On Linnik and Selberg's Conjecture about Sums of Kloosterman Sums

Peter Sarnak[1,2] and Jacob Tsimerman[1]

[1] Department of Mathematics, Princeton University, Princeton, NJ
[2] Institute for Advanced Study, Princeton, NJ

Dedicated to Y. Manin on the Occasion of his 70th Birthday

1 Statements

In [Li] Linnik and later Selberg [Se] put forth far reaching Conjectures concerning cancellations due signs of Kloosterman Sums. Both point to the connection between this and the theory of modular forms. For example, the generalization of their conjectures to sums over arithmetic progressions imply the general Ramanujan Conjectures for $GL_2/\mathbb{Q}$. Selberg exploited this connection to give a nontrivial bound towards his well known "$\lambda_1 \geq \frac{1}{4}$" conjecture. Later Kuznetzov [Ku] used his summation formula, which gives an explicit relation between these sums and the spectrum of $GL_2/\mathbb{Q}$ modular forms, to prove a partial result towards Linnik's Conjecture. Given that we now have rather strong bounds towards the Ramanaujan Conjectures for $GL_2/\mathbb{Q}$ [L-R-S], [Ki-Sh], [Ki-Sa] it seems timely to revisit the Linnik and Selberg Conjectures.

The Kloosterman Sum $S(m,n;c)$, for $c \geq 1$, $m \geq 1$ and $n \neq 0$, is defined by

$$S(m,n;c) = \sum_{\substack{x \bmod c \\ x\bar{x} \equiv 1(c)}} e\left(\frac{mx + n\bar{x}}{c}\right) \tag{1}$$

(here $e(z) = e^{2\pi i z}$).

It follows from Weil's bound [Wei] for $S(m,n;p)$, p prime, and elementary considerations that

$$|S(m,n;c)| \le \tau(c)\,(m,n,c)^{1/2}\sqrt{c} \tag{2}$$

where $\tau(c)$ is the number of divisors of c. Linnik's Conjecture asserts that for $\epsilon > 0$ and $x \ge \sqrt{|n|}$

$$\sum_{c \le x} \frac{S(1,n;c)}{c} \ll_{\epsilon} x^{\epsilon}\,. \tag{3}$$

Note that from (2) we have that

$$\sum_{c \le x} \frac{|S(m,n;c)|}{c} \ll_{\epsilon} \sqrt{x}\,(x(m,n))^{\epsilon}\,. \tag{4}$$

On the other hand for m,n fixed, Michel [Mi] shows that

$$\sum_{c \le x} \frac{|S(m,n;c)|}{c} \gg_{\epsilon} x^{\frac{1}{2}-\epsilon}\,, \tag{5}$$

and hence if (3) is true it represents full "square-root" cancellation due to the signs of the Kloosterman Sums.

Selberg puts forth the much stronger conjecture, which has been replicated in many places and which reads;

For $x \ge (m,n)^{1/2}$,

$$\sum_{c \le x} \frac{S(m,n;c)}{c} \ll_{\epsilon} x^{\epsilon}\,. \tag{6}$$

As stated, this is false (see Section 2). It needs to be modified to incorporate an "ϵ-safety valve" in all parameters and not only in x and (m,n). For example, the following modification seems okay

$$\sum_{c\,\leq\,x} \frac{S(m,n;c)}{c} \ll_{\epsilon} (|mn|x)^{\epsilon}\,. \tag{7}$$

We call the range $x \geq \sqrt{|mn|}$ the Linnik range and $x \leq \sqrt{|mn|}$ the Selberg range. Obtaining nontrivial bounds (i.e., a power saving beyond (4)) for the sums (3) and (7) in the Selberg range is quite a bit more difficult.

By a smooth dyadic sum of the type (7) we mean

$$\sum_{c} \frac{S(m,n;c)}{c}\, F\left(\frac{c}{x}\right) \tag{8}$$

where $F \in C_0^\infty(\mathbb{R}^+)$ is of compact support and where the estimates for (8) as x, m, n vary, are allowed to depend on F. Summation by parts shows that an estimate for the left hand side of (7) will give a similar one for (8), but not conversely. In particular, estimates for (7) imply bounds for non-dyadic sums $\sum_{x\,\leq\,c\,\leq\,x+h} \frac{S(m,n;c)}{c}$ with $x^{\beta} \leq h \leq x$ and $\beta < 1$. We note that for many applications understanding smooth dyadic sums is good enough and in the Linnik range these smooth dyadic sums are directly connected to the Ramanujan Conjectures, see Deshouilliers and Iwaniec [D-I] and Iwaniec and Kowalski [I-K pp. 415-418].

Let $\pi \cong \bigotimes_v \pi_v$ be an automorphic cuspidal representation of $GL_2(\mathbb{Q})\backslash GL_2(\mathbb{A})$ with a unitary central character. Here v runs over all primes p and ∞. Let $\mu_1(\pi_v)$, $\mu_2(\pi_v)$ denote the Satake parameters of π_v at an unramified place v of π. We normalize as in [Sa] and use the results described there. For $0 \leq \theta < \frac{1}{2}$ we denote by H_θ the hypothesis; For any π as above

$$|\Re\,(\mu_j(\pi_v))\,| \,\leq\, \theta\,,\; j \,=\, 1, 2\,. \tag{9}$$

Thus H_0 is the Ramanujan-Selberg Conjecture and H_θ is known for $\theta = \frac{7}{64}$ ([Ki-Sa]).

