $\Gamma(\ell)|\rho_f(n)|^2$ is bounded on average with respect to ℓ .

Suppose $g(x)$ satisfies (2.8) and (2.9). Then, one can show that

$$(2.22) \quad \sum_{\ell} \ell |\tilde{g}(\ell)| < \infty .$$

Multiplying (2.20) through by $\pi^{-1}(\ell - 1)\tilde{g}(\ell - 1)e^{\pi i \ell / 2}$ and summing over ℓ on the left side we get

$$\sum_{\substack{\ell \geq 2 \\ \ell \equiv k \pmod{2}}} \pi^{-1} e^{\pi i \ell / 2} \tilde{g}(\ell - 1) \Gamma(\ell) \sum_{f \in \mathcal{B}_\ell} \bar{\rho}_f(m) \rho_f(n) = e^{\pi i k / 2} \mathcal{N}_{\tilde{g}}(m, n),$$

say, where

$$(2.23) \quad \mathcal{N}_{\tilde{g}}(m, n) = \sum_{\substack{\ell \geq 2 \\ \ell \equiv k \pmod{2}}} \check{g}(\ell) \Gamma(\ell) \sum_{f \in \mathcal{B}_\ell} \bar{\rho}_f(m) \rho_f(n)$$

with

$$(2.24) \quad \check{g}(\ell) = \pi^{-1} \tilde{g}(\ell - 1) e^{\pi i (\ell - k) / 2} .$$

On the other hand, we obtain $\delta_{m,n} I + \mathcal{K}_{g^*}(m, n)$ (see (2.7)), where

$$(2.25) \quad I = \sum_{\substack{\ell \geq 2 \\ \ell \equiv k \pmod{2}}} \pi^{-1} (\ell - 1) \tilde{g}(\ell - 1) e^{\pi i \ell / 2} .$$

Since $0 \leq k < 2$ the condition $\ell \geq 2$ means $\ell = 2a + k$ with a running over the positive integers. Then, by the formula $2\nu x^{-1} J_\nu(x) = J_{\nu-1}(x) + J_{\nu+1}(x)$ we find that $e^{-\pi i k / 2} I = I_k(g)$ is the same as in (2.14). Hence we conclude

Proposition 2.4. *Let $g(x)$ be a smooth function on $[0, \infty)$ which satisfies (2.8) and (2.9). Then for $m, n \geq 1$ we have*

$$(2.26) \quad \gamma_k \mathcal{K}_{g^*}(m, n) = \mathcal{N}_{\tilde{g}}(m, n) + \delta_{m,n} I_k(g).$$

Now, adding (2.26) to (2.14) we gladly observe that the diagonal terms $\delta_{m,n} I_k(g)$ cancel out and we arrive at the complete formula.

Theorem 2.5. *Let $g(x)$ be a smooth function on $[0, \infty)$ which satisfies (2.8) and (2.9). Then for $m, n \geq 1$ we have*

$$(2.27) \quad \gamma_k \mathcal{K}_g(m, n) = \mathcal{L}_{\tilde{g}}(m, n) + \mathcal{M}_{\tilde{g}}(m, n) + \mathcal{N}_{\tilde{g}}(m, n).$$

The decay conditions (2.9) for the test function $g(x)$ can be considerably weakened, for it suffices that the above series and integrals converge absolutely. Indeed, this can be justified by mollifying $g(x)$ so as to satisfy (2.9) and then taking the limit which removes the mollifier. Therefore, the conditions on $g(x)$ sufficient for the formula (2.27) to hold are (2.8) together with the following estimates:

$$(2.28) \quad \tilde{g}(2it) \ll t^{-2-\varepsilon} \operatorname{ch}(\pi t), \quad t \geq 1,$$

$$(2.29) \quad \tilde{g}(\ell) \ll \ell^{-2-\varepsilon}, \quad \ell \geq 1.$$

The first estimate (2.28) implies (2.5) and the second estimate (2.29) implies (2.22).

3 Cauchy's Inequality for Kloosterman Sums

Applying Cauchy's inequality to the right side of (2.27) we obtain

$$(3.1) \quad |\mathcal{K}_g(m, n)|^2 \leq R(m)R(n)$$

where

$$(3.2) \quad R(m) = \mathcal{L}_{|\check{g}|}(m, m) + \mathcal{M}_{|\check{g}|}(m, m) + \mathcal{N}_{|\check{g}|}(m, m).$$

If one could find a legitimate function $G(x)$, let us call it a majorant of $g(x)$, such that

$$(3.3) \quad |\hat{g}(t)| \leq \hat{G}(t), \quad \text{for } t \geq 0 \quad \text{and for } 0 < it \leq \frac{1}{2}|k-1|$$

and also

$$(3.4) \quad |\check{g}(\ell)| \leq \check{G}(\ell), \quad \text{for } \ell \geq 2, \quad \ell \equiv k \pmod{2},$$

then

$$\begin{aligned} R(m) &\leq \gamma_k \mathcal{K}_G(m, m), \\ R(n) &\leq \gamma_k \mathcal{K}_G(n, n). \end{aligned}$$

These estimates would yield the neat Cauchy-like inequality

$$|\mathcal{K}_g(m, n)|^2 \leq |\mathcal{K}_G(m, m)| |\mathcal{K}_G(n, n)|.$$

In practice it is not difficult to construct a nice majorant $G(x)$ of $g(x)$ which satisfies (3.3) and (3.4), apart from a single value of ℓ for which (3.4) fails because $\check{G}(\ell)$ is negative at this exceptional point. However this failure can be easily overcome as follows. We have

$$\mathcal{N}_{|\check{g}|}(m, m) \leq \mathcal{N}_{\check{G}}(m, m) + (|\check{g}(\ell)| - \check{G}(\ell)) \Gamma(\ell) \sum_{f \in \mathcal{B}_\ell} |\rho_f(m)|^2$$

by borrowing and returning the troubled term (see (2.23)). Let c be any number such that

$$(3.5) \quad (|\check{g}(\ell)| + |\check{G}(\ell)|) (\ell - 1) \leq c.$$

Then by Proposition 2.3 for $m = n$, the correcting term above is bounded by

$$c (1 \pm 2\pi\gamma_k \mathcal{K}_{J_{\ell-1}}(m, m))$$

with the sign \pm chosen according to $e^{-\pi i \ell / 2} = \pm \gamma_k$. This gives

$$R(m) \leq \gamma_k \mathcal{K}_G(m, m) \pm 2\pi c \gamma_k \mathcal{K}_{J_{\ell-1}}(m, m) + c = \gamma_k \mathcal{K}_H(m, n) + c$$

where

$$(3.6) \quad H(x) = G(x) \pm 2\pi c J_{\ell-1}(x).$$

Hence we conclude the following true Cauchy inequality:

$$(3.7) \quad |\mathcal{K}_g(m, n)|^2 \leq |\gamma_k \mathcal{K}_H(m, m) + c| |\gamma_k \mathcal{K}_H(n, n) + c|.$$

We conclude this section with several remarks. First of all, the failure (see the next section) in finding a total majorant $G(x)$, that is without an exceptional point, seems to be an intrinsic issue. Do not neglect a number of other properties which $G(x)$ must satisfy, such as those which warrant the absolute convergence of sums of Kloosterman sums as well as the spectral sums over the Fourier coefficients. Therefore, an inequality of type (3.7) without the positive constant c seems to be unrealistic (because somewhere the diagonal term $\delta_{m,n} I_k(g)$ bubbles under the surface!).

Next, one can enquire as to what is the benefit of having the Cauchy inequality (3.7). Our point is that a general sum of Kloosterman sums $K(m, n; c)$ is majorized by those with $m = n$ and the latter, $K(m, m; c)$, can be treated directly without an appeal to the theory of automorphic forms; well, at least in the case of the theta multiplier, which is our primary interest in this work. Although $K(m, n; c)$ with $m = n$ can be evaluated very explicitly we must say the job does not stop here. This only allows us to transform $\mathcal{K}_G(m, m)$ into bilinear forms with Kloosterman type fractions and the estimation for such bilinear forms is the life-saving ingredient. For this we borrow the relevant results from [DFI2], which in turn were obtained by a double amplification technique.

4 Construction of the Majorant Function

We are looking for a smooth function $G(x)$ on $[0, \infty)$ with

$$(4.1) \quad G(0) = G'(0) = 0$$

which satisfies

$$(4.2) \quad \tilde{G}(2it) \ll t^{-2-\varepsilon} \text{ch}(\pi t) \quad \text{if } t \geq 1,$$

$$(4.3) \quad \tilde{G}(\ell) \ll \ell^{-2-\varepsilon} \quad \text{if } \ell \geq 1 .$$

At the same time we also want $\hat{G}(t)$ and $\check{G}(\ell)$ to be positive and not too small in the ranges

$$(4.4) \quad t \geq 0 \quad \text{or} \quad 0 < it \leq \frac{1}{2}|k-1| ,$$

$$(4.5) \quad \ell \geq 2, \quad \ell \equiv k \pmod{2} ,$$

respectively. We do however exclude one value of ℓ for which $\check{G}(\ell)$ can be negative.

We search among functions of the type

$$(4.6) \quad G(x) = x^{-\alpha} J_{\beta}(x), \quad \text{with } \alpha > 0, \beta > \alpha + 1$$

where α, β will be further restricted in due course. First we compute

$$(4.7) \quad \tilde{G}(s) = \int_0^{\infty} J_s(x) J_{\beta}(x) x^{-\alpha-1} dx .$$

By (6.574.2) of [GR] and using the functional equation for the last gamma factor in the denominator of (6.574.2) we get

$$(4.8) \quad \tilde{G}(s) = -\pi^{-1} 2^{-\alpha-1} \Gamma(\alpha+1) P(s) \sin \frac{\pi}{2}(s-\beta+\alpha)$$

where

$$(4.9) \quad P(s) = \Gamma\left(\frac{s+\beta-\alpha}{2}\right) \Gamma\left(\frac{-s+\beta-\alpha}{2}\right) \Gamma\left(\frac{s+\beta+\alpha}{2}+1\right)^{-1} \Gamma\left(\frac{-s+\beta+\alpha}{2}+1\right)^{-1}$$

provided $\text{Re } s > \alpha - \beta$. Note that $P(s)$ is even. Next, applying the functional equation for the first gamma factor in the denominator of (6.574.2) we get another formula

$$(4.10) \quad \tilde{G}(s) = \pi^{-1} 2^{-\alpha-1} \Gamma(\alpha+1) Q(s) \sin \frac{\pi}{2}(s-\beta-\alpha) ,$$

where

$$(4.11) \quad Q(s) = \Gamma\left(\frac{s+\beta-\alpha}{2}\right) \Gamma\left(\frac{s-\beta-\alpha}{2}\right) \Gamma\left(\frac{s+\beta+\alpha}{2}+1\right)^{-1} \Gamma\left(\frac{s-\beta+\alpha}{2}+1\right)^{-1} .$$

By the bound

$$(4.12) \quad \Gamma(\sigma + it) \asymp t^{\sigma - \frac{1}{2}} e^{-\pi t/2}, \quad t \geq 1,$$

we deduce

$$P(2it) \asymp t^{-2(\alpha+1)}$$

which implies (4.2) by (4.8). Similarly by the bound $\Gamma(s+\mu)\Gamma(s+\nu)^{-1} \asymp s^{\mu-\nu}$ for μ, ν, s real, s large, we deduce

$$Q(\ell) \asymp \ell^{-2(\alpha+1)}$$

which implies (4.3) by (4.10), provided ℓ is large. Then (4.3) holds for all $\ell \geq 1$ by continuity and compactness.

