

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

**ScienceDirect**

Advances in Mathematics 226 (2011) 3550–3578

---



---

**ADVANCES IN  
Mathematics**


---



---

[www.elsevier.com/locate/aim](http://www.elsevier.com/locate/aim)

# The second moment of $GL(3) \times GL(2)$ $L$ -functions, integrated <sup>☆</sup>

Matthew P. Young

*Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA*

Received 19 February 2010; accepted 15 October 2010

Available online 30 October 2010

Communicated by Ravi Vakil

---

## Abstract

We consider the family of Rankin–Selberg convolution  $L$ -functions of a fixed  $SL(3, \mathbb{Z})$  Maass form with the family of Hecke–Maass cusp forms on  $SL(2, \mathbb{Z})$ . We estimate the second moment of this family of  $L$ -functions with a “long” integration in  $t$ -aspect. These  $L$ -functions are distinguished by their high degree (12) and large conductors (of size  $T^{12}$ ).

© 2010 Elsevier Inc. All rights reserved.

*Keywords:* Moments;  $L$ -functions; Maass forms; Automorphic forms; Voronoi formula; Kuznetsov formula; Spectral large sieve

---

## 1. Introduction

In this paper we study the second moment of the Rankin–Selberg  $L$ -functions  $L(\phi \times u_j, \frac{1}{2} + it)$  of a fixed Hecke–Maass form  $\phi$  on  $SL(3, \mathbb{Z})$  convolved with the family  $u_j$  of Hecke–Maass cusp forms on  $SL(2, \mathbb{Z})$  as well as with the twists by  $n^{it}$ . This family is “large” as measured in a variety of ways: there are  $T^3$  elements in the family, each having degree 12 and conductor of size  $T^{12}$ . For comparison, the classical large sieve can estimate the eighth moment of the family of classical Dirichlet  $L$ -functions of modulus  $q \leq Q$  (having  $Q^2$  elements of degree 8 and conductors of size  $Q^8$ ). The various  $GL(2)$  large sieve type inequalities also generally allow for degree 8  $L$ -functions with similarly-sized conductors, so that one can make a case that this family is significantly larger than others appearing in the literature. For use in applications, it

---

<sup>☆</sup> Work supported by NSF grant DMS-0758235.  
E-mail address: [myoung@math.tamu.edu](mailto:myoung@math.tamu.edu).

is highly desirable to have control over large families of harmonics, as they produce stronger detectors of arithmetical functions; see the introduction of [3].

Another way to motivate interest in this particular family is that the first moment (at the central point  $s = 1/2$ ) was recently used by X. Li [17,18] to show subconvexity for a self-dual  $GL(3)$   $L$ -function in  $t$ -aspect (amongst other things). The self-duality is crucially used to impose nonnegativity of the central values. In order to use moments to study non-self-dual forms, as well as Rankin–Selberg convolutions at points other than  $s = 1/2$ , it seems necessary to study the second moment. However, this approach has substantial new difficulties. In particular, the second moment of this family at the central point has prohibitively large conductors (of size  $T^{12}$  compared to  $T^2$  elements in the family, a sixth power). However, one can enlarge the size of the family without substantially growing the size of the conductors by twisting by  $n^{it}$  with  $t$  almost as large as the spectral parameter. This procedure then brings the problem into the presumably more feasible range where the conductor is the fourth power of the size of the family. Even so, the conductors of the family are still very large so that estimating this moment requires a substantial amount of cancellation. In fact, the main difficulty is showing simultaneous cancellation in the twists by the Hecke–Maass forms as well as by  $n^{it}$ . Many of the methods in the literature used to show cancellation in the Maass form aspect are incompatible with the  $t$ -aspect integration.