Concerning (7) the main result is still that of Kuznetzov [Ku] who using his formula together with the elementary fact that $\lambda_1(SL_2(\mathbb{Z})\backslash\mathbb{H}) \geq \frac{1}{4}$ showed (the barrier $\frac{1}{6}$ is discussed at the end of this paper):

Theorem 1 (Kuznetzov) *Fix m and n then*

$$\sum_{c \,\leq\, x} \frac{S(m,n;c)}{c} \ll_{m,n} x^{1/6}(\log x)^{1/3}\,.$$

One can ask for similar bounds when c varies over an arithmetic progression. For this one applies the Kuznetzov formula for $\Gamma = \Gamma_0(q)$ in place of $SL_2(\mathbb{Z})$ (see [D-I]) or one can use the softer method in [G-S]. What is needed is the $v = \infty$ version of H_θ with $\theta \leq \frac{1}{6}$. This was first established in [Ki-Sh] and leads to

Theorem 1′. *Fix m, n, q, then*

$$\sum_{\substack{c \,\equiv\, 0(q) \\ c \,\leq\, x}} \frac{S(m,n;c)}{c} \ll_{m,n,q} x^{\frac{1}{6}}\,(\log x)^{1/3}\,.$$

Consider now what happens if m and n are allowed to be large as asked by Linnik and Selberg. We examine the case $q = 1$, the general case with q also varying deserves a similar investigation. We also assume that $mn > 0$, the $mn < 0$ case is similar except that some results such as Theorem 4 are slightly weaker since when $mn > 0$ we can exploit H_0 which is known for the Fourier coefficients of holomorphic-cusp forms [De]. The analogue of Theorem 1 that we seek is

$$\sum_{c \leq x} \frac{S(m,n;c)}{c} \ll_{\epsilon} \left(x^{\frac{1}{6}} + (mn)^{\frac{1}{6}}\right)(mnx)^{\epsilon}\,. \tag{10}$$

As with Theorem 1 we show below that the exponent $\frac{1}{6}$ in the mn aspect constitutes a natural barrier. We will establish (10) assuming the general Ramanujan Conjectures for $GL_2/\mathbb{Q}$ (i.e assuming H_0) and unconditionally we come close to proving it.

Theorem 2 *Assuming H_θ we have*

$$\sum_{c \le x} \frac{S(m,n;c)}{c} \ll_{\epsilon} \left(x^{\frac{1}{6}} + (mn)^{\frac{1}{6}} + (m+n)^{1/8}(mn)^{\theta/2} \right) (xmn)^{\epsilon} .$$

Corollary 3 *Assuming $H_{\frac{1}{12}}$ (and afortiori H_0), (10) is true. Unconditionally Theorem 2 is true with $\theta = 7/64$ and in particular (10) is true if m and n are close in the sense that $n^{5/43} \le m \le n$ or $m^{5/43} \le n \le m$.*

The proof of Theorem 2 uses Kuznetzov's formula to study the dyadic sums

$$\sum_{x \le c \le 2x} \frac{S(m,n;c)}{c} ,$$

see Section 2. The dyadic pieces with $x = (mn)^{1/3}$ or smaller are estimated trivially using (2) and give the $(mn)^{1/6}$ term in (10). For $x \le (mn)^{1/3}$ we don't have any nontrivial bound for the sum even in smooth dyadic form. Indeed applying the Kuznetzov formula to such smooth sums in range $x \le (mn)^{1/2-\delta}$ for some $\delta > 0$, one finds that the main terms on the spectral side localize in the transitional range for the Bessel functions of large order and argument. The analysis becomes a delicate one with the Airy function. In this $mn > 0$ case the main terms only involve Fourier coefficients of holomorphic modular forms. This indicates that one should be able to obtain the result backwards using the Petersson formula [I-K] together with the smooth k-averaging technique [I-L-S pp. 102]. Indeed this is so and we give the details of both methods in Section 2. Another bonus that comes from this holomorphic localization is that we can appeal to Deligne's Theorem, that is H_0, for these forms.

Theorem 4 *Fix $F \in C_0^{\infty}(0,\infty)$ and $\delta > 0$. For $mn > 0$ and $x \le (mn)^{\frac{1}{2}-\delta}$*

$$\sum_{c} \frac{S(m,n;c)}{c} F\left(\frac{c}{x}\right) \ll_{F,\delta} \frac{\sqrt{mn}}{x} .$$

The bound in Theorem 4 is nontrivial only in the range $x \ge (mn)^{1/3}$. To extend this to smaller x requires establishing cancellations in short spectral sums for Fourier coefficients of modular forms (see equation (34)) which is an interesting challange. In any case this analysis of the smooth dyadic sums explains the $(mn)^{1/6}$ barrier in (10).

2 Proofs

As we remarked at the beginning, the "randomness" in the signs of the Kloosterman sums in the form (7) implies the general Ramanujan Conjectures for $GL_2/\mathbb{Q}$. Since there seems to be no complete proof of this in the literature we give one here. For the archimedian q-aspect, that is the $\lambda_1 \geq 1/4$ Conjecture, this follows from Selberg [Se] since his 'zeta function' $Z(m,n,s)$ has poles with nonzero residues in $\Re(s) > \frac{1}{2}$ if there are exceptional eigenvalues (i.e. $\lambda < 1/4$). For the case of Fourier coefficients of holomorphic modular forms the implication is derived in [Mu pp. 240-242]. So we focus on the Fourier coefficients of Maass forms and restrict to level $q = 1$ (the case of general q is similar). We apply Kuznetzov's formula with $m = n$ for $\Gamma = SL_2(\mathbb{Z})$, see [I-K pp. 409] for the notation;

$$\sum_{j=1}^{\infty} |\rho_j(n)|^2 \, \frac{h(t_j)}{\cosh(\pi t_j)} + \frac{1}{4\pi} \int_{-\infty}^{\infty} |\tau(n,t)|^2 \, \frac{h(t)}{\cosh(\pi t)} \, dt$$

$$= g_0 + \sum_{c=1}^{\infty} \frac{S(m,n;c)}{c} \, g\left(\frac{4\pi n}{c}\right) . \tag{11}$$