Next we are going to compute $\hat{G}(t)$ and $\check{G}(\ell)$ in the ranges (4.4) and (4.5) respectively. For these we need precise expressions to be able to see the positivity and good lower bounds. First we deal with $\hat{G}(t)$ which is given by (2.13). To this end we use (4.8). Since $P(s)$ is even we need the difference

$$\begin{aligned} \cos \pi \left(\frac{k}{2} + it \right) \sin \pi \left(\frac{\alpha - \beta}{2} + it \right) - \cos \pi \left(\frac{k}{2} - it \right) \sin \pi \left(\frac{\alpha - \beta}{2} - it \right) \\ = i \operatorname{sh}(2\pi t) \cos \frac{\pi}{2} (k + \alpha - \beta) . \end{aligned}$$

Hence, (2.13) and (4.8) yield

$$(4.13) \quad \hat{G}(t) = A(t) \cos \frac{\pi}{2} (k + \alpha - \beta) ,$$

where

$$(4.14) \quad A(t) = \pi^{-1} 2^{-\alpha} \Gamma(\alpha + 1) P(2it) D(t) \operatorname{ch}(\pi t) .$$

It is clear that $P(2it)$ and $D(t) \operatorname{ch}(\pi t)$ are positive in the whole range (4.4) and

$$P(2it) \asymp (|t| + 1)^{-2(\alpha+1)}, \quad D(t) \operatorname{ch}(\pi t) \asymp (|t| + 1)^k .$$

Hence

$$(4.15) \quad A(t) \asymp (|t| + 1)^{k-2(\alpha+1)},$$

and the sign of $\hat{G}(t)$ is that of $\cos \frac{\pi}{2} (k + \alpha - \beta)$.

Next, we deal with $\check{G}(\ell)$ which is given by (2.24). To this end we use (4.10). Writing $\ell = 2a + k$ we get

$$\begin{aligned}\check{G}(\ell) &= \pi^{-1}(-1)^a \check{G}(\ell - 1) \\ &= \pi^{-2} 2^{-\alpha-1} \Gamma(\alpha + 1) Q(\ell - 1) \sin \frac{\pi}{2}(\beta + \alpha - k - 1)\end{aligned}$$

where

$$Q(\ell - 1) = \frac{\Gamma\left(a + \frac{k-1+\beta-\alpha}{2}\right) \Gamma\left(a + \frac{k-1-\beta-\alpha}{2}\right)}{\Gamma\left(a + \frac{k-1+\beta+\alpha}{2} + 1\right) \Gamma\left(a + \frac{k-1-\beta+\alpha}{2} + 1\right)}.$$

Clearly, $Q(\ell - 1)$ is positive if a is large enough and the sign of $\check{G}(\ell)$ is that of $\sin \frac{\pi}{2}(\beta + \alpha - k - 1)$. There is no choice of $\alpha > 0$, $\beta > \alpha + 1$ for which $\check{G}(\ell) > 0$ if $a = 1$, that is for $\ell = k + 2$, but there are many easy choices if $a \geq 2$. For simplicity we assume $0 < k < 2$ and take

$$(4.16) \quad 0 < \alpha < k, \quad \beta = k + 1,$$

getting

$$(4.17) \quad \check{G}(\ell) = \frac{\Gamma(\alpha + 1)}{\pi^2 2^{\alpha+1}} \frac{\Gamma(a + k - \frac{1}{2}\alpha) \Gamma(a - 1 - \frac{1}{2}\alpha)}{\Gamma(a + k + 1 + \frac{1}{2}\alpha) \Gamma(a + \frac{1}{2}\alpha)} \sin \frac{\pi\alpha}{2}.$$

This is positive if $a \geq 2$, that is for $\ell \geq k + 4$, and

$$(4.18) \quad \check{G}(\ell) \asymp \ell^{-2(\alpha+1)}.$$

Recall that the sign of $\hat{G}(t)$ is that of $\cos \frac{\pi}{2}(k + \alpha - \beta) = \sin \frac{\pi\alpha}{2} > 0$. Gathering the above results we conclude the following.

Proposition 4.1. *Let $0 < \alpha < k$ and*

$$(4.19) \quad G(x) = x^{-\alpha} J_{k+1}(x).$$

Then, the conditions (4.1), (4.2), (4.3) hold. Moreover, for t in (4.4) and ℓ in (4.5), except for $\ell = k + 2$, the functions $\hat{G}(t)$, $\check{G}(\ell)$ are positive and satisfy

$$(4.20) \quad \hat{G}(t) \asymp (|t| + 1)^{k-2(\alpha+1)},$$

$$(4.21) \quad \check{G}(\ell) \asymp \ell^{-2(\alpha+1)}.$$

Note that for our choice $\beta = k+1$, $\check{G}(\ell)$ is given by (4.17) and the function $\hat{G}(t)$ given by (4.13), (4.14), (4.9) and (2.12), becomes

$$(4.22) \quad \hat{G}(t) = \frac{\Gamma(\alpha + 1)}{\pi^3 2^\alpha} \{\dots\} \operatorname{ch}(\pi t) \sin \frac{\pi \alpha}{2},$$

where

$$\{\dots\} = \frac{\Gamma\left(\frac{k+1-\alpha}{2} + it\right) \Gamma\left(\frac{k+1-\alpha}{2} - it\right) \Gamma\left(\frac{k+1}{2} + it\right) \Gamma\left(\frac{k+1}{2} - it\right)}{\Gamma\left(\frac{k+1+\alpha}{2} + 1 + it\right) \Gamma\left(\frac{k+1+\alpha}{2} + 1 - it\right)}.$$

5 The Test Functions

Our goal here is to pick up fairly general test functions $g(x)$ for which $G(x)$ given in (4.19) is a majorant, so that we can reduce the problem of estimating the sums $\mathcal{K}_g(m, n)$ to $\mathcal{K}_G(m, m)$ and $\mathcal{K}_G(n, n)$ by means of the Cauchy inequality (3.7). For all the applications in mind we are satisfied with the following choices:

Lemma 5.1. *Let $g(x)$ be a smooth function supported on $X \leq x \leq 2X$ with $X > 0$, such that*

$$(5.1) \quad |g(x)| \leq 1, \quad x^2 |g''(x)| \leq 1.$$

Then, we have

$$(5.2) \quad \hat{g}(t) \ll (|t| + 1)^{k - \frac{5}{2}}, \quad t \text{ real},$$

$$(5.3) \quad \hat{g}(t) \ll X^r + X^{-r}, \quad -r \leq 2it \leq r,$$

$$(5.4) \quad \check{g}(\ell) \ll X^\sigma \ell^{-1-\sigma}, \quad \ell \geq 4,$$

where $1 \leq \sigma \leq 2$ and the implied constants are absolute.

We remark that the bound (5.4) is equivalent to $\ell^{-3} X \min(\ell, X)$ which for $\ell \asymp X$ becomes X^{-1} , and this cannot be improved in general.

Proof. First we treat (5.3) by direct arguments. For $-r \leq \nu \leq r$ we have

$$J_\nu(x) \ll x^r + x^{-r}$$

This follows by the power series expansion (8.440) of [GR] if $0 < x \leq 1$ and by the trivial bound $J_\nu(x) \ll 1$ if $x \geq 1$. Hence

$$\tilde{g}(\nu) = \int_0^\infty J_\nu(x)g(x)x^{-1} dx \ll X^r + X^{-r}$$

and (5.3) follows by (2.12). \square

For the proof of (5.2) and (5.4) we consider the Mellin transform

$$M_g(s) = \int_0^\infty g(x)x^{s-1} dx.$$

Clearly $M_g(s)$ is entire. By partial integration

$$M_g(s) = \frac{1}{s(s+1)} \int_0^\infty g''(x)x^{s-1} dx.$$

Using both integrals we derive by (5.1) the following estimates:

$$|M_g(s)| \leq \frac{2^{\sigma-1}}{\sigma} X^\sigma \min\left(1, \frac{1}{|s(s+1)|}\right) \ll X^\sigma (|s| + 1)^{-2}$$

where $\sigma = \operatorname{Re} s$ and the implied constant depends only on σ . By Mellin inversion

$$g(x) = \frac{1}{2\pi i} \int_{(\sigma)} M_g(s)x^{-s} ds.$$

Hence, if $-\frac{3}{2} < \sigma < \operatorname{Re} \nu$ we get by (6.561.14) of [GR]

$$\begin{aligned} \tilde{g}(\nu) &= \frac{1}{2\pi i} \int_{(\sigma)} M_g(s) \int_0^\infty J_\nu(x)x^{-s-1} dx \\ (5.5) \quad &= \frac{1}{2\pi i} \int_{(\sigma)} M_g(s) 2^{-s-1} \Gamma\left(\frac{\nu-s}{2}\right) \Gamma\left(\frac{\nu+s}{2} + 1\right)^{-1} ds. \end{aligned}$$

For $\ell \geq 4$ the formula (5.5) with $1 \leq \sigma \leq 2$ yields

$$\tilde{g}(\ell - 1) \ll X^\sigma \int_{(\sigma)} \left| s^{-2} \Gamma\left(\frac{\ell-s-1}{2}\right) \Gamma\left(\frac{\ell+s+1}{2}\right)^{-1} \right| |ds| \ll X^\sigma \ell^{-\sigma-1}$$

because $\Gamma(\sigma)\Gamma(\sigma)^{-1} \asymp \ell^{-\sigma-1}$ by (4.12). This shows (5.4) by (2.24).

