Invent. math. 143, 221–248 (2001)
Digital Object Identifier (DOI) 10.1007/s002220000104

Inventiones mathematicae

Bounds for automorphic L–functions. III

W. Duke[1,⋆], **J.B. Friedlander**[2,⋆⋆], **H. Iwaniec**[1,⋆⋆⋆]

[1] Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA
[2] Department of Mathematics, University of Toronto, 100 St. George Street, M5S 3G3 Canada

Oblatum 27-III-2000 & 25-V-2000
Published online: 11 October 2000 –

1 Introduction

We continue our study of GL_2 L–functions with the aim of providing upper bounds for their order of magnitude. As is familiar it suffices to provide such bounds on the critical line and, both for the sake of applications and for the ideas involved, we are most interested in breaking the convexity bound and this with respect to the conductor. In this paper we are interested primarily in L–functions attached to characters of the class group of the imaginary quadratic field $K = \mathbb{Q}(\sqrt{-D})$. We are motivated by our paper [DFI4]. That work was not included in the current series because the class group L–functions are treated there directly. They may however be viewed as L–functions associated to cusp forms of weight 1, level D and character (the nebentypus)

$$\chi_{\mathrm{D}}(n) = \left(\frac{-D}{n}\right) , \tag{1.1}$$

the Kronecker symbol (we assume throughout that $-D$ is a fundamental discriminant).

In this paper we focus on this larger framework and consider L–functions for cusp forms

$$f(z) = \sum_{1}^{\infty} \lambda_f(n) n^{\frac{k-1}{2}} e(nz) \tag{1.2}$$

⋆ Supported by NSF grant DMS-98-01642 and by the Clay Mathematics Institute

⋆⋆ Supported by NSERC grant A5123 and by an NEC grant to the Institute for Advanced Study

⋆⋆⋆ Supported by the Ambrose Monell Foundation and the Hansmann Membership by grants to the Institute for Advanced Study and by NSF grant DMS-98-01642

of weight k, level D and any primitive character $\chi \pmod D$. For a technical reason we assume that $k \geqslant 3$ which helps us to resolve easily the convergence problems for various series and integrals (see further comments later). The full space of such cusp forms $S_k\left(\Gamma_0(D), \chi_D\right)$ has a unique finite basis, say $\mathcal{F}$, consisting of primitive forms. To $f \in \mathcal{F}$ given by the Fourier expansion (1.2) we attach the L–function

$$L(s, f) = \sum_1^\infty \lambda_f(n) n^{-s} \tag{1.3}$$

and the completed product

$$\Lambda(s, f) = \left(\frac{\sqrt{D}}{2\pi}\right)^s \Gamma\left(s + \frac{k-1}{2}\right) L(s, f) . \tag{1.4}$$

As shown by Hecke [H] this is an entire function and satisfies the functional equation (see Sect. 3)

$$\Lambda(s, f) = \varepsilon_f \Lambda(1 - s, \bar{f}) . \tag{1.5}$$

The Riemann hypothesis is expected to hold for every such L–function. Amongst its many consequences is the Lindelöf type bound

$$L(s, f) \ll (k|s|D)^\varepsilon \tag{1.6}$$

for $\operatorname{Re} s = \frac{1}{2}$, where ε is any positive number and the implied constant depends only on ε. Of course, we are far from proving anything like this in the near future, however from the Phragmen-Lindelöf principle and the functional equation one easily obtains the "convexity" bound

$$L(s, f) \ll \left(k^2|s|^2 D\right)^{\frac{1}{4}+\varepsilon} . \tag{1.7}$$

For many applications it is sufficient to improve this bound in the D aspect alone. Moreover, although the exponent $\frac{1}{4}$ is rather weak, it is only required to replace it by any smaller number. Thus, our priority in writing this paper will be for simplification of the arguments rather than attainment of the sharpest improvement.

A classical example of breaking the convexity bound in the conductor aspect is due to Burgess [B]. He showed that for Dirichlet L–functions with any character $\chi \pmod D$ one has

$$L(s, \chi) \ll |s|^A D^{\frac{3}{16}+\varepsilon} \tag{1.8}$$

while the convexity bound gives only $\frac{1}{4}$ in place of $\frac{3}{16}$. Burgess found an ingenious method for transforming short character sums into high moments of complete sums for which estimates were available and, in particular, the Riemann hypothesis for curves over a finite field (Weil's theorem) furnishes a strong bound.

In [FI] an alternative method was given for breaking the convexity barrier. Although it produced a quantitatively weaker result in the Burgess case, it turned out to be possible to apply this new method more generally, in particular to GL_2 automorphic L–functions in various aspects. For example in [DFI3] we proved that

$$L(s, f) \ll D^{\frac{1}{4}-\frac{1}{192}+\varepsilon} \tag{1.9}$$

for $f \in S_k(\Gamma_0(D))$, k even, $\operatorname{Re} s = \frac{1}{2}$, $\varepsilon > 0$, the implied constant depending on k, s and ε.

In this paper we treat the analogous problem when the cusp form $f \in S_k(\Gamma_0(D), \chi)$ transforms in accordance with a multiplier given by a primitive character χ of conductor equal to the level D. Our first result is

Theorem 1.1 *Let $k \geqslant 3$ and D squarefree. Let $\chi \pmod D$ be a primitive character with $\chi(-1) = (-1)^k$. Let $\mathcal{F}$ be the Hecke basis of $S_k(\Gamma_0(D), \chi)$ and $L(s, f)$ the L–function associated to $f \in \mathcal{F}$. Let $\mathbf{c} = (c_\ell)$ be any sequence of complex numbers with $c_\ell = 0$ if ℓ has a prime divisor $< z$. Then for $\operatorname{Re} s = \frac{1}{2}$ we have*

$$\sum_{f\in\mathcal{F}} \Big| \sum_{\ell\leqslant L} c_\ell \lambda_f(\ell) \Big|^2 |L(s, f)|^4 \ll \left(\|\mathbf{c}\|^2 + \|\mathbf{c}\|_1^2 z^{-1}\right) |s|^6 D^{1+\varepsilon} \tag{1.10}$$

for

$$L = D^\alpha \quad \textit{with} \quad \alpha = 1/13(48)^2 = .0000333\ldots \tag{1.11}$$

where $\|\mathbf{c}\|$ denotes the ℓ_2–norm and $\|\mathbf{c}\|_1$ denotes the ℓ_1–norm, the implied constant depending on ε and k.

Remarks. The condition that c_ℓ are supported on numbers free of small prime divisors is introduced only for technical simplification; advantage of this is taken only once, in the derivation of Corollary 5.2 from Proposition 5.1. This condition does not affect the application we have in mind, nevertheless, with extra effort one could show that (1.10) holds for any numbers c_ℓ, and without the second term on the right side.

A theorem of this type is an essential ingredient in our method, the amplification method, for bounding an L–function; see the survey article [F]. Specifically we have by positivity

$$\left|\mathcal{A}_f(L)\right|^2 |L(s, f)|^4 \ll \left(\|\mathbf{c}\|^2 + \|\mathbf{c}\|_1^2 z^{-1}\right) |s|^6 D^{1+\varepsilon} \tag{1.12}$$

where $\mathcal{A}_f(L)$ is the "amplifier"

$$\mathcal{A}_f(L) = \sum_{\ell\leqslant L} c_\ell \lambda_f(\ell) . \tag{1.13}$$

Taking the trivial amplifier of length $L = 1$, we just recover the convexity bound for $L(s, f)$. The idea to improve on this goes by choosing the coefficients c_ℓ so as to amplify the contribution to the left side of (1.12) coming from the individual f whose L–function we are seeking to bound. A natural choice for this would be $c_\ell = \overline{\lambda}_f(\ell)$, expecting that

$$\mathcal{A}_f(L) = \sum_{\ell \leqslant L} \left|\lambda_f(\ell)\right|^2 \gg L \tag{1.14}$$

whereas

$$\|\mathbf{c}\|^2 = \sum_{\ell \leqslant L} \left|\lambda_f(\ell)\right|^2 \ll L \ . \tag{1.15}$$

Choosing $L = D^{4\alpha}$ would then yield (ignore $\|\mathbf{c}\|_1^2 z^{-1}$ for the sake of illustration) the bound $L(s, f) \ll D^{\frac{1}{4}-\alpha+\varepsilon}$. This scheme works in the case of Dirichlet L–functions and would work here as well if we could prove the lower bound (1.14). Unfortunately this escapes us, at least for the above "maximal" choice of c_ℓ. The trouble is that $\lambda_f(\ell)$ could be very small very often. Fortunately we are able to derive weaker, but non-trivial, lower bounds for other choices of c_ℓ. We do know that for prime p the Hecke eigenvalues satisfy the relation

$$\lambda_f^2(p) - \lambda_f(p^2) = \chi(p) \ . \tag{1.16}$$

This for $p \nmid D$ shows that at least one of the two terms has absolute value $\geqslant \frac{1}{2}$, hence both cannot be small. To take advantage of this we choose

$$c_\ell = \begin{cases} \lambda_f(p)\overline{\chi}(p), & \text{if } \ell = p, \ \ \frac{1}{2}\sqrt{L} < p \leqslant \sqrt{L} \\ -\overline{\chi}(p), & \text{if } \ell = p^2, \ \frac{1}{2}\sqrt{L} < p \leqslant \sqrt{L} \end{cases} \tag{1.17}$$

and zero otherwise. This choice give

$$\mathcal{A}_f(L) = \sharp\left\{p; \ \ p \nmid D, \ \frac{1}{2}\sqrt{L} < p \leqslant \sqrt{L}\right\} \asymp \sqrt{L}(\log L)^{-1} \tag{1.18}$$

whereas $\|\mathbf{c}\| \ll \sqrt{L}(\log L)^{-1}$ and $\|\mathbf{c}\|_1 \ll \sqrt{L}(\log L)^{-1}$ by the Deligne bound $\left|\lambda_f(p)\right| \leqslant 2$. The numbers c_ℓ are supported on integers having no prime divisors $< z = \frac{1}{2}\sqrt{L}$. Hence we obtain by (1.12) and (1.18) (with the aid of the estimate $\alpha > 2^{-17}$)

Theorem 1.2 *Let $k \geqslant 3$, D squarefree and $\chi \pmod D$ a primitive character with $\chi(-1) = (-1)^k$. Then, for any Hecke cusp form $f \in S_k(\Gamma_0(D), \chi)$, and* $\operatorname{Re} s = \frac{1}{2}$, *we have*

$$L(s, f) \ll |s|^2 D^{\frac{1}{4}-\alpha} \tag{1.19}$$

with $\alpha = 1/2^{18}$ and the implied constant depends only on k.