Here h is entire and decays faster than $(1+|t|)^{-2-\delta}$, with $\delta > 0$ as $t \to \pm\infty$, and

$$g_0 = \frac{1}{\pi^2} \int_{-\infty}^{\infty} r\, h(r) \, \tan h(\pi r) \, dr$$

and

$$g(x) = \frac{2i}{\pi} \int_{-\infty}^{\infty} J_{2ir}(x) \, \frac{r\, h(r)}{\cosh(\pi r)} \, dr \, .$$

We want to prove that for a fixed $j = j_0$, we have

$$\rho_{j_0}(n) = O_\epsilon\left(|n|^\epsilon\right) \text{ as } |n| \to \infty \, . \tag{12}$$

It is well known how to deduce the various forms of H_0 for ϕ_{j_0} from (12). We choose h with $h(t) \geq 0$ for $t \in \mathbb{R}$ and $h(t_{j_0}) \geq 1$ and so that $g \in C^\infty(0,\infty)$ is rapidly decreasing at ∞ and g vanishes at $x = 0$ to order 1. Now let $|n| \to \infty$ in (11). Since $|\tau(n,t)| = O_\epsilon(|n|^\epsilon)$ we have

$$|\rho_{j_0}(n)|^2 \ll |n|^\epsilon + \left| \sum_{c=1}^{\infty} \frac{S(n,n;c)}{c}\, g\left(\frac{4\pi n}{c}\right) \right| .$$

If we assume (7) then we can estimate the sum on c by summing by parts and we obtain the desired bound (12).

We turn next to (6) and show that as stated it is false. Let x be a large integer. For each prime p in $(x, 2x)$ there is an integer a_p such that

$$\frac{S(1,a_p;p)}{\sqrt{p}} \geq \frac{1}{10} \tag{13}$$

This follows from the easily verified identities

$$\left.\begin{aligned} \frac{1}{p} \sum_{a(p)} \frac{S(1,a;p)}{\sqrt{p}} &= 0\,, \\ \frac{1}{p} \sum_{a(p)} \left(\frac{S(1,a;p)}{\sqrt{p}}\right)^2 &= 1 - \frac{1}{p} \end{aligned}\right\} . \tag{14}$$

We note in passing that the asymptotics of any moment $\frac{1}{p} \sum_{a} \left(\frac{S(1,a;p)}{\sqrt{p}}\right)^m$, in the limit as $p \to \infty$, have been determined by Katz [Ka]. The m^{th} moment being that of the Sato-Tate measure. Now let n be an integer satisfying $n \equiv 0$ $(\mathrm{mod}\ (x!)^2)$ and $n \equiv a_p (\mathrm{mod}\ p)$ for $x < p < 2x$. Such an n can be chosen since $(x!)^2$ and the p's are all relatively prime to each other. For $0 < c \leq x$, $n \equiv 0$ $(\mathrm{mod}\ c)$ and hence $S(1,n;c) = \mu(c)$ the Mobius function. For $x < c < 2x$ and c not a prime, we have $S(1,n;c) = S(1,n;c_1c_2)$ with $1 < c_1, c_2 < x$. Hence $n \equiv 0 (\mathrm{mod}\ c)$ and $S(1,n;c) = \mu(c)$. It follows that

$$\sum_{c \le 2x} \frac{S(1,n;c)}{c} = \sum_{c \le x} \frac{\mu(c)}{c} + \sum_{\substack{x < c \le 2x \\ c \text{ not prime}}} \frac{\mu(c)}{c} + \sum_{x < p < 2x} \frac{S(1,a_p;c)}{c}$$

$$\ge \frac{1}{10} \sum_{x < p < 2x} \frac{1}{\sqrt{p}} + O(\log x)$$

$$\ge \frac{x^{1/2}}{\log x} + O(\log x)\,. \tag{15}$$

Hence (6) is false.

Of course the n constructed above is very large and the right hand side in (12) is certainly $O_\epsilon(n^\epsilon)$ for any $\epsilon > 0$. Thus with the n^ϵ in (7) this is no longer a counter example.

We turn next to Theorem 4, that is the analysis of the smooth dyadic sums in the Selberg range $x \le (mn)^{1/3}$. We need Kuznetzov's formula [I-K]. Let $f \in C_0^\infty(\mathbb{R}^+)$ and

$$M_f(t) = \frac{\pi i}{\sinh(2\pi t)} \int_0^\infty (J_{2it}(x) - J_{-2it}(x)\, f(x)\, \frac{dx}{x} \tag{16}$$

and

$$N_f(k) = \frac{4(k-1)|}{(4\pi i)^k} \int_0^\infty J_{k-1}(x)\, f(x)\, \frac{dx}{x}\,. \tag{17}$$

Let $f_{j,k}(z)$, $j = 1, \ldots, \dim S_k(\Gamma)$, be an orthonormal basis of Hecke eigenforms for the space of holomorphic cusp forms of weight k for $\Gamma = SL_2(\mathbb{Z})$. Let $\psi_{j,k}(n)$ denote the corresponding Fourier coefficients normalized by

$$f_{j,k}(z) = \sum_{k=1}^\infty \psi_{j,k}(m)\, m^{\frac{k-1}{2}}\, e(mz)\,. \tag{18}$$

Let $\tau(m,t)$ be the n^{th} Fourier coefficient of the unitary Eisenstein series $E(z, \frac{1}{2} + it)$. Then for $mn > 0$

$$\sum_c \frac{S(m,n;c)}{c} f\left(\frac{4\pi\sqrt{mn}}{c}\right) =$$

$$\sum_j M_f(t_j)\,\overline{\rho_j(m)}\,\rho_j(n) + \frac{1}{4\pi}\int_{-\infty}^{\infty} M_f(t)\,\bar{\tau}(m,t)\,\tau(n,t)\,dt$$

$$+ \sum_{k\equiv 0(2)} N_f(k) \sum_{1\le j\le \dim\, S_k(\Gamma)} \overline{\psi_{j,k}(m)}\,\psi_{j,k}(n)\,. \qquad (19)$$