Now let $\nu = 2it$, t real. Write $s = \sigma + 2iu$ in (5.5) with $-1 \leq \sigma < 0$ giving

$$\begin{aligned} \tilde{g}(2it) &\ll X^\sigma \int_{-\infty}^{\infty} (1+u^2)^{-1} \left| \Gamma\left(-\frac{\sigma}{2} + i(t-u)\right) \Gamma\left(\frac{\sigma}{2} + 1 + i(t+u)\right)^{-1} \right| du \\ &\asymp X^\sigma \int_{-\infty}^{\infty} (1+u^2)^{-1} A(t,u)^{-\frac{\sigma+1}{2}} \exp\frac{\pi}{2}(|t+u| - |t-u|) du \end{aligned}$$

by (4.12), where $A(t,u) = (1+|t-u|)(1+|t+u|) \gg 1+|t|$. Hence,

$$\begin{aligned} (5.6) \quad \tilde{g}(2it) &\ll X^\sigma (1+|t|)^{-\frac{\sigma+1}{2}} \int_{-\infty}^{\infty} (1+u^2)^{-1} \exp(\pi \min(|t|, |u|)) du \\ &\ll \left(\frac{X}{\sqrt{1+|t|}}\right)^\sigma (1+|t|)^{-\frac{5}{2}} \exp(\pi|t|). \end{aligned}$$

Another estimate comes by moving the contour of integration in (5.5) to the $-\sigma$ line. The pole at $s = \nu$ has residue $M_g(\nu)/2^\nu \Gamma(\nu+1)$ which is bounded by

$$(5.7) \quad (1+|t|)^{-\frac{5}{2}} \exp(\pi|t|)$$

while the integral on the $-\sigma$ line is estimated by (5.6), but with σ changed to $-\sigma$. Combining both estimates we deduce that $\tilde{g}(2it)$ is bounded by (5.7). Hence (5.2) follows by (2.12).

Put $\sigma = 1 + 2\alpha$ in (5.4) with $0 < \alpha \leq \frac{1}{4}$. Then Proposition 4.1 and Lemma 5.1 show that our test function $g(x)$ is “majorized” by $G(x)$ up to a scaling factor

$$(5.8) \quad X^r + X^{-r} + X^{1+2\alpha}$$

where $r = |k-1|$, $0 < \alpha \leq \frac{1}{4}$, $\alpha < k$. The last component $X^{1+2\alpha}$ looks excessive even when we choose α arbitrarily small. However, in applications we shall use (5.7) only when X is not large, covering the range of large X by “ad hoc” arguments.

6 Kloosterman - Salié–Weyl Sums

From now on we take $k = \frac{1}{2}$. An important example of an automorphic form of weight one half is the standard theta series

$$\theta(z) = \sum_{-\infty}^{\infty} e(n^2 z) = 1 + 2 \sum_1^{\infty} e(n^2 z)$$

on the group $\Gamma_0(4)$. This has multiplier given by

$$(6.1) \quad \vartheta(\gamma) = \bar{\varepsilon}_d \left(\frac{c}{d} \right) \quad \text{if } \gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix}$$

for $c \equiv 0 \pmod{4}$ and $(c, d) = 1$. Recall that

$$(6.2) \quad \varepsilon_d = \begin{cases} 1 & \text{if } d \equiv 1 \pmod{4}, \\ i & \text{if } d \equiv 3 \pmod{4} \end{cases}$$

and $\left(\frac{c}{d} \right)$ denotes the Jacobi symbol if $d > 1$. If $d < -1$ we have $c \neq 0$ and

$$\left(\frac{c}{d} \right) = \frac{c}{|c|} \left(\frac{c}{-d} \right).$$

For $d = \pm 1$ we have $c = 0$ and $\left(\frac{0}{1} \right) = \left(\frac{0}{-1} \right) = 1$.

The corresponding Kloosterman sums are

$$(6.3) \quad K(m, n; c) = \sum_{d \pmod{c}}^* \varepsilon_d \left(\frac{c}{d} \right) e \left(\frac{\bar{d}m + dn}{c} \right)$$

defined for $c > 0$, $c \equiv 0 \pmod{4}$, where $\bar{d}d \equiv 1 \pmod{c}$. They are closely related (via quadratic reciprocity) with the Salié sums

$$(6.4) \quad S(m, n; q) = \sum_{d \pmod{q}} \left(\frac{d}{q} \right) d \left(\frac{\bar{d}m + dn}{q} \right).$$

We need these sums only for $q \geq 1$, q odd, as defined above. By Lemma 2 of [I1] the Kloosterman sum factors as follows:

$$(6.5) \quad K(m, n; c) = K'(m\bar{q}, n\bar{q}; r) S(m\bar{r}, n\bar{r}; q)$$

where $c = qr$ with $(q, r) = 1$, $r \equiv 0 \pmod{4}$. Here $K'(*, *; r)$ is the Kloosterman sum (6.3) with ε_d replaced by ε_d^q .

The Salié sums can be evaluated explicitly and then estimated trivially, giving (cf [IK, Corollary 11.12])

$$(6.6) \quad |S(m, n; q)| \leq (m, n, q)^{\frac{1}{2}} q^{\frac{1}{2}} \tau(q).$$

One can also evaluate the Kloosterman sum to modulus $r = 2^\alpha$, $\alpha \geq 2$, and show that

$$(6.7) \quad |K(m, n; r)| \leq 4(m, n, r)^{\frac{1}{2}} r^{\frac{1}{2}}.$$

For the proof of (6.7) we can assume that $\alpha \geq 5$, otherwise the result is trivial. First consider $(m, n, r) = 1$. Put $\beta = \lfloor \frac{\alpha}{2} \rfloor$, $\gamma = \lfloor \frac{\alpha+1}{2} \rfloor$ and $d = a(1 + 2^\gamma b)$ getting

$$\begin{aligned} K(m, n; r) &= \sum_{a \pmod{2^\gamma}} \varepsilon_a \left(\frac{2}{a}\right)^\alpha \sum_{b \pmod{2^\beta}} e\left(\frac{\bar{a}(1 - 2^\gamma b)m + a(1 + 2^\gamma b)n}{c}\right) \\ &= 2^\beta \sum_{\substack{a \pmod{2^\gamma} \\ \bar{a}m \equiv an \pmod{2^\beta}}} \varepsilon_a \left(\frac{2}{a}\right)^\alpha e\left(\frac{\bar{a}m + an}{c}\right). \end{aligned}$$

The number of choices for a is at most four which proves (6.7) if $(m, n, r) = 1$. In the general case, put $(m, n, r) = r_0$, $m_1 = m/r_0$, $n_1 = n/r_0$, $r_1 = r/r_0$. If $4 \mid r_1$ then

$$|K(m, n; r)| = r_0 |K(m_1, n_1; r_1)| \leq 4r_0 r_1^{\frac{1}{2}}$$

which proves (6.7). If $r_1 = 1, 2$ then (6.7) follows by the trivial bound $|K(m, n; r)| \leq r/2$.

Now, by (6.5), (6.6), (6.7) we obtain

Lemma 6.1. *For $c \equiv 0 \pmod{4}$ we have*

$$(6.8) \quad |K(m, n; c)| \leq (m, n, c)^{\frac{1}{2}} c^{\frac{1}{2}} \tau(c).$$

The Kloosterman sums are also related to the Weyl sums. Precisely, for $c \equiv 0 \pmod{4}$ and $D \equiv 1 \pmod{4}$ we have (see Proposition 1 of [DIT])

$$(6.9) \quad W_{2h}(D; c) = (1 - i) \sum_{\substack{4vw=c \\ uw=h}} K(u^2, D; 4v) / (2, v) \sqrt{v}.$$

Note that $1 - i = e^{-\pi i/4} \sqrt{2}$ and $e^{-\pi i/4} = e^{-\pi i k/2} = \gamma_k$ is the same factor which we have attached to the Kloosterman sums in Section 2.

One can express the Weyl sum (1.1) for any modulus $c \geq 1$ and any frequency h by the somewhat special ones (6.9). Suppose $D \equiv 1 \pmod{4}$. Then we have

$$(6.10) \quad W_h(D; c) = \frac{1}{2(2, c)} W_{4h}(D; 4c).$$

Combining (6.9) and (6.10) we obtain

Lemma 6.2. *Let $D \equiv 1 \pmod{4}$. Then for any $c \geq 1$ and any h we have*

$$(6.11) \quad W_h(D; c) = \frac{1 - i}{2(2, c)} \sum_{\substack{vw=c \\ uw=2h}} K(u^2, D; 4v) / (2, v) \sqrt{v}.$$

7 Sums of Salié Sums

We need the Salié sums (6.4) only for $m = n$. To avoid confusion with our previous notation we change the letters and recall the definition

$$(7.1) \quad S(a, a; c) = \sum_{d \pmod{c}} \left(\frac{d}{c} \right) e \left(\frac{a(\bar{d} + d)}{c} \right),$$

where c is any positive, odd number. There is a very simple formula for such sums (see (12.43) of [IK]):

$$(7.2) \quad S(a, a; c) = \varepsilon_c c^{\frac{1}{2}} \left(\frac{a}{c} \right) \sum_{\substack{mn=c \\ (m,n)=1}} e \left(2a \left(\frac{\bar{m}}{n} - \frac{\bar{n}}{m} \right) \right).$$

Hence, one can see the sign variation of $S(a, a; c)$ pretty well; nevertheless showing that it causes significant cancellation in sums over the modulus c is by no means easy. The effect of such cancellation is measured in Theorem 4 of [DFI2]. We are going to use a slightly generalized version of this result in which the coefficient a is replaced by $a\bar{q}$ with \bar{q} standing for the multiplicative inverse of q modulo c . Although q is relatively small in our applications, its

inverse \bar{q} makes a difference in the behaviour of $S(a\bar{q}, a\bar{q}; c)$ as c runs over large segments. Fortunately, the proof of Theorem 4 of [DFI2] requires only a little modification to accommodate this factor \bar{q} . Here is the result in the form we need.

Theorem 7.1. *For positive integers a, q, r with aq not a square and $8 \mid r$, and for $(b, r) = 1$, we have*

$$(7.3) \quad \sum_{\substack{c \leq x \\ (c, aq) = 1 \\ c \equiv b \pmod{r}}} c^{-\frac{1}{2}} S(a\bar{q}, a\bar{q}; c) \ll qr^{\frac{1}{5}} (x+a)^{\frac{47}{118}} x^{\frac{35}{59} + \varepsilon}$$

where the implied constant depends only on ε .

Note, by (7.2), that if a is a square the Salié sum $S(a, a; c)$ does not change sign significantly, so the condition aq not a square in (7.3) cannot be neglected. Actually, in Theorem 4 of [DFI2] we gave a precise expression for the extra term which occurs when aq is a square.