Of special interest to us is the case of the Kronecker symbol $\chi = \chi_{\mathrm{D}}$. Note that since $D > 0$, k must be odd, so we are only missing the case $k = 1$. The cusp forms of weight one are special in many respects (algebraically and analytically) and we intend to cover this case in a separate work. For $k \geqslant 3$ the space $S_k\left(\Gamma_0(D), \chi_{\mathrm{D}}\right)$ contains as a small but prominent subspace $\Theta_k(D)$ spanned by theta series formed from Hecke characters for the imaginary quadratic field $K = \mathbb{Q}(\sqrt{-D})$. The Hecke characters in question are those multiplicative functions ψ, defined on the ideals, which satisfy

$$\psi\big((\beta)\big) = \left(\frac{\beta}{|\beta|}\right)^{k-1} \quad \text{for every } \beta \in O,\ \beta \neq 0 \tag{1.20}$$

(we assume that $D > 3$ so the ring of integers O has two units ± 1 and (1.20) is well defined because $k-1$ is even). Given any such character the others are obtained on multiplying it by the characters of the class group $\mathcal{C}l(K)$. Therefore we have $h(-D) = \sharp\mathcal{C}l(K)$ such characters, where the class number h satisfies

$$D^{\frac{1}{2}-\varepsilon} \ll h(-D) \ll D^{\frac{1}{2}} \log D\,. \tag{1.21}$$

We denote the set of these characters by $\mathcal{H}_k(D)$. For any $\psi \in \mathcal{H}_k(D)$ we have the corresponding theta series

$$\begin{aligned}\theta_\psi(z) &= \sum_{\mathfrak{a}\subset O} \psi(\mathfrak{a})(N\mathfrak{a})^{\frac{k-1}{2}} e(zN\mathfrak{a})\\ &= \sum_{1}^{\infty} \lambda_\psi(n) n^{\frac{k-1}{2}} e(nz)\end{aligned} \tag{1.22}$$

where

$$\lambda_\psi(n) = \sum_{N\mathfrak{a}=n} \psi(\mathfrak{a})\,. \tag{1.23}$$

These are linearly independent Hecke cusp forms in $S_k\left(\Gamma_0(D), \chi_{\mathrm{D}}\right)$ so that $\dim \Theta_k(D) = h(-D)$ whereas

$$\dim S_k\left(\Gamma_0(D), \chi_{\mathrm{D}}\right) \asymp \prod_{p|D}(p+1) \tag{1.24}$$

is much larger. Here we have abused notation a little by writing $\lambda_\psi(n)$ in place of $\lambda_f(n)$ with $f = \theta_\psi$. We shall also write $L(s, \psi)$ in place of $L(s, f)$ for $f = \theta_\psi$. In particular Theorem 1.2 gives

Theorem 1.3 *Let ψ be a Hecke character for $K = \mathbb{Q}(\sqrt{-D})$ as above. Then for* $\operatorname{Re} s = \frac{1}{2}$

$$L(s, \psi) \ll |s|^2 D^{\frac{1}{4}-\alpha} \tag{1.25}$$

with $\alpha = 1/2^{18}$ and where the implied constant depends only on k.

In the proof of Theorem 1.1 essential use is made of the spectral completeness of the family $\mathcal{F} = \mathcal{F}_k(D)$. The family $\mathcal{H} = \mathcal{H}_k(D)$, although much smaller, still enjoys a different completeness due to the orthogonality of characters of the finite abelian group $\mathcal{C}l(K)$. Therefore one should be able to prove the analogous result

$$\sum_{\psi \in \mathcal{H}} \Big| \sum_{N\mathfrak{a} \leqslant L} c_{\mathfrak{a}} \psi(\mathfrak{a}) \Big|^2 |L(s, \psi)|^2 \ll \|\mathbf{c}\|^2 |s|^2 D^{\frac{1}{2}+\varepsilon} . \tag{1.26}$$

Here $c_{\mathfrak{a}}$ are any complex numbers supported on ideals composed of primes of degree one. In the case of $k = 1$ the bound (1.26) was completely established in [DFI4] (apart from the factor $|s|^2$), and that method clearly works for any odd $k \geqslant 1$. Note that here it is the square rather than the fourth power of the L–function. Nevertheless, taking the trivial amplifier one would recover the convexity bound for each $L(s, \psi)$. However when attempting the use of (1.26) to break convexity we are again faced with the problem of choosing $c_{\mathfrak{a}}$ so that the amplifier is relatively large. In this case, due to the above restriction on the support of $c_{\mathfrak{a}}$, one needs to produce many prime ideals of first degree and small norm. This is an important and difficult problem. For special discriminants, but not for all, this problem was resolved in [DFI4]. For example one finds there the bound

$$L(s, \psi) \ll D^{\frac{1}{4}-\alpha+\varepsilon} \tag{1.27}$$

for those D with no prime divisors $> D^{\alpha^2}$, where $\alpha = 1/1156$.

The fact that the results of [DFI4] do not cover every D was the motivation for this work. To explain the reason why the new method succeeds in avoiding the problem described above (of finding small splitting primes) we begin by observing that (1.26) is deduced in [DFI4] from a non-trivial bound for the sum

$$\sum_{\psi \in \mathcal{H}} \lambda_\psi(\ell) |L(s, \psi)|^2 . \tag{1.28}$$

For simplicity of this discussion only we take ℓ prime (and $k = 1$). Recalling (1.23) for $\lambda_\psi(\ell)$ we see that if ℓ does not split then $\lambda_\psi(\ell) = 0$ for every ψ so there is no possibility of cancellation in (1.28) for such ℓ. On the other hand, if we knew there were relatively small primes ℓ which do split in K then for these there would be considerable cancellation, saving a factor $\ell^{-\frac{1}{2}}$. In similar fashion Theorem 1.1 could be deduced from a bound for the sum

$$\sum_{f \in \mathcal{F}} \lambda_f(\ell) |L(s, f)|^4 \tag{1.29}$$

(actually for technical reasons we shall replace the L–functions by partial sums). Here there will always be cancellation. Even if ℓ does not split, and hence $\lambda_\psi(\ell) = 0$ for all $\psi \in \mathcal{H}$, these $f = \theta_\psi$ now make up only a part of the total and do not inhibit cancellation in the full sum over $f \in \mathcal{F}$.

In many respects the arguments of this paper follow the path established in the previous members of this series [DFI1] and [DFI3]. An obstinate obstruction to the adoption of those however lies in the presence of the non-trivial nebentypus χ. As in the earlier papers, almost at the outset we encounter Kloosterman sums which after Poisson summation degenerate to the simpler Ramanujan sums $S(h, 0; c)$ with h of the type

$$h = \det \begin{pmatrix} m_1 & m_2 \\ n_1 & n_2 \end{pmatrix} = m_1 n_2 - m_2 n_1 \ . \tag{1.30}$$

We need to count the solutions of this equation with considerable precision when, in the critical range, the variables are all about the size $\sqrt{D}$ (or slightly different by factors coming from the amplifier). In [DFI3] this led to the quadratic divisor problem, which already had some history and we treated in [DFI2] producing an asymptotic formula uniform for h in a wide range. In the current work this problem is much more difficult because the variables in the lower row are now weighted by the character values $\chi(n_1), \overline{\chi}(n_2)$. Since the size of the ranges for n_1, n_2 is relatively small (in terms of the conductor) we are not able to exploit the fact that these are characters and have to find results which hold for general complex coefficients. This general determinant problem was solved in [DFI6] but the main work for this is the paper [DFI5]. Both of these papers were completely motivated by the current work. Certainly they have other applications. For example, the main result of [DFI5] gave new bounds for sums of Salié sums and thereby a more direct proof of a subconvexity bound for Fourier coefficients of half–integral weight cusp forms; see [I1] for the original proof. We are hopeful that these ideas will find still other applications elsewhere.

An interesting aspect (to the authors in any case) occurred when the main idea for the solution of the general determinant equation [DFI6] turned out to be a completely different application of the amplification method in an unexpected setting. This occurs in the proof of Theorem 2 of [DFI5] where, briefly speaking, we introduced a complete set of multiplicative characters as companions to the invisible trivial character and amplified the contribution of the latter. We are hopeful that the amplification method has a bright future more generally.

Indeed, the amplification method has now begun to also be exploited in a number of other works; see [KMV] for a very nice recent example of such a work. Meanwhile, Burgess's ideas have now also entered into the ring of GL_2 theory. Specifically these have been applied in [FoIw] to break the convexity bound for the Hecke L–functions of $K = \mathbb{Q}(\sqrt{-D})$ which come from cusp forms of level D^2 and non-trivial nebentypus.

The resolution of the problems treated in [DFI5-6] still left behind a number of thorny issues and, as might already be guessed from the previous paragraphs, the present work has evolved over a period of years. During this time we benefited from the hospitality of both of our home universities and from frequent visits to MSRI Berkeley and to IAS Princeton. Who can tell the extent to which our work was enhanced by the scenery at the former and the cuisine at the latter?

2 A brief account of $S_k(\Gamma_0(D), \chi)$

In this section we collect basic results about classical automorphic forms which are needed in this paper. They all are standard and can be found with proofs in [I2]. Throughout $k \geqslant 3$, $D \geqslant 3$, $\Gamma = \Gamma_0(D)$ is the group of matrices $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$ with $c \equiv 0 \pmod{D}$, so

$$\nu(D) = [\Gamma_0(1) : \Gamma_0(D)] = D \prod_{p|D} \left(1 + \frac{1}{p}\right) , \tag{2.1}$$

and $\chi \pmod{D}$ is a primitive character such that

$$\chi(-1) = (-1)^k . \tag{2.2}$$

Let $\mathbb{H}$ denote the upper-half plane consisting of complex numbers $z = x + iy$ with $y > 0$ acted on by the linear fractional transformations

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

Let $\mathcal{A}_k(\Gamma, \chi)$ be the linear space of functions $f : \mathbb{H} \to \mathbb{C}$ which satisfy

$$f(\gamma z) = \chi(d)(cz + d)^k f(z) .$$

Let $S_k(\Gamma, \chi)$ be the subspace of cusp forms. This is a finite dimensional Hilbert space with respect to the Petersson inner product

$$\langle f, g \rangle = \int_{\Gamma \backslash \mathbb{H}} f(z)\bar{g}(z) y^{k-2} dx\, dy . \tag{2.3}$$

Let $\mathcal{F}$ be an orthonormal basis of $S_k(\Gamma, \chi)$. Every $f \in \mathcal{F}$ has a Fourier series expansion

$$f(z) = \sum_{1}^{\infty} a_f(n) n^{\frac{k-1}{2}} e(nz) \tag{2.4}$$

with complex coefficients $a_f(n)$. The celebrated formula of Petersson asserts that for any $m, n \geqslant 1$

$$\begin{aligned} (4\pi)^{1-k} \Gamma(k-1) \sum_{f \in \mathcal{F}} \bar{a}_f(m) a_f(n) = \\ \delta_{mn} + 2\pi i^{-k} \sum_{c \equiv 0 (D)} c^{-1} S_\chi(m, n; c) J_{k-1}\left(\frac{4\pi}{c}\sqrt{mn}\right) \end{aligned} \tag{2.5}$$

where δ_{mn} is the Kronecker diagonal symbol, c runs over positive integral multiples of D, $S_\chi(m,n;c)$ is the Kloosterman sum given by

$$S_\chi(m,n;c) = \sum_{ad\equiv 1 (\mathrm{mod}\, c)} \chi(a) e\left(\frac{am+dn}{c}\right) \tag{2.6}$$

and $J_{k-1}(x)$ is the Bessel function. Since $J_{k-1}(x) \ll x$ for $k \geqslant 3$ and

$$\left|S_\chi(m,n;c)\right| \leqslant (m,n,c)^{\frac{1}{2}} c^{\frac{1}{2}} \tau(c) \tag{2.7}$$

the series on the right-hand side of (2.5) converges absolutely.

From now on we assume that $\mathcal{F}$ is the Hecke basis. Since the character $\chi \pmod D$ is primitive every $f \in \mathcal{F}$ is a primitive cusp form. This means f is an eigenfunction of all the Hecke operators

$$(T_n f)(z) = \frac{1}{\sqrt{n}} \sum_{ad=n} \chi(a) \left(\frac{a}{d}\right)^{\frac{k}{2}} \sum_{b (\mathrm{mod}\, d)} f\left(\frac{az+b}{d}\right) . \tag{2.8}$$

Thus for any n (not necessarily prime to the level D) we have

$$T_n f = \lambda_f(n) f \tag{2.9}$$

for some complex numbers $\lambda_f(n)$. Note that we have normalized T_n so that the Ramanujan conjecture (proved by P. Deligne) is the bound

$$|\lambda_f(n)| \leqslant \tau(n) \tag{2.10}$$

where $\tau(n)$ denotes the divisor function. The eigenvalues $\lambda_f(n)$ enjoy the multiplicativity property

$$\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \chi(d) \lambda_f(mnd^{-2}) . \tag{2.11}$$

They are not always real numbers, even if $\chi \pmod D$ is a real character. Precisely we have

$$\lambda_f(n) = \chi(n) \overline{\lambda}_f(n) \quad \text{if } (n,D) = 1 . \tag{2.12}$$

If n is not prime to D then the relation (2.12) does not hold. For every $p \mid D$ we have $\lambda_f(p^\ell) = \lambda_f(p)^\ell$ with

$$|\lambda_f(p)| = 1 . \tag{2.13}$$

Every primitive form is automatically an eigenfunction of the involution operator $\overline{W} : S_k(\Gamma,\chi) \to S_k(\Gamma,\chi)$ which is defined by $\overline{W} = KW$, where

$$(Wf)(z) = (z\sqrt{D})^{-k} f(-1/zD)$$

$$(Kf)(z) = \bar{f}(-\bar{z}) .$$

Note that $W : S_k(\Gamma, \chi) \to S_k(\Gamma, \overline{\chi})$ and $K : S_k(\Gamma, \overline{\chi}) \to S_k(\Gamma, \chi)$. The operator K acts on the Fourier series (2.4) by conjugating its coefficients, thus K is not linear over $\mathbb{C}$, nor is $\overline{W}$. As we have said and as follows from the multiplicity-one property of Hecke operators

$$\overline{W} f = \eta_f f \qquad \text{if } f \in \mathcal{F} \tag{2.14}$$

for a complex number η_f with $|\eta_f| = 1$ (which follows from $\overline{W}^2 = 1$ and $\overline{W}(\eta f) = \bar{\eta}\,\overline{W} f$ for $\eta \in \mathbb{C}$). The eigenvalue η_f is given by

$$\eta_f = \lambda_f(D)\tau_\chi D^{-\frac{1}{2}} \tag{2.15}$$

where τ_χ is the Gauss sum

$$\tau_\chi = \sum_{x (\mathrm{mod}\, D)} \chi(x) e\left(\frac{x}{D}\right) . \tag{2.16}$$

The Fourier coefficients $a_f(n)$ of a primitive form f are proportional to the Hecke eigenvalues $\lambda_f(n)$,

$$a_f(n) = a_f(1)\lambda_f(n) . \tag{2.17}$$

Hence putting

$$\omega_f = (4\pi)^{1-k}\Gamma(k-1)|a_f(1)|^2 \tag{2.18}$$

we can write the Petersson formula (2.5) as

Lemma 2.1 *Let $\mathcal{F}$ be the Hecke basis of $S_k(\Gamma, \chi)$. Then for any $m, n \geqslant 1$*

$$\sum_{f \in \mathcal{F}} \omega_f \bar{\lambda}_f(m)\lambda_f(n) = \tag{2.19}$$

$$\delta_{mn} + 2\pi i^{-k} \sum_{c \equiv 0 (D)} c^{-1} S_\chi(m, n; c) J_{k-1}\left(\frac{4\pi}{c}\sqrt{mn}\right) .$$

Assuming $\| f \| = 1$ we have (cf. Proposition 13.1 of [I2])

$$|a_f(1)|^{-2} = (4\pi)^{-k}\Gamma(k) \operatorname{vol}(\Gamma \backslash \mathbb{H}) \operatorname*{res}_{s=1} L(s, f \otimes \bar{f}) \tag{2.20}$$

where $L(s, f \otimes \bar{f})$ is the Rankin-Selberg L–function

$$L(s, f \otimes \bar{f}) = \sum_{1}^{\infty} |\lambda_f(n)|^2 n^{-s} . \tag{2.21}$$

This has the Euler product $L(s, f \otimes \bar{f}) = \prod_p L_p(s, f \otimes \bar{f})$ with

$$L_p(s, f \otimes \bar{f}) = \left(1 - p^{-s}\right)^{-1} \tag{2.22}$$

if $p \mid D$ and

$$L_p(s, f \otimes \bar{f}) = \left(1 + p^{-s}\right) \\ \left(1 - |\alpha_f(p)|^2 p^{-s}\right)^{-1} \left(1 - p^{-s}\right)^{-1} \left(1 - |\beta_f(p)|^2 p^{-s}\right)^{-1} \tag{2.23}$$

if $p \nmid D$. Moreover, for a suitable factor $L_\infty(s, f \otimes \bar{f})$, the product may be completed to $\Lambda(s, f \otimes \bar{f}) = L_\infty(s, f \otimes \bar{f}) L(s, f \otimes \bar{f})$ (cf. Theorem 2.2 of [L]) which is holomorphic on $\mathbb{C}$ except for simple poles at $s = 1, 0$, and satisfies the functional equation

$$\Lambda(s, f \otimes \bar{f}) = \Lambda(1 - s, f \otimes \bar{f}) \ . \tag{2.24}$$

In general the local factor at $p = \infty$ is quite involved, but if D is squarefree then

$$L_\infty(s, f \otimes \bar{f}) = (2\pi)^{-s} \zeta(2s) \Gamma(s) \Gamma(s + k - 1) \prod_{p \mid D} (p^s + 1). \tag{2.25}$$

Define the symmetric square L–function by means of the Euler product $L(s, \mathrm{sym}^2 f) = \prod_p L_p(s, \mathrm{sym}^2 f)$ with the local factors given by

$$L_p(s, \mathrm{sym}^2 f) = \left(1 - p^{-s}\right)^{-1} \tag{2.26}$$

if $p \mid D$ and

$$L_p(s, \mathrm{sym}^2 f) = \\ \left(1 - |\alpha_f(p)|^2 p^{-s}\right)^{-1} \left(1 - p^{-s}\right)^{-1} \left(1 - |\beta_p|^2 p^{-s}\right)^{-1} \tag{2.27}$$

if $p \nmid D$. This can be also written in terms of the Hecke eigenvalues

$$L_p(s, \mathrm{sym}^2 f) = \\ \left(1 - \overline{\chi}(p) \lambda_f(p^2) p^{-s} + \overline{\chi}(p) \lambda_f(p^2) p^{-2s} - p^{-3s}\right)^{-1} \tag{2.28}$$

(note that $\overline{\chi}(p) \lambda_f(p^2)$ is real). At the infinite place we define

$$L_\infty(s, \mathrm{sym}^2 f) = \left(D/2\pi^{3/2}\right)^s \Gamma\left(\tfrac{s+1}{2}\right) \Gamma(s + k - 1) \tag{2.29}$$

and we put $\Lambda(s, \mathrm{sym}^2 f) = L_\infty(s, \mathrm{sym}^2 f) L(s, \mathrm{sym}^2 f)$. Note that

$$\Lambda(s, f \otimes \bar{f}) = \tfrac{1}{2} \pi^{-\frac{s+1}{2}} \Gamma\left(\tfrac{s}{2}\right) \zeta(s) \Lambda(s, \mathrm{sym}^2 f) \tag{2.30}$$

by the duplication formula $\Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{s+1}{2}\right) = \sqrt{\pi}2^{1-s}\Gamma(s)$. Therefore the functional equations for $L(s, f \otimes \bar{f})$ and $\zeta(s)$ yield the functional equation for $L(s, \mathrm{sym}^2 f)$, namely

$$\Lambda(s, \mathrm{sym}^2 f) = \Lambda(1-s, \mathrm{sym}^2 f) . \tag{2.31}$$

Moreover it is known that $\Lambda(s, \mathrm{sym}^2 f)$ is entire (essentially due to G. Shimura, cf. [S]). From (2.30) we obtain

$$\operatorname*{res}_{s=1} L(s, f \otimes \bar{f}) = \frac{D}{\nu(D)} \frac{L(1, \mathrm{sym}^2 f)}{2\pi\zeta(2)} . \tag{2.32}$$

Moreover we have $\mathrm{vol}(\Gamma \setminus \mathbb{H}) = \frac{\pi}{3}[\Gamma_0(1) : \Gamma_0(D)] = \frac{\pi}{3}\nu(D)$. Inserting these values into (2.20) we get

$$|a_f(1)|^{-2} = \pi^{-2}(4\pi)^{-k}\Gamma(k)DL(1, \mathrm{sym}^2 f) . \tag{2.33}$$

Finally we get by (2.33) and (2.18)

$$\omega_f = \frac{24\pi\zeta(2)}{(k-1)DL(1, \mathrm{sym}^2 f)} . \tag{2.34}$$

Using (2.10) one can show that

$$L(1, \mathrm{sym}^2 f) \ll (\log kD)^3 . \tag{2.35}$$

Hence

$$\omega_f \gg (kD)^{-1}(\log kD)^{-3} \tag{2.36}$$

where the implied constant is absolute. This bound is quite precise because the average value of ω_f is asymptotically $1/\dim S_k(\Gamma, \chi)$. Indeed it follows from the Petersson formula (2.19) that

$$\sum_{f \in \mathcal{F}} \omega_f = 1 + O(D^{-1}) . \tag{2.37}$$

We shall require the following extension of (2.37).

Lemma 2.2 *For any complex numbers a_n we have*

$$\sum_{f \in \mathcal{F}} \omega_f \Big| \sum_{n \leqslant N} a_n \lambda_f(n) \Big|^2 = \left\{1 + O(ND^{-1})\right\} \sum_{n \leqslant N} |a_n|^2 . \tag{2.38}$$

Proof. Follows verbatim the proof of Theorem 5.7 of [I2] which gave the result for $\chi = 1$.

3 The L–functions

To each $f \in \mathcal{F}$ we associate the Hecke L function

$$L(s, f) = \sum_{1}^{\infty} \lambda_f(n) n^{-s} . \tag{3.1}$$

By (2.11) this has the Euler product

$$\begin{aligned} L(s, f) &= \prod_p \left(1 - \lambda_f(p) p^{-s} + \chi(p) p^{-2s}\right)^{-1} \\ &= \prod_p \left(1 - \alpha_f(p) p^{-s}\right)^{-1} \left(1 - \beta_f(p) p^{-s}\right)^{-1} \end{aligned} \tag{3.2}$$

with $\alpha_f(p) + \beta_f(p) = \lambda_f(p)$ and $\alpha_f(p)\beta_f(p) = \chi(p)$. Hecke showed that $L(s, f)$ is entire and satisfies the functional equation

$$\Lambda(s, f) = \varepsilon_f \Lambda(1 - s, \bar{f}) \tag{3.3}$$

where $\bar{f} = K f$ and $\Lambda(s, f)$ is the completed product

$$\Lambda(s, f) = \left(\frac{\sqrt{D}}{2\pi}\right)^s \Gamma\left(s + \frac{k-1}{2}\right) L(s, f) . \tag{3.4}$$

This follows by applying the Mellin transform to (2.14) for $z = iy$. One gets (3.3) with

$$\varepsilon_f = i^k \overline{\lambda}_f(D) \bar{\tau}_\chi D^{-\frac{1}{2}} . \tag{3.5}$$

However, all we need to know about ε_f in this paper is that

$$|\varepsilon_f| = 1 . \tag{3.6}$$

Using the functional equation (3.3) we shall represent $L(s, f)$ on the critical line $\mathrm{Re}\, s = \frac{1}{2}$ in terms of rapidly converging series, essentially equivalent to partial sums of length $|s|\sqrt{D}$. To this end we choose a function $G(u)$ which is holomorphic in $|\mathrm{Re}\, u| \leqslant 1$ such that

$$\begin{aligned} G(u) &= G(-u) \\ G(0) &= 1 \\ G(u) &\ll (1 + |u|)^{-k} e^{\frac{\pi}{2}|u|} . \end{aligned} \tag{3.7}$$

Consider the integral

$$I(s, f) = \frac{1}{2\pi i} \int_{(1)} \Lambda(s + u, f) G(u) u^{-1} du .$$

Moving the integration to the line $\mathrm{Re}\, u = -1$ and applying (3.3) we get

$$\Lambda(s, f) = I(s, f) + \varepsilon_f I(1-s, \bar{f})\ . \tag{3.8}$$

On the other hand, introducing the Dirichlet series (3.1) and integrating termwise we obtain

$$I(s, f) = \sum_1^\infty \lambda_f(n) \frac{1}{2\pi i} \int_{(1)} \left(\frac{\sqrt{D}}{2\pi n} \right)^{s+u} \Gamma\left(s + u + \tfrac{k-1}{2}\right) G(u) u^{-1} du\ .$$

Inserting this into (3.8) and dividing by $(\sqrt{D}/2\pi)^s \Gamma\left(s + \frac{k-1}{2}\right)$ we arrive at the desired representation:

Lemma 3.1 *For s with $\mathrm{Re}\, s = \frac{1}{2}$ we have*

$$L(s, f) = \sum_1^\infty \lambda_f(n) n^{-s} V_s\left(\frac{2\pi n}{\sqrt{D}}\right) + \varepsilon_f(s) \sum_1^\infty \bar{\lambda}_f(n) n^{s-1} V_{1-s}\left(\frac{2\pi n}{\sqrt{D}}\right) \tag{3.9}$$

where

$$\varepsilon_f(s) = \varepsilon_f \left(\frac{\sqrt{D}}{2\pi}\right)^{1-2s} \frac{\Gamma(1-s+\kappa)}{\Gamma(s+\kappa)}, \tag{3.10}$$

$\kappa = \frac{k-1}{2}$, *and $V_s(y)$ is given by the Mellin integral*

$$V_s(y) = \frac{1}{2\pi i} \int_{(1)} \frac{\Gamma(s+u+\kappa)}{\Gamma(s+\kappa)} \frac{G(u)}{u} y^{-u} du\ . \tag{3.11}$$

Remark. For $\mathrm{Re}\, s = \frac{1}{2}$ we have $|\varepsilon_f(s)| = 1$.

Moving the integration to the line $\mathrm{Re}\, s = -\kappa$ we deduce that

$$V_s(y) = 1 + O\left(\left(\frac{y}{|s|}\right)^\kappa\right) \tag{3.12}$$

where the implied constant depends on κ, that is on k. In applications of Lemma 3.1 we need to control the growth of $V_s(y)$ and its derivatives for large y. To this end we choose

$$G(u) = \left(\cos \frac{\pi u}{A}\right)^{-A} \tag{3.13}$$

where A is a large integer, $A \geqslant 3$.

Lemma 3.2 *For any integer $a \geqslant 0$ we have*

$$V_s^{(a)}(y) \ll \left(\frac{|s|}{y}\right)^a \left(1 + \frac{y}{|s|}\right)^{-A} \tag{3.14}$$

the implied constant depending only on a, A and k.

Proof. Move the integration to the line $\operatorname{Re} u = A$ and differentiate a times getting

$$V_s^{(a)}(y) \ll \int_{(A)} \left| \frac{\Gamma(s+u+\kappa)}{\Gamma(s+\kappa)} \left(y \cos \frac{\pi u}{A}\right)^{-A} \left(\frac{u}{y}\right)^{a} \frac{du}{u} \right| .$$

By Stirling's formula we derive that $u^a \Gamma(s+u+\kappa)/\Gamma(s+\kappa)$ is bounded by

$$|s|^{-\kappa}|s+u|^{\kappa+A}|u|^a \exp \frac{\pi}{2} (|\operatorname{Im} s| - |\operatorname{Im}(s+u)|) \ll |s|^{a+A} \exp\left(\frac{\pi}{2}|u|\right) .$$

Hence we deduce that $V_s^{(a)}(y)$ is bounded by

$$\left(\frac{|s|}{y}\right)^{a+A} \int_{(A)} \left| \left(\cos \frac{\pi u}{A}\right)^{-A} \exp\left(\frac{\pi}{2}|u|\right) \frac{du}{u} \right| \ll \left(\frac{|s|}{y}\right)^{a+A} .$$

This yields (3.14) if $y \geqslant |s|$. If $y < |s|$ and $a > 0$, then we move the integration to the line $\operatorname{Re} u = 0$ and estimate as above getting (3.14). If $y < |s|$ and $a = 0$ then (3.12) is more precise than (3.14).

4 Preliminary estimates of $L(s, f)$

Applying a smooth partition of unity we derive by (3.9) and (3.14) that

$$L(s, f) \ll \sum_N |G_f(N)| N^{-\frac{1}{2}} \left(1 + \frac{N}{|s|\sqrt{D}}\right)^{-A} \tag{4.1}$$

where $G_f(N)$ are sums of type

$$G_f(N) = \sum_n \lambda_f(n) g(n) \tag{4.2}$$

with $g(x)$ a smooth function supported on $[N, 2N]$ for $N = 2^{\nu/2}$, $\nu \geqslant -1$, such that

$$g^{(a)}(x) \ll \left(\frac{|s|}{N}\right)^{a} \tag{4.3}$$

for all $a \geqslant 0$, the implied constant depending only on a, A and k. By Hölder's inequality

$$L(s, f)^4 \ll \sum_N |G_f(N)|^4 N^{-2} \left(1 + \frac{N}{|s|\sqrt{D}}\right)^{-4A} \log 2N . \tag{4.4}$$

Here, and in (4.1), the implied constant depends only on A and k.

Fix the complex numbers c_ℓ. Put

$$A_f(L) = \sum_{\ell \leqslant L} c_\ell \lambda_f(\ell) \ . \tag{4.5}$$

By (2.10) and the Cauchy inequality we get

$$A_f(L) \ll \|\mathbf{c}\| L^{\frac{1}{2}} (\log 2L)^{\frac{3}{2}} \tag{4.6}$$

where $\|\mathbf{c}\|$ denotes the ℓ_2–norm. This trivial estimate, of course, cannot be improved for a given general f (apart from the logarithmic factor). Our goal is to establish non-trivial estimates for the averages

$$\mathcal{D}(L, N) = \sum_{f \in \mathcal{F}} \omega_f \left|A_f(L)\right|^2 \left|G_f(N)\right|^4 \ . \tag{4.7}$$

Using the multiplicativity of Hecke eigenvalues one easily derives by Lemma 2.2 the following general result

$$\mathcal{D}(L, N) \ll \|\mathbf{c}\|^2 \left(1 + D^{-1} L N^2\right) N^{2+\varepsilon} \tag{4.8}$$

where ε is any positive number, the implied constant depending on ε and k. This would be sufficient if $LN^2 \leqslant D$, but not so in the crucial range $N \asymp D^{\frac{1}{2}}$ in which case the factor $|A_f(L)|^2$ has only a trivial effect on the bound (4.8). Exploiting the smooth sum of Hecke eigenvalues $G_f(N)$ we shall improve on (4.8), provided that L is sufficiently small. For technical simplifications we restrict the sum (4.5) to numbers free of small prime divisors.

Proposition 4.1 *Suppose c_ℓ are supported on positive integers having no prime divisors $< z$. Then we have*

$$\mathcal{D}(L, N) \ll \tag{4.9}$$
$$\left\{\|\mathbf{c}\|^2 (1 + L^{13} D^{-\theta}) + \|\mathbf{c}\|_1^2 z^{-1}\right\} N^2 \left(1 + D^{-1} N^2\right)^2 |s|^2 D^\varepsilon$$

where $\theta = (48)^{-2}$, and ε is any positive number, the implied constant depending on ε and k.

Choosing $L = D^{\theta/13}$ the bound (4.9) simplifies to

$$\mathcal{D}(L, N) \ll \left(\|\mathbf{c}\|^2 + \|\mathbf{c}\|_1^2 z^{-1}\right) N^2 \left(1 + D^{-1} N^2\right)^2 |s|^2 D^\varepsilon \ . \tag{4.10}$$

Next we derive by (4.4) and (4.10) that

$$\sum_{f \in \mathcal{F}} \omega_f |A_f(L)|^2 |L(s, f)|^4 \ll \left(\|\mathbf{c}\|^2 + \|\mathbf{c}\|_1^2 z^{-1}\right) |s|^6 D^\varepsilon \ . \tag{4.11}$$

Finally we remove the spectral weights ω_f by applying the lower bound (2.36) and as a result obtain Theorem 1.1.

5 The sums $\mathcal{N}_f(\ell)$

It remains to prove Proposition 4.1 and for this we spend the rest of the paper. We begin by making suitable arrangements in the double sum

$$\left|G_f(N)\right|^2 = \sum_{n_1} \sum_{n_2} \lambda_f(n_1)\overline{\lambda}_f(n_2)g(n_1)\bar{g}(n_2)\ .$$

Here we write n_2 as δn_2 with $\delta \mid D^\infty$ and $(n_2, D) = 1$. Then we have $\overline{\lambda}_f(\delta n_2) = \overline{\lambda}_f(\delta)\overline{\chi}(n_2)\overline{\lambda}_f(n_2)$ by (2.12) and

$$\lambda_f(n_1)\overline{\lambda}_f(\delta n_2) = \overline{\lambda}_f(\delta)\overline{\chi}(n_2) \sum_{d|(n_1,n_2)} \chi(d)\lambda_f\left(n_1 n_2 d^{-2}\right)$$

by (2.11). Thus $\left|G_f(N)\right|^2$ is equal to

$$\sum_{\delta|D^\infty} \overline{\lambda}_f(\delta) \sum_d \chi(d) \sum_{n_1} \sum_{n_2} \lambda_f(n_1 n_2)\overline{\chi}(n_2)g(dn_1)\bar{g}(\delta d n_2)\ .$$

Hence, by Cauchy's inequality (use also (2.13))

$$\left|G_f(N)\right|^4 \ll (DN)^\varepsilon \sum_{\delta|D^\infty} \sqrt{\delta} \sum_{(d,D)=1} d\ \left|\sum_n \lambda_f(n)\sigma(n,\chi)\right|^2 \tag{5.1}$$

where

$$\sigma(n,\chi) = \sum\sum_{n_1 n_2 = n} g(dn_1)\bar{g}(\delta d n_2)\overline{\chi}(n_2) \tag{5.2}$$

for the relevant δ, d and g.

Having in mind some other applications in the future we consider slightly more general sums of type

$$\mathcal{N}_f(\ell) = \sum_n \lambda_f(\ell n)\sigma_F(n,\chi) \tag{5.3}$$

where $F(x_1, x_2)$ is a function of Schwartz class on $\mathbb{R}^2$ and

$$\sigma_F(n,\chi) = \sum_{n_1 n_2 = n} F(n_1, n_2)\overline{\chi}(n_2)\ . \tag{5.4}$$

In particular the innermost sum in (5.1) agrees with $\mathcal{N}_f(1)$ for the test function $F(x_1, x_2) = g(dx_1)\bar{g}(\delta d x_2)$. In this case

$$\operatorname{supp} F \subset [X_1, 2X_1] \times [X_2, 2X_2] \tag{5.5}$$

with $dX_1 = \delta dX_2 = N$ and the partial derivatives satisfy (see (4.3))

$$F^{(\alpha_1,\alpha_2)} \ll |s|^{\alpha_1+\alpha_2} X_1^{-\alpha_1} X_2^{-\alpha_2} \ . \tag{5.6}$$

From now on F is any smooth function satisfying (5.5) and (5.6) with $X_1, X_2 \geqslant \frac{1}{2}$ for any $\alpha_1, \alpha_2 \geqslant 0$, the implied constant depending on α_1, α_2. Put

$$P = 1 + (X_1 + X_2)^2 D^{-1} \ . \tag{5.7}$$

We think of X_1, X_2 both close to $D^{\frac{1}{2}}$ so P is small. Our goal is

Proposition 5.1 *For any $\ell \geqslant 1$ and $X_1, X_2 \geqslant \frac{1}{2}$ we have*

$$\sum_{f\in\mathcal{F}} \omega_f \overline{\mathcal{N}_f}(\ell)\mathcal{N}_f(1) \ll \left(\ell^{-\frac{1}{2}} + \ell^6 D^{-\theta}\right) X_1 X_2 P^2 |s|^2 D^{\varepsilon} \tag{5.8}$$

where $\theta = (48)^{-2}$ and ε is any positive number, the implied constant depending on ε.

We do not require ℓ to be free of small prime divisors, however for technical simplifications in the following corollary we express the result in terms of the greatest prime factor of ℓ, denoted by $p(\ell)$ (by convention $p(1) = 1$).

Corollary 5.2 *For any $\ell \geqslant 1$ and $X_1, X_2 \geqslant \frac{1}{2}$ we have*

$$\sum_{f\in\mathcal{F}} \omega_f \lambda_f(\ell) \left|\mathcal{N}_f(1)\right|^2 \ll \left(\frac{1}{\sqrt{\ell}} + \frac{1}{p(\ell)} + \frac{\ell^6}{D^\theta}\right) X_1 X_2 P^2 |s|^2 D^{\varepsilon} \tag{5.9}$$

for any $\varepsilon > 0$, the implied constant depending on ε.

Proof. We write

$$\lambda_f(\ell)\mathcal{N}_f(1) = \sum_m \sum_{d|(\ell,m)} \chi(d)\lambda_f(\ell m d^{-2})\sigma_F(m,\chi) \ .$$

If $d > 1$ then $d \geqslant p(\ell)$ so the contribution of such terms to the left side of (5.9) is bounded by (apply Lemma 2.2)

$$O\left(\frac{X_1X_2}{p(\ell)}\left(1 + \frac{X_1X_2}{D}\right)(\log 5X_1X_2)^3\right) \tag{5.10}$$

and this is absorbed by the right side of (5.9). From $d = 1$ we get $\mathcal{N}_f(\ell)$ so the contribution of such terms is bounded by (5.8). Clearly the right side of (5.8) is absorbed by the right side of (5.9). This completes the proof of Corollary 5.2.

Now Proposition 4.1 follows from (5.9) and (5.1) on using the multiplicativity of the Hecke eigenvalue $\lambda_f(\ell)$. Therefore, it remains to prove Proposition 5.1.

6 Applications of the Petersson and the Poisson formulas

We begin the proof of Proposition 5.1 by applying the Petersson formula (2.19); this gives

$$\sum_{f\in\mathcal{F}} \omega_f \overline{\mathcal{N}}_f(\ell)\mathcal{N}_f(1) = R_0 + 2\pi i^{-k} \sum_{c\equiv 0(D)} c^{-1} R_c \tag{6.1}$$

where

$$R_0 = \sum_m \bar{\sigma}_F(m,\chi)\sigma_F(\ell m,\chi) \ll \ell^{-1} X_1 X_2 D^{\varepsilon} \tag{6.2}$$

and

$$R_c = \sum_m \sum_n \bar{\sigma}_F(m,\chi)\sigma_F(n,\chi) S_\chi(\ell m, n; c) J_{k-1}\left(\frac{4\pi}{c}\sqrt{\ell m n}\right). \tag{6.3}$$

Next we transform R_c by means of the following involution.

Proposition 6.1 *Let* $c > 0$, $c \equiv 0 \pmod D$ *and* $(c,d) = 1$. *Let* F *be a Schwartz function on* $\mathbb{R}^2$. *Then*

$$\sum_n \sigma_F(n,\chi) e\left(\frac{dn}{c}\right) = \chi(-d) c \sum_n \sigma_G(n,\chi) e\left(-\frac{\bar{d}n}{c}\right) \tag{6.4}$$

where

$$G(y_1,y_2) = \iint F(cx_1, cx_2) e(-x_1 y_2 - x_2 y_1) dx_1 dx_2 \,. \tag{6.5}$$

Proof. Split into classes modulo c and apply the Poisson summation formula as in Lemma 9.2 of [DI].

Now we open the Kloosterman sum $S_\chi(\ell m, n; c)$ in (6.3) and execute the summation in n, but not in m, by means of (6.4). Consequently the Kloosterman sum degenerates to the Ramanujan sum $S(\ell m - n, 0; c)$, and we obtain

$$R_c = \chi(-1) c \sum_m \sum_n \bar{\sigma}_F(m,\chi)\sigma_G(n,\chi) S(\ell m - n, 0; c) \tag{6.6}$$

where G is the following integral transform of F (note the added Bessel function):

$$G(y_1,y_2) = \iint F(cx_1, cx_2) J_{k-1}(4\pi\sqrt{\ell m x_1 x_2}) e(-x_1 y_2 - x_2 y_1) dx_1\, dx_2. \tag{6.7}$$

An important feature of the new expression (6.6) for R_c (aside from the fact that Ramanujan sums are simpler than Kloosterman sums) is that the "dual" variable n appears with the original m in an additive form rather than in the multiplicative form as it was in (6.3). We now split the sum (6.6) in accordance with $\ell m - n = h$, say

$$R_c = \sum_h R_c(h) \tag{6.8}$$

and we shall treat each $R_c(h)$ separately.

7 The singular contribution

For $h = 0$ we get $S(0, 0; c) = \varphi(c)$ and

$$\begin{aligned} R_c(0) &= \chi(-1)\varphi(c)c \sum_m \bar{\sigma}_F(m, \chi)\sigma_G(\ell m, \chi) \\ &= \chi(-1)\varphi(c)c \sum_{a_1 a_2 \ell = b_1 b_2} \chi(a_2)\overline{\chi}(b_2)\overline{F}(a_1, a_2)G(b_1, b_2). \end{aligned}$$

Combining the terms for b_1, b_2 with those for $-b_1, -b_2$ we can replace $G(b_1, b_2)$ by

$$\begin{aligned} &\chi(-1)G_k(b_1, b_2) = \chi(-1)G(b_1, b_2) + G(-b_1, -b_2) \\ &= \iint F(cx_1, cx_2)\, J_{k-1}\big(4\pi\sqrt{\ell a_1 a_2 x_1 x_2}\big) E_k(b_1 x_2 + b_2 x_1) dx_1\, dx_2 \end{aligned} \tag{7.1}$$

where $E_k(x) = e(x) + (-1)^k e(-x)$, that is

$$E_k(x) = \begin{cases} 2\cos 2\pi x, & \text{if } k \text{ even}, \\ 2i \sin 2\pi x, & \text{if } k \text{ odd}. \end{cases}$$

We obtain

$$R_c(0) = (-1)^k \varphi(c)c \sum_{a_1 a_2 \ell = b_1 b_2} \chi(a_2)\overline{\chi}(b_2)\overline{F}(a_1, a_2)G_k(b_1, b_2) \tag{7.2}$$

where a_1, a_2, b_1, b_2 run over positive integers. Have in mind that $G_k(b_1, b_2)$ depends on $a_1 a_2$ and c by way of the integrated functions $F(cx_1, cx_2)$ and $J_{k-1}\left(4\pi\sqrt{\ell a_1 a_2 x_1 x_2}\right)$.

We execute the summation over c by means of the following

Lemma 7.1 *Let F be smooth, compactly supported on $\mathbb{R}^+$. Then*

$$\sum_{c \equiv 0(D)} \varphi(c)F(c) = \frac{1}{\zeta(2)\nu(D)} \int tF(t)dt + \int \xi_D(t)dtF(t)$$

where

$$\xi_D(t) = \frac{\varphi(D)}{D} \sum_{(d,D)=1} \frac{\mu(d)}{d} \left\{\frac{t}{dD}\right\} \ll \log\left(1 + \frac{t}{D}\right).$$

Proof. This follows from the Euler-Maclaurin formula, cf. [DFI3].

Summing $c^{-1}R_c(0)$ over $c \equiv 0 (\mathrm{mod}\ D)$ we obtain by Lemma 7.1

$$\sum_{c\equiv 0(D)} \varphi(c)F(cx_1, cx_2) =$$
$$\frac{1}{\zeta(2)\nu(D)} \int tF(tx_1, tx_2)dt + \int \xi_D(t)\frac{\partial}{\partial t}\left(tF(tx_1, tx_2)\right) dt\ .$$

Integrating the first integral over x_1, x_2 and changing these variables by a factor t we arrive at the integral of $F(x_1, x_2)$ against

$$\int_0^\infty J_{k-1}\big(4\pi t\sqrt{\ell a_1 a_2 x_1 x_2}\big) E_k\left((b_1x_2 + b_2x_1)t\right) t^{-1}dt\ .$$

However this last integral vanishes by the orthogonality formula

$$\int_0^\infty J_{k-1}(2\pi at)E_k(bt)t^{-1}dt = 0$$

which is valid for $b \geqslant a \geqslant 0$ (see (6.693.1) and (6.693.2) of [GR]). This is reminiscent of (55) of [DFI3]. In our case

$$b = b_1x_2 + b_2x_1 \geqslant 2\sqrt{b_1b_2x_1x_2} = 2\sqrt{\ell a_1a_2x_1x_2} = a\ .$$

Therefore we are left with

$$\sum_{c\equiv 0(D)} c^{-1}R_c(0) = \sum_{a_1a_2\ell = b_1b_2} \chi(a_2)\overline{\chi}(b_2)\overline{F}(a_1, a_2)I(a_1a_2; b_1, b_2)$$

where

$$I(m; b_1, b_2) = \iiint \xi_D(t)\frac{\partial}{\partial t}\left(tF(tx_1, tx_2)\right) J_{k-1}\big(4\pi\sqrt{\ell m x_1x_2}\big)$$
$$E_k(b_2x_1 + b_1x_2)dx_1\, dx_2\, dt\ .$$

Change x_2 into x/x_1 and use the estimate

$$\int \frac{\partial}{\partial t}\left(tF(tx_1, tx/x_1)\right) e(b_2x_1 + b_1x/x_1)dx_1 \ll |s|^2(1 + b_1b_2x)^{-\frac{1}{4}}$$

(see Lemma 11.3 of [DFI3]). Note that $b_1b_2 = a_1a_2\ell \asymp X_1X_2\ell$. A similar bound holds for the Bessel function,

$$J_{k-1}\big(4\pi\sqrt{\ell m\, x}\big) \ll \big(1 + \sqrt{\ell m\, x}\big)^{-\frac{1}{2}} \ll (1 + X_1X_2\ell x)^{-\frac{1}{4}}\ .$$

Note that $t^2x \asymp X_1X_2$ by the support of $F(tx_1, tx/x_1)$. From the combination of these two estimates we deduce that

$$I(m; b_1, b_2) \ll |s|^2X_1X_2 \int_0^\infty \log\Big(1 + \frac{t}{D}\Big)\big(t + \sqrt{\ell}\, X_1X_2\big)^{-1}t^{-1}dt$$
$$\ll |s|^2\ell^{-\frac{1}{2}} \log D\ .$$

Hence we conclude that

$$\sum_{c\equiv 0(D)} c^{-1}R_c(0) \ll \ell^{-\frac{1}{2}}X_1X_2|s|^2D^{\varepsilon}\ . \tag{7.3}$$

This bound is absorbed by the right side of (5.8), thus completing the goal for this section of estimating the singular contribution.

It is appropriate to insert here some explanation of our plans for the estimation of the non-singular contribution which will take place in the next two sections. Let $h \neq 0$. We have

$$R_c(h) = \chi(-1)c\, S(h,0;c)V(h) \tag{7.4}$$

where

$$V(h) = \sum\sum_{\ell m - n = h} \bar{\sigma}_F(m,\chi)\sigma_G(n,\chi)\ . \tag{7.5}$$

Opening $\sigma_F(m,\chi)$ and $\sigma_G(n,\chi)$ we can see $\ell m - n = h$ as a kind of determinant equation. We shall estimate the sum $V(h)$ separately for each $h \neq 0$. The main tool for this is Theorem 1 of [DFI6] which was derived for this purpose although it deals with a more general determinant equation (the character values can be replaced by arbitrary but bounded complex numbers). Actually we shall need to combine Theorem 1 of [DFI6] with a few estimates of a more direct nature. The required combination will be performed in the next section.

8 Representations by the determinant

Let F be a smooth function on $\mathbb{R}\times\mathbb{R}\times\mathbb{R}^+\times\mathbb{R}^+$ such that

$$a_1^{\alpha_1}a_2^{\alpha_2}b_1^{\beta_1}b_2^{\beta_2}\frac{\partial^{(\alpha_1,\alpha_2,\beta_1,\beta_2)}}{\partial a_1^{\alpha_1}\partial a_2^{\alpha_2}\partial b_1^{\beta_1}\partial b_2^{\beta_2}}F(a_1,a_2;b_1,b_2) \ll \tag{8.1}$$

$$Z^{\alpha_1+\alpha_2+\beta_1+\beta_2}\left(1+\frac{|a_1|}{A}\right)^{-4}\left(1+\frac{|a_2|}{A}\right)^{-4}\left(1+\frac{b_1}{B}\right)^{-4}\left(1+\frac{b_2}{B}\right)^{-4}$$

where $A, B, Z \geqslant 1$, the implied constant depending on $\alpha_1,\alpha_2,\beta_1,\beta_2$ alone. Let $\gamma_{b_1}, \delta_{b_2}$ be complex numbers for $b_1, b_2 > 0$. Put

$$V(h) = \sum\sum\sum\sum_{a_1b_2-a_2b_1=h}\gamma_{b_1}\delta_{b_2}F(a_1,a_2;b_1,b_2)\ ,$$

$$W(h) = \sum\sum_{(b_1,b_2)|h}\gamma_{b_1}\delta_{b_2}\frac{(b_1,b_2)}{b_1b_2}\int F\left(\frac{x}{b_2},\ \frac{x-h}{b_1}\ ; b_1, b_2\right)dx\ .$$

In the sequel we shall denote the above integral by $I(b_1,b_2)$.

Theorem 8.1 *Let* $|\gamma_{b_1}| \leqslant 1$ *and* $|\delta_{b_2}| \leqslant 1$. *For any* $h \neq 0$ *we have*

$$V(h) = W(h) + O\left(\tau(h)\left(1+\frac{|h|}{AB}\right)^{-2}\left(Z^8 A^{-1} B^{\frac{47}{48}}\right)^{\frac{1}{24}}(AB)^{1+\varepsilon}\right) \tag{8.2}$$

with any $\varepsilon > 0$, *the implied constant depending on* ε.

Proof. By applying a smooth partition of unity on $\mathbb{R}^4$ we may assume that F is supported in one of the following sets:

$$\begin{aligned}
\mathcal{B}_1 &= [-1,1] \times \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+,\\
\mathcal{B}_2 &= \mathbb{R} \times [-1,1] \times \mathbb{R}^+ \times \mathbb{R}^+,\\
\mathcal{B}^\sigma &= \sigma_1[X_1, 4X_1] \times \sigma_2[X_2, 4X_2] \times [Y_1, 4Y_1] \times [Y_2, 4Y_2],
\end{aligned}$$

where $\sigma = (\sigma_1, \sigma_2) = (\pm, \pm)$ and X_1, X_2, Y_1, Y_2 take values $2^n \geqslant \frac{1}{2}$.

If F is supported in $\mathcal{B}_1$ then the left side of (8.2) is

$$\begin{aligned}
V(h) &= \sum_{-a_2 b_1 = h} \gamma_{b_1}\delta_{b_2} F(0, a_2; b_1, b_2)\\
&\ll B \sum_{ab=h}\left(1+\frac{|a|}{A}\right)^{-4}\left(1+\frac{b}{B}\right)^{-4} \ll B\tau(h)\left(1+\frac{|h|}{AB}\right)^{-4}.
\end{aligned}$$

Next, the integral $I(b_1, b_2)$ in $W(h)$ is bounded by

$$\left(1+\frac{b_1}{B}\right)^{-4}\left(1+\frac{b_2}{B}\right)^{-4}\int_{-\infty}^{\infty}\left(1+\frac{|x|}{b_2}\right)^{-4}\left(1+\frac{|x-h|}{Ab_1}\right)^{-4} dx.$$

Here we have

$$\begin{aligned}
&\left(1+\frac{b_1}{B}\right)\left(1+\frac{b_2}{B}\right)\left(1+\frac{|x|}{b_2}\right)\left(1+\frac{|x-h|}{Ab_1}\right)\\
&\qquad \geqslant \left(1+\frac{|x|}{B}\right)\left(1+\frac{|x-h|}{AB}\right) \geqslant 1+\frac{|h|}{AB}.
\end{aligned}$$

Hence the main term in (8.2) satisfies

$$\begin{aligned}
W(h) &\ll \left(1+\frac{|h|}{AB}\right)^{-2}\sum\sum_{(b_1,b_2)|h}\frac{(b_1,b_2)}{b_1}\left(1+\frac{b_1}{B}\right)^{-2}\left(1+\frac{b_2}{B}\right)^{-2}\\
&\ll B\tau(h)\left(1+\frac{|h|}{AB}\right)^{-2}.
\end{aligned}$$

These estimates are absorbed by the error term in (8.2) showing that Theorem 8.1 is trivial if F is supported in $\mathcal{B}_1$. Similarly we see that Theorem 8.1 is trivial if F is supported in $\mathcal{B}_2$.

Now suppose F is supported in the positive box $\mathcal{B} = \mathcal{B}^{++} = [X_1, 4X_1] \times [X_2, 4X_2] \times [Y_1, 4Y_1] \times [Y_2, 4Y_2]$. First we estimate both $V(h)$ and $W(h)$ trivially using the bound $F \ll T^{-4}$ where

$$T = T(\mathcal{B}) = \left(1 + \frac{X_1}{A}\right)\left(1 + \frac{X_2}{A}\right)\left(1 + \frac{Y_1}{B}\right)\left(1 + \frac{Y_2}{B}\right).$$

Hence

$$\begin{aligned} V(h) &\ll T^{-4} \mid \{(a_1, a_2, b_1, b_2) \in \mathcal{B}\,; \quad a_1 b_2 - a_2 b_1 = h\} \mid \\ &\ll T^{-4} \left(1 + \frac{|h|}{X_1 Y_2 + X_2 Y_1}\right)^{-2} \min(X_1 Y_2,\ X_2 Y_1)\ (X_1 X_2 Y_1 Y_2)^{\varepsilon} \end{aligned}$$

and, because $T\left(1 + \frac{|h|}{X_1Y_2 + X_2Y_1}\right) \geqslant 1 + \frac{|h|}{AB}$, it follows that

$$V(h) \ll T^{-2} \left(1 + \frac{|h|}{AB}\right)^{-2} (X_1 X_2 Y_1 Y_2)^{\frac{1}{2}+\varepsilon}.$$

Similarly it follows that

$$W(h) \ll T^{-2} \left(1 + \frac{|h|}{AB}\right)^{-2} (X_1 X_2 Y_1 Y_2)^{\frac{1}{2}+\varepsilon}.$$

We apply these estimates for $V(h)$ and $W(h)$ when the box $\mathcal{B}$ does not satisfy

$$\begin{aligned} \Delta^2 A < X_1, X_2 < \Delta^{-1} A \\ \Delta^2 B < Y_1, Y_2 < \Delta^{-1} B \end{aligned} \tag{8.3}$$

where $\Delta > 0$ will be chosen later. We obtain

$$V(h) = W(h) + O\left(\frac{\Delta}{T}\left(1 + \frac{|h|}{AB}\right)^{-2} (AB)^{1+\varepsilon}\right). \tag{8.4}$$

Next we apply Theorem 1 of [DFI6] getting

$$\begin{aligned} V(h) = W(h) + O\Bigg(&\left(1 + \frac{|h|}{X_1 Y_2 + X_2 Y_1}\right)^{-2} \left(\frac{X_1 Y_2}{X_2 Y_1} + \frac{X_2 Y_1}{X_1 Y_2}\right)^{\frac{19}{8}} \\ &Z^8 (Y_1 Y_2)^{\frac{7}{8}} (Y_1 + Y_2)^{\frac{11}{48}} (X_1 X_2 Y_1 Y_2)^{\varepsilon}\Bigg). \end{aligned}$$

To be precise Theorem 1 of [DFI6] requires the variables a_1, a_2 to be separated from b_1, b_2, however this can be accomplished for any $F(a_1, a_2; b_1, b_2)$ which satisfies (8.1) at the cost of an extra factor Z^5 in the error term. This explains why we have above Z^8 rather than $Z^{19/8}$ as in [DFI6]. Moreover the first factor $(1+|h|/(X_1Y_2+X_2Y_1))^{-2}$ is redundant, but will be needed later on. We use the above formula if $\mathcal{B}$ does satisfy (8.3) getting

$$V(h) = W(h) + O\left(\Delta^{-23}T^{-1}Z^8B^{\frac{95}{48}}(AB)^{\varepsilon}\left(1+\frac{|h|}{AB}\right)^{-2}\right). \tag{8.5}$$

We equalize the error terms in (8.4) and (8.5) by choosing

$$\Delta = \left(Z^8A^{-1}B^{\frac{47}{48}}\right)^{\frac{1}{24}}. \tag{8.6}$$

This choice makes the bound (8.4) valid for any positive box $\mathcal{B}$. Similarly one can show that (8.4) holds for any box of type $\mathcal{B}^{\sigma}$ with $\sigma = (\pm,\pm)$. Summing over the boxes $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}^{\sigma}$ we complete the proof of Theorem 8.1 since

$$\sum_{\mathcal{B}} T(\mathcal{B}) \ll (\log 2A)^2(\log 2B)^2 \ll (AB)^{\varepsilon}.$$

For special coefficients we can estimate the main term $W(h)$ in (8.2) successfully. We are interested in the coefficients given by

$$\begin{array}{llll} \gamma_{b_1} = \chi_1(b_1/\ell_1) & \text{if } \ell_1 \mid b_1\,, & \gamma_{b_1} = 0 & \text{if } \ell_1 \nmid b_1\,, \\ \delta_{b_2} = \chi_2(b_2/\ell_2) & \text{if } \ell_2 \mid b_2\,, & \delta_{b_2} = 0 & \text{if } \ell_2 \nmid b_2\,, \end{array} \tag{8.7}$$

where $\chi_1 \pmod{D_1}$, $\chi_2 \pmod{D_2}$ are non-trivial Dirichlet characters. In this case we write

$$W(h) = \sum_{\delta d|h}\mu(\delta)d \sum\sum_{\substack{b_1\equiv 0([\delta d,\ell_1])\\ b_2\equiv 0([\delta d,\ell_2])}} \chi_1(b_1)\chi_2(b_2)\frac{I(b_1,b_2)}{b_1b_2}.$$

Trivially $I(b_1,b_2) \ll A\min(b_1,b_2) \leqslant A\sqrt{b_1b_2}$, but the condition (8.1) implies that

$$\begin{aligned} &b_1^{\beta_1}b_2^{\beta_2}\frac{\partial^{\beta_1+\beta_2}}{\partial b_1^{\beta_1}\partial b_2^{\beta_2}}I(b_1,b_2) \\ &\ll Z^{\beta_1+\beta_2}A\sqrt{b_1b_2}\left(1+\frac{b_1}{B}\right)^{-2}\left(1+\frac{b_2}{B}\right)^{-2}\left(1+\frac{|h|}{AB}\right)^{-2}. \end{aligned}$$

Hence, applying Burgess's estimate (see [B])

$$\sum_{b\leqslant B}\chi(b)b^{-\frac{1}{2}} \ll D^{\frac{3}{16}+\varepsilon},$$

which holds for any non-trivial character $\chi \pmod{D}$, one derives

Proposition 8.2 *If the coefficients are given by (8.7) then*

$$W(h) \ll \tau(h)\left(1+\frac{|h|}{AB}\right)^{-2}Z^2A(D_1D_2)^{\frac{3}{16}+\varepsilon} \tag{8.8}$$

for any $\varepsilon > 0$, the implied constant depending on ε.

9 The non-singular contribution

Now we return to $R_c(h)$ for $h \neq 0$ which is given by (7.4) and (7.5). Opening $\sigma_F(m, \chi)$ and $\sigma_G(n, \chi)$ we obtain

$$V(h) = \sum\sum\sum\sum_{\ell a_1 b_2 - a_2 b_1 = h} \chi(b_2)\overline{\chi}(b_1)\overline{F}(a_1, b_2)G(a_2, b_1)$$

with $G(a_2, b_1)$ being the integral transform

$$\iint F(cx_1, cx_2)\, J_{k-1}\big(4\pi\sqrt{\ell a_1 b_2 x_1 x_2}\big) e(-x_1 b_1 - x_2 a_2) dx_1\, dx_2\ .$$

This is a sum of the type considered in Theorem 8.1 for

$$F(a_1, a_2; b_1, b_2) = \overline{F}(a_1, \ell^{-1} b_2)$$
$$\iint F(cx_1, cx_2)\, J_{k-1}(4\pi\sqrt{a_1 b_2 x_1 x_2}) e(-x_1 b_1 - x_2 a_2) dx_1\, dx_2\ .$$

We have $F(a_1, a_2; b_1, b_2) \ll c^{-2} X_1 X_2$ by a trivial estimation. The condition (8.1) is satisfied (after scaling by the factor $c^{-2} X_1 X_2$) with $Z = 2|s|$, $A = (c + \sqrt{\ell}\, X_1 X_2) X_2^{-1}$ and $B = (c + \sqrt{\ell}\, X_1 X_2) X_1^{-1}$. Moreover our coefficients are given by (8.7) with $\ell_1 = 1$, $\ell_2 = \ell$, $\chi_1 = \overline{\chi}$, $\chi_2 = \chi$. Combining Theorem 8.1 and Proposition 8.2 we get

$$V(h) \ll \tau(h)|s|^2 c^{\varepsilon - 2} X_1 X_2 \left(1 + \frac{|h| X_1 X_2}{(c + \sqrt{\ell}\, X_1 X_2)^2}\right)^{-2}$$
$$\left\{(c + \sqrt{\ell}\, X_1 X_2) X_2^{-1} D^{\frac{3}{8}} + (c + \sqrt{\ell}\, X_1 X_2)^2 (X_1 X_2)^{-1} \Big(X_1^{-1} X_2^{\frac{47}{48}}\Big)^{\frac{1}{24}}\right\}$$

and so, applying the bound $|S(h, 0; c)| \leqslant (h, c)$ then summing over all $h \neq 0$ we conclude that $R_c^* = R_c - R_c(0)$ satisfies

$$R_c^* \ll |s|^2 c^{\varepsilon - 1} (c + \sqrt{\ell}\, X_1 X_2)^2 \Big\{(c + \sqrt{\ell}\, X_1 X_2) X_1^{-1} D^{\frac{3}{8}}$$
$$+ (c + \sqrt{\ell}\, X_1 X_2)^2 (X_1 X_2)^{-1} \Big(X_1^{-1} X_2^{\frac{47}{48}}\Big)^{\frac{1}{24}}\Big\}.$$

This result is not useful if c is very large. We shall apply it for $c \leqslant C$, say, with some $C \geqslant \sqrt{\ell}\, X_1 X_2$ getting

$$\sum_{\substack{c \leqslant C \\ c \equiv 0 (D)}} c^{-1} R_c^* \ll \frac{|s|^2 C^{3+\varepsilon}}{D^2 X_1 X_2} \left\{ D^{\frac{3}{8}} X_1 + C \Big(X_1^{-1} X_2^{\frac{47}{48}}\Big)^{\frac{1}{24}} \right\}. \qquad (9.1)$$

10 An elementary estimate for R_c^*

For large c we can do better by using the original expression (6.3).

Lemma 10.1 *Suppose $F(x_1, x_2)$ is bounded and supported in the box $[X_1, 2X_2] \times [X_2, 2X_2]$ with $X_1, X_2 \geqslant \frac{1}{2}$. Then*

$$R_c \ll \left(\frac{\sqrt{\ell}\, X_1 X_2}{c}\right)^{k-1} (c + \ell X_1 X_2)\, X_1 X_2\, (\log 5X_1 X_2)^5 . \tag{10.1}$$

Proof. Open the Kloosterman sum $S_\chi(\ell m, n; c)$ and the Bessel function $J_{k-1}\big(\frac{4\pi}{c}\sqrt{\ell mn}\big)$ by means of the integral representation

$$J_{k-1}(x) = \frac{1}{2\pi i} \int_{(\sigma)} \frac{\Gamma(-s)}{\Gamma(s+k)} \left(\frac{x}{2}\right)^{2s+k-1} ds$$

with $\sigma = (\log 5X_1 X_2)^{-1}$ (see (8.412.4) of [GR]). Then apply

$$\sideset{}{^*}\sum_{d(\operatorname{mod} c)} \Bigg|\sum_{m \leqslant M} \alpha_m e\Big(\frac{d\ell m}{c}\Big)\Bigg| \, \Bigg|\sum_{n \leqslant N} \beta_n e\Big(\frac{\bar{d}\, n}{c}\Big)\Bigg| \leqslant \|\alpha\|\, \|\beta\| (c + \ell M)^{\frac{1}{2}} (c + N)^{\frac{1}{2}}$$

which follows by Cauchy's inequality and the orthogonality of additive characters. This leads to (10.1).

We also need a bound for $R_c(0)$. To this end we use the trivial estimate $G_k(b_1, b_2) \ll c^{-2} X_1 X_2 (c^{-1}\sqrt{\ell}\, X_1 X_2)^{k-1}$, giving

$$R_c(0) \ll \big(c^{-1}\sqrt{\ell}\, X_1 X_2\big)^{k-1} X_1 X_2 (\log 5X_1 X_2)^2.$$

This is absorbed by the right side of (10.1); therefore Lemma 10.1 holds also for $R_c^* = R_c - R_c(0)$.

Assume $k \geqslant 3$ and $C \geqslant \ell X_1 X_2$. By (10.1) for R_c^* we derive

$$\sum_{\substack{c > C \\ c \equiv 0 (D)}} c^{-1} R_c^* \ll \ell (CD)^{-1} (X_1 X_2)^3 (\log 5X_1 X_2)^5 . \tag{10.2}$$

11 Proof of Proposition 5.1. Conclusion

Adding (7.3), (9.1) and (10.2) we get

$$\sum_{c \equiv 0 (D)} c^{-1} R_c \ll \ell^{-\frac{1}{2}} X_1 X_2 |s|^2 D^\varepsilon + \ell (CD)^{-1} (X_1 X_2)^3 (\log 5X_1 X_2)^5$$
$$+ \Big\{ D^{\frac{3}{8}} X_1 + C \Big(X_1^{-1} X_2^{\frac{47}{48}} \Big)^{\frac{1}{24}} \Big\} D^{-2} (X_1 X_2)^{-1} |s|^2 C^{3+\varepsilon}$$

where C is any number $\geqslant \ell X_1 X_2$. We choose $C = \ell^{\frac{3}{2}} X_1 X_2$ and estimate some parts by using $X_1, X_2 < \sqrt{PD}$ to deduce that

$$\sum_{c \equiv 0(D)} c^{-1} R_c \ll \left(\ell^{-\frac{1}{2}} + \ell^6 D^{-\theta}\right) X_1 X_2 (|s|P)^2 D^{\varepsilon} \tag{11.1}$$

where $\theta = (48)^{-2}$. By virtue of (6.2) being absorbed by the first term on the right side of (11.1) this also gives a bound for the sum (6.1). Hence (5.8) follows.

References

[B] D.A. Burgess, On character sums and L–series II, Proc. London Math. Soc. **13**, (1963) 253–256

[DI] J.-M. Deshouillers, H. Iwaniec, Kloosterman sums and Fourier coefficients of cusp forms, Invent. math. **70**, (1982/83) 219–288

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

[DFI2] W. Duke, J. Friedlander, H. Iwaniec, A quadratic divisor problem, Invent. math. **115**, (1994) 209–217

[DFI3] W. Duke, J. Friedlander, H. Iwaniec, Bounds for automorphic L–functions II, Invent. math. **115**, (1994) 219–239; erratum, ibid. **140**, (2000) 227–242

[DFI4] W. Duke, J. Friedlander, H. Iwaniec, Class group L–functions, Duke Math. J. **79**, (1995) 1–56

[DFI5] W. Duke, J. Friedlander, H. Iwaniec, Bilinear forms with Kloosterman functions, Invent. math. **128**, (1997) 23–43

[DFI6] W. Duke, J. Friedlander, H. Iwaniec, Representations by the determinant and mean values of L–functions, Sieve Methods, Exponential Sums, and their Applications in Number Theory (Cardiff 1995), 109–115, London Math. Soc. Lecture Note Ser. 237, Cambridge Univ. Press, Cambridge 1997

[FoIw] E. Fouvry, H. Iwaniec, A subconvexity bound for Hecke L–functions, to appear in Ann. Scient. de l'École Normale Supérieure

[F] J.B. Friedlander, Bounds for L–functions, Proc. Int. Cong. Math. (Zürich 1994), 363–373, Birkhäuser (Basel) 1995

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

[GR] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series, and Products, Academic Press (London) 1965

[H] E. Hecke, Über eine neue Art von Zetafunktionen, Math. Zeit. **6**, (1920) 11–51

[I1] H. Iwaniec, Fourier coefficients of modular forms of half–integral weight, Invent. math. **87**, (1987) 385–401

[I2] H. Iwaniec, Topics in Classical Automorphic Forms, Graduate Studies in Math. Vol. 17, Amer. Math. Soc. (Providence) 1997

[KMV] E. Kowalski, P. Michel, J. VanderKam, Rankin–Selberg L–functions in the level aspect, preprint, 1999

[L] W. Li, Newforms and functional equations, Math. Ann. **212**, (1975) 285–315

[S] G. Shimura, On the holomorphy of certain Dirichlet series, Proc. London Math. Soc. **31**, (1975) 79–98