For $(mn)^\delta \le Y \le (mn)^{1/2}$ we apply (19) with

$$f_Y(x) = f_0\left(\frac{x}{Y}\right) \qquad (20)$$

and $f_0 \in C_0^\infty(\mathbb{R} > 0)$ fixed. The left hand side of the formula is

$$\sum_c \frac{S(m,n;c)}{c} f_0\left(\frac{4\pi\sqrt{mn}}{cY}\right) = \sum_c \frac{S(m,n;c)}{c} F_0\left(\frac{c}{X}\right)$$

where $F_0(\xi) = f_0(1/\xi)$ and $X = \frac{4\pi\sqrt{mn}}{Y}$, which is what we are interested in when Y is large. One can analyze $N_{f_Y}(t)$ and $M_{f_Y}(t)$ as $Y \longrightarrow \infty$ using the known asymptotics of the Bessel functions $J_{it}(x)$ and $J_{k-1}(x)$ for x and t large. Using repeated integrations by parts one can show that the main term comes from the holomorphic forms only, with their contribution coming from the transition range;

$$J_k(Yx)\,, \text{ with } |Yx-k| \ll k^{1/3}\,. \qquad (21)$$

Using the approximations [Wa pp. 249] by the Airy function;

$$J_k(x) \sim \frac{1}{\pi}\left(\frac{2(k-x)}{3x}\right)^{1/2} K_{\frac{1}{3}}\left(\frac{2^{3/2}(k-x)^{3/2}}{3x^{1/2}}\right) \text{ for } k-k^{1/3} \ll x < k \quad (22)$$

and

$$J_k(x) \sim \frac{1}{3}\left(\frac{2(x-k)}{x}\right)^{1/2}\left(J_{1/3}+J_{-1/3}\right)\left(\frac{2^{3/2}(x-k)^{3/2}}{3x^{1/2}}\right)$$
$$\text{for } k < x \ll k + k^{1/3} \tag{23}$$

and keeping only the leading terms of all asymptotics one finds that

$$\sum_c \frac{S(m,n;c)}{c} f_0\left(\frac{4\pi\sqrt{mn}}{cY}\right)$$
$$\sim \frac{1}{\pi}\left(\int_{-\infty}^{\infty} Ai(\xi)\,d\xi\right)\sum_k \frac{1}{k} f_0\left(\frac{k}{Y}\right)\frac{4(k-1)!}{(4\pi i)^k}\sum_{1\leq j\leq \dim S_\Gamma(\Gamma)} \overline{\psi_{j,k}}\,(m)\,\psi_{j,k}(n)\,. \tag{24}$$

Here $Ai(\xi) = \int\limits_{-\infty}^{\infty} \cos(t^3 + \xi t)\,dt$, is the Airy function.

The above derivation is tedious and complicated but once we see the form of the answer and especially that it involves only holomorphic modular forms, another approach suggests itself and fortunately it is simpler. We start with the Peterson formula [I-K].

$$\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\sum_{j=i,\dots,\dim\ S_k(\Gamma)} \psi(m)\,\overline{\psi_j(n)}$$
$$= \delta(m,n) + 2\pi\, i^k \sum_{c=1}^{\infty}\frac{S(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

Setting

$$\rho_f(n) = \left(\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\right)^{1/2}\psi_f(n) \tag{25}$$

we have

$$\sum_{f\in H_k(\Gamma)} \overline{\rho_f(m)}\,\rho_f(n) = \delta(m,n) = 2\pi i^k \sum_{c=1}^{\infty}\frac{S(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right) \tag{26}$$

where $H_k(\Gamma)$ denotes any orthonormal basis of $S_k(\Gamma)$ and in particular the Hecke basis which we choose.

In this case

$$\rho_f(m) \;=\; \rho_f(1)\,\lambda_f(m) \tag{27}$$

where $|\lambda_f(m)| \;\leq\; \tau(m)$, by Deligne's proof of the holomorphic Ramanujan Conjectures [De]. With this normalization (30) reads

$$\sum_c \frac{S(m,n;c)}{c}\, f_0\left(\frac{4\pi\sqrt{mn}}{cY}\right) \sim \;(\mathsf{const})\sum_k f_0\left(\frac{k}{Y}\right) \sum_{f\in H_k(\Gamma)} \overline{\rho_f(m)}\,\rho_f(n)\,. \tag{28}$$

From (26) with $m = n = 1$ we have that for k large

$$\sum_{f\in H_k(\Gamma)} |\rho_f(1)|^2 \;=\; 1 \;+\; \text{small}\,. \tag{29}$$

We use the k-averaging formula in [I-L-S pp.102] which gives for $K \;\geq\; 1$, $x > 0$ and $h \in C_0^\infty(\mathbb{R} > 0)$ that

$$h_k(x) \;:=\; \sum_{k\equiv 0(2)} h\left(\frac{k-1}{K}\right)\, J_{k-1}(x) \;=\; -iK\int\limits_{-\infty}^{\infty} \hat{h}(tK)\,\sin\left(x\,\sin\,2\pi t\right) dt\,. \tag{30}$$

From this it follows easily that if $x \geq K^{1+\epsilon_0}$ then for any fixed $N \geq 1$,

$$\sum_{k\,\equiv\,0(2)} h\left(\frac{k-1}{K}\right)\, J_{k-1}\,(x) \underset{N}{\ll} K^{-N}\,. \tag{31}$$