The proof of (7.3) follows closely on the lines of the proof of Theorem 4 of [DFI2] so we feel only the need to explain a few modifications. We begin with the formula (7.2) with a replaced by $a\bar{q}$. Using the relations

$$\frac{\bar{A}}{B} + \frac{\bar{B}}{A} \equiv \frac{1}{AB} \pmod{1}$$

for A, B positive co-prime integers, and

$$\frac{\overline{qm}}{n} + \frac{\overline{mn}}{q} \equiv \frac{\bar{m}}{qn} \pmod{1},$$

we obtain

$$S(a\bar{q}, a\bar{q}; c) = \varepsilon_c c^{\frac{1}{2}} \left(\frac{aq}{c}\right) \sum_{\substack{mn=c \\ (m,n)=1}} e\left(4a \frac{\overline{qm}}{n}\right) e\left(2a \frac{\overline{mn}}{q} - \frac{2a}{qc}\right).$$

Next, in order to separate the variables m, n in the exponential $e(2a \overline{mn}/q)$ we could split m, n into residue classes modulo q , but it is more economical to apply Gauss sums:

$$e\left(2a \frac{\overline{mn}}{q}\right) = \frac{1}{\varphi(q)} \sum_{\chi \pmod{q}} \bar{\chi}(2a) \chi(mn) \tau(\chi).$$

This application is not really essential because we can afford to lose a power of q . Note that the other part of the exponential satisfies

$$e\left(\frac{-2a}{qc}\right) = 1 + O\left(\frac{a}{qmn}\right)$$

and can be incorporated by partial summation or other techniques of separation of variables. This issue was addressed at the end of Section 3 of [DFI2].

Now, we are in a position to apply Theorem 3 of [DFI2] with M replaced by qM . The result will be worse than Theorem 3 of [DFI2] by a factor of $q^{\frac{1}{2}}q^{\frac{1}{2}} = q$, the first because of this change of variable m to qm and the second because of the presence and the size of the Gauss sum. As in [DFI2] this modified Theorem 3 applies to the proof of Theorem 4 except when one of M, N is small and in these extreme cases we can use the Polya–Vinogradov inequality in the same fashion as on page 39 of [DFI2]. Here we lose slightly more in the q aspect, namely $(q^{\frac{1}{2}} \log 2q)q^{\frac{1}{2}}$, the first because the relevant modulus is larger by a factor q and the second due to the Gauss sums. This leads to the current version of Theorem 4 of [DFI2] as stated here in Theorem 7.1 (well, ignore $\log 2q$ which can be assumed to be $\ll x^\varepsilon$, since otherwise the result is trivial!).

8 Direct Estimation of $\mathcal{K}_g(m, m)$

Recall that $\mathcal{K}_g(m, n)$ is defined by (2.1) as the sum of the Kloosterman sums $K(m, n; c)$ over the modulus $c > 0$, which is the lower-left entry of the group elements $\gamma = \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \Gamma_0(q)$ with $4 \mid q$. For the theta multiplier $\vartheta(\gamma)$ the Kloosterman sums $K(m, n; c)$ are given by (6.3) for every $c > 0$, $c \equiv 0 \pmod{q}$ and

$$(8.1) \quad \mathcal{K}_g(m, n) = \sum_{c \equiv 0 \pmod{q}} \frac{1}{c} K(m, n; c) g\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

In this section we treat $\mathcal{K}_g(m, n)$ easily by applying the estimates for individual Kloosterman sums given in Lemma 6.1. Suppose that the test

function satisfies

$$(8.2) \quad |g(x)| \leq \min(x, x^{-\frac{1}{2}}) .$$

Hence, $|g(x)| \leq x^{\frac{1}{2}+\varepsilon}$ for any $0 < \varepsilon \leq \frac{1}{2}$ and

$$|\mathcal{K}_g(m, n)| \leq (4\pi\sqrt{mn})^{\frac{1}{2}+\varepsilon} \sum_{c \equiv 0(q)} \tau(c) c^{-1-\varepsilon} (m, n, c)^{\frac{1}{2}} .$$

Here, the sum is bounded by

$$\frac{\tau(q)}{q} (m, n, q)^{\frac{1}{2}} \sum_{c=1}^{\infty} \tau(c) c^{-1-\varepsilon} (m, n, c)^{\frac{1}{2}} .$$

The last sum is bounded by

$$\sum_{\delta|(m,n)} \tau(\delta) \delta^{-\frac{1}{2}} \sum_{c=1}^{\infty} \tau(c) c^{-1-\varepsilon} = \prod_{p|(m,n)} \left(1 - \frac{1}{\sqrt{p}}\right)^{-2} \zeta(1+\varepsilon)^2 .$$

Choosing $\varepsilon = 1/\log(8mn)$ we conclude

Lemma 8.1. *Let m, n, q be positive integers with $4 \mid q$. Suppose $g(x)$ satisfies (8.2). Then*

$$(8.3) \quad \mathcal{K}_g(m, n) \ll \frac{\tau(q)}{q} (m, n, q)^{\frac{1}{2}} \prod_{p|(m,n)} \left(1 - \frac{1}{\sqrt{p}}\right)^{-2} (mn)^{\frac{1}{4}} (\log 2mn)^2 ,$$

where the implied constant is absolute.

In particular the function $H(x)$ given by (3.6) with $\ell = \frac{5}{2}$ and (4.19) with $k = \frac{1}{2}$, $0 < \alpha < \frac{1}{2}$, satisfies

$$(8.4) \quad H(x) \ll \min\left(x, x^{-\frac{1}{2}}\right)$$

so Lemma 8.1 is applicable, giving

$$(8.5) \quad \mathcal{K}_H(m, m) \ll \frac{\tau(q)}{q} (m, q)^{\frac{1}{2}} \prod_{p|m} \left(1 - \frac{1}{\sqrt{p}}\right)^{-2} m^{\frac{1}{2}} (\log 2m)^2 ,$$

where the implied constant is absolute.

9 Estimation of $\mathcal{K}_F(n, n)$

Now, we give a more delicate treatment of sums of the Kloosterman sums $K(n, n; c)$, one which takes advantage of the sign variations. As we mentioned earlier, this requires n not to be a square. For simplicity we assume that

$$(9.1) \quad n \text{ is squarefree, } n > 1 .$$

First, we are going to estimate the sum

$$(9.2) \quad T(x) = \sum_{\substack{c \leq x \\ c \equiv 0 \pmod{q}}} c^{-\frac{1}{2}} K(n, n; c).$$

Here we need to factor the Kloosterman sums into the Salié sums which appear in Theorem 7.1. To this end we write the modulus uniquely as cr with

$$(9.3) \quad (c, nq) = 1, \quad r \mid (nq)^\infty, \quad r \equiv 0 \pmod{q}.$$

Note that nr can be a square only if $r = ns^2$. According to (6.5) we have

$$K(n, n; cr) = K'(n\bar{c}, n\bar{c}; r) S(n\bar{r}, n\bar{r}; c).$$

On the right side the Kloosterman sum $K'(n\bar{c}, n\bar{c}; r)$ does depend on c , but only on its residue class modulo r . Hence

$$\begin{aligned} T(x) &= \sum_{r \leq x} r^{-\frac{1}{2}} \sum_{b \pmod{r}}^* K'(n\bar{b}, n\bar{b}; r) \sum_{\substack{c \leq x/r \\ (c, n)=1 \\ c \equiv b \pmod{r}}} c^{-\frac{1}{2}} S(n\bar{r}, n\bar{r}; c) \\ &= T'(x) + T''(x) + T'''(x) \end{aligned}$$

say, where $T'(x)$ is the partial sum over $r \leq R$, $T''(x)$ is the partial sum over $R < r \leq x$ such that nr is not a square and $T'''(x)$ is the partial sum over $R < r \leq x$ such that nr is a square.

For the sum of Salié sums in $T'(x)$ we use Theorem 7.1 and for the Kloosterman sum we use the trivial bound, getting

$$\begin{aligned} T'(x) &\ll \sum_{r \leq R} r^{\frac{3}{2}} r^{\frac{6}{5}} \left(\frac{x}{r} + n\right)^{\frac{47}{118}} \left(\frac{x}{r}\right)^{\frac{35}{59} + \varepsilon} \\ &\ll \left(1 + \frac{n}{x}\right)^{\frac{47}{118}} x^{1 - \frac{1}{118} + \varepsilon} \sum_{r \leq R} r^{1243/590}. \end{aligned}$$

Since, for $\alpha \geq 0$, $y \geq 1$ we have

$$\sum_{\substack{r \leq y \\ r|(nq)^\infty}} r^\alpha \leq y^{\alpha+\varepsilon} \prod_{p|nq} (1-p^{-\varepsilon})^{-1} \leq y^{\alpha+\varepsilon} \tau(nq) \prod_{p^\varepsilon < 2} (1-p^{-\varepsilon})^{-1}$$

with any $\varepsilon > 0$, we get

$$(9.4) \quad T'(x) \ll \tau(nq)x^\varepsilon \left(1 + \frac{n}{x}\right)^{\frac{47}{118}} x^{1-\frac{1}{118}} R^{1243/590} .$$

Next we estimate $T''(x)$ by using the following bounds for the individual sums: $|S(n\bar{r}, n\bar{r}; c)| \leq \tau(c)\sqrt{c}$ and $|K'(n\bar{b}, n\bar{b}; r)| \leq r$. We get

$$(9.5) \quad T''(x) \ll \sum_{R < r \leq x} r^{-\frac{1}{2}} x \log x \ll \tau(nq)x^{1+\varepsilon} R^{-\frac{1}{2}} .$$

Finally, we estimate $T'''(x)$ as follows:

$$T'''(x) \ll \sum_r''' r^{-\frac{1}{2}} r \sum_{c \leq x/r} \tau(c) \ll \left(\sum_r''' r^{-\frac{1}{2}}\right) x \log x .$$

Here \sum_r''' runs over $r = ns^2$ with $s|(nq)^\infty$. This gives

$$(9.6) \quad T'''(x) \ll n^{-\frac{1}{2}} \prod_{p|nq} \left(1 + \frac{1}{p}\right) x \log x .$$

We choose R which equalizes the bounds (9.4), (9.5):

$$(9.7) \quad R = \left(1 + \frac{n}{x}\right)^{-\frac{235}{1538}} x^{\frac{5}{1538}} .$$

This gives

$$T'(x) + T''(x) \ll \tau(nq)x^\varepsilon \left(1 + \frac{n}{x}\right)^{\frac{235}{2 \cdot 1538}} x^{1-\frac{5}{2 \cdot 1538}} \ll \left(1 + \frac{n}{x}\right)^{\frac{1}{13}} x^{1-\frac{1}{616}} .$$

Adding in (9.6) and sacrificing the exponents a bit, we obtain:

Lemma 9.1. *For n squarefree we have*

$$(9.8) \quad T(x) \ll \left(1 + \frac{n}{x}\right)^{\frac{1}{13}} x^{1-\frac{1}{616}} + n^{-\frac{1}{2}} x(\log x)^2 ,$$

where the implied constant is absolute.

We remark that, in the above arguments we tacitly assumed that $q \leq x$ and $1 < n \leq x^2$, as we may, since otherwise (9.8) is trivial.