**Theorem 1.1.** *We have*

$$\int_{-T^{1-\varepsilon}}^{T^{1-\varepsilon}} \sum_{\substack{T < t_j \leq 2T \\ u_j \text{ even}}} \left| L\left(u_j \times \phi, \frac{1}{2} + it\right) \right|^2 dt \ll T^{3+\varepsilon}. \tag{1.1}$$

**Remarks.**

- The conductor of  $|L(u_j \times \phi, \frac{1}{2} + it)|^2$  is  $\asymp T^{12}$ , so that the convexity bound is recovered above using the method of Heath-Brown (Lemma 3 of [8]).
- An easy consequence of Theorem 1.1 is that for “almost all”  $t_j$  in the family,

$$\int_{-T^{1-\varepsilon}}^{T^{1-\varepsilon}} \left| L\left(u_j \times \phi, \frac{1}{2} + it\right) \right|^2 dt \ll T^{1+\varepsilon},$$

in the sense that for any fixed  $\varepsilon > 0$ , the number of  $t_j$  for which this bound is not satisfied is  $o(T^2)$ .

- The reason  $t$  is slightly smaller than  $T$  is to avoid the intricate dependence of the conductor on  $t_j$ , though for  $t = t_j$  see the companion paper [25]. With some extra work it is likely that one could extend the left-hand side of (1.1) to  $|t| \leq T, t_j \leq T$ , but we have not carried out the details.
- The method of proof can also handle the analogous twists in the weight aspect by classical holomorphic modular forms on the full modular group; see Section 8 of [15].
- The combined  $t$ -integral and spectral sum is reminiscent of Sarnak’s work on the fourth moment of Größencharacter  $L$ -functions [22].
- Diaconu, Garrett, and Goldfeld [5] have generalized the method of Good [7] to capture quantities of the form (1.1), but with certain weight functions (depending on  $t_j$ ) in the integral.

It is difficult to asymptotically evaluate these weight functions, so it is unknown what this implies about (1.1).

- A. Venkatesh posed this problem during a problem session at the October 2006 AIM workshop on the subconvexity problem.

Theorem 1.1 potentially represents progress towards subconvexity for these  $L$ -functions. If one could shrink the family by any significant amount and still obtain a Lindelöf-consistent upper bound then one would obtain a subconvex bound. This would require another source of cancellation. (Perhaps from solving the shifted convolution problem for a fixed  $GL(3)$  Maass form?) An additional problem is that one has to abandon the use of the large sieve inequality (a crucial ingredient in our proof) yet reclaim its substantial savings effect.

A natural way to attack this problem is with a hybrid large sieve inequality of the form

$$\int_{-U}^U \sum_{t_j \leq T} \left| \sum_{n \leq N} a_n \lambda_j(n) n^{it} \right|^2 dt \ll \Delta(N, T, U) (NTU)^\varepsilon \sum_{n \leq N} |a_n|^2. \tag{1.2}$$

A simple application of Iwaniec’s spectral large sieve [11] shows  $\Delta(N, T, U) \ll U(N + T^2)$ , and one would like to replace this by  $N + UT^2$ . However, this appears to be an extremely difficult problem, and in fact in this generality it essentially implies the Ramanujan–Pettersson conjecture for Maass forms! (To see this, take  $U$  to be a large fixed power of  $N$  to pick out the diagonal only on the left-hand side and choose  $a_n$  to select  $n = N$  only, showing  $|\lambda_j(N)|^2 \ll T^2(TN)^\varepsilon \ll_j N^\varepsilon$ .) One might even consider a simpler problem where  $U$  has restricted size (say  $U \leq T$ ) and  $N$  is large with respect to  $T$ . Even this seems to be a difficult and highly interesting problem (in my opinion). The state of affairs here for  $GL(2)$  harmonics is quite different than for  $GL(1)$  (multiplicative characters), where we do have the essentially optimal result of Gallagher [4]

$$\int_{-U}^U \sum_{q \leq Q} \frac{q}{\phi(q)} \sum_{\chi \pmod q}^* \left| \sum_{n \leq N} a_n \chi(n) n^{it} \right|^2 dt \ll (N + UQ^2) \sum_{n \leq N} |a_n|^2. \tag{1.3}$$