While for $0 \;\leq\; x \;\leq\; K^{1+\epsilon_0}$ and for any fixed N

$$h_k(x) = h_0\left(\frac{x}{K}\right) + \frac{1}{K^2}\,h_1\left(\frac{x}{K}\right)\cdots + \frac{1}{K^{2N}}\,h_N\left(\frac{x}{K}\right) + O_N\left(K^{-2N-1}\right) \tag{32}$$

where $h_0(\xi) = h(\xi)$ and $h_1, \ldots, h_N$ involve derivatives of $h(\xi)$ and are also in $C_0^\infty(\mathbb{R} > 0)$. From (26) we have that

$$\sum_{k\equiv 0(2)} i^k\,h\left(\frac{k-1}{K}\right)\sum_{f\in H_k(\Gamma)}\overline{\rho_f(m)}\,\rho_f(n)$$

$$= \sum_{k\equiv 0(2)}(i)^k\,h\left(\frac{k-1}{K}\right)\delta(m,n) - 2\pi\sum_{c=1}^{\infty}\frac{S(m,n;c)}{c}\,h_K\left(\frac{4\pi\sqrt{mn}}{c}\right). \tag{33}$$

We assume that $(mn)^\delta \le K \le \sqrt{mn}$ with $\delta > 0$ and fixed. Then for N large enough (depending on δ) we have from (27) (28) and (29) that

$$\sum_{j=0}^{N} K^{-2j}\sum_{c}\frac{S(m,n;c)}{c}\,h_j\left(\frac{4\pi\sqrt{mn}}{cK}\right) = \sum_{k\equiv 0(2)} i^k\,h\left(\frac{k-1}{K}\right)\sum_{f\in H_k(\Gamma)}\overline{\rho_f(m)}\,\rho_f(n) + O(1)\,. \tag{34}$$

Thus the lead term $j = 0$ in (34) recovers the asymptotics (24) and in a much more precise form. We can estimate the right hand side of (34) as being at most

$$RHS \le \sum_{k}|h|\left(\frac{k-1}{K}\right)\sum_{f\in H_k(\Gamma)}|\lambda_f(m)\lambda_f(n)|\,|\rho_f(1)|^2 \ll K\,\tau(m)\,\tau(n)\,, \tag{35}$$

by (27) and (29). It follows that

$$\sum_{j=0}^{N} K^{-2j}\sum_{c}\frac{S(m,n;c)}{c}\,h_j\left(\frac{4\pi\sqrt{mn}}{cK}\right) \ll K\,\tau(m)\,\tau(n)\,. \tag{36}$$

From this it follows by estimating the $j \geq 1$ sums trivially that

$$\sum_c \frac{S(m,n;c)}{c} h\left(\frac{4\pi\sqrt{mn}}{cK}\right) \ll K\,\tau(m)\,\tau(n) + \frac{(mn)^{1/4}}{K^{5/2}} \tag{37}$$

Using this bound which is now valid for $h_1, \ldots, h_N$ we can feed it back into (36) to get

$$\sum_c \frac{S(m,n;c)}{c} h\left(\frac{4\pi\sqrt{mn}}{Kc}\right)$$
$$\ll K\,\tau(m)\,\tau(n) + K^{-2}\left(K\,\tau(m)\,\tau(n) + \frac{(mn)^{1/4}}{K^{5/2}}\right) \ll K\,\tau(m)\,\tau(n) + \frac{(mn)^{1/4}}{K^{5/2}}\,. \tag{38}$$

Replicating this iteration a finite number of times yields that for $(mn)^{\delta} \leq K \leq \sqrt{mn}$,

$$\sum_c \frac{S(m,n;c)}{c} h\left(\frac{4\pi\sqrt{mn}}{Kc}\right) \ll K\,\tau(m)\,\tau(n) \tag{39}$$

or what is the same thing

$$\sum_c \frac{S(m,n;c)}{c} F\left(\frac{c}{x}\right) \ll_{\epsilon} \frac{(mn)^{1/2}\,\tau(m)\,\tau(n)}{x}\,, \quad \text{for} \quad x \leq (mn)^{1/2-\delta} \tag{40}$$

This completes the proof of Theorem 4. □

Finally we turn to the proof of Theorem 2. The dependence on mn in Kuznetzov's argument was examined briefly in Huxley [Hu]. In order for us to bring in H_θ effectively we first examine the dyadic sums $\sum_{x \leq c \leq 2x} \frac{S(m,n;c)}{c}$ for x in various ranges.

Proposition 5 *Assume H_θ, then for $\epsilon > 0$*

$$\sum_{x \leq c \leq 2x} \frac{S(m,n;c)}{c} \ll_{\epsilon} (mnx)^{\epsilon}\left(x^{1/6} + \frac{\sqrt{mn}}{x} + (m+n)^{1/8}(mn)^{\theta/2}\right).$$

Theorem 2 follows from this Proposition by breaking the sum $1 \le c \le x$ into at most $\log x$ dyadic pieces $y \le c \le 2y$ with $(mn)^{1/3} \le y \le x$ and estimating the initial segment $1 \le c \le \min((mn)^{1/3}, x)$ using (4). We conclude by proving Proposition 5. In the formula (19) we choose the test function f to be $\phi(t)$ depending on mn, x and a parameter $x^{1/3} \le T \le x^{2/3}$ to be chosen later. ϕ is smooth on $(0, \infty)$ taking values in $[0, 1]$ and satisfying

(i) $\phi(t) = 1$ for $\frac{a}{2x} \le t \le \frac{a}{x}$ where $a = 4\pi\sqrt{mn}$.

(ii) $\phi(t) = 0$ for $t \le \frac{a}{2x+2T}$ and $t \ge \frac{a}{x-T}$.

(iii) $\phi'(t) \ll \left(\frac{a}{x-T} - \frac{a}{x}\right)^{-1}$

(iv) ϕ and ϕ' are piecewise monotone on a fixed number of intervals.