By Lemma 9.1 we deduce (essentially by partial summation):

Lemma 9.2. *Let $n \geq 1$ squarefree and $q \equiv 0 \pmod{4}$. Let $F(x)$ be a smooth function on \mathbb{R}^+ such that*

$$(9.9) \quad |F(x)| \leq \min\left(x, x^{-\frac{1}{2}}\right) ,$$

$$(9.10) \quad |F'(x)| \leq \min\left(1, x^{-\frac{1}{2}}\right) .$$

Then, we have

$$(9.11) \quad \mathcal{K}_F(n, n) \ll n^{\frac{1}{2}-\frac{1}{665}} ,$$

where the implied constant is absolute.

Proof. We split

$$\mathcal{K}_F(n, n) = \sum_{c \equiv 0 \pmod{q}} c^{-1} K(n, n; c) F\left(\frac{4\pi n}{c}\right) = \sum_{c \leq C} + \sum_{c > C} ,$$

where $1 \leq C \leq n$ will be chosen later. In the first sum over $c \leq C$ we apply the individual estimates for $K(n, n; c)$ getting

$$(9.12) \quad \sum_{c \leq C} \ll \sum_{c \leq C} \tau(c) (n, c)^{\frac{1}{2}} c^{-\frac{1}{2}} (c/n)^{\frac{1}{2}} \ll \tau(n) n^{-\frac{1}{2}} C \log 2n .$$

For the sum over $c > C$ we write

$$\sum_{c > C} = \int_C^\infty x^{-\frac{1}{2}} F\left(\frac{4\pi n}{x}\right) dT(x)$$

and integrate by parts, showing this is bounded by

$$C^{-\frac{1}{2}} \left| F\left(\frac{4\pi n}{C}\right) T(C) \right| + \int_C^\infty \left(\frac{1}{x^{\frac{3}{2}}} \left| F\left(\frac{4\pi n}{x}\right) \right| + \frac{4\pi n}{x^{\frac{3}{2}}} \left| F'\left(\frac{4\pi n}{x}\right) \right| \right) |T(x)| dx .$$

In the segment $x \geq n$ the largest contribution comes from $x \asymp n$ (see (9.9) and (9.10)) which is of size

$$n^{-\frac{1}{2}} n^{1-\frac{1}{616}} = n^{\frac{1}{2}-\frac{1}{616}} .$$

In the segment $C < x < n$ we get by (9.9), (9.10), the bound

$$n^{\frac{1}{2}} \int_C^n x^{-2} |T(x)| dx .$$

Now the largest contribution comes from $x \asymp C$, which is of size

$$(9.13) \quad n^{\frac{1}{2}} C^{-1} \left(1 + \frac{n}{C}\right)^{\frac{1}{13}} C^{1-\frac{1}{616}} \asymp n^{\frac{1}{2}} \left(\frac{n}{C}\right)^{\frac{1}{13}} C^{-\frac{1}{616}} .$$

We choose C which equalizes (9.12), (9.13); precisely we take $C = n^{1-13/8637}$. This yields (9.11) because $13/8637 > 1/665$. \square

Note that the function $H(x)$ satisfies the conditions (9.9), (9.10) up to an absolute constant factor, so in particular Lemma 9.2 yields

$$(9.14) \quad \mathcal{K}_H(n, n) \ll n^{\frac{1}{2}-\frac{1}{665}} .$$

10 Proof of Theorem 1.2

Let $H(x)$ be the function given by (3.6), (4.19) for $k = \frac{1}{2}$, $0 < \alpha < \frac{1}{2}$, $\ell = \frac{5}{2}$, that is

$$(10.1) \quad H(x) = x^{-\alpha} J_{\frac{3}{2}}(x) \pm 2\pi c J_{\frac{3}{2}}(x) .$$

By the Cauchy inequality for sums of Kloosterman sums (3.7) and by estimates (8.5), (9.14) we obtain

$$\mathcal{K}_g(m, n) \ll \left(X^{-\frac{1}{4}} + X^{\frac{1}{2}+\alpha} \right) (mn)^{\frac{1}{4}} n^{-\frac{1}{1330}} \prod_{p|m} \left(1 + \frac{1}{\sqrt{p}}\right) \log 2m$$

(see the comments about the scaling factor (5.8)). This proves Theorem 1.2.

11 Proof of Theorem 1.1

By Lemma 6.2 we express the sum (1.2) as

$$\mathcal{W}_h(D) = \frac{1-i}{2} \sum_{uw=2h} \sum_{v \equiv 0 \pmod{(q,(q,w))}} f(vw)K(u^2, D; 4v)/(2, v)(2, vw)\sqrt{v} .$$

Split the sum over v into odd and even numbers and arrange as follows:

$$\begin{aligned} \sum_v &= \frac{1}{(2, w)} \sum_{v \equiv 0 \pmod{(q,(q,w))}} f(vw)K(u^2, D; 4v)/\sqrt{v} \\ &+ \left(\frac{1}{4} - \frac{1}{(2, w)}\right) \sum_{\substack{v \equiv 0 \pmod{(q,(q,w))} \\ v \text{ even}}} f(vw)K(u^2, D; 4v)/\sqrt{v} . \end{aligned}$$

Theorem 1.2 applies to both sums with the test functions f, g related by

$$\frac{f(yw)}{\sqrt{y}} = y^{-1}g\left(\frac{4\pi u\sqrt{D}}{4y}\right) .$$

Since $uw = 2h$, we have

$$f(y) = \left(\frac{w}{y}\right)^{\frac{1}{2}} g\left(\frac{2\pi h\sqrt{D}}{y}\right) .$$

Whereas $f(y)$ is supported on $Y \leq y \leq 2Y$, the function $g(x)$ is supported on $X \leq x \leq 2X$ with $X = \pi h\sqrt{D}Y^{-1}$. Moreover $g(x)$ satisfies (1.7) up to a scaling factor of size $(Y/w)^{\frac{1}{2}}$. Therefore, (1.8) gives

$$\sum_v \ll \left(\frac{Y}{w}\right)^{\frac{1}{2}} \left[\left(\frac{Y}{h\sqrt{D}}\right)^{\frac{1}{4}} + \left(\frac{h\sqrt{D}}{Y}\right)^{\frac{1}{2}+\varepsilon} \right] (u^2 D)^{\frac{1}{4}} D^{-\frac{1}{1330}} \tau(u^2) \log 2u .$$

Hence,

$$\begin{aligned} \mathcal{W}_h(D) &\ll Y^{\frac{1}{2}} \left[\left(\frac{Y}{h\sqrt{D}}\right)^{\frac{1}{4}} + \left(\frac{h\sqrt{D}}{Y}\right)^{\frac{1}{2}+\varepsilon} \right] h^{\frac{1}{2}} D^{\frac{1}{4}-\frac{1}{1330}} \tau(h^2) \log 2h \\ &\ll \left(Y + h\sqrt{D}\right)^{\frac{3}{4}} h^{\frac{1}{4}} D^{\frac{1}{8}-\frac{1}{1331}} , \end{aligned}$$

because we can assume that $h \leq D$, since otherwise (1.4) is trivial.

12 Proof of Theorem 1.3

Theorem 1.3 follows from Theorem 1.1 by Fourier analysis which is by now fairly standard. It clearly suffices to prove (12.2) when the counting function $N(x; \alpha, \beta)$ runs over the dyadic interval $x < c \leq 2x$ in place of $1 \leq c \leq x$, which modification we assume for the proof without changing the notation.

To accelerate the convergence of Fourier series we first consider a smoothed version of $N(x; \alpha, \beta)$, namely

$$N(f; \phi) = \sum_{c \equiv 0 \pmod{q}} f(c) \sum_{\substack{b \pmod{c} \\ b^2 \equiv D \pmod{c}}} \phi(b/c) ,$$

where $f(y)$ is as in Theorem 1.1 with $Y \asymp x$ and $\phi(t)$ is a smooth compactly supported function, such that

$$\int |\phi'''(t)| dt \leq 1 .$$

Therefore, its Fourier transform

$$\hat{\phi}(h) = \int \phi(t) e(ht) dt$$

satisfies $|\hat{\phi}(h)| \leq (2\pi|h|)^{-3}$, on integrating by parts three times. Applying Poisson summation and then (1.4), we derive

$$\begin{aligned} N(f; \phi) &= \sum_h \hat{\phi}(h) \sum_{c \equiv 0 \pmod{q}} f(c) W_h(D; c) \\ &= \hat{\phi}(0) \sum_{c \equiv 0 \pmod{q}} f(c) \rho(c) + O\left((x + \sqrt{D})^{\frac{3}{4}} D^{\frac{1}{8} - \frac{1}{1331}}\right) , \end{aligned}$$

where $\rho(c)$ is the number of roots $b \pmod{c}$ of the congruence $b^2 \equiv D \pmod{c}$.

Next, we require the sum of $f(c)\rho(c)$. We shall evaluate this by means of the L -function

$$L(s) = \sum_{c \equiv 0 \pmod{q}} \rho(c) c^{-s} ,$$

which can be expressed in terms of the zeta-function of the field $\mathbb{Q}(\sqrt{D})$. We postpone the relevant computations to the end of this section. By (12.3) and a standard contour integration we derive that

$$\sum_{c \equiv 0 \pmod{q}} f(c)\rho(c) = \hat{f}(0)R + O(x^{\frac{1}{2}}D^{\frac{3}{16}+\varepsilon}) ,$$

where R is the residue of $L(s)$ given by (12.4) and the error term comes from the Burgess bound for $L(s, \chi_D)$ on the critical line. Today, better bounds are available (see [CI]), but these are not needed for our purpose here. However, the convexity bound $L(s, \chi_D) \ll (|s|D)^{1/4}$ would not be sufficiently strong. We conclude that

$$(12.1) \quad N(f; \phi) = \hat{f}(0)\hat{\phi}(0)R + O((x + \sqrt{D})^{\frac{3}{4}}D^{\frac{1}{8}-\frac{1}{1331}}) .$$

Now, we are going to replace the smoothing factors by sharp cuts. First, let $N(f; [\alpha, \beta])$ denote the sum in which $\phi(b/c)$ is replaced by the condition $\alpha \leq b/c \leq \beta$. The characteristic function of the interval $[\alpha, \beta]$ can be majorized and minorized by functions $\phi(t)$ such that

$$\begin{aligned} \hat{\phi}(0) &= \int \phi(t) dt = \beta - \alpha + O(\Delta) , \\ \int |\phi'''(t)| dt &\ll \Delta^{-3} , \end{aligned}$$

where $\Delta > 0$ is at our disposal. Hence, re-normalizing $\phi(t)$ by the factor Δ^3 , we find

$$N(f; [\alpha, \beta]) = \hat{f}(0)(\beta - \alpha + O(\Delta)) R + O(\Delta^{-3}(x + \sqrt{D})x^{-2/1331}) .$$

Note that $\hat{f}(0) \ll x$. We choose $\Delta = x^{-1/2662}(1 + \sqrt{D}/x)$ getting

$$N(f; [\alpha, \beta]) = \hat{f}(0)(\beta - \alpha) R + O(\Delta x \log D) .$$

Next, the characteristic function of the interval $(x, 2x]$ can be majorized and minorized by functions $f(y)$ such that

$$\hat{f}(0) = \int f(y) dy = x + O(z) ,$$

$$f''(y) \ll z^{-2}, \quad \text{supp } f \subset [\frac{1}{2}x, 4x],$$

where $0 < z \leq x$ is at our disposal. Hence, re-normaizing $f(y)$ by the factor $(z/x)^2$ we obtain

$$N(x; \alpha, \beta) = (x + O(z)) (\beta - \alpha)R + O((x/z)^2 \Delta x \log D) .$$

We choose $z = x\Delta^{1/3}$ (the condition $z \leq x$ can be ignored by a trivial estimation), getting

$$(12.2) \quad N(x; \alpha, \beta) = (\beta - \alpha) R x + O(x^{1-1/7986}(1 + \sqrt{D}/x)^{1/12} \log D) .$$

This completes the proof of Theorem 1.3, modulo the demonstration of (12.3) and (12.4).