The difficulty of estimating (1.2) for general coefficients  $a_n$  is a barrier in our problem, which requires good estimates when the  $a_n$ ’s are specialized to be coefficients of a  $GL(3)$   $L$ -function. However, we have more available tools for these specific choices of coefficients, and in particular the  $GL(3)$  Voronoi summation formula plays a key role. X. Li [17,18] showed how this summation formula can be very powerful in the study of this family, but an attempt to directly generalize her approach on the first moment meets extreme difficulties with the second moment (see the first few sentences of Section 6 below).

In our companion paper [25], we considered this family of  $L$ -functions at the special point  $\frac{1}{2} + it_j$ . There are some similarities between the two problems but each requires substantially different ideas. In particular, the analog of one of the key ideas in [25] (namely, Poisson summation in the variable  $a$  modulo  $b$  in (6.2) below) is not used here as it ultimately seemed not to be substantially helpful, despite some promising hints. Another major difference between these works is that in [25] we were able to appeal to some large sieve inequalities due to [19] (after [2]), which we improved further, while here we could not use anything of the form (1.2). Furthermore, the conductors of the family of  $T^2$  elements in [25] have size  $T^6$  (the cube of the number

of elements in the family as opposed to the fourth power appearing here), which has the effect that the  $GL(3)$  Voronoi summation formula is relatively less powerful in this article than in the companion. Indeed, the dual sum after Voronoi summation had essentially no length! (Though it should be stressed that this feature only occurred due to our improvement on the relevant large sieve inequality.) This paper is independent of [25] (at the cost of some repetition).

## 2. Notation

See [6] for the material and notation on  $GL(3)$  Maass forms. Suppose  $\phi$  is a Maass form for  $SL(3, \mathbb{Z})$  of type  $(\nu_1, \nu_2) \in \mathbb{C}^2$  which is an eigenfunction of all the Hecke operators. The Godement–Jacquet  $L$ -function associated to  $\phi$  is

$$L(\phi, s) = \sum_{n=1}^{\infty} \frac{A(1, n)}{n^s} = \prod_p (1 - A(1, p)p^{-s} + A(p, 1)p^{-2s} - p^{-3s})^{-1}.$$

Here  $A(m, n)$  are the Fourier coefficients normalized as in [6]. In particular,  $A(1, 1) = 1$  and  $|A(m, n)|^2$  are constant on average (see Remark 12.1.8 of [6]). The dual Maass form  $\tilde{\phi}$  is of type  $(\nu_2, \nu_1)$  and has  $A(n, m) = \overline{A(m, n)}$  as its  $(m, n)$ -th Fourier coefficient, whence  $L(\tilde{\phi}, s) = \sum_{n=1}^{\infty} \overline{A(n, 1)}n^{-s}$ . Letting

$$\Gamma_{\nu_1, \nu_2}(s) = \pi^{-3s/2} \Gamma\left(\frac{s + 1 - 2\nu_1 - \nu_2}{2}\right) \Gamma\left(\frac{s + \nu_1 - \nu_2}{2}\right) \Gamma\left(\frac{s - 1 + \nu_1 + 2\nu_2}{2}\right),$$

the functional equation for  $L(\phi, s)$  reads

$$\Gamma_{\nu_1, \nu_2}(s)L(\phi, s) = \Gamma_{\nu_2, \nu_1}(1 - s)L(\tilde{\phi}, 1 - s). \tag{2.1}$$