For ϕ chosen this way we have that

$$\left|\sum_{c=1}^{\infty} \frac{S(m,n;c)}{c} \phi\left(\frac{4\pi\sqrt{mn}}{c}\right) - \sum_{x \le c \le 2x} \frac{S(m,n;,c)}{c}\right| \le \sum_{\substack{x-T \le c \le x \\ 2x \le c \le 2x+T}} \left|\frac{S(m,n;c)}{c}\right|$$

$$\ll_{\epsilon} \frac{(mn)^{\epsilon}}{\sqrt{x}} \sum_{\substack{x-T \le c \le x \\ 2x \le c \le 2x+T}} \tau(c) \ll \frac{(mn)^{\epsilon}\, T \log x}{\sqrt{x}} . \tag{41}$$

where we have used (2) and the mean value bound for the divisor function. We estimate $\sum\limits_{c=1}^{\infty} \frac{S(m,n;c)}{c} \phi\left(\frac{4\pi\sqrt{mn}}{c}\right)$ using (19) and to this end, according to (16) we first estimate $\hat{\phi}(r) = \cosh(\pi r)\, M_{\phi}(r)$. We follow Kuznetzov keeping track of the dependence on mn.

The following asymptotic expansion of the Bessel function is uniform see [Du];

$$J_{ir}(y) = \frac{c\, e^{irw(y/r) + \frac{\pi r}{2}}}{\sqrt{y^2 + r^2}} \left(1 + \frac{a_1}{\sqrt{r^2 + y^2}} \dots\right)$$

where c, a_1 are constants and

$$w(s) = \sqrt{1 + s^2} + \log\left(\frac{1}{s} - \sqrt{\frac{1}{s^2} + 1}\right) . \tag{42}$$

We analyze the leading term, the lower order terms are treated similarly and their contribution is smaller. For $|r| \leq 1$ we have

$$\hat{\phi}(r) \ll |r|^{-2} \tag{43}$$

as is clear from the Taylor expansion for J_{ir} when $\frac{a}{x} \leq 1$ and from (42) otherwise. So assume that $r \geq 1$ (or ≤ -1) and making the substitution $y = rs$ we are reduced to bounding

$$r^{-1/2} \int_0^\infty \frac{e^{irw(s)}}{(s^2+1)^{1/4}} \, \phi(rs) \, \frac{ds}{s} \tag{44}$$

$$= r^{-3/2} \int_0^\infty \left(e^{irw(s)} rw'(s) \right) \frac{\phi(rs)}{w'(s)\, s(s^2+1)^{1/4}} \, . \tag{45}$$

Now $w'(s)$ is bounded away from zero uniformly on $(0, \infty)$ and approaches 2 as $s \to \infty$ and behaves like $\frac{1}{s}$ as $s \to 0$. We may then apply the following easily proven mean value estimate: If F and G are defined on $[A, B]$ with G monotonic and taking values in $[0, 1]$ then

$$\left| \int_A^B Fx)\, G(x)\, dx \right| \leq 2 \sup_{A \leq B \leq C} \left| \int_A^C F(x) dx \right| . \tag{46}$$

This yields that the quantity in (45) is at most $O(r^{-3/2})$ and hence that

$$\hat{\phi}(r) \ll r^{-3/2} \text{ for } |r| \geq 1 \, . \tag{47}$$

For r large we seek a better bound which one gets by integration by parts in (45). This gives

$$-r^{-3/2} \int_0^\infty e^{irw(s)} \, \frac{d}{ds} \left(\frac{\phi(rs)}{w'(s)s} \right) ds$$

$$= O(r^{-5/2}) + r^{-3/2} \int_0^\infty \frac{e^{irw(s)}\, \theta'(s)}{(s^2+1)^{1/4} s w'(s)} \, ds$$

where the first term follows as in (47) and $\theta(s) = \phi(rs)$. Applying the mean value estimate to the last integral and using property (iii) of ϕ yields

$$\hat{\phi}(r) \ll \frac{x}{T} r^{-5/2} . \tag{48}$$

The bounds (47) and (48) are the same as those obtained by Kuznetzov only now they are uniform in x and nm.

For the term involving $\overline{\tau(m,t)}$ in (19) and our choice of test function we have

$$\int_{-\infty}^{\infty} \frac{\hat{\phi}(r)}{|\rho(1 + 2ir)|^2} \, dr \ll 1 \tag{49}$$

which follows immediately from (43).

The term in (19) involving the k-sum over Fourier coefficients of holomorphic forms and is handled by using the Ramanujan Conjecture for these. We find that this sum is bounded by

$$(mn)^{\epsilon} \sum_{\substack{k \equiv 0(2) \\ k > 2}} 4(k-1) \, |N_{\phi}(k)| \, . \tag{50}$$

Now for $x \geq \sqrt{mn}$ it is immediate from the decay of the Bessel function at small argument that $N_{\phi}(k) \ll 1/k!$. Hence the sum in (50) is $O_{\epsilon}((mn)^{\epsilon})$ which is a lot smaller than the upper bounds that we derive for the right hand side of (19).

For $x \leq \sqrt{mn}$ we need to investigate the transition ranges for the Bessel functions $J_k(y)$. That is, the ranges $y \leq k - k^{1/3}$, $k - k^{1/3} \leq y \leq k + k^{1/3}$ and $y \geq k + k^{1/3}$. We invoke the formula for the leading term asymptotic behavior these in each region, see ([Ol] for uniform asymptotics which allows one to connect the ranges). For our ϕ we break the integral (16) defining $N_{\phi}(k)$ into the corresponding ranges. In the range $(0, k - k^{1/3})$ the Bessel function is exponentially small and so the contribution is negligible. On the transitional range we use (22) and bound the integrand in absolute value. The contribution from this part to $(k-1) \, N_{\phi}(k)$ is $O(1)$ and since there are $O\left(\frac{\sqrt{mn}}{x}\right)$ values of k for which the transitional range is present we conclude that:

The contribution to the sum (50) from the transitional range is

$$O\left(\frac{\sqrt{mn}}{x} \, (mn)^{\epsilon}\right) . \tag{51}$$

We are left with the contribution to (50) from the range $y \geq k + k^{1/3}$ in the integral defining $N_\phi(k)$. In this range we have

$$J_k(ks) \;\sim\; \frac{e^{ikW(s)}}{\sqrt{k}(s^2-1)^{1/4}} \tag{52}$$

where

$$W((s) \;=\; \sqrt{s^2-1} \;-\; \arctan\sqrt{s^2-1}\,. \tag{53}$$

In particular

$$W'(s) \;=\; \frac{\sqrt{s^2-1}}{s}\,. \tag{54}$$

Changing variables for the integral in the range in question leads one to considering

$$\int\limits_{1+k^{-2/3}}^{\infty} \frac{\phi(ks)\, e^{ikW(s)}}{\sqrt{k}\, s(s^2-1)^{1/4}}\, ds\,. \tag{55}$$

We argue as with $\hat{\phi}(r)$ and the elementary mean-value estimate. It is enough to bound

$$\int\limits_{1+k^{-2/3}}^{c} \frac{e^{ikW(s)}}{s\sqrt{k}(k^2-1)^{1/4}}\, ds \tag{56}$$

independent of c.

Multiply and divide by $kW'(s)$ and integrate by parts. The boundary terms are $e^{ikW(s)}\big/\left(k^{3/2}(s^2-1^{3/4}\right)$ evaluated at $1+k^{2/3}$ and c. Hence they are $O(1/k)$. The resulting new integral is

$$-k^{-3/2} \int\limits_{1+k^{-2/3}}^{c} \frac{3s\, e^{ikW(s)}}{2(s^2-1)^{7/4}}\, ds \tag{57}$$

Now bound this trivially by estimating the integrand in absolute value. This also gives a contribution of $O(1/k)$. The number of k's for which this range intersects the support of ϕ is again $O\left(\frac{\sqrt{mn}}{x}\right)$. It follows that; the contribution to the k sum from this range is

$$O\left((mn)^{\epsilon}\,\frac{\sqrt{mn}}{x}\right)\,. \tag{58}$$

That is we have shown that

$$\sum_{k\,\equiv\,0(2)} N_{\phi}(k) \sum_{1\le j\le \dim s_k(\Gamma)} \overline{\psi_{j,k}(m)}\,\psi_{j,k(n)} \ll (mn)^{\epsilon}\left(1+\frac{\sqrt{mn}}{x}\right)\,. \tag{59}$$

(49) and (59) give us the desired bounds for the last two terms in (19). In order to complete the analysis we must estimate the first term on the right hand side of (19). It is here that we invoke H_θ. Consider the dyadic sums

$$\sum_{A\le t_j\le 2A} \frac{\rho_j(n)\,\overline{\rho_j(m)}}{\cosh\,\pi t_j}\,\hat{\phi}(t_j) \tag{60}$$

One can treat these in two ways. Firstly we can use H_θ directly from which it follows (for a Hecke basis of Maass forms)

$$|\rho_j(n)| \le \tau(n)\,n^{\theta}|\rho_j(1)|\,. \tag{61}$$

Hence

$$\sum_{A\le t_j\le 2A}\left|\frac{\rho_j(n)\,\overline{\rho_j(m)}}{\cosh\,\pi t_j}\right| \le \tau(n)\,\tau(m)(nm)^{\theta} \sum_{A\le t_j\le 2A}\frac{|\rho_j(1)|^2}{\cosh\,\pi t_j} \tag{62}$$

We recall Kuznetzov's meal value estimate:

$$\sum_{t_j\le y}\frac{|\rho_j(n)|^2}{\cosh\,\pi t_j} = \frac{y^2}{\pi} + O_{\epsilon}\left(y\log y + yn^{\epsilon} + n^{\frac{1}{2}+\epsilon}\right) \tag{63}$$

Applying (63) with $n=1$ in (62) yields

$$\left|\sum_{A\le t_j\le 2A}\frac{\rho_j(n)\,\overline{\rho_j(m)}}{\cosh\,\pi t_j}\right| \ll_{\epsilon} (nm)^{\theta+\epsilon}\,A^2\,. \tag{64}$$

Alternatively we can estimate (60) directly using (63) via Cauchy-Schwartz and obtain

$$\sum_{A\le t_j\le 2A} \left|\frac{\rho_j(n)\,\overline{\rho_j(m)}}{\cosh(\pi t_j)}\right| \le \left(\sum_{A}^{2A}\frac{|\rho_j(n)|^2}{\cosh \pi t_j}\right)^{1/2}\left(\sum_{A}^{2A}\frac{|\rho_j(m)|^2}{\cosh \pi t_j}\right)^{1/2}$$
$$\ll_{\epsilon} (A + m^{1/4+\epsilon})\,(A + n^{1/4+\epsilon})\,. \tag{65}$$