12.1 Computing $L(s)$

We now give the computations on $L(s)$ required for the above proof. These will be needed as well in the next section. Recall that

$$L(s) = \sum_{c \equiv 0 \pmod{q}} \rho(c) c^{-s} ,$$

with $\rho(c)$ being the number of roots $b \pmod{c}$ of $b^2 \equiv D \pmod{c}$, where D is an odd, positive, fundamental discriminant and $\chi = \chi_D = (D/n)$.

We have:

$$\begin{aligned} \rho(p^\alpha) &= 1 + \chi(p) \text{ for } \alpha \geq 1 \text{ if } p \nmid 2D , \\ \rho(p) &= 1, \quad \rho(p^\alpha) = 0 \text{ for } \alpha \geq 2 \text{ if } p \mid D , \\ \rho(4) &= 2, \quad \rho(2^\alpha) = 2(1 + \chi(2)) \text{ for } \alpha \geq 3 . \end{aligned}$$

Let $4 \mid q$ and $(q, D) = 1$. Then, the local factors of $L(s)$ are:

For $p \mid D$,

$$L_p(s) = 1 + p^{-s} = (1 - p^{-2s})(1 - p^{-s})^{-1} ;$$

for $p \nmid qD$,

$$L_p(s) = 1 + (1 + \chi(p))p^{-s}(1 - p^{-s})^{-1} = (1 + \chi(p)p^{-s})(1 - p^{-s})^{-1} ;$$

for $p^\beta || q$, $p \neq 2$, $\beta \geq 1$,

$$L_p(s) = (1 + \chi(p))p^{-\beta s}(1 - p^{-s})^{-1} ;$$

for $2^\beta || q$, $\beta \geq 3$,

$$L_2(s) = 2(1 + \chi(2))2^{-\beta s}(1 - 2^{-s})^{-1} ;$$

for $2^2 || q$,

$$L_2(s) = 2^{1-2s} + 2(1 + \chi(2))2^{-3s}(1 - 2^{-s})^{-1} = 2^{1-2s}(1 + \chi(2)2^{-s})(1 - 2^{-s})^{-1} .$$

Hence, we have

$$L(s) = 2\zeta(s)q^{-s} \prod_{p|D} (1 - p^{-2s}) \prod_{p \nmid 4} (1 + \chi(p)p^{-s}) \prod_{p|q}''' (1 + \chi(p)) ,$$

where the superscript $'''$ indicates that the factor $1 + \chi(2)$ must be replaced $1 + \chi(2)2^{-s}$ if $4 || q$. Hence

$$L(s) = \frac{2\zeta(s)L(s, \chi)}{q^s \zeta(2s)} \prod_{p|q} (1 + \chi(p)p^{-s})^{-1} \prod_{p|q}''' (1 + \chi(p)) .$$

Moreover, we have

$$\rho(q) = 2 \prod_{p|q}'' (1 + \chi(p)) ,$$

where the superscript $''$ indicates that the factor $1 + \chi(2)$ must be omitted if $4 || q$. Hence, we conclude:

Lemma 12.2. *Let $D > 1$, $D \equiv 1 \pmod{4}$, D squarefree, $4|q$, $(q, D) = 1$. Then,*

$$(12.3) \quad L(s) = \frac{\zeta(s)L(s, \chi)}{\zeta(2s)} \frac{\rho(q)}{q^s} \prod_{p|q}' (1 + \chi(p)p^{-s})^{-1} ,$$

where the superscript $'$ indicates that the factor $1 + \chi(2)2^{-s}$ must be omitted if $4 || q$.

It is clear from (12.3) that $L(s)$ is holomorphic in $\text{Res} \geq \frac{1}{2}$ except for a simple pole at $s = 1$ with residue

$$R = \text{res}_{s=1} L(s) = \frac{6}{\pi^2} L(1, \chi_D) \frac{\rho(q)}{q} \prod'_{p|q} \left(1 + \frac{\chi(p)}{p}\right)^{-1}.$$

Naturally, we want $\rho(q) \neq 0$. This means that:

$$D \equiv 1 \pmod{8} \text{ if } 8|q,$$

and

$$\chi(p) = 1 \text{ if } p|q, p \neq 2.$$

Assuming these conditions we find that

$$(12.4) \quad R = \text{res}_{s=1} L(s) = \frac{3}{\pi^2} L(1, \chi_D) c_q 2^{\omega(q)} q^{-1} \prod'_{p|q} \left(1 + \frac{1}{p}\right)^{-1},$$

where $c_q = 2 + (2, q/4)$, that is $c_q = 3$ if $4||q$ and $c_q = 4$ if $8|q$.

13 Distribution of Roots of Binary Quadratic Forms

Now we proceed to the proofs of the asymptotic formula (1.18) and of Theorem 1.4. These are really corollaries of the results in Section 12. Since the arguments are very much the same, we shall be brief.

First, let us modify slightly the asymptotic formula (12.2). For a smooth function $W(u, v)$ compactly supported on $\mathbb{R}^+ \times \mathbb{R}^+$ we consider the sum

$$(13.1) \quad \mathcal{W}(D) = \sum_{b^2 - 4ac = D} W\left(\frac{a}{\sqrt{D}}, \frac{b}{\sqrt{D}}\right).$$

Writing this in the form

$$\mathcal{W}(D) = \sum_a \sum_{b^2 \equiv D \pmod{4a}} W\left(\frac{a}{\sqrt{D}}, \frac{b}{\sqrt{D}}\right),$$

we see a similarity to $N(f; \phi)$ in Section 12. In fact both sums agree for the test function

$$W(u, v) = f(4u\sqrt{D})\phi(v/4u).$$

Of course, not every $W(u, v)$ can be factored as above, nevertheless the result (12.2) remains true with the main term $\hat{f}(0)\hat{\phi}(0)$ replaced by the integral

$$\iint W(y/4\sqrt{D}, ty/\sqrt{D}) dy dt = W\sqrt{D},$$

where

$$(13.2) \quad W = \iint W(u, v)u^{-1} du dv.$$

In our case $q = 4$, so the residue given by (12.4) becomes

$$(13.3) \quad R = 3\pi^{-2}L(1, \chi_D)$$

and (12.2) becomes

$$(13.4) \quad \mathcal{W}(D) = RW\sqrt{D} + O(D^{\frac{1}{2} - \frac{1}{1331}}).$$

It is important that we have saved a power of D in the error term of (13.4) because, up to the current state of knowledge, we can only claim the lower bound (due to C.L. Siegel) $L(1, \chi_D) \gg D^{-\varepsilon}$. Thus, the asymptotic formula (13.4) is meaningful; it implies

$$(13.5) \quad \mathcal{W}(D) \sim \frac{3}{\pi^2}W\sqrt{D}L(1, \chi_D), \quad \text{as } D \rightarrow \infty,$$

provided $W \neq 0$.

Next, notice that $|\mathcal{H}(D; \lambda)|$ is just the sum (13.1) for the test function $W(u, v)$ being the characteristic function of the set

$$0 < v < 1, \quad 1 - v < 2u < 1 + v, \quad 1 < \frac{v+1}{2u} < \lambda.$$

These conditions can be written in the following form:

$$(13.6) \quad |2u - 1| < v < \min(1, 2\lambda u - 1).$$

Note that the set (13.6) is not empty if and only if

$$(13.7) \quad \frac{1}{\lambda+1} < u < 1.$$

Because the characteristic function of the set (13.6) is not smooth, one cannot apply (13.5) directly to it. However, we can smooth it out by modifying (increasing or decreasing) it within a set of asymptotically shrinking measure. This gives us (by positivity) an upper and a lower bound for $|\mathcal{H}(D; \lambda)|$, both asymptotically equal. Hence the formula (13.5) is justified also for the characteristic function of (13.6). In this case we compute (13.2) as follows:

$$\begin{aligned} W &= \int_{1/(\lambda+1)}^1 [\min(1, 2\lambda u - 1) - |2u - 1|] \frac{du}{u} \\ &= \int_{1/\lambda}^1 \frac{du}{u} + \int_{1/(\lambda+1)}^{1/\lambda} (2\lambda u - 1) \frac{du}{u} - \int_{1/2}^1 (2u - 1) \frac{du}{u} - \int_{1/(\lambda+1)}^{1/2} (1 - 2u) \frac{du}{u} \\ &= \log \lambda + 2\lambda \left(\frac{1}{\lambda} - \frac{1}{\lambda+1} \right) - \log \frac{\lambda+1}{\lambda} - 1 + \log 2 - \log \frac{\lambda+1}{2} + 1 - \frac{2}{\lambda+1}. \end{aligned}$$

Hence,

$$W = 2 \log \frac{2\lambda}{\lambda+1} \rightarrow 2 \log 2 \quad \text{as } 1 < \lambda \rightarrow \infty.$$

This completes the proof of (1.18) and (1.20).

