The orthogonality of Hecke eigenvalues

Henryk Iwaniec and Xiaoqing Li

Abstract

In this paper, we study the orthogonalities of Hecke eigenvalues of holomorphic cusp forms. An asymptotic large sieve with an unusually large main term for cusp forms is obtained. A family of special vectors formed by products of Kloosterman sums and Bessel functions is constructed for which the main term is exceptionally large. This surprising phenomenon reveals an interesting fact: that Fourier coefficients of cusp forms favor the direction of products of Kloosterman sums and Bessel functions of compatible type.

1. Introduction

The characters on residue classes modulo q (the additive and the multiplicative characters) enjoy the following orthogonality properties:

$$\frac{1}{q} \sum_{a(\mathrm{mod}\, q)} e\left(\frac{a(m-n)}{q}\right) = \begin{cases} 1 & \text{if } m \equiv n(\mathrm{mod}\, q) \\ 0 & \text{otherwise} \end{cases} \tag{1.1}$$

$$\frac{1}{\phi(q)} \sum_{\chi(\mathrm{mod}\, q)} \chi(m)\bar{\chi}(n) = \begin{cases} 1 & \text{if } m \equiv n(\mathrm{mod}\, q) \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

the latter for $(mn, q) = 1$. In analytic number theory, these properties manifest in the following forms

$$\frac{1}{q} \sum_{a(\mathrm{mod}\, q)} \left|\sum_{n \leqslant N} a_n e\left(\frac{an}{q}\right)\right|^2 = \left(1 + O\left(\frac{N}{q}\right)\right) \sum_{n \leqslant N} |a_n|^2 \tag{1.3}$$

$$\frac{1}{\phi(q)} \sum_{\chi(\mathrm{mod}\, q)} \left|\sum_{n \leqslant N} a_n \chi(n)\right|^2 = \left(1 + O\left(\frac{N}{q}\right)\right) \sum_{\substack{n \leqslant N \\ (n,q)=1}} |a_n|^2 \tag{1.4}$$

where a_n are arbitrary complex numbers. Owing to the considerable cancellation of terms which these formulas reveal, we can say that a fixed sequence $\alpha = (a_n)$ is almost orthogonal to almost all characters.

Much stronger effects of orthogonality are observed for primitive characters to various moduli (conductors), specifically we have

$$\sum_{q \leqslant Q} \sum_{a(\mathrm{mod}\, q)}^{*} \left|\sum_{n \leqslant N} a_n e\left(\frac{an}{q}\right)\right|^2 \leqslant (Q^2 + N)\|\alpha\|^2 \tag{1.5}$$

$$\int_{-R}^{R} \sum_{q \leqslant Q} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, q)} \left| \sum_{n \leqslant N} a_n e\left(\frac{an}{q}\right) n^{ir} \right|^2 dr \ll (Q^2 R + N)\|\alpha\|^2 \qquad (1.6)$$

$$\sum_{q \leqslant Q} \sideset{}{^*}\sum_{\chi \,(\mathrm{mod}\, q)} \left| \sum_{n \leqslant N} a_n \chi(n) \right|^2 \leqslant (Q^2 + N)\|\alpha\|^2 \qquad (1.7)$$

where the superscript * restricts the summation by $(a, q) = 1$ and χ primitive, respectively, and $\|\alpha\|$ is the l_2-norm

$$\|\alpha\|^2 = \sum_{n \leqslant N} |a_n|^2. \qquad (1.8)$$

These results are called *large sieve inequalities* because of their historical relevance to the evolution of the large sieve theory (cf. [Bom74, Lin41, Mon71]).

It is a critical feature of (1.5) and (1.7) being inequalities rather than more precise approximate formulas such as (1.3) and (1.4). In fact, Q^2 is not the right number of characters in the family, asymptotically this is $3\pi^{-2}Q^2$ in (1.5), so the contribution of the diagonal terms in (1.5) is smaller than one-third of $Q^2\|\alpha\|^2$. This discrepancy in the constants comes from the proofs which use the duality principle (stating that the norm of a linear operator acting on Banach spaces equals the norm of its adjoint). For general coefficients $\alpha = (a_n)$ the upper bounds (1.5), (1.7) cannot be improved by a constant factor. However, with some restrictions on the coefficients (practically acceptable) one can establish the expected asymptotic large sieve formula. Of course, the duality principle must be abandoned, and this makes the problem much harder.

Over the last 25 years analytic number theory has been driven by the spectral theory of automorphic forms. The Fourier coefficients of cusp forms, or preferably the eigenvalues of Hecke operators acting on cusp forms, play a role similar to characters, and they enjoy more orthogonality. In this paper we work in the Hilbert space $S_k(\Gamma, \chi)$ of cusp forms of weight $k \geqslant 3$ for the group $\Gamma = \Gamma_0(q)$ and the nebentypus character $\chi\,(\mathrm{mod}\, q)$. An analog of (1.3) and (1.4) for this space is

$$\sum_{f \in \mathcal{F}_\chi} \left| \sum_{n \leqslant N} a_n \psi_f(n) \right|^2 = \left(1 + O\left(\frac{N}{q}\right)\right)\|\alpha\|^2 \qquad (1.9)$$

where $\mathcal{F}_\chi$ is any orthonormal basis of $S_k(\Gamma, \chi)$ and $\psi_f(n)$ are normalized Fourier coefficients of f, the implied constant depending only on k (the left-hand side of (1.9) is basis independent, see the next section and [Iwa97, Theorem 5.7]).

In this paper we expand (1.9) by averaging over the characters $\chi\,(\mathrm{mod}\, q)$ of compatible parity, so our family of cusp forms $\mathcal{F} = \bigcup_\chi \mathcal{F}_\chi$ is significantly larger. In other words, we work in the space

$$S_k(\Gamma_1(q)) = \bigoplus_{\chi \,(\mathrm{mod}\, q)} S_k(\Gamma_0(q), \chi).$$

We are exclusively interested in the level aspect, so k is fixed while q is considered to be large. While each basis $\mathcal{F}_\chi$ has about q elements, more precisely we have $|\mathcal{F}_\chi| \sim \frac{1}{12}(k-1)\nu(q)$, where

$$\nu(q) = [\Gamma_0(1) : \Gamma_0(q)] = q \prod_{p|q}(1 + p^{-1}),$$

the whole family $\mathcal{F}$ has about q^2 elements, more precisely

$$|\mathcal{F}| = \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} |\mathcal{F}_\chi| \sim \frac{1}{24}(k-1)q^2 \prod_{p|q}(1 - p^{-2}).$$

One could choose $\mathcal{F}$ to consist of primitive cusp forms (of levels dividing q which are common eigenfunctions of all Hecke operators). This special choice is popular in advanced arithmetic; however,

it has no advantage for our considerations. In fact, our conclusions will sound nicer if $\mathcal{F}$ is not fixed. We shall see that whatever a basis is, it must contain many cusp forms whose coefficients are pointing towards some distinguished vectors.

Specifically, set

$$\mathcal{L}_f(\alpha) = \sum_m a_m \psi_f(m),$$

$$\mathcal{P}_{ht}(\alpha) = \sum_n a_n S(h\bar{q}, n; t) J_{k-1}\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right),$$

our main theorem is the following.

THEOREM 1.1. *Let q be prime, $N \geqslant q$, $T = Nq^{-1}$ and $1 \leqslant H \leqslant T$. Then for any complex vectors $\alpha = (a_n)$ with $N < n \leqslant 2N$ we have*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi(\mathrm{mod}\,q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2 = \frac{1}{q} \sum_{\substack{1 \leqslant t \leqslant T \\ (t,q)=1}} \left(\frac{2\pi}{t}\right)^2 \sum_{1 \leqslant h \leqslant H} |\mathcal{P}_{ht}(\alpha)|^2 + O\left(N^\varepsilon\left(\frac{N}{q^2} + \sqrt{\frac{N}{qH}}\right)\right)\|\alpha\|^2$$

(1.10)

with any $\varepsilon > 0$, $S(m, n; c)$ is the Kloosterman sum defined by

$$S(m, n; t) = \sum_{a\bar{a} \equiv 1(\mathrm{mod}\,t)} e\left(\frac{am + \bar{a}n}{c}\right)$$

(1.11)

and $J_{k-1}(x)$ is the J-Bessel function of order $k-1$, the implied constant depending only on k and ε.

Our main theorem provides an asymptotic large sieve formula for the family $\mathcal{F}$. This formula does reveal that the orthogonality of sequences of Fourier coefficients of cusp forms is stronger within the larger family $\mathcal{F}$, but surprisingly enough this is not perfectly comparable with the classical large sieve inequalities (1.5) and (1.7). The asymptotic formula in question contains a series of main terms. The situation does not improve even if we restrict to primitive forms (for example, take q prime and k odd), the large main term is still there. From the main term one can construct vectors $\alpha = (a_n)$ having large projections on the coefficients of many cusp forms in the basis. The results are fascinating because these special vectors (see (13.1)) are built out of Kloosterman sums and Bessel functions. Hence, we conclude that the Fourier coefficients of cusp forms tend to point in the directions of the Kloosterman–Bessel products to some degree.

The applications of our main theorem would include the estimate of the eighth moment of L-functions, zero density estimates of L-functions, etc. These could be worked out along the lines in [Mon71].

2. A trace formula for the Fourier coefficients of cusp forms

Throughout this work $k \geqslant 3$, $q \geqslant 3$ and $\Gamma = \Gamma_0(q)$, which is a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ with the condition that $c \equiv 0(\mathrm{mod}\,q)$ for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q)$. The group $\Gamma_0(q)$ acts on the upper half-plane $\mathbb{H}$ by linear fractional transformations

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

Let $\chi(\mathrm{mod}\,q)$ be a Dirichlet character and $S_k(\Gamma, \chi)$ be the linear space of cusp forms $f : \mathbb{H} \to \mathbb{C}$ which satisfy

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

543

Since $-1 \in \Gamma$, the space $S_k(\Gamma, \chi)$ is zero unless χ satisfies the parity condition

$$\chi(-1) = (-1)^k \tag{2.1}$$

which we henceforth assume to hold. An important fact is that $S_k(\Gamma, \chi)$ is a finite-dimensional Hilbert space equipped with the Petersson's inner product

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

Let $\mathcal{F}_\chi$ be an orthonormal basis of $S_k(\Gamma, \chi)$. We have

$$|\mathcal{F}_\chi| = \dim S_k(\Gamma, \chi) \asymp kq \prod_{p \mid q} \left(1 + \frac{1}{p}\right). \tag{2.2}$$

Every $f \in \mathcal{F}_\chi$ has a Fourier series expansion

$$f(z) = \sum_{n \geqslant 1} \hat{f}(n)e(nz) \tag{2.3}$$

with some complex coefficients $\hat{f}(n)$, where as usual $e(z) = e^{2\pi i z}$. These coefficients satisfy the Ramanujan bound (proved by P. Deligne)

$$\hat{f}(n) \ll \tau(n)n^{(k-1)/2}$$

where the implied constant depends on k and q. This holds with the implied constant if f is a Hecke normalized primitive cusp form. However, we prefer to work with the spectrally normalized coefficients

$$\psi_f(n) = (\Gamma(k-1)/(4\pi n)^{k-1})^{\frac{1}{2}} \hat{f}(n). \tag{2.4}$$

One can show that (see (2.10))

$$\sum_{f \in \mathcal{F}_\chi} |\psi_f(n)|^2 = 1 + O\left(\frac{n}{q}\right) \tag{2.5}$$

where the implied constant depends only on k. This tells us that the normalized coefficients $\psi_f(n)$ have an average size of about $1/\sqrt{\nu(q)}$.

Let $S_k(\Gamma_1(q))$ be the space of holomorphic cusp forms for the group

$$\Gamma_1(q) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{q} \right\}.$$

In this section, we derive a kind of trace formula for $S_k(\Gamma_1(q))$ from that for $S_k(\Gamma_0(q), \chi)$.

For any positive integers m, n we introduce the trace of the Fourier coefficients (of forms in $\mathcal{F}_\chi$)

$$\Delta_\chi(m, n) = \sum_{f \in \mathcal{F}_\chi} \bar{\psi}_f(m)\psi_f(n). \tag{2.6}$$

This satisfies the following formula of Petersson (cf. [Iwa97])

$$\Delta_\chi(m, n) = \delta(m, n) + \sigma_\chi(m, n) \tag{2.7}$$

where $\delta(m, n)$ is the diagonal symbol of Kronecker and

$$\sigma_\chi(m, n) = 2\pi i^{-k} \sum_{c \equiv 0 \pmod{q}} c^{-1} S_\chi(m, n; c) J_{k-1}\left(\frac{4\pi}{c}\sqrt{mn}\right). \tag{2.8}$$

Here $S_\chi(m, n; c)$ is the Kloosterman sum defined by

$$S_\chi(m, n; c) = \sum_{a\bar{a} \equiv 1 \pmod{c}} \bar{\chi}(a)e\left(\frac{am + \bar{a}n}{c}\right). \tag{2.9}$$

Using trivial estimates $|S_\chi(m, n; c)| \leqslant c$ and (A.1) it follows from (2.7) that

$$\Delta_\chi(m, n) = \delta(m, n) + O(q^{-1}\sqrt{mn}) \tag{2.10}$$

where the implied constant depends only on k.

The main object of our investigation is $\Delta(m, n)$-the trace of Fourier coefficients of forms in the whole family $\mathcal{F}$. This is obtained by averaging $\Delta_\chi(m, n)$ over the characters, precisely

$$\Delta(m, n) = \frac{2}{\phi(q)} \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} \bar{\psi}_f(m)\psi_f(n). \tag{2.11}$$

Summing both sides of (2.7) over the characters of relevant parity we obtain

$$\Delta(m, n) = \delta(m, n) + \sigma(m, n) \tag{2.12}$$

where $\sigma(m, n)$ is the sum (2.8) with the Kloosterman sums $S_\chi(m, n; c)$ replaced by their average

$$\mathcal{S}(m, n; c) = \frac{2}{\phi(q)} \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} S_\chi(m, n; c). \tag{2.13}$$

Using trivial estimates $|\mathcal{S}(m, n; c)| \leqslant 2cq^{-1}$ and (A.1) it follows from (2.12) that

$$\Delta(m, n) = \delta(m, n) + O(q^{-2}\sqrt{mn}) \tag{2.14}$$

where the error term is much smaller than that in (2.10). However, we are going to establish considerably better estimates for general bilinear forms in $\Delta(m, n)$ with respect to m, n essentially saving extra factor $\sqrt{mn}$.

To treat both cases (k even or k odd) uniformly, we use the operator $\mathcal{K}$ defined by

$$\mathcal{K}f = i^{-k}f + i^k\bar{f} = 2\Re(i^{-k}f). \tag{2.15}$$

Note that if g is real then $\mathcal{K}(gf) = g\mathcal{K}f$. For the exponential function $e(x) = e^{2\pi ix}$ we have

$$\mathcal{K}e(x) = (-1)^{[k/2]} \begin{cases} 2\cos(2\pi x) & \text{if } k \text{ even}, \\ 2\sin(2\pi x) & \text{if } k \text{ odd}. \end{cases}$$

By opening the Kloosterman sum $S_\chi(m, n; c)$ in (2.13) we get

$$\frac{1}{\phi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\, c)} (\chi(a) + (-1)^k\chi(-a))e\left(\frac{am + \bar{a}n}{c}\right).$$

Hence, by the orthogonality of characters (1.2) we find that

$$\mathcal{S}(m, n; c) = i^k\mathcal{K} \sideset{}{^*}\sum_{\substack{a(c) \\ a\equiv 1(q)}} e\left(\frac{am + \bar{a}n}{c}\right).$$

Now we put $c = qr$ and $a = 1 + qy$ with y running modulo r. Let $(y, r) = s$, so $r = st$ and $a = 1 + qsx$ with x running modulo t, $(x(1 + qsx), t) = 1$. We get

$$\sideset{}{^*}\sum_a = \sum_{qst=c} \sum_{\substack{x \,(\mathrm{mod}\, t) \\ (x(1+qsx),t)=1}} e\left(\frac{m(1 + qsx) + n\overline{(1 + qsx)}}{qst}\right).$$

Next, writing $\bar{a} \equiv 1 - (a - 1)\bar{a}\,(\mathrm{mod}\, t)$ for $a = 1 + qsx$ and changing x to $\bar{x}$, we arrive at

$$\mathcal{S}(m, n; c) = i^k\mathcal{K}e\left(\frac{m + n}{c}\right) \sum_{qst=c} V_{qs}(m, n; t) \tag{2.16}$$

545

where

$$V_d(m, n; t) = \sum_{\substack{x \,(\mathrm{mod}\, t) \\ (x(x+d), t)=1}} e\left(\frac{m\bar{x} - n\overline{x+d}}{t}\right). \tag{2.17}$$

Note that $V_d(m, n; t)$ is essentially a Kloosterman sum; in particular, we have

$$V_d(m, n; t) = e\left(-(m + n)\frac{\bar{d}}{t}\right) S(m\bar{d}, n\bar{d}; t) - 1, \quad \text{if } (d, t) = 1, t'. \tag{2.18}$$

If $t | d$ we have the Ramanujan sum $V_0(m, n; t) = S(0, m - n; t)$. Gathering the above results we obtain the following primary version of the Petersson formula for the trace $\Delta(m, n)$ in $S_k(\Gamma_1(q))$.

Lemma 2.1. *For any positive integers m, n we have*

$$\Delta(m, n) = \delta(m, n) + \sigma(m, n) \tag{2.19}$$

where

$$\sigma(m, n) = \mathcal{K} \sum_s \sum_t \frac{2\pi}{qst} V_{qs}(m, n; t) e\left(\frac{m + n}{qst}\right) J_{k-1}\left(\frac{4\pi\sqrt{mn}}{qst}\right) \tag{2.20}$$

and $V_{qs}(m, n; t)$ is given by (2.17).

In the next section we shall elaborate on $\sigma(m, n)$ further by applying Fourier analysis.

3. An asymptotic trace formula

Changing the order of summation we write (2.20) as

$$\sigma(m, n) = \sum_t \frac{2\pi}{qt} \sigma(m, n; t) \tag{3.1}$$

where

$$\sigma(m, n; t) = \mathcal{K} \sum_s s^{-1} V_{qs}(m, n; t) e\left(\frac{m + n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) \tag{3.2}$$

and $J(x) = J_{k-1}(x)$ is the Bessel function of order $k - 1$. We are going to execute the summation over s by Poisson's formula. Before this application we insert a redundant factor $\eta(s)$ which is a smooth function on $\mathbb{R}^+$ with $\eta(s) = 0$ if $0 < s < \frac{1}{4}$, $0 \leqslant \eta(s) \leqslant 1$ if $\frac{1}{4} \leqslant s \leqslant \frac{1}{2}$ and $\eta(s) = 1$ if $s > \frac{1}{2}$. We get

$$\sigma(m, n; t) = \mathcal{K} \sum_{z(t)} V_{qz}(m, n; t) \sum_{s \equiv z(t)} \frac{\eta(s)}{s} e\left(\frac{m + n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right)$$

$$= \mathcal{K} \sum_h \frac{1}{t} \sum_{z(t)} V_{qz}(m, n; t) e\left(\frac{-hz}{t}\right) \int_0^\infty \frac{\eta(s)}{s} e\left(\frac{hs}{t} + \frac{m + n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) ds. \tag{3.3}$$

Inserting the expression (A.5) for the Bessel function we obtain two Fourier integrals with the phase function

$$\frac{hs}{t} + \frac{m \pm 2\sqrt{mn} + n}{qst}.$$

None of these has a stationary point if $|h| \geqslant 64(m + n)q^{-1}$ (the linear term hs/t dominates by a considerable factor). Integrating by parts A times, we find that

$$\int_0^\infty \frac{\eta(s)}{s} e\left(\frac{hs}{t} + \frac{m + n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) ds \ll \left(1 + \frac{|h|}{t}\right)^{-A}$$

546

provided $|h| \geqslant 64(m+n)q^{-1}$, where the implied constant depends only on k, A and the cut-off function η. This is very small if $|h|$ is slightly larger than t, because A is arbitrary. Choose any $H = H(t)$ satisfying

$$H \geqslant t + 64(m+n)q^{-1}. \tag{3.4}$$

Actually we shall require later that H is somewhat larger. The above estimate allows us to remove the summation over $|h| \geqslant H$ in (3.3) at the cost of a small error term. We do it smoothly by inserting a cut-off factor $\omega(|h|/H)$, where $\omega(x)$ is a smooth function on $\mathbb{R}^+$ with $\omega(x) = 1$ if $0 \leqslant x \leqslant 1$ and $\omega(x) = 0$ if $x > 2$. We obtain

$$\sigma(m,n;t) = \sum_h \omega\left(\frac{|h|}{H}\right) \frac{1}{t} \sum_{z(t)} V_{qz}(m,n,t) e\left(\frac{-hz}{t}\right)$$
$$\times \mathcal{K} \int_0^\infty \frac{\eta(s)}{s} e\left(\frac{hs}{t} + \frac{m+n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) ds + \sigma'(m,n;t) \tag{3.5}$$

where

$$\sigma'(m,n;t) \ll t \sum_{h>H} \left(\frac{t}{h}\right)^A \ll tH\left(\frac{t}{H}\right)^A \tag{3.6}$$

by the trivial estimation $|V_d(m,n;t)| \leqslant t$. Note that we moved the operator $\mathcal{K}$ to the integral because the preceding terms are real.

Next we extend the integral to all $s > 0$. The complete integral equals

$$\mathcal{K} \int_0^\infty e\left(\frac{hs}{t} + \frac{m+n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) \frac{ds}{s} = 2\pi J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right) \tag{3.7}$$

if $h > 0$ and it vanishes if $h \leqslant 0$ (see (A.9)). The complementary integral is

$$\mathcal{K} \int_0^{\frac{1}{2}} \frac{\xi(s)}{s} e\left(\frac{hs}{t} + \frac{m+n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) ds \tag{3.8}$$

where $\xi(s) = 1 - \eta(s)$, so $\xi(s) = 1$ if $0 < s < \frac{1}{4}$, $0 \leqslant \xi(s) \leqslant 1$ if $\frac{1}{4} \leqslant s \leqslant \frac{1}{2}$ and $\xi(s) = 0$ if $s > \frac{1}{2}$.

Now we proceed to the sum (3.5) with the complementary integral (3.8). Here we interchange the integration over $0 < s < \frac{1}{2}$ with the summation over $|h| \leqslant 2H$, and execute the latter by Poisson's formula

$$\sum_h \omega\left(\frac{|h|}{H}\right) e\left(\frac{s-z}{t}h\right) = H \sum_l \hat{\omega}\left(\left(l + \frac{z-s}{t}\right)H\right), \tag{3.9}$$

where $\hat{\omega}(y)$ is the Fourier transform of $\omega(|x|)$. This satisfies

$$\hat{\omega}(y) \ll (1 + |y|)^{-A} \tag{3.10}$$

for any $A \geqslant 0$, the implied constant depending on A and ω. Hence, we see that the sum (3.9) is very small unless $l + z/t = 0$, so $z \equiv 0(\mathrm{mod}\, t)$, in which case it is equal to $H\hat{\omega}(Hs/t)$ up to a small error term. Precisely, the error term is much less than $H(t/H)^A$. This error term contributes to $\sigma(m,n;t)$ a quantity, say $\sigma''(m,n;t)$, which satisfies (use that $J(x) \in L^1(\mathbb{R}^+, x^{-1}\,dx)$)

$$\sigma''(m,n;t) \ll tH\left(\frac{t}{H}\right)^A. \tag{3.11}$$

We are left with the condition $l + z/t = 0$ which yields

$$S(0, m-n;t)\mathcal{K} \int_0^\infty H\hat{\omega}\left(\frac{Hs}{t}\right) \frac{\xi(s)}{st} e\left(\frac{m+n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) ds. \tag{3.12}$$

547

Next we are going to simplify the above integral. Note that $\hat{\omega}(Hs/t)$ is very small if s is slightly larger than tH^{-1}, see (3.10). Because $\xi(s) = 1$ if $s < \frac{1}{4}$, we can remove $\xi(s)$ in (3.12) up to a term, say $\sigma'''(m, n; t)$, which satisfies

$$\sigma'''(m, n; t) \ll H \int_{\frac{1}{4}}^{\infty} \left(1 + \frac{Hs}{t}\right)^{-A} \frac{ds}{s} \ll H \left(\frac{t}{H}\right)^A. \tag{3.13}$$

Finally, we are left with the integral

$$\mathcal{K} \int_0^{\infty} H\hat{\omega}\left(\frac{Hs}{t}\right) e\left(\frac{m+n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) \frac{ds}{s}. \tag{3.14}$$

By Plancherel's formula $\langle \hat{F}, G \rangle = \langle F, \hat{G} \rangle$ this integral equals

$$\mathcal{K} \int_{-\infty}^{\infty} \omega\left(\frac{|h|}{H}\right) \left(\int_0^{\infty} e\left(\frac{hs}{t} + \frac{m+n}{qst}\right) J\left(\frac{4\pi\sqrt{mn}}{qst}\right) \frac{ds}{s}\right) dh$$

$$= 2\pi \int_0^{\infty} \omega\left(\frac{h}{H}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right) dh$$

as in (3.7) (see (A.9)).

Collecting the above results we obtain the following formula

$$\sigma(m, n; t) = \sigma^+(m, n; t) - \sigma^-(m, n; t) + \sigma^{\infty}(m, n; t) \tag{3.15}$$

where

$$\sigma^+(m, n; t) = \frac{2\pi}{t} \sum_{h=1}^{\infty} \omega\left(\frac{h}{H}\right) W(h, m, n; t) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right) \tag{3.16}$$

$$\sigma^-(m, n; t) = \frac{2\pi}{t} S(0, m-n; t) \int_0^{\infty} \omega\left(\frac{h}{H}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right) dh \tag{3.17}$$

and $\sigma^{\infty}(m, n; t)$ is the remainder term which satisfies

$$\sigma^{\infty}(m, n; t) \ll tH\left(\frac{t}{H}\right)^A. \tag{3.18}$$

The leading term $\sigma^+(m, n; t)$ contains the exponential sum

$$W(h, m, n; t) = \sum_{z(\bmod t)} V_{qz}(m, n; t) e\left(-\frac{hz}{t}\right) \tag{3.19}$$

where $V_d(m, n; t)$ is given by (2.17). Hence, $W(h, m, n; t)$ opens to

$$W(h, m, n; t) = \sum_{\substack{x(t) \\ }} \sum_{\substack{z(t) \\ (x(x+qz),t)=1}} e\left(\frac{m\bar{x} - n\overline{x+qz} - hz}{t}\right). \tag{3.20}$$

Note that $W(h, m, n; t)$ is real, the property which we have already used to move the operator $\mathcal{K}$ in (3.5) to the integral over s.

It is a wonderful fact that in most cases the exponential sum $W(h, m, n; t)$ factors into Kloosterman sums in the same fashion as the integral (3.7) factors into Bessel functions. The resemblance of the Kloosterman sums to the Bessel functions is striking (compare (3.7) and (3.21)). Precisely in all cases we have the following lemma.

548

LEMMA 3.1. *For h, m, n, t integers, $t > 0$, we have*

$$W(h, m, n; t) = \sum_{\substack{a \,(\mathrm{mod}\, t) \\ aq \equiv h \,(\mathrm{mod}\, t)}} S(a, m; t) S(a, n; t). \tag{3.21}$$

Proof. The double sum (3.20) equals

$$\sum_{x \,(\mathrm{mod}\, t)} \sum_{\substack{y \,(\mathrm{mod}\, t) \\ (xy, t) = 1}} \sum_{\substack{z \,(\mathrm{mod}\, t) \\ y \equiv x + qz \,(\mathrm{mod}\, t)}} e\left(\frac{m\bar{x} - n\bar{y} - hz}{t}\right)$$

$$= \frac{1}{t} \sum_{a \,(\mathrm{mod}\, t)} \sum_{x} \sum_{y} \sum_{z} e\left(\frac{m\bar{x} - n\bar{y} - hz + a(x - y + qz)}{t}\right)$$

$$= \sum_{\substack{a \,(\mathrm{mod}\, t) \\ aq \equiv h \,(\mathrm{mod}\, t)}} \sum_{x} \sum_{y} e\left(\frac{m\bar{x} - n\bar{y} + a(x - y)}{t}\right)$$

and this is the right-hand side of (3.21). □

Remarks. The summation over the classes $a \,(\mathrm{mod}\, t)$ in (3.21) is quite short. Precisely, the condition $aq \equiv h \,(\mathrm{mod}\, t)$ implies $(q, t) | h$ and $a \equiv h_1 \bar{q}_1 \,(\mathrm{mod}\, t_1)$ where $h_1 = h/(q, t)$, $q_1 = q/(q, t)$, $t_1 = t/(q, t)$, therefore the number of classes is exactly (q, t).

COROLLARY 3.1. *If $(q, t) = 1$, then*

$$W(h, m, n; t) = S(h\bar{q}, m; t) S(h\bar{q}, n; t). \tag{3.22}$$

COROLLARY 3.2. *If $t = qt_1$ with $(q, t_1) = 1$, then $h = qh_1$ and*

$$W(h, m, n; t) = qS(0, m - n; q) S(h_1 \bar{q}^2, m; t_1) S(h_1 \bar{q}^2, n; t_1).$$

Proof. We have

$$S(a, m; t) S(a, n; t) = S(a\bar{t}_1^2, m; q) S(a\bar{t}_1^2, n; q) S(a\bar{q}^2, m; t_1) S(a\bar{q}^2, n; t_1)$$

and $a \equiv h_1 + bt_1$ with b running freely modulo q. This yields the Ramanujan sum $S(0, m - n; q)$ and the other two factors (Kloosterman sums). □

Inserting (3.21) into (3.16) we get

$$\sigma^+(m, n; t) = \frac{2\pi}{t} \sum_{h=1}^{\infty} \omega\left(\frac{h}{H}\right) \sum_{\substack{a \,(\mathrm{mod}\, t) \\ aq \equiv h \,(\mathrm{mod}\, t)}} S(a, m; t) S(a, n; t) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right). \tag{3.23}$$

Now we are ready to state our first main result.

THEOREM 3.1. *Let $\omega(x)$ be a smooth function on $\mathbb{R}^+$ with $\omega(x) = 1$ if $0 \leqslant x \leqslant 1$ and $\omega(x) = 0$ if $x > 2$. For any positive integers m, n we have*

$$\Delta(m, n) = \delta(m, n) + \sigma^+(m, n) - \sigma^-(m, n) + \sigma^\infty(m, n) \tag{3.24}$$

549

where

$$\sigma^+(m,n) = \frac{1}{q}\sum_t \left(\frac{2\pi}{t}\right)^2 \sum_{h=1}^{\infty} \omega\left(\frac{h}{H}\right) \sum_{\substack{a(\mathrm{mod}\,t) \\ aq\equiv h(\mathrm{mod}\,t)}} S(a,m;t)S(a,n;t)J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right),$$

(3.25)

$$\sigma^-(m,n) = \frac{1}{q}\sum_t \left(\frac{2\pi}{t}\right)^2 S(0,m-n;t)\int_0^{\infty} \omega\left(\frac{h}{H}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right)\,dh \qquad (3.26)$$

and

$$\sigma^{\infty}(m,n) \ll \frac{1}{q}\int_1^{\infty} \left(\frac{t}{H(t)}\right)^A t\,dt. \tag{3.27}$$

This holds with any $H = H(t) \geqslant t + 64(m+n)q^{-1}$ and any constant $A \geqslant 0$ for which the integral (3.27) converges. The implied constant in the estimation (3.27) depends only on k, A and the cut-off function ω.

Proof. The main terms in (3.24) come from (2.19), (3.1), (3.15), (3.23) and (3.17), and the estimation of the remainder term (3.27) comes from (3.18). $\qquad\square$

Of course, our intention is to choose the cut-off parameter $H = H(t)$ as small as possible to make the sums $\sigma^+(m,n)$, $\sigma^-(m,n)$ as short as possible and in the same time to have a sufficiently good bound for the remainder term $\sigma^{\infty}(m,n)$.

Examples. Let $T \geqslant (m+n)/2q$ and $0 < \varepsilon \leqslant 1$. Choose

$$H(t) \geqslant 128(t+T)t^{\varepsilon}. \tag{3.28}$$

Then (3.24) holds with $\sigma^+(m,n)$, $\sigma^-(m,n)$ as in (3.25), (3.26) and the remainder term satisfies

$$\sigma^{\infty}(m,n) \ll q^{-1}T^{-A} \tag{3.29}$$

with any $A > 0$, the implied constant depends only on ε, k, A and the cut-off function ω. Indeed, if $A > 2\varepsilon^{-1}$, the integral in (3.27) is bounded by

$$\int_0^{\infty} \left(\frac{t}{H(t)}\right)^A t\,dt \leqslant (\varepsilon A - 2)^{-1}T^{2-\varepsilon A}.$$

This gives (3.29) with $\varepsilon A - 2$ in place of A. Changing the constant A completes the proof.

4. Bilinear forms of the trace

The orthogonality of the Fourier coefficients $\psi_f(m)$ of the cusp forms $f \in \mathcal{F}$ is best expressed quantitatively in the context of bilinear forms

$$\mathcal{B}(\alpha,\beta) = \sum_m \sum_n \bar{a}_m b_n \Delta(m,n) \tag{4.1}$$

where a_m, b_n are arbitrary complex numbers for $1 \leqslant m, n \leqslant N$ and $\alpha = (a_m)$, $\beta = (b_n)$ denote the corresponding vectors. By (2.11) this decomposes into a sum of products of general linear forms $\mathcal{L}_f(\alpha)$, recall

$$\mathcal{L}_f(\alpha) = \sum_m a_m \psi_f(m), \tag{4.2}$$

550

that is

$$\mathcal{B}(\alpha, \beta) = \frac{2}{\phi(q)} \sum_{\substack{\chi (\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} \bar{\mathcal{L}}_f(\alpha) \mathcal{L}_f(\beta). \tag{4.3}$$

We shall evaluate $\mathcal{B}(\alpha, \beta)$ asymptotically by means of the trace formula (3.24). First we write, by (2.12),

$$\mathcal{B}(\alpha, \beta) = \langle \alpha, \beta \rangle + \mathcal{C}(\alpha, \beta) \tag{4.4}$$

where

$$\langle \alpha, \beta \rangle = \sum_n \bar{a}_n b_n \tag{4.5}$$

and

$$\mathcal{C}(\alpha, \beta) = \sum_m \sum_n \bar{a}_m b_n \sigma(m, n). \tag{4.6}$$

Then, according to (3.24), this splits into

$$\mathcal{C}(\alpha, \beta) = \mathcal{C}^+(\alpha, \beta) - \mathcal{C}^-(\alpha, \beta) + \mathcal{C}^\infty(\alpha, \beta). \tag{4.7}$$

The last form $\mathcal{C}^\infty(\alpha, \beta)$ can be estimated directly by using (3.29). For example, if $H(t)$ satisfies (3.28) with $T \geqslant Nq^{-1}$, we get

$$\mathcal{C}^\infty(\alpha, \beta) \ll \|\alpha\|\|\beta\| Nq^{-1}T^{-A}. \tag{4.8}$$

One can also estimate trivially $\mathcal{C}^+(\alpha, \beta), \mathcal{C}^-(\alpha, \beta)$ obtaining a non-trivial bound for $\mathcal{B}(\alpha, \beta)$ if the vectors α, β are not very long relative to q. However, we shall do better by exploiting specific structures of these forms. By (3.25) the bilinear form $\mathcal{C}^+(\alpha, \beta)$ decomposes into a sum of products of a new kind of linear forms

$$\mathcal{P}_{aht}(\alpha) = \sum_m a_m S(a, m; t) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) \tag{4.9}$$

where $J(x) = J_{k-1}(x)$ is the Bessel function and $S(a, m; t)$ is the Kloosterman sum with $a(\mathrm{mod}\, t)$ satisfying

$$aq \equiv h(\mathrm{mod}\, t). \tag{4.10}$$

Precisely, we have

$$\mathcal{C}^+(\alpha, \beta) = \frac{1}{q} \sum_t \left(\frac{2\pi}{t}\right)^2 \sum_{h=1}^\infty \omega\left(\frac{h}{H}\right) \sum_{\substack{a(\mathrm{mod}\, t) \\ aq \equiv h(\mathrm{mod}\, t)}} \mathcal{P}_{aht}(\alpha) \mathcal{P}_{aht}(\beta).$$

If $(q, t) = 1$, then (4.10) determines $a(\mathrm{mod}\, t)$ uniquely, namely

$$a \equiv h\bar{q}(\mathrm{mod}\, t). \tag{4.11}$$

In this case (a major case) we get linear forms (4.9) in the following terms

$$S(h\bar{q}, m; t) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right). \tag{4.12}$$

We have already seen in the context of Petersson's formula that the Bessel function $J(4\pi\sqrt{mn}/c)$ is a natural companion of the Kloosterman sum $S(m, n; c)$. Our new products (4.12) offer extra support for this natural association. It becomes even more convincing if one realizes that the products (4.12) emerged from

$$S(m, n; c) J\left(\frac{4\pi\sqrt{mn}}{c}\right) \tag{4.13}$$

551

by an application of Fourier analysis, so they are not derived by purely arithmetical substitutions. All of these observations suggest that the products (4.12) as well as (4.13) deserve their own names. Perhaps these should be called *absolute Kloosterman sums*. Moreover, in view of the structure and relations between our linear forms one can think of the absolute Kloosterman sums (4.12) as dual objects to the Fourier coefficients $\psi_f(m)$.

Similarly by (3.26) the bilinear form $\mathcal{C}^-(\alpha, \beta)$ decomposes into a sum of products of linear forms

$$\mathcal{Q}_{aht}(\alpha) = \sum_m a_m e\left(\frac{am}{t}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hm}{q}}\right) \tag{4.14}$$

where $(a, t) = 1$ (after opening the Ramanujan sum $S(0, m - n; t)$). Note that

$$\sum_{\substack{d(\bmod t) \\ (d,t)=1}} \bar{\mathcal{Q}}_{dht}(\alpha)\mathcal{Q}_{dht}(\beta) = \frac{1}{t} \sum_{a(\bmod t)} \bar{\mathcal{P}}_{aht}(\alpha)\mathcal{P}_{aht}(\beta). \tag{4.15}$$

5. Bilinear forms $\mathcal{C}(\alpha, \beta; t)$

According to (3.1), the bilinear form (4.6) splits into

$$\mathcal{C}(\alpha, \beta) = \sum_t \frac{2\pi}{qt}\mathcal{C}(\alpha, \beta; t), \tag{5.1}$$

say, where

$$\mathcal{C}(\alpha, \beta; t) = \sum_m \sum_n \bar{a}_m b_n \sigma(m, n; t). \tag{5.2}$$

Furthermore, according to (3.24) this splits into

$$\mathcal{C}(\alpha, \beta; t) = \mathcal{C}_H^+(\alpha, \beta; t) - \mathcal{C}_H^-(\alpha, \beta; t) + \mathcal{C}_H^\infty(\alpha, \beta; t) \tag{5.3}$$

where by (3.16) and (3.21) we have

$$\mathcal{C}_H^+(\alpha, \beta; t) = \frac{2\pi}{t} \sum_1^\infty \omega\left(\frac{h}{H}\right) \sum_{\substack{a(\bmod t) \\ aq\equiv h}} \bar{\mathcal{P}}_{aht}(\alpha)\mathcal{P}_{aht}(\beta) \tag{5.4}$$

and by (3.17), after opening the Ramanujan sum, we have

$$\mathcal{C}_H^-(\alpha, \beta; t) = \frac{2\pi}{t} \int_0^\infty \omega\left(\frac{h}{H}\right) \sum_{\substack{d(\bmod t) \\ (d,t)=1}} \bar{\mathcal{Q}}_{dht}(\alpha)\mathcal{Q}_{dht}(\beta) \, dh. \tag{5.5}$$

Here $H = H(t)$ is a cut-off parameter subject to $H \geqslant t + 128Nq^{-1}$. The remainder term satisfies

$$\mathcal{C}_H^\infty(\alpha, \beta; t) \ll \|\alpha\|\|\beta\|NHt\left(\frac{t}{H}\right)^A \tag{5.6}$$

with any exponent $A > 0$, where the implied constant depends on k, A and the cut-off function ω. This follows from (3.18) and the Cauchy inequality

$$\sum_{m\leqslant N} \sum_{n\leqslant N} |a_m b_n| \leqslant \|\alpha\|\|\beta\|N. \tag{5.7}$$

If t is large the original expression (3.2) for $\sigma(m, n; t)$ is quite convenient to produce quickly good bounds for the bilinear forms $\mathcal{C}(\alpha, \beta; t)$. On the other hand, the dual forms (5.4), (5.5) are more effective for estimations in smaller t, because due to their structure one can apply the duality

principle (as in the large sieve theory). Furthermore, the dual forms of small moduli represent the true main terms in asymptotic evaluation of $\mathcal{C}(\alpha, \beta)$. Having recognized these features we make a further splitting

$$\mathcal{C}(\alpha, \beta) = \mathcal{C}_T(\alpha, \beta) + \mathcal{D}_T(\alpha, \beta) \tag{5.8}$$

$$\mathcal{C}_T(\alpha, \beta) = \sum_{t \leqslant T} \frac{2\pi}{qt} \mathcal{C}(\alpha, \beta; t) \tag{5.9}$$

and $\mathcal{D}_T(\alpha, \beta)$ is the remainder sum over $t > T$.

6. Estimation of $\mathcal{D}_T(\alpha, \beta)$

Recall that

$$\mathcal{D}_T(\alpha, \beta) = \sum_{t > T} \frac{2\pi}{qt} \mathcal{C}(\alpha, \beta; t). \tag{6.1}$$

We choose

$$T \geqslant N q^{-1}. \tag{6.2}$$

Inserting (2.17) into (3.2) we derive

$$|\mathcal{D}_T(\alpha, \beta)| \leqslant \sum_{s > 0} \sum_{t > T} \frac{4\pi}{qst} \sum_{\substack{x (\mathrm{mod}\, t) \\ (x(x+qs),t)=1}} \left| \sum_m \sum_n \bar{a}_m b_n e\left(\frac{m\bar{x} - n\overline{(x+qs)}}{t} + \frac{m+n}{qst} \right) J\left(\frac{4\pi\sqrt{mn}}{qst} \right) \right|. \tag{6.3}$$

Because the argument of the Bessel function $J(4\pi\sqrt{mn}/qst)$ and the arguments of the exponentials $e(m/qst), e(n/qst)$ are small, we can separate the variables m, n, t (at no cost) by applying the power series expansions of the exponential function and (A.2). Having done this, we can apply the large sieve inequality (1.5) (use Cauchy's inequality and that $k \geqslant 3$, so $J_{k-1}(y) \ll y^2$) obtaining

$$\sum_{t > T} t^{-1} \sum_{x(t)} \left| \sum_m \sum_n \right| \ll \frac{T^2 + N}{T} \left(\frac{N}{qst} \right)^2 \|\alpha\| \|\beta\|.$$

Summing over s and simplifying the results by (6.3), we conclude as follows.