With these we have

$$\left|\sum_j \hat{\phi}(t_j)\,\frac{\rho_j(m)\,\overline{\rho_j(n)}}{\cosh \pi t_j}\right| \le \sum_j \frac{|\hat{\phi}(t_j)\,\rho_j(n)\,\rho_j(m)|}{\cosh \pi t_j} \tag{66}$$

and breaking this into dyadic pieces applying (47), (48), (64) or (65) one obtains

$$\sum_{A}^{2A}\left|\hat{\phi}(t_j)\,\frac{\rho_j(n)\,\overline{\rho_j(m)}}{\cosh \pi t_j}\right| \ll (mn)^{\epsilon}$$
$$\min\left(1,\frac{x}{TA}\right)\,\min\left(\sqrt{A}(nm)^{\theta},\ \sqrt{A}+\left(n^{\frac14}+m^{1/4}\right)\,A^{-1/2}+(mn)^{\frac14}\,A^{-3/2}\right) \tag{67}$$

$$\ll (mn)^{\epsilon}\min\left(1,\frac{x}{TA}\right)\,\left(\sqrt{A}+\left(m^{1/8}+n^{1/8}\right)\,(mn)^{\theta/2}\right). \tag{68}$$

Hence,

$$\sum_{A}^{2A}\left|\hat{\phi}(t_j)\frac{\rho_j(n)\overline{\rho_j(m)}}{\cosh \pi t_j}\right| \ll (mn)^{\epsilon}\left(\left(m^{1/8}+n^{1/8}\right)(mn)^{\theta/2}+\min\left(\sqrt{A},\frac{x}{T\sqrt{A}}\right)\right) \tag{69}$$

Combining the dyadic pieces yields

$$\sum_j \hat{\phi}(t_j)\,\frac{\rho_j(n)\,\overline{\rho_j(m)}}{\cosh \pi t_j} \ll (mn)^{\epsilon}\,\left(\left(m^{1/8}+n^{1/8}\right)\,(mn)^{\theta/2} + \frac{\sqrt{x}}{T}\right)\,. \tag{70}$$

Putting this together with (41) and (51) in (19) yields

$$\sum_{x\le c\le 2x} \frac{S(m,n;c)}{c} \ll (xmn)^\epsilon \left(\frac{T}{\sqrt{x}} + \frac{\sqrt{mn}}{x} + \left(m^{1/8}+n^{1/8}\right)(nm)^{\theta/2} + \sqrt{\frac{x}{T}} \right).$$

Finally choosing $T = x^{2/3}$ yields

$$\sum_{x\le c\le 2x} \frac{S(m,n;c)}{c} \ll (xmn)^\epsilon \left(x^{1/6} + \frac{\sqrt{mn}}{x} + \left(m^{1/8}+n^{1/8}\right)(mn)^{\theta/2} \right). \tag{71}$$

This completes the proof of Proposition 5.

In Theorem 4 we explained the $(mn)^{1/6}$ in Theorem 2. The $x^{1/6}$ barrier is similar, that is in the proof of Proposition 5 if we want to go beyond the exponent $1/6$ (ignoring the mn dependence) we would need to capture cancellations in sums of the type

$$\sum_{t_j \sim x^{1/3}} \frac{|\rho_j(1)|^2}{\cosh \pi t_j} x^{itj}. \tag{72}$$

This appears to be quite difficult. A similar feature appears with the exponent of $1/3$ in the remainder term in the hyperbolic circle problem which has resisted improvements (see [L-P] and [Iw]).

Acknowledgement: We thank D. Hejhal for his comments on an earlier draft of this paper.

3 References

[D-I] J. Deshouillers and H. Iwaniec, *Inv. Math.*, **70**, (1982), 219-288.

[De] P. Deligne, *Publ. Math. IHES*, **43**, (1974), 273-307.

[Du] T.M. Dunster, "Bessel functions of purely imaginary order with applications to second order linear differential equations," *Siam Jnl. Math. Anal.*, **21**, No. 4, (1990).

[G-S] D. Goldfeld and P. Sarnak, *Inv. Math.*, **71**, (1983), 243-250.

[Hu] M.N. Huxley, "Introduction to Kloosertmania" in Banach Center Publications, Vol. 17, (1985), Polish Scientific Publishers.

[Iw] H. Iwaniec, "Introduction to the spectral theory of automorphic forms," *AMS*, (2002).

[I-K] H. Iwaniec and F. Kowalski, "Analytic Number Theory," *AMS*, Coll. Publ., Vol. 53, (2004).

[I-L-S] H. Iwaniec, W. Luo and P. Sarnak, *Publ. IHES*, **91**, (2001), 55-131.

[Ka] N. Katz, "Gauss Sums, Kloosterman Sums and Monodromy," P.U. Press, (1988).

[Ki-Sh] H. Kim and F. Shahidi, *Annals of Math.*, **155**, no. 3 (2002), 837-893.

[Ki-Sa] H. Kim and P. Sarnak, *JAMS*, **16**, (2003), 175-181.

[Ku] N. Kuznetzov, *Mat. Sb. (N.S)*, **111**, (1980), 334-383, *Math. USSR-Sb.*, **39**, (1981), 299-342.

[L-P] P. Lax and R. Phillips, *Jnl Funct. Anal.*, **46**, (1982), 280-350.

[Li] Y. Linnik, *Proc. ICM Stockholm*, (1962), 270-284.

[L-R-S] W. Luo, Z. Rudnick and P. Sarnak, *GAFA* **5**, (1995), 387-401.

[Mi] P. Michel, *Inv. Math.*, **121**, (1995), 61-78.

[Mu] R. Murty, in "Lectures on Automorphic L-functions", Fields Institute Monographs, vol. 20, (2004).

[Ol] F.W.J. Olver, *Phil. Trans. Royal Soc. Ser A*, **247**, (1954), 328-368.

[Sa] P. Sarnak, "Notes on the Generalized Ramanujan Conjectures," *Clay Math. Proc.*, Vol. 4, (2005), 659-685.

[Se] A. Selberg, *Proc. Symp.*, *Pure Math (AMS)*, **8**, (1965), 1-8.

[Wa] G. Watson, "A treatise on the Theory of Bessel Functions," Cambridge Press, (1966).

[We] A. Weil, *Proc. Nat. Acad. Sci.*, U.S.A., **34**, (1948), 204-207.