14 Application to Cycle Integrals

The proof of Theorem 1.5 is based on the following identity (1.24) of [DIT]:

$$Tr_D(j_m) = -24 \sigma_1(m) Tr_D(1) + D^{-1/2} Z_m(D)$$

where

$$(14.1) \quad Z_m(D) = \sum_{c \equiv 0 \pmod{4}} W_{2m}(D; c) \sin\left(\frac{4\pi m \sqrt{D}}{c}\right)$$

and the question of a non-trivial bound was raised. This would be a special case of our sum (1.2) for

$$(14.2) \quad f(y) = \sin\left(\frac{4\pi m\sqrt{D}}{y}\right),$$

except that such a function is oscillatory and not compactly supported. To overcome this obstacle we use a smooth partition of unity into dyadic segments. In fact, for $c \leq C$ with an appropriate choice, we estimate trivially, getting

$$Z_m(D) \ll C \log C + \sum_{\ell=0}^{\infty} \left| \sum_{c \equiv 0 \pmod{4}} W_{2m}(D; c) f_\ell(c) \right|,$$

where $f_\ell(y) = f(y)F_\ell(y)$, and $F_\ell(y)$ is a constituent of the partition, supported on $Y_\ell < y < 2Y_\ell$ with $Y_\ell = 2^{\ell/2}C$. Note that

$$y^j f_\ell^{(j)}(y) \ll \frac{m\sqrt{D}}{Y_\ell} \left(1 + \frac{m\sqrt{D}}{Y_\ell}\right)^2, \quad j = 1, 2,$$

so our Theorem 1.1 applies with $f_\ell(y)$ normalized by the above bounding factor. Therefore, we obtain

$$\begin{aligned} Z_m(D) &\ll C \log C + \sum_{\ell=0}^{\infty} \frac{m\sqrt{D}}{Y_\ell} \left(1 + \frac{m\sqrt{D}}{Y_\ell}\right)^2 (Y_\ell + m\sqrt{D})^{\frac{3}{4}} m^{\frac{1}{4}} D^{\frac{1}{8} - \frac{1}{1331}} \\ &\ll C \log C + \frac{m\sqrt{D}}{C} \left(1 + \frac{m\sqrt{D}}{C}\right)^2 (C + m\sqrt{D})^{\frac{3}{4}} m^{\frac{1}{4}} D^{\frac{1}{8} - \frac{1}{1331}}. \end{aligned}$$

We choose $C = m\sqrt{D} D^{-1/4 \cdot 1331}$, getting

Theorem 14.1. *Let $m \geq 1$, $D \geq 1$, $D \equiv 1 \pmod{4}$, D squarefree. Then*

$$(14.3) \quad Z_m(D) \ll m D^{\frac{1}{2} - \frac{1}{5325}},$$

where the implied constant is absolute.

15 General Cusps and Prime Points

In the previous sections we have purposely chosen to present our main results in somewhat less generality than we might have, in order to avoid some

ancillary arguments which would have impeded the flow of the main ideas. Although sufficiently broad to capture our previous applications, for the problem of prime points on the sphere we need some further developments. In this final section we present the new arguments needed to complete those results. These new ingredients are mainly of an arithmetic nature whereas the analytic aspects remain unchanged.

We now consider two cusps \mathfrak{a} , \mathfrak{b} of Γ which are singular for the multiplier ϑ . Let $\sigma_{\mathfrak{a}}$, $\sigma_{\mathfrak{b}}$ be the corresponding scaling matrices (these are unique up to translations $\begin{pmatrix} 1 & t \\ & 1 \end{pmatrix}$, $t \in \mathbb{R}$, from the right side). Let c run over positive real numbers such that $\begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1} \Gamma \sigma_{\mathfrak{b}}$. For every such c and integers n , m the Kloosterman sum is defined by

$$(15.1) \quad K_{\mathfrak{ab}}(n, m; c) = \sum_{\tau} \bar{\vartheta}(\sigma_{\mathfrak{a}} \tau \sigma_{\mathfrak{b}}^{-1}) e\left(\frac{\alpha n + \delta m}{c}\right),$$

where $\tau = \begin{pmatrix} \alpha & * \\ c & \delta \end{pmatrix}$ runs over the double cosets $B \setminus \sigma_{\mathfrak{a}}^{-1} \Gamma \sigma_{\mathfrak{b}} / B$ with respect to B (the group of translations by integers). This means that α/c and δ/c run modulo one.

The Fourier expansions of an automorphic form f at cusp \mathfrak{a} is just the Fourier–Whittaker series for $f(\sigma_{\mathfrak{a}} z)$; it runs over integer frequencies and has the same shape at every singular cusp, but of course the coefficients are different, depending on \mathfrak{a} and the choice of $\sigma_{\mathfrak{a}}$. In fact, the choice of $\sigma_{\mathfrak{a}}$ has only a mild effect. Precisely, the Fourier coefficients change by twisting with an additive character $e(\xi n)$, $\xi \in \mathbb{R}$, while the Kloosterman sum changes by a factor $e(\nu n + \mu m)$ with ν , μ real numbers independent of the varying modulus c .

Now, for positive integers n , m we consider

$$(15.2) \quad \mathcal{K}_{\mathfrak{ab}g}(n, m) = \sum_c c^{-1} K_{\mathfrak{ab}}(n, m; c) g\left(\frac{4\pi\sqrt{nm}}{c}\right)$$

which is a generalization of the sum (2.1). The summation formula (2.27) now holds for $\mathcal{K}_{\mathfrak{ab}g}(n, m)$ with the spectral sums \mathcal{L} , \mathcal{M} , \mathcal{N} having the Fourier coefficients of the basic automorphic forms expanded at cusps \mathfrak{a} , \mathfrak{b} respectively. Therefore, the Cauchy inequality (3.7) generalizes as follows:

$$(15.3) \quad |\mathcal{K}_{\mathfrak{ab}g}(m, n)|^2 \leq \left(|\mathcal{K}_{\mathfrak{aa}H}(n, n)| + c \right) \left(|\mathcal{K}_{\mathfrak{bb}H}(m, m)| + c \right),$$

where the function $H(x)$ and the constant c are the same as in (3.7).

In Sections 8 and 9 we specialized to the congruence group $\Gamma = \Gamma_0(q)$ and the cusp $\mathfrak{a} = \mathfrak{b} = \infty$. In this section we take the level to be

$$(15.4) \quad q = rs \quad \text{with } (r, s) = 1, 4|s$$

and the cusps $\mathfrak{a} = \infty$, $\mathfrak{b} = 1/s$ with scaling matrices

$$(15.5) \quad \sigma_{\mathfrak{a}} = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad \sigma_{\mathfrak{b}} = \begin{pmatrix} \sqrt{r} & \\ s\sqrt{r} & 1/\sqrt{r} \end{pmatrix}.$$

For background material, see Section 2.4 of [I2], Section 4.2 of [I3] and Section 2 of [DI].

For the pair of cusps \mathfrak{a} , \mathfrak{a} we sum over the double cosets of our group $\sigma_{\mathfrak{a}}^{-1} \Gamma \sigma_{\mathfrak{a}} = \Gamma$ getting the standard Kloosterman sum (6.3)

$$(15.6) \quad K_{\mathfrak{a}\mathfrak{a}}(n, n; c) = K(n, n; c)$$

of modulus $c = \gamma q$ with γ any positive integer. For the pair of cusps \mathfrak{b} , \mathfrak{b} we sum over the double cosets of the conjugate group

$$\sigma_{\mathfrak{b}}^{-1} \Gamma \sigma_{\mathfrak{b}} = \left\{ \begin{pmatrix} \alpha & \beta/r \\ \gamma q & \delta \end{pmatrix}; \alpha, \beta, \gamma, \delta \in \mathbb{Z} \right\}$$

with $\alpha, \beta, \gamma, \delta$ restricted by

$$\alpha\delta - \beta\gamma s = 1, \quad (\alpha + \gamma)(\delta - \gamma) \equiv 1 \pmod{r}.$$

Hence, the Kloosterman sum $K_{\mathfrak{b}\mathfrak{b}}(m, m; c)$ also has modulus $c = \gamma q$. This sum satisfies the same bound (6.8) as does the standard Kloosterman sum $K(m, m; c)$.

By the above observations it follows that the bound (9.14) remains true for $\mathcal{K}_{\mathfrak{a}\mathfrak{a}H}(n, n)$ and the bound (8.5) remains true for $K_{\mathfrak{b}\mathfrak{b}}(m, m; c)$. Consequently, by Cauchy's inequality (15.3) we conclude that the bound (1.8) remains true for $\mathcal{K}_{\mathfrak{a}\mathfrak{b}g}(n, m)$ under exactly the same conditions as in Theorem 1.2.

Next, we are going to express the Kloosterman sums $K_{\mathfrak{a}\mathfrak{b}}(n, m; c)$ for the pair of cusps \mathfrak{a} , \mathfrak{b} in terms of the standard Kloosterman sums. Here, we sum

over double cosets of the set (not a group)

$$\sigma_{\mathbf{a}}^{-1} \Gamma \sigma_{\mathbf{b}} = \left\{ \begin{pmatrix} \alpha\sqrt{r} & \beta/\sqrt{r} \\ \gamma s\sqrt{r} & \delta/\sqrt{r} \end{pmatrix}; \alpha, \beta, \gamma, \delta \in \mathbb{Z} \right\}$$

with $\alpha, \beta, \gamma, \delta$ restricted by

$$\alpha\delta - \beta\gamma s = 1, \quad \delta \equiv \gamma \pmod{r}.$$

Hence, the Kloosterman sum $K_{\mathbf{ab}}(n, m; c)$ has modulus $c = \gamma s\sqrt{r}$ with γ any positive integer, $(\gamma, r) = 1$. For τ in $\sigma_{\mathbf{a}}^{-1} \Gamma \sigma_{\mathbf{b}}$ we compute

$$\sigma_{\mathbf{a}} \tau \sigma_{\mathbf{b}}^{-1} = \begin{pmatrix} \alpha\sqrt{r} & \beta/\sqrt{r} \\ \gamma s\sqrt{r} & \delta/\sqrt{r} \end{pmatrix} \begin{pmatrix} 1/\sqrt{r} & \\ -s\sqrt{r} & \sqrt{r} \end{pmatrix} = \begin{pmatrix} \alpha - \beta s & \beta \\ (\gamma - \delta)s & \delta \end{pmatrix}.$$

Hence, for our theta multiplier we get

$$\vartheta(\sigma_{\mathbf{a}} \tau \sigma_{\mathbf{b}}^{-1}) = \bar{\varepsilon}_{\delta} \left(\frac{(\gamma - \delta)s}{\delta} \right) = \bar{\varepsilon}_{\delta} \left(\frac{\gamma}{\delta} \right).$$

Given $c = \gamma s\sqrt{r}$ with $(\gamma, r) = 1$ we have

$$K_{\mathbf{ab}}(n, m; c) = \sum_{\substack{\delta \pmod{\gamma sr} \\ \delta \equiv \gamma \pmod{r}}}^* \varepsilon_{\delta} \left(\frac{\gamma}{\delta} \right) e \left(\frac{\bar{\delta}n}{\gamma s} + \frac{\delta m}{\gamma sr} \right),$$

because $\alpha\delta \equiv 1 \pmod{\gamma s}$. Write $\delta \equiv \gamma s\bar{s} + xr\bar{r}$, where x runs over primitive classes modulo γs . Then $\varepsilon_{\delta} = \varepsilon_x$, $(\gamma/\delta) = (\gamma/x)$, $\bar{\delta} \equiv \bar{x} \pmod{\gamma s}$. Hence,

$$K_{\mathbf{ab}}(n, m; c) = e \left(\frac{m\bar{s}}{r} \right) \sum_{x \pmod{\gamma s}}^* \varepsilon_x \left(\frac{\gamma}{x} \right) e \left(\frac{\bar{x}n}{\gamma s} + \frac{x\bar{r}m}{\gamma s} \right).$$

The last sum is the Kloosterman sum $K(n, m\bar{r}; \gamma s)$. Hence, by the symmetry $K(a, b; c) = K(b, a; c)$, we have

$$K_{\mathbf{ab}}(n, m; c) = e \left(\frac{m\bar{s}}{r} \right) K(m\bar{r}, n; \gamma s).$$

Therefore, our sum (15.2) becomes

$$(15.7) \quad \mathcal{K}_{\mathbf{abg}}(n, m) = \frac{1}{\sqrt{r}} e \left(\frac{m\bar{s}}{r} \right) \sum_{\substack{c \equiv 0 \pmod{s} \\ (c, r) = 1}} c^{-1} K(m\bar{r}, n; \gamma c) g \left(\frac{4\pi\sqrt{mn}}{c\sqrt{r}} \right).$$

To conform to our earlier style we remove the factor \sqrt{r} in the argument of the test function $g(x)$ by re-scaling its support. Then, we conclude the following generalization of Theorem 1.2.