Let  $(u_j)$  be an orthonormal basis of Hecke–Maass cusp forms on  $SL(2, \mathbb{Z})$  with corresponding Laplace eigenvalues  $\frac{1}{4} + t_j^2$ . Let  $\lambda_j(n)$  be the Hecke eigenvalue of the  $n$ -th Hecke operator for the form  $u_j$ . Since the Hecke operators on  $GL(2)$  are self-adjoint, the  $\lambda_j(n)$ 's are real. Then  $L(u_j, s) = \sum_{n=1}^{\infty} \lambda_j(n)n^{-s}$  satisfies a functional equation relating to  $L(u_j, 1 - s)$ . To be clear, the Hecke operators are normalized so that the Ramanujan–Petersson bound is  $|\lambda_j(p)| \leq 2$ . Say that the  $n$ -th Fourier coefficient of  $u_j$  is  $\rho_j(n)$ , so that  $\rho_j(n) = \rho_j(1)\lambda_j(n)$ . With the definition

$$\alpha_j = \frac{|\rho_j(1)|^2}{\cosh(\pi t_j)}, \tag{2.2}$$

then  $t_j^{-\varepsilon} \ll \alpha_j \ll t_j^\varepsilon$  due to [9] and [12].

As explained in Chapter 12.2 of [6], the Rankin–Selberg convolution of  $\phi$  and  $u_j$  is

$$L(u_j \times \phi, s) = \sum_{m, n=1}^{\infty} \frac{\lambda_j(n)A(m, n)}{(m^2 n)^s}. \tag{2.3}$$

The completed  $L$ -function associated to  $L(u_j \times \phi, s)$ , for  $u_j$  even, takes the form

$$\Lambda(u_j \times \phi, s) = \pi^{-3s} \Gamma\left(\frac{s - it_j - \alpha}{2}\right) \Gamma\left(\frac{s - it_j - \beta}{2}\right) \Gamma\left(\frac{s - it_j - \gamma}{2}\right) \\ \times \Gamma\left(\frac{s + it_j - \alpha}{2}\right) \Gamma\left(\frac{s + it_j - \beta}{2}\right) \Gamma\left(\frac{s + it_j - \gamma}{2}\right) L(u_j \times \phi, s), \quad (2.4)$$

where  $\alpha = -\nu_1 - 2\nu_2 + 1$ ,  $\beta = -\nu_1 + \nu_2$ , and  $\gamma = 2\nu_1 + \nu_2 - 1$  (see Theorem 12.3.6 of [6] for the explicit gamma factors). Then this Rankin–Selberg convolution has a holomorphic continuation to  $s \in \mathbb{C}$  and satisfies the functional equation

$$\Lambda(u_j \times \phi, s) = \Lambda(u_j \times \tilde{\phi}, 1 - s). \quad (2.5)$$

### 3. Basic tools

#### 3.1. Approximate functional equation

We shall use an approximate functional equation to represent the values of  $L$ -functions. Write  $\lambda_{u_j \times \phi}(n)$  for the coefficient of  $n^{-s}$  in the Dirichlet series (2.3) for  $L(u_j \times \phi, s)$ . Then Theorem 5.3 of [13] says for any  $X > 0$ ,

$$L\left(u_j \times \phi, \frac{1}{2} + it\right) = \sum_n \frac{\lambda_{u_j \times \phi}(n)}{n^{\frac{1}{2} + it}} V_{\frac{1}{2} + it}(n/X) + \epsilon_{t,t_j} \sum_n \frac{\lambda_{u_j \times \tilde{\phi}}(n)}{n^{\frac{1}{2} - it}} V_{\frac{1}{2} - it}^*(nX), \quad (3.1)$$

where  $V_s(y)$  and  $V_s^*(y)$  are certain explicit smooth functions, and  $\epsilon_{t,t_j}$  is a certain complex number of absolute value 1. More precisely,

$$V_{\frac{1}{2} + it}(y) = \frac{1}{2\pi i} \int_{(3)} y^{-s} \frac{\gamma(\frac{1}{2} + it + s)}{\gamma(\frac{1}{2} + it)} \frac{G(s)}{s} ds, \quad (3.2)$$

where  $\Lambda(u_j \times \phi, s) = \gamma(s)L(u_j \times \phi, s)$  and  $G(s)$  is an entire function with rapid decay in the imaginary direction. Here  $V_{\frac{1}{2} - it}^*$  has a similar form to  $V_{\frac{1}{2} + it}$  but with  $\gamma(s)$  replaced by  $\gamma^*(s)$ , where  $\Lambda(u_j \times \tilde{\phi}, s) = \gamma^*(s)L(u_j \times \tilde{\phi}, s)$ .

#### 3.2. The large sieve

The classical large sieve inequality for Farey fractions states

$$\sum_{b \leq B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_{N \leq m < N+M} a_m e\left(\frac{xm}{b}\right) \right|^2 \leq (B^2 + M) \sum_{N \leq m < N+M} |a_m|^2. \quad (3.3)$$

For our purposes we require an additional oscillatory integral in the spirit of [4], but we could not find the following result in the literature.

**Lemma 3.1.** *Let  $f(y)$  be a continuously differentiable function on  $[N, N + M]$  such that  $f'$  does not vanish. Let  $X = \sup_{y \in [N, N+M]} \frac{1}{|f'(y)|}$ . Then for any complex numbers  $b_m$ ,*

$$\int_{-T}^T \sum_{b \leq B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_{N \leq m < N+M} b_m e\left(\frac{xm}{b}\right) e(tf(m)) \right|^2 dt \ll (B^2 T + X) \sum_{N \leq m < N+M} |b_m|^2. \tag{3.4}$$

We reproduce the proof appearing in [25].

**Proof of Lemma 3.1.** By the change of variables  $t \rightarrow tT$ , it suffices to consider the case  $T = 1$ . Let  $g$  be a nonnegative Schwartz function such that  $g(x) \geq 1$  for  $|x| \leq 1$ , such that  $\hat{g}$  has compact support. See [24] for a nice survey on such functions as well as some ideas relevant in this proof. Then for any sequence of complex numbers  $c_m$ , we have

$$\int_{-1}^1 \left| \sum_m c_m e(tf(m)) \right|^2 dt \leq \sum_{m,n} c_m \bar{c}_n \hat{g}(f(m) - f(n)). \tag{3.5}$$

Since  $\hat{g}$  is compactly supported, we must have  $|f(m) - f(n)| \ll 1$ . By the mean-value theorem,  $|f(m) - f(n)| \geq |m - n| \inf_y |f'(y)|$ , so  $|m - n| \ll X$ . Dissect the sum over  $m$  and  $n$  into boxes  $I \times J$  of sidelength  $\ll \min(M, X)$  so that the only relevant boxes  $I \times J$  have  $I$  and  $J$  either equal or adjacent (“nearby”, say). Thus the right-hand side of (3.5) equals

$$\sum_{I,J \text{ nearby}} \sum_{(m,n) \in I \times J} c_m \bar{c}_n \hat{g}(f(m) - f(n)). \tag{3.6}$$

Having enforced the condition that  $I$  and  $J$  are nearby, we then reverse the Fourier transform to express it in terms of  $g$ , getting that (3.6) equals to

$$\int_{-\infty}^{\infty} g(t) \sum_{I,J \text{ nearby}} \sum_{(m,n) \in I \times J} c_m e(tf(m)) \overline{c_n e(tf(n))} dt. \tag{3.7}$$

By Cauchy’s inequality, (3.7) is

$$\ll \int_{-\infty}^{\infty} g(t) \sum_I \left| \sum_{m \in I} c_m e(tf(m)) \right|^2 dt.$$

Specializing this to  $c_m = b_m e(\frac{xm}{b})$  and summing over  $x$  and  $b$  appropriately gives that the left-hand side of (3.4) (with  $T = 1$ ) is

$$\ll \int_{-\infty}^{\infty} g(t) \sum_I \sum_{b \leq B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_{m \in I} b_m e\left(\frac{xm}{b}\right) e(tf(m)) \right|^2 dt. \tag{3.8}$$

By (3.3) with  $a_m = b_m e(tf(m))$ , we complete the proof noting that (3.8) is

$$\ll \int_{-\infty}^{\infty} g(t) \sum_I (B^2 + \min(M, X)) \sum_{m \in I} |b_m|^2 \ll (B^2 + X) \sum_{N \leq m < N+M} |b_m|^2. \quad \square$$

### 3.3. Kuznetsov formula

Our tool for summing over the spectrum is the following.

**Lemma 3.2 (Kuznetsov).** *Suppose that  $h$  is holomorphic in the region  $|\text{Im}(r)| \leq \frac{1}{2} + \delta$  and satisfies  $h(r) = h(-r)$  and  $|h(r)| \ll (1 + |r|)^{-2-\delta}$  for some  $\delta > 0$ . Then*

$$\begin{aligned} & \sum_j \alpha_j \lambda_j(m) \lambda_j(n) h(t_j) + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sigma_{2ir}(m) \sigma_{2ir}(n)}{(mn)^{ir} |\zeta(1 + 2ir)|^2} h(r) dr \\ &= \pi^{-2} \delta_{m=n} \int_{-\infty}^{\infty} r \tanh(\pi r) h(r) dr + \sum_{c=1}^{\infty} \frac{S(m, n; c)}{c} \check{h}\left(\frac{4\pi\sqrt{mn}}{c}\right), \end{aligned}$$

where  $\alpha_j$  is defined by (2.2) and

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

## 4. Initial cleaning

Throughout this article it is very convenient to refer to functions  $f$  satisfying the following bounds

$$x^k f^{(k)}(x) \ll_{k,C} \left(1 + \frac{|x|}{Y}\right)^{-C}, \tag{4.1}$$

for each  $k, C \geq 0$ , and some parameter  $Y$ .

The following technical lemma gives a pointwise upper bound on an  $L$ -function with a weight function that only loosely depends on its parameters.

**Lemma 4.1.** *Let  $\varepsilon > 0$ , suppose that  $T^\varepsilon \leq U \leq \Delta \leq T^{1-\varepsilon}$ , and  $|t| \leq U, T < t_j \leq T + \Delta$ . Then there exist finitely many functions  $W_k$  independent of  $t$  and  $t_j$  (but depending on  $U$  and  $T$ ) satisfying (4.1) with  $Y = 1$ , such that for some fixed interval  $[X_0, X_1] \subset (0, \infty)$ , we have*

$$\left| L\left(u_j \times \phi, \frac{1}{2} + it\right) \right|^2 \ll \sum_k \int_{X_0}^{X_1} \left| \sum_n \frac{\lambda_{u_j \times \phi}(n)}{n^{\frac{1}{2}+it}} W_k(n/(XT^3)) \right|^2 dX + O(T^{-200}). \tag{4.2}$$

**Corollary 4.2.** Let  $\varepsilon > 0$ , suppose that  $T^\varepsilon \leq U \leq \Delta \leq T^{1-\varepsilon}$ , and  $|t| \leq U$ ,  $T < t_j \leq T + \Delta$ . Let

$$\mathcal{M}(T, U, \Delta) = \int_{-U}^U \sum_{\substack{T < t_j \leq T + \Delta \\ u_j \text{ even}}} \left| L\left(u_j \times \phi, \frac{1}{2} + it\right) \right|^2 dt. \tag{4.3}$$

Then there exists  $W$  independent of  $t$  and  $t_j$  (but depending on  $U$  and  $T$ ) satisfying (4.1) with  $Y = 1$  such that

$$\mathcal{M}(T, U, \Delta) \ll \int_{-U}^U \sum_{\substack{T < t_j \leq T + \Delta \\ u_j \text{ even}}} \left| \sum_n \frac{\lambda_{u_j \times \phi}(n)}{n^{\frac{1}{2} + it}} W(n/T^3) \right|^2 dt + O(T^{-100}). \tag{4.4}$$

To prove this corollary, take the  $k$ -sum and  $X$ -integral outside the  $t_j$ -sum and  $t$ -integral, bound it