LEMMA 6.1. *Let $\alpha = (a_m), \beta = (b_n)$ be sequences of complex numbers for $1 \leqslant m, n \leqslant N$. We have*

$$\mathcal{D}_T(\alpha, \beta) \ll \left(1 + \frac{N}{q^2} \right) \frac{N}{qT} \|\alpha\| \|\beta\| \tag{6.4}$$

where the implied constant depends only on k.

Remark. The results also hold for $k = 2$, but we assumed $k \geqslant 3$ to simplify the arguments. However, the case $k = 1$ (cusp forms of weight one) is a completely different matter, see the introduction of [DFI02].

We are left with the bilinear form $\mathcal{C}_T(\alpha, \beta)$ which we are going to evaluate by means of the dual forms. Inserting (5.3) into (5.8) we write

$$\mathcal{C}(\alpha, \beta) = \mathcal{C}_{HT}^+(\alpha, \beta) - \mathcal{C}_{HT}^-(\alpha, \beta) + \mathcal{C}_{HT}^\infty(\alpha, \beta) \tag{6.5}$$

where

$$\mathcal{C}_{HT}^+(\alpha, \beta) = \sum_{t \leqslant T} \frac{2\pi}{qt} \mathcal{C}_H^+(\alpha, \beta; t) \tag{6.6}$$

553

and $\mathcal{C}_{HT}^-(\alpha, \beta), \mathcal{C}_{HT}^\infty(\alpha, \beta)$ defined accordingly (see the local forms $\mathcal{C}_H^+(\alpha, \beta; t), \mathcal{C}_H^-(\alpha, \beta; t), \mathcal{C}_H^\infty(\alpha, \beta; t)$ in (5.4), (5.5), (5.6)). The above decomposition holds for any cut-off parameters H, T satisfying

$$H \geqslant 129T, \quad T \geqslant Nq^{-1}. \tag{6.7}$$

7. Estimation of $\mathcal{C}_{HT}^\infty(\alpha, \beta)$

From (5.6) with H, T satisfying (6.7) we get

$$\mathcal{C}_{HT}^\infty(\alpha, \beta) \ll \|\alpha\| \|\beta\| \frac{NHT}{q} \left(\frac{T}{H} \right)^A \tag{7.1}$$

where A is any positive number, the implied constant depending only on k, A and the cut-off function ω. Choosing H a bit larger than T and A sufficiently large we can make the bound (7.1) very small. For example, if

$$H \geqslant T^{1+\varepsilon}, \quad T \geqslant Nq^{-1}, \tag{7.2}$$

then (7.1) yields

$$\mathcal{C}_{HT}^\infty(\alpha, \beta) \ll \|\alpha\| \|\beta\| \tag{7.3}$$

where the implied constant depends only on k, ε and the cut-off function ω.

Evaluation of the leading terms $\mathcal{C}_{HT}^+(\alpha, \beta), \mathcal{C}_{HT}^-(\alpha, \beta)$ requires more attention. This will be carried out after next section where we are going to establish estimates of the large sieve type (1.5) in which the characters are replaced by products of Kloosterman sums and Bessel functions (so-called absolute Kloosterman sums).

8. Asymptotic formula for bilinear forms

We can state now our second main result. Recall that $\mathcal{L}_f(\alpha), \mathcal{P}_{aht}(\alpha), \mathcal{Q}_{aht}(\alpha)$ are the linear forms in (4.2), (4.9), (4.14), respectively. Putting together (4.4), (5.8), (6.4), (7.3) we obtain the following.

THEOREM 8.1. *Let $\alpha = (a_m), \beta = (b_n)$ be sequences of complex numbers for $1 \leqslant m, n \leqslant N$. We have*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi (\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} \bar{\mathcal{L}}_f(\alpha) \mathcal{L}_f(\beta)$$

$$= \langle \alpha, \beta \rangle + \frac{1}{q} \sum_{t \leqslant T} \left(\frac{2\pi}{t} \right)^2 \sum_h \omega \left(\frac{h}{H} \right) \sum_{aq \equiv h(t)} \bar{\mathcal{P}}_{aht}(\alpha) \mathcal{P}_{aht}(\beta)$$

$$- \frac{1}{q} \sum_{t \leqslant T} \left(\frac{2\pi}{t} \right)^2 \int_{-\infty}^\infty \omega \left(\frac{h}{H} \right) \sum_{a(t)}^* \bar{\mathcal{Q}}_{aht}(\alpha) \mathcal{Q}_{aht}(\beta)\, dh + O\left(\left(1 + \frac{N}{q^2} \right) \frac{N}{qt} \|\alpha\| \|\beta\| \right) \tag{8.1}$$

where $T \geqslant Nq^{-1}, H \geqslant T^{1+\varepsilon}$ and $\omega(x)$ is a smooth function supported on $[0, 2]$ with $\omega(x) = 1$ if $0 \leqslant x \leqslant 1$. In the error term the implied constant depends only on ε, k and the cut-off function ω (ε is any positive number which is arbitrarily small).

We are particularly interested in the formula (8.1) for equal vectors, in which case it becomes the following.

COROLLARY 8.1. *Let $\alpha = (a_n)$ be a sequence of complex numbers for $1 \leqslant n \leqslant N$, we have*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2 = \|\alpha\|^2 + \frac{1}{q} \sum_{t \leqslant T} \left(\frac{2\pi}{t}\right)^2 \sum_h \omega\left(\frac{h}{H}\right) \sum_{aq \equiv h(t)} |\mathcal{P}_{aht}(\alpha)|^2$$

$$- \frac{1}{q} \sum_{t \leqslant T} \left(\frac{2\pi}{t}\right)^2 \int_{-\infty}^{\infty} \omega\left(\frac{h}{H}\right) \sideset{}{^*}\sum_{a(t)} |\mathcal{Q}_{aht}(\alpha)|^2 \, dh$$

$$+ O\left(\left(1 + \frac{N}{q^2}\right) \frac{N}{qT} \|\alpha\|^2\right) \tag{8.2}$$

subject to the condition of Theorem 8.1.

Remarks. The cut-off factors $\omega(h/H)$ in (8.1) and (8.2) can be removed and replaced by the condition $0 < h \leqslant H$, up to an extra error term which is relatively small. We have not done this at this point in order to save a special feature of the present error term, namely that it vanishes as T goes to infinity. Therefore, (8.1), (8.2) are truly asymptotic formulas.

9. Large sieves for the linear forms $\mathcal{Q}_{aht(\alpha)}$

If we were only interested in an upper bound of the large sieve type for $\mathcal{L}_f(\alpha)$, then the sum of $|\mathcal{Q}_{aht}(\alpha)|^2$ on the right-hand side of (8.2) could be ignored, because its contribution is negative. Nevertheless, we are going to estimate this sum to show that it is somewhat smaller than what comes out of the linear forms $\mathcal{P}_{aht}(\alpha)$. To simplify the arguments, from now on we assume that the vector α is supported in a dyadic interval

$$\alpha = (a_n), \quad N < n \leqslant 2N. \tag{9.1}$$

Let $H, N, T \geqslant 1$. First we estimate

$$\int_H^{2H} \sum_{T < t \leqslant 2T} \sideset{}{^*}\sum_{a(\bmod t)} |\mathcal{Q}_{aht}(\alpha)|^2 \, dh. \tag{9.2}$$

Recall that

$$\mathcal{Q}_{aht}(\alpha) = \sum_n a_n e\left(\frac{an}{t}\right) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right).$$

Changing the variable of integration $h = (xt)^2 q N^{-1}$ we find that (9.2) is bounded by

$$\frac{4H}{X} \int_X^{3X} \sum_{T < t \leqslant 2T} \sideset{}{^*}\sum_{a(\bmod t)} \left|\sum_n a_n e\left(\frac{an}{t}\right) J\left(4\pi x\sqrt{\frac{n}{N}}\right)\right|^2 dx \tag{9.3}$$

where

$$X = \frac{1}{2T}\sqrt{\frac{HN}{q}}. \tag{9.4}$$

By (A.5), this is less than

$$\frac{H}{X} \int_X^{3X} \sum_{T < t \leqslant 2T} \sideset{}{^*}\sum_{a(t)} \left|\sum_n a_n \left(\frac{N}{n}\right)^{\frac{1}{4}} W\left(2x\sqrt{\frac{n}{N}}\right) e\left(\frac{an}{t} + 2x\sqrt{\frac{n}{N}}\right)\right|^2 \frac{dx}{x} \tag{9.5}$$

555

where $W(y)$ is a nice smooth function whose derivatives satisfy $y^\nu W^{(\nu)}(y) \ll 1$. Using the Mellin transform technique to separate the variables in $W(2x\sqrt{n/N})$, we estimate the above by

$$\frac{H}{X^2} \int_X^{3X} \sum_{T < t \leqslant 2T} \sum_{a(t)}^* \left| \sum_n b_n e\left(\frac{an}{t} + 2x\sqrt{\frac{n}{N}} \right) \right|^2 dx \qquad (9.6)$$

where b_n are complex numbers with $b_n \ll |a_n|$, hence also $\|\beta\| \ll \|\alpha\|$, where the implied constant is absolute.

Next we split the sum over n into short segments of length $M = NX^{-1}$ and write

$$\sum_n b_n e\left(\frac{an}{t} + 2x\sqrt{\frac{n}{N}} \right) = \sum_L e\left(2x\sqrt{\frac{L}{N}} \right) \sum_{L < n \leqslant L+M} b_n e\left(\frac{an}{t} + 2x\frac{\sqrt{n} - \sqrt{L}}{\sqrt{N}} \right)$$

where $L = N + lM, 0 \leqslant l \leqslant X$. Note that $2x(\sqrt{n} - \sqrt{l})/\sqrt{N}$ is bounded by one, so we can separate the variables in $e(2x(\sqrt{n} - \sqrt{l})/\sqrt{N})$ at no cost by applying its power series expansion. Hence, we deduce that (9.6) is bounded by

$$\frac{H}{X^2} \int_X^{3X} \sum_{T < t \leqslant 2T} \sum_{a(t)}^* \left| \sum_L e\left(2x\sqrt{\frac{L}{N}} \right) \sum_{L < n \leqslant L+M} b_n(L) e\left(\frac{an}{t} \right) \right|^2 dx \qquad (9.7)$$

where $b_n(L) \ll |b_n|$. First smoothing the integration, then squaring out and integrating in x, we find that (9.7) is bounded by

$$\frac{H}{X} \sum_{L_1} \sum_{L_2} \left(1 + \left(\frac{L_1 - L_2}{M} \right)^2 \right)^{-1} \sum_{T < t \leqslant 2T} \sum_{a(t)}^* \left| \sum_{L_1 < n \leqslant L_1+M} \right| \left| \sum_{L_2 < n \leqslant L_2+M} \right|.$$

Now by the large sieve inequality (1.5), this is estimated by

$$HX^{-1}(T^2 + M) \sum_L \sum_{L < n \leqslant L+M} |b_n(L)|^2 \ll HX^{-1}(T^2 + M)\|\alpha\|^2.$$

Inserting (9.4) and $M = NX^{-1}$ we conclude as follows.

LEMMA 9.1. *For $H, N, T \geqslant 1$ we have*

$$\int_H^{2H} \sum_{T < t \leqslant 2T} \sum_{a(t)}^* \left| \sum_n a_n e\left(\frac{an}{t} \right) J\left(\frac{4\pi}{t} \sqrt{\frac{hn}{q}} \right) \right|^2 dh \ll qT^2 \left(1 + T\sqrt{\frac{H}{qN}} \right) \|\alpha\|^2 \qquad (9.8)$$

where the implied constant depends only on k.

COROLLARY 9.1. *Let $H, N, T \geqslant 1$ and $\alpha = (a_n)$ be any sequence of complex numbers for $N < n \leqslant 2N$. Then*

$$\frac{1}{q} \sum_{t \leqslant T} \left(\frac{2\pi}{t} \right)^2 \int_0^{2H} \sum_{a(t)}^* |\mathcal{Q}_{aht}(\alpha)|^2 \, dh \ll \left(\log 2HT + T\sqrt{\frac{H}{qN}} \right) \|\alpha\|^2 \qquad (9.9)$$

where the implied constant depends only on k.

Applying (9.9) to (8.2) we derive the following.

THEOREM 9.1. *Let $\alpha = (a_n)$ be a sequence of complex numbers for $N < n \leqslant 2N$. Then*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi(\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2 = \frac{1}{q} \sum_{t \leqslant T} \left(\frac{2\pi}{t} \right)^2 \sum_{h \leqslant H} \sum_{aq \equiv h(t)} |\mathcal{P}_{aht}(\alpha)|^2 + O\left(\left(1 + \frac{N}{q^2} \right) N^\varepsilon \|\alpha\|^2 \right) \qquad (9.10)$$

where $T = N/q, H = (N/q)^{1+\varepsilon}$ with any $\varepsilon > 0$, the implied constant depending only on k and ε.

556

Remark. Directly, one obtains (9.10) with the cut-off factors $\omega(h/H)$ rather than with the restriction $h \leqslant H$ which disagrees only in the segment $H \leqslant h \leqslant 2H$. However, the terms for h in the segment $H \leqslant h \leqslant 2H$ are absorbed by the existing error term (this can be seen by playing with two different functions ω), so we removed them for simplicity.

10. Eliminating the incidental terms

Next we are going to remove the terms in (9.10) with $q|t$. We shall show that these terms contribute less than the existing error term. To this end, it is easier to estimate the corresponding original terms in (2.20)

$$\sigma_q(m,n) = \mathcal{K} \sum_s \sum_t \frac{2\pi}{q^2 st} V_{qs}(m,n;qt) e\left(\frac{m+n}{q^2 st}\right) J_{k-1}\left(\frac{4\pi\sqrt{mn}}{q^2 st}\right). \tag{10.1}$$

Recall that $V_{qs}(m,n;qt)$ is given by (2.17). Our goal is to estimate the bilinear form

$$\mathcal{C}_q(\alpha,\beta) = \sum_{m \leqslant N} \sum_{n \leqslant N} \bar{a}_m b_n \sigma_q(m,n). \tag{10.2}$$

For separating the variables in $J_{k-1}(4\pi\sqrt{mn}/q^2 st)$ we use the Mellin integral representation (A.7). Then, by (1.3), we derive

$$\mathcal{C}_q(\alpha,\beta) \ll \sum_s \sum_t \frac{1}{q^2 st}\left(\frac{N}{q^2 st}\right)^\sigma (qt+N)\|\alpha\|\|\beta\|$$

where σ can be chosen with $\varepsilon \leqslant \sigma \leqslant k-1-\varepsilon$ at will. We take $\sigma = \varepsilon$ if $t \leqslant Nq^{-1}$ and $\sigma = 1+\varepsilon$ otherwise, and obtain the following.

LEMMA 10.1. *For any $\alpha = (a_m), \beta = (b_n)$ with $1 \leqslant m,n \leqslant N$ we have*

$$\mathcal{C}_q(\alpha,\beta) \ll q^{-2} N^{1+\varepsilon}\|\alpha\|\|\beta\| \tag{10.3}$$

where the implied constant depends only on k and ε.

Lemma 10.1 shows indeed that the terms in (9.10) with $t \equiv 0 (\mathrm{mod}\, q)$ can be ignored. Assuming that q is prime, we are left with the terms for $(t,q) = 1$. In this case $aq \equiv h (\mathrm{mod}\, t)$ has the unique solution $a \equiv h\bar{q} (\mathrm{mod}\, t)$ and we denote the corresponding linear form $\mathcal{P}_{aht}(\alpha)$ simply by $\mathcal{P}_{ht}(\alpha)$; recall

$$\mathcal{P}_{ht}(\alpha) = \sum_n a_n S(h\bar{q},n;t) J_{k-1}\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right). \tag{10.4}$$

Theorem 9.1 implies the following (our third main theorem).

THEOREM 10.1. *Let q be prime and $\alpha = (a_n)$ be any sequence of complex numbers for $N < n \leqslant 2N$. Then*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi(\mathrm{mod}\, q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2 = \frac{1}{q} \sum_{\substack{t \leqslant T \\ (t,q)=1}} \left(\frac{2\pi}{t}\right)^2 \sum_{h \leqslant H} |\mathcal{P}_{ht}(\alpha)|^2 + O\left(\left(1 + \frac{N}{q^2}\right) N^\varepsilon \|\alpha\|^2\right) \tag{10.5}$$

where $T = H = N/q$, the implied constant depending only on k and ε.

Remarks. Actually, we have established (10.5) only with $T = Nq^{-1}$ and $H = (Nq^{-1})^{1+\varepsilon}$. However, we shall see soon (Proposition 12.1) that the contribution of the terms t, h with $t \leqslant Nq^{-1} \leqslant h \leqslant (Nq^{-1})^{1+\varepsilon}$ is absorbed in the error term of (10.5), so we excluded this range.

557

The principal character $\chi_0(\bmod q)$ is present in (10.5) if k is even. In this case it contributes (see (1.9))

$$\frac{2}{\phi(q)} \sum_{f \in \mathcal{F}_{\chi_0}} |\mathcal{L}_f(\alpha)|^2 = \frac{2}{\phi(q)}\left(1 + O\left(\frac{N}{q}\right)\right)\|\alpha\|^2. \tag{10.6}$$

This contribution is absorbed by the error term in (10.5). Therefore, (10.5) holds without the principal character. Because q is prime, the remaining characters are primitive, and the basis $\mathcal{F}_\chi$ of $S_k(\Gamma_0(q), \chi)$ can be made of primitive forms.

11. Large sieve for absolute Kloosterman sums (direct approach)

We are interested in estimation of the following forms

$$\mathcal{A}(H, T, N) = \sum_{H \leqslant h < 2H} \sum_{\substack{T \leqslant t < 2T \\ (t,q)=1}} |\mathcal{P}_{ht}(\alpha)|^2 \tag{11.1}$$

for any $H, T, N \geqslant 1$, where $\mathcal{P}_{ht}(\alpha)$ is the linear form (10.4). These are building blocks of the right-hand side of (10.5). We shall see that $\mathcal{A}(H, T, N)$ contribute more than the error term in (10.5). Our goal is to take off a few of these terms somewhat compromising the error term.

We begin by estimating $\mathcal{A}(H, T, N)$ with a direct approach. First we separate the variables in the Bessel functions $J_{k-1}(4\pi\sqrt{hn/qt^2})$ by means of (A.7). We get

$$\mathcal{A}(H, T, N) \leqslant \int_{-\infty}^{\infty} \sum_h \sum_t \left|\sum_n a_n n^{ir} S(h\bar{q}, n; t)\right|^2 \frac{X}{X^2 + |r|^2}\, dr$$

where $X = T^{-1}(HN/q)^{1/2}$. Next we expand the summation over $H \leqslant h < 2H$ into complete segments of length t getting

$$\mathcal{A}(H, T, N) \leqslant \int_{-\infty}^{\infty} \sum_t (H + t) \sum_{d(\bmod t)}^{*} \left|\sum_n a_n n^{ir} e\left(\frac{dn}{t}\right)\right|^2 \frac{X}{X^2 + |r|^2}\, dr.$$

Then by the hybrid large sieve inequality (1.6) we derive

$$\mathcal{A}(H, T, N) \ll (H + T)(RT^2 + N)\frac{X}{X^2 + R^2}\|\alpha\|^2$$

for some $R > 0$, the worst bound being at $R \asymp X$. Hence, we conclude as follows.

LEMMA 11.1. *For any $H, T, N \geqslant 1$ we have*

$$\mathcal{A}(H, T, N) \ll (H + T)T\left(T + \sqrt{\frac{qN}{H}}\right)\|\alpha\|^2 \tag{11.2}$$

where the implied constant depends only on k.

12. Large sieve for absolute Kloosterman sums (dual approach)

We give another estimation for $\mathcal{A}(H, T, N)$ using the duality principle. Because a_n are arbitrary numbers, it suffices to estimate the dual form

$$\mathcal{A}^*(H, T, N) = \sum_n \left|\sum_h \sum_t \gamma(h, t) S(h\bar{q}, n; t) J_{k-1}\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right)\right|^2 \tag{12.1}$$

558

where $\gamma(h,t)$ are arbitrary complex numbers for $H \leqslant h < 2H$, $T \leqslant t < 2T$ with $(t,q) = 1$. Our estimate must depend on the $\gamma(h,t)$ only by way of being proportional to the l_2-norm

$$\|\gamma\|^2 = \sum_h \sum_t |\gamma(h,t)|^2, \tag{12.2}$$

otherwise the result would be useless for applications to the original form $\mathcal{A}(H,T,N)$.

Before opening the square we enlarge and smooth the outer summation

$$\mathcal{A}^*(H,T,N) \leqslant \sum_n g\left(\frac{n}{N}\right)\left|\sum_h \sum_t \gamma(h,t)S(h\bar{q},n;t)J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right)\right|^2.$$

Recall that $J(x)$ denotes $J_{k-1}(x)$. Here $g(x)$ is a smooth function supported on $[\frac{1}{2}, \frac{5}{2}]$. Now this equals

$$\sum_{h_1} \sum_{t_1} \sum_{h_2} \sum_{t_2} \gamma(h_1,t_1)\bar{\gamma}(h_2,t_2) \sideset{}{^*}\sum_{d_1(t_1)} \sideset{}{^*}\sum_{d_2(t_2)} e\left(\frac{h_1\overline{qd_1}}{t_1} - \frac{h_2\overline{qd_2}}{t_2}\right)$$

$$\times \sum_n g\left(\frac{n}{N}\right)e\left(n\left(\frac{d_1}{t_1} - \frac{d_2}{t_2}\right)\right)J\left(\frac{4\pi}{t_1}\sqrt{\frac{h_1 n}{q}}\right)J\left(\frac{4\pi}{t_2}\sqrt{\frac{h_2 n}{q}}\right).$$

The terms with $t_1 = t_2, d_1 = d_2$ contribute

$$\mathcal{A}_0^*(H,T,N) = \sum_n g\left(\frac{n}{N}\right) \sum_t \sideset{}{^*}\sum_{d(t)} \left|\sum_h \gamma(h,t)e\left(\frac{dn}{t}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right)\right|^2.$$

Before squaring out we expand the summation over $d(\mathrm{mod}\,t), (d,t) = 1$ to all residue classes modulo t getting

$$\mathcal{A}_0^*(H,T,N) \leqslant \sum_t t \sum_{h_1} \sum_{h_2 \equiv h_1(t)} |\gamma(h_1,t)\gamma(h_2,t)|\left|\sum_n g\left(\frac{n}{N}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{h_1 n}{q}}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{h_2 n}{q}}\right)\right|.$$

Assuming $N \geqslant 128HT^{-2}q^{-1}$ (this is very weak condition) we find that the sum over n is equal to the corresponding integral

$$\int_0^\infty g\left(\frac{x}{N}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{h_1 x}{q}}\right)J\left(\frac{4\pi}{t}\sqrt{\frac{h_2 x}{q}}\right)dx \tag{12.3}$$

up to an error term $O(N^{-A})$ which is negligible (see (A.5)). Moreover, the integral (12.3) has no stationary phase, so by (A.5) this is

$$\ll N\left(1 + \frac{|\sqrt{h_1} - \sqrt{h_2}|}{T}\sqrt{\frac{N}{q}}\right)^{-2}\min\left(1, T\sqrt{\frac{q}{HN}}\right)$$

$$\asymp T\sqrt{\frac{qN}{H}}\left(1 + \left(\frac{h_1 - h_2}{T}\right)^2\frac{N}{qH}\right)^{-1}\left(1 + \frac{qT^2}{TN}\right)^{-\frac{1}{2}}.$$

Then, given h_1 the resulting sum over $h_2 \equiv h_1(\mathrm{mod}\,t)$ yields

$$\sum_{h_2}\left(1 + \left(\frac{h_1 - h_2}{T}\right)^2\frac{N}{qH}\right)^{-1} \ll 1 + \sqrt{\frac{qH}{N}}.$$

Hence, we conclude that the terms $t_1 = t_2, d_1 = d_2$ contribute at most

$$\mathcal{A}_0^*(H,T,N) \ll qT^2\left(1 + \frac{N}{qH}\right)^{\frac{1}{2}}\left(1 + \frac{qT^2}{HN}\right)^{-\frac{1}{2}}\|\gamma\|^2. \tag{12.4}$$

559

We are left with the terms satisfying

$$\frac{d_1}{t_1} \neq \frac{d_2}{t_2} \pmod{1}. \tag{12.5}$$

As above the sum over n is equal to the corresponding integral

$$\int_0^\infty g\left(\frac{x}{N}\right) e\left(x \left\|\frac{d_1}{t_1} - \frac{d_2}{t_2}\right\|\right) J\left(\frac{4\pi}{t_1}\sqrt{\frac{h_1 x}{q}}\right) J\left(\frac{4\pi}{t_2}\sqrt{\frac{h_2 x}{q}}\right) dx \tag{12.6}$$

up to an error term $O(N^{-A})$ which is negligible. Now, however, one may encounter a stationary phase, unless we assume that the linear part in the exponential dominates over the square root parts in the exponentials coming from the Bessel functions (see (A.5)). Note that

$$\left\|\frac{d_1}{t_1} - \frac{d_2}{t_2}\right\| \geqslant \frac{1}{t_1 t_2} \geqslant \frac{1}{4T^2}.$$

Therefore, in order to miss the stationary point, we assume that N is sufficiently large, say

$$N \geqslant \left(R + 128\frac{H}{q}\right)T^2 \tag{12.7}$$

where $R \geqslant 1$ is at our disposal. Then the integral (12.6) is much less than NR^{-A} for any $A \geqslant 0$ by partial integration A times. Hence, the terms (12.5) contributes to $\mathcal{A}^*(H,T,N)$ at most

$$\sum_{h_1}\sum_{t_1}\sum_{h_2}\sum_{t_2} |\gamma(h_1,t_1)\gamma(h_2,t_2)| t_1 t_2 NR^{-A} \ll HT^3 NR^{-A}\|\gamma\|^2.$$

Adding this bound to (12.4) we obtain a bound for $\mathcal{A}_0^*(H,T,N)$. Finally applying the duality principle we deduce the following bound for the original sum $\mathcal{A}(H,T,N)$.

LEMMA 12.1. *Let $H,T,R \geqslant 1$ and N satisfies (12.7). Then*

$$\mathcal{A}(H,T,N) \ll \left\{ qT^2\left(1 + \frac{N}{qH}\right)^{\frac{1}{2}}\left(1 + \frac{qT^2}{HN}\right)^{-\frac{1}{2}} + HT^3 NR^{-A} \right\}\|\alpha\|^2 \tag{12.8}$$

where A is any positive number, and the implied constant depends only on k, A.

The condition (12.7) can be relaxed by a device which is reminiscent of that used in [FV73]. We begin with the following observation

$$\sum_{\substack{P < p \leqslant 2P \\ p \nmid t}} \log p \gg P \quad \text{if } P \gg \log t.$$

Hence,

$$\mathcal{A}(H,T,N) \ll \frac{1}{P} \sum_{P < p \leqslant 2P} (\log p) \sum_h \sum_{(t,p)=1} \left| \sum_n a_n S(h\bar{q},n;t) J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right) \right|^2.$$

Given p we can write $S(h\bar{q},n;t) = S(h\overline{qp},np;t)$ and

$$J\left(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\right) = J\left(\frac{4\pi}{t}\sqrt{\frac{hnp}{qp}}\right).$$

Hence, the resulting sum is of the same type as $\mathcal{A}(H,T,N)$, but with q replaced by qp, N replaced by Np and the sequence $\alpha = (a_n)$ replaced by the lacunary sequence (a_{nq}) of the same l_2-norm. Lemma 12.1 is applicable in the modified situation provided

$$NP \geqslant \left(R + 128\frac{H}{qP}\right)T^2 \tag{12.9}$$

getting

$$\mathcal{A}(H,T,N) \ll \left\{ qPT^2 \left(1 + \frac{N}{qH}\right)^{\frac{1}{2}} \left(1 + \frac{qT^2}{HN}\right)^{-\frac{1}{2}} + HT^3 NPR^{-A} \right\} \|\alpha\|^2. \tag{12.10}$$

We take

$$P = \frac{RT^2}{N} + 12T\sqrt{\frac{H}{qN}} + \log T \tag{12.11}$$

which satisfies (12.9). To simplify we assume that

$$1 \leqslant T \leqslant \frac{N}{q}, \quad 1 \leqslant H \leqslant \left(\frac{N}{q}\right)^{1+\varepsilon}. \tag{12.12}$$

Then we take $R = N^\varepsilon$, so

$$P \ll \left(1 + \frac{T}{q}\right) N^\varepsilon$$

$$\left(1 + \frac{T}{q}\right)\left(1 + \frac{qT^2}{HN}\right)^{-\frac{1}{2}} \ll \left(1 + \frac{HN}{q^3}\right)^{\frac{1}{2}}$$

$$\left(1 + \frac{N}{qH}\right)^{\frac{1}{2}} \left(1 + \frac{HN}{q^2}\right)^{\frac{1}{2}} \ll \left(1 + \frac{N}{q^2} + \sqrt{\frac{N}{qH}}\right) N^\varepsilon.$$

Hence, (12.10) implies the following.

PROPOSITION 12.1. *Let H, T, N satisfy (12.12) and $\alpha = (a_n)$ be a sequence of complex numbers for $N < n \leqslant 2N$. Then*

$$\mathcal{A}(H,T,N) \ll qT^2 \left(1 + \frac{N}{q^2} + \sqrt{\frac{N}{qH}}\right) N^\varepsilon \|\alpha\|^2 \tag{12.13}$$

where the implied constant depends only on k and ε.

Applying (12.13) to (10.5) we derive our Theorem 1.1.

COROLLARY 12.1. *Let q be prime and $N \geqslant q$. Then for any $\alpha = (a_n)$ with $N < n \leqslant 2N$ we have*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi (\mathrm{mod}\ q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2 \ll N^\varepsilon \left(\frac{N}{q^2} + \sqrt{\frac{N}{q}}\right) \|\alpha\|^2 \tag{12.14}$$

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

The component $\sqrt{N/qH}$ in the error term of (1.10) dominates if

$$HN \leqslant q^3. \tag{12.15}$$

In the next section we shall see that this component cannot be removed.

13. Quantitative study of the asymptotic large sieve

In this section, we will see that the upper bound in (1.10) is optimal and the component $\sqrt{N/qH}$ cannot be removed. A bunch of special vectors can be constructed for which the main term in (1.10) is exceptionally large.

Given $t_0 \geqslant 1$ and $h_0 \geqslant 1$ we consider the vector $\alpha_0 = (a_0(n))$ with

$$a_0(n) = S(h_0\bar{q}, n; t_0) J_{k-1}\left(\frac{4\pi}{t_0}\sqrt{\frac{h_0 n}{q}}\right) \tag{13.1}$$

561

for $N < n \leqslant 2N$. For this vector we show the following asymptotic formula for the dual liner form $\mathcal{P}_{h_0 t_0}(\alpha_0)$ (recall that $\mathcal{P}_{ht}(\alpha)$ is given by (10.4)).

Proposition 13.1. *Let $h_0 \leqslant Nq^{-1}$, then*

$$\mathcal{P}_{h_0 t_0}(\alpha_0) = \frac{\sqrt{2}-1}{2\pi^2}\phi(t_0)t_0\sqrt{\frac{qN}{h_0}} + O\left(\frac{qt_0{}^3\tau(t_0)}{h_0}\right) \tag{13.2}$$

where the implied constant is absolute.

Proof. We have

$$\mathcal{P}_{h_0 t_0}(\alpha_0) = \sum_{N<n\leqslant 2N} \alpha_0(n)^2$$

$$= \sum_{N<n\leqslant 2N} S^2(h_0\bar{q}, n; t_0)J^2\left(\frac{4\pi}{t_0}\sqrt{\frac{h_0 n}{q}}\right)$$

$$= \frac{t_0}{4\pi^2}\sqrt{\frac{q}{h_0}}\sum_{N<n\leqslant 2N} n^{-1/2}S^2(h_0\bar{q}, n; t_0)\left\{1 - \Im e\left(\frac{4}{t_0}\sqrt{\frac{h_0 n}{q}} - \frac{k}{2}\right)\right\} + O(qh_0^{-1}t_0^3\tau(t_0))$$

by (A.4), where the error term is obtained by Weil's bound for the Kloosterman sum $S(h_0\bar{q}, n; t_0) \ll (h_0, n, t_0)^{\frac{1}{2}}t_0^{\frac{1}{2}}\tau(t_0)$. Opening the Kloosterman sum, we get

$$\mathcal{P}_{h_0 t_0}(\alpha_0) = \frac{t_0}{4\pi^2}\sqrt{\frac{q}{h_0}}\sum_{d_1(t_0)}^* \sum_{d_2(t_0)}^* e\left(\frac{h_0\bar{q}(\bar{d}_1 - \bar{d}_2)}{t_0}\right)\sum_{N<n\leqslant 2N} n^{-1/2}(1 - \Im e(f(n))) + O(qh_0^{-1}t_0^3\tau(t_0))$$

where

$$f(x) = x\frac{d_1 - d_2}{t_0} + \frac{4}{t_0}\sqrt{\frac{h_0 x}{q}}.$$

We choose d_1, d_2 such that $|d_1 - d_2| < \frac{1}{2}t_0$. Then

$$|f'(x)| < t_0^{-1}\left(|d_1 - d_2| + 2\sqrt{\frac{h_0}{qN}}\right) \leqslant \frac{1}{2} + \frac{2}{qt_0} < \frac{3}{4}$$

$$f'(x) \gg t_0^{-1}\left(|d_1 - d_2| + \sqrt{\frac{h_0}{qN}}\right).$$

Therefore, the sum over n can be replaced by the corresponding integral with a small error term, precisely

$$\sum_{N<n\leqslant 2N} = \int_N^{2N} x^{-1/2}(1 - \Im e(f(x)))\,dx + O(N^{-1/2}).$$

If $d_1 = d_2$, we get

$$\int_N^{2N} x^{-1/2}\left(1 - \Im e\left(\frac{4}{t_0}\sqrt{\frac{h_0 x}{q}}\right)\right)dx = 2(\sqrt{2}-1)\sqrt{N} + O\left(t_0\sqrt{\frac{q}{h_0}}\right)$$

and if $d_1 \neq d_2$, we get

$$\int_N^{2N} \ll \frac{t_0}{|d_1 - d_2|\sqrt{N}}.$$

Hence, we derive (13.2) by trivial summation over d_1, d_2. This completes the proof of Proposition 13.1. $\qquad\square$

Using (13.2) and (1.10) for the vector (13.1) we derive by positivity that

$$\frac{2}{\phi(q)} \sum_{\substack{\chi(\mathrm{mod}\,q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha_0)|^2 \geqslant \left[\frac{\phi(t_0)}{2t_0}\sqrt{\frac{N}{qh_0}} + O\left(\frac{t_0\tau(t_0)}{h_0}\right)\right]\|\alpha_0\|^2$$

$$+ O\left(N^\varepsilon\left(\frac{N}{q^2} + \sqrt{\frac{N}{qH}}\right)\right)\|\alpha_0\|^2 \tag{13.3}$$

provided $t_0 \leqslant Nq^{-1}$ and $h_0 < H \leqslant Nq^{-1}$. In order to make the error terms negligible we take $H = Nq^{-1}$ and assume the following conditions

$$qt_0^2N^{-1+\varepsilon} \leqslant h_0 \leqslant \min\left(\frac{N}{q}, \frac{q^3}{N}\right)N^{-\varepsilon}. \tag{13.4}$$

COROLLARY 13.1. *Let $1 \leqslant q \leqslant N^{1-\varepsilon}$ and let h_0 satisfy (13.4). Then for the vector $\alpha_0 = (a_0(n))$ given by (13.1) we have*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi(\mathrm{mod}\,q) \\ \chi(-1)=(-1)^k}} \sum_{f \in \mathcal{F}_\chi} |\mathcal{L}_f(\alpha_0)|^2 \geqslant \frac{\phi(t_0)}{2t_0}\sqrt{\frac{N}{qh_0}}(1 + O(N^{-\varepsilon}))\|\alpha_0\|^2$$

where the implied constant depends only on ε.

The last result shows that for many characters, we cannot get square root cancellation in the linear form $\mathcal{L}_f(\alpha_0)$ which would be consistent with the Riemann hypothesis. This is quite surprising phenomenon because it shows that the Fourier coefficients of f point towards the direction of families of vectors formed by products of Kloosterman sums, and Bessel functions. A similar phenomenon appears in [ILS00, Appendix C].

Appendix A. Properties of Bessel functions

In this paper we referred to several properties of the Bessel function $J_{k-1}(x)$ with $k \geqslant 3$. First we have the following bound

$$J_{k-1}(x) \ll \min(x^{-\frac{1}{2}}, x^2), \tag{A.1}$$

$$J_{k-1}(2x) = \sum_{l=0}^{\infty}(-1)^l\frac{x^{2l+k-1}}{l!(l+k-1)!} \tag{A.2}$$

and from the asymptotic formula

$$J_{k-1}(x) = \left(\frac{2}{\pi x}\right)^{\frac{1}{2}}\cos\left(x + \frac{\pi}{4} - \frac{\pi k}{2}\right) + O(x^{-\frac{3}{2}}). \tag{A.3}$$

This gives

$$J_{k-1}^2(\pi x) = \frac{1}{\pi^2 x}\left(1 - \Im e\left(x - \frac{k}{2}\right) + O\left(\frac{1}{x}\right)\right). \tag{A.4}$$

563

Moreover, we have the following expression

$$J_{k-1}(2\pi x) = \pi^{-1}x^{-\frac{1}{2}}\Re W(2\pi x)e\left(x - \frac{k}{4} + \frac{1}{8}\right) \tag{A.5}$$

with

$$W(2\pi x) = \frac{1}{\Gamma(k-\frac{1}{2})}\int_0^\infty e^{-u}u^{k-\frac{3}{2}}\left(1 + \frac{iu}{4\pi x}\right)^{k-\frac{3}{2}}du$$

(see [Wat44, p. 206]). We only need estimates for $W(x)$ and its derivatives. It is easy to check that $x^a W^{(a)}(x) \ll 1$ for any $a \geqslant 0$, and the implied constant depends on k and a.

Next we use the Mellin integral representation

$$J_{k-1}(x) = \frac{1}{2\pi i}\int_{(\sigma)}\gamma(r)x^r\,dr, \quad 0 < \sigma < k-1 \tag{A.6}$$

with

$$\gamma(r) = 2^{r-1}\Gamma\left(\frac{k-r-1}{2}\right)\Gamma^{-1}\left(\frac{k+r+1}{2}\right).$$

This represents the Bessel function $J_{k-1}(x)$ by the power function x^r which is useful for separating variables, particularly when x is a product of many independent parameters. However as stated above we lose control of the size of the function. We recover this control by moving the line of integration $\Re r = \sigma$ to the contour $\mathcal{L}_X$ with $\Re r = \sigma = -1$ if $|\Im r| \leqslant X$ and $\Re r = \sigma = 1$ if $|\Im r| > X$. Then write

$$J_{k-1}(x) = \frac{1}{2\pi i}\int_{\mathcal{L}_X}\gamma_X(r)\left(\frac{x}{X}\right)^r\,dr \tag{A.7}$$

with $\gamma_X(r) = \gamma(r)X^r$. For $r \in \mathcal{L}_X$ we have $\gamma_X(r) \ll X(X+|r|)^{-2}$. Hence,

$$\int_{\mathcal{L}_X}|\gamma_X(r)||dr| \ll 1. \tag{A.8}$$

The fundamental role in this paper is played by the following product formula (see [Obe72])

$$\mathcal{K}\int_0^\infty e((\alpha+\beta)x + \gamma x^{-1})J_{k-1}(4\pi\sqrt{\alpha\beta}x)\frac{dx}{x} = 2\pi J_{k-1}(4\pi\sqrt{\alpha\gamma})J_{k-1}(4\pi\sqrt{\beta\gamma}) \tag{A.9}$$

for any $\alpha, \beta, \gamma > 0$. If $\alpha, \beta > 0$ and $\gamma \leqslant 0$, then the left-hand side vanishes, the property of which is also quite important.

References

Bom74   E. Bombieri, *Le Grand Crible dans la Théorie Analytique des Nombres*, Astérisque, vol. 18 (Société mathématique de France, Paris, 1974).

DFI02   W. Duke, J. B. Friedlander and H. Iwaniec, *The subconvexity problem for Artin L functions*. Invent. Math. **149** (2002), 489–577.

FV73   M. Forti and C. Viola, *On large sieve type estimates for Dirichlet series operator*, Proc. Symp. Pure Math. **24** (1973), 31–49.

Iwa97   H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17 (American Mathematical Society, Providence, RI, 1997).

ILS00   H. Iwaniec, W. Luo and P. Sarnak, *Low lying zeros of families of L-functions*, Publ. Math. Inst. Hautes Études Sci. **91** (2000), 55–131.

Lin41   Yu. V. Linnik, *The large sieve*, Dokl. Akad. Nauk. SSSR **30** (1941), 292–294 (Russian).

Mon71   H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Mathematics, vol. 227 (Springer, Berlin, 1971).

Obe72   F. Oberhettinger, *Tables of Bessel transforms* (Springer, Berlin, 1972).

Wat44   G. N. Watson, *A treatise on the theory of Bessel functions* (Cambridge University Press, Cambridge 1944).

Henryk Iwaniec   iwaniec@math.rutgers.edu
Department of Mathematics, Rutgers University, New Brunswick, NJ 08854, USA

Xiaoqing Li   xl29@math.buffalo.edu
Department of Mathematics, Columbia University, MC 4423, New York, NY 10027, USA
Current address: Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14260, USA