Theorem 15.1. *Let $m \geq 1$, $n \geq 1$, $s \geq 4$, $4|s$, $(r, s) = 1$, n squarefree. Let $g(x)$ be a smooth function supported on $X \leq x \leq 2X$ with $X > 0$ such that*

$$|g(x)| \leq 1, \quad x^2|g''(x)| \leq 1.$$

Then we have

$$(15.8) \quad \sum_{\substack{c \equiv 0 \pmod{s} \\ (c,r)=1}} c^{-1} K(m\bar{r}, n; \gamma c) g\left(\frac{4\pi\sqrt{mn}}{c}\right) \\ \ll r^{\frac{5}{8}} \left(X^{-\frac{1}{4}} + X^{\frac{1}{2}+\varepsilon}\right) (mn)^{\frac{1}{4}} n^{-1/1330} \tau(m) \log 2m,$$

where the implied constant depends only on ε .

The parameters r, s in (15.8) offer numerous possibilities for modifications and applications, especially to sums of Weyl sums. In a fashion similar to the derivation (in Section 11) of Theorem 1.1 from Theorem 1.2, we can derive its following generalization.

Theorem 15.2. *Let $h, r, s \geq 1$, $(r, s) = 1$, $D \geq 1$, $D \equiv 1 \pmod{4}$, D squarefree. Let $f(y)$ be a smooth function supported on $Y \leq y \leq 2Y$ with $Y \geq 1$, such that*

$$(15.9) \quad |f(y)| \leq 1, \quad y^2|f''(y)| \leq 1.$$

Then

$$(15.10) \quad \sum_{\substack{c \equiv 0 \pmod{s} \\ (c,r)=1}} f(c) W_{h\bar{r}}(D; c) \ll h^{\frac{1}{4}} r^{\frac{5}{4}} (Y + h\sqrt{D})^{\frac{3}{4}} D^{\frac{1}{8} - \frac{1}{1331}},$$

where the implied constant is absolute.

It is obvious that Theorem 15.2 implies Proposition 14.16 of [FI] and hence Theorem 1.6.

16 Negative Discriminants

Our assumption that D is positive is not at all essential. Most of the arguments in this paper work as well for negative fundamental discriminants. Actually, some of the arguments are simpler in their analytic aspects. For $m, n \geq 1$ we consider

$$\mathcal{K}_g(m, -n) = \sum_{c \equiv 0 \pmod{q}} c^{-1} K(m, -n; c) g\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

where $g(x)$ is the same test function as before. Then, a spectral decomposition like (2.27) holds but missing the last term which comes from the contribution of the holomorphic forms. Of course, the integral transform $\hat{g}(t)$ needs to be replaced by another transform, say $\check{g}(t)$ which is of the same kind and satisfies similar estimates (see [DI] where $\mathcal{K}_g(m, -n)$ is considered for $k = 0$). Now the Cauchy inequality (3.7) becomes

$$|\mathcal{K}_g(m, -n)|^2 \leq |\gamma_k \mathcal{K}_H(m, m) + c| |\gamma_k \mathcal{K}_H(-n, -n) + c|.$$

Here $H(x)$ is given by (3.6) with $G(x)$ being a majorant of $g(x)$ in the sense that $|\check{g}(t)| \leq \hat{G}(t)$ in place of (3.3) and $\check{G}(\ell) \geq 0$ in place of (3.4), except for one $\ell \geq 2, \ell \equiv k \pmod{2}$ for which $|\check{G}(\ell)|(\ell - 1) \leq c$ in place of (3.5). Note that, although the holomorphic cusp forms do not appear in the spectral decomposition of $\mathcal{K}_g(m, -n)$, one still needs them for $\mathcal{K}_H(m, m)$ and $\mathcal{K}_H(-n, -n)$.

The other parts of the work need only a little attention when tracking the sign changes. Having gone through these points one arrives at the following statements.

STATEMENT A: Assuming the conditions of Theorem 1.2, the sum $\mathcal{K}_g(m, -n)$ satisfies the bound (1.8).

STATEMENT B: Let $h \geq 1, q \geq 1$ and D be a negative fundamental discriminant. Let $f(y)$ be as in (1.3). Then the sum of Weyl sums (1.2) satisfies (1.4), but with D replaced by $|D|$.

Acknowledgements: Research of W. D. is supported in part by NSF Grant DMS-10-01527, that of J. F. by NSERC grant A5123, and that of H. I. by NSF Grant DMS-08-02246. The authors express thanks to the Institute for Advanced Study in Princeton where, during its March 2010 Workshop on Analytic Number Theory, the groundwork for the project was laid.

References

- [Bu] D. A. Buell, *Binary Quadratic Forms: Classical Theory and Modern Computations*, Springer-Verlag (New York), 1989.
- [By] V. A. Bykovskii, Spectral expansions of certain automorphic functions and their number-theoretic applications, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI)*, 134, (1984) 15–33 (translated in *J. Soviet Math*, 36, (1987) 8–21).
- [Ch] T. Chelluri, *Equidistribution of Roots of Quadratic Congruences*, thesis, Rutgers University (New Brunswick) 2004.
- [CI] J. B. Conrey and H. Iwaniec, The cubic moment of central values of automorphic L -functions, *Ann. Math.*, 151, (2000) 1175–1216.
- [DI] J.-M. Deshouillers and H. Iwaniec, Kloosterman sums and Fourier coefficients of cusp forms, *Invent. Math.*, 70 (1982/83) 219–288.
- [D1] W. Duke, Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.*, 92, (1988) 73–90.
- [D2] W. Duke, Modular functions and the uniform distribution of CM points, *Math. Ann.*, 334, (2006) 241–252.
- [DFI1] W. Duke, J. B. Friedlander and H. Iwaniec, Equidistribution of roots of a quadratic congruence to prime moduli, *Ann. Math.*, 141, (1995) 423–441.
- [DFI2] W. Duke, J. B. Friedlander and H. Iwaniec, Bilinear forms with Kloosterman fractions, *Invent. Math.*, 128, (1997) 23–43.

- [DFI3] W. Duke, J. B. Friedlander and H. Iwaniec, The subconvexity problem for Artin L -functions, *Invent. Math.*, 149, (2002) 489–577.
- [DIT] W. Duke, Ö. Imamoglu and Á. Tóth, Cycle integrals of the j -function and mock modular forms, *Ann. Math.*, to appear.
- [DS] W. Duke and R. Schulze-Pillot, Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids, *Invent. Math.*, 99, (1990) 49–57.
- [FI] J. B. Friedlander and H. Iwaniec, *Opera de Cribro*, Colloq. Pub. 57 Amer. Math. Soc. (Providence) 2010.
- [GR] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 4th ed., Academic Press, (New York) 1965.
- [He] D. A. Hejhal, The Selberg Trace Formula for $\mathrm{PSL}(2, R)$ (I), *Lecture Notes in Mathematics*, Vol. 548 Springer-Verlag, (Berlin) 1976, (II), *ibid.* Vol. 1001, 1983.
- [I1] H. Iwaniec, Fourier coefficients of modular forms of half-integral weight, *Invent. Math.*, 87 (1987) 385–401,
- [I2] H. Iwaniec, *Spectral methods of automorphic forms* (2nd ed.), *Grad. Studies in Math.*, 53, Amer. Math. Soc., (Providence) 2002.
- [I3] H. Iwaniec, *Topics in classical automorphic forms*, *Grad. Studies in Math.*, 17, Amer. Math. Soc., (Providence) 1997.
- [IK] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Colloq. Pub. 53 Amer. Math. Soc. (Providence) 2004.
- [Ku] N. V. Kuznetsov, The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums, *Mat. Sb.*, 111 (153), (1980) 334–383.
- [Li] Yu. V. Linnik, *Ergodic Properties of Algebraic Fields*, *Ergeb. Math. Grenz.* 45, Springer-Verlag (New York) 1968.
- [Ma] H. Maass, Über eine neue Art von nichtanalytischen automorphen Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen, *Math. Ann.*, 121, (1949), 141–183.

- [Ms] R. Masri, The asymptotic distribution of traces of cycle integrals of the j -function, preprint (2010).
- [MV] P. Michel and A. Venkatesh, Equidistribution, L -functions and ergodic theory. Proc ICM Madrid 2006, Vol. II, 421–457.
- [Pr] N. V. Proskurin, On general Kloosterman sums, Zap. Nauchn. Sem. St. Petersburg. Otdel. Mat. Inst. Steklov. (POMI), 302, (2003) 107–134, 200.
- [Ro] W. Roelcke, Über die Wellengleichung bei Grenzkreisgruppen erster Art, S.-B. Heidelberger Akad. Wiss. Math.-Nat. Kl, (1953/55) 159–267.
- [Sa] P. Sarnak, Class numbers of indefinite binary quadratic forms, J. Number Theory, 15, (1982) 229–247.
- [Se] A. Selberg, Collected Papers. Vol. I, Springer-Verlag (Berlin) 1989.
- [Sk] B. F. Skubenko, The asymptotic distribution of integers on a hyperboloid of one sheet and ergodic theorems, Izv. Akad. Nauk SSSR, Ser. Mat. 26, (1962) 721–752.
- [TV] A. I. Vinogradov and L. A. Tahtadzjan, Linnik-Skubenko asymptotics, Dokl. Akad. Nauk SSSR, 253, (1980) 777–780.
- [We] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, Math. Ann., 77, (1916) 313–352.

Department of Mathematics, UCLA, Los Angeles CA 90095, USA

Department of Mathematics, University of Toronto, M5S 2E4 Canada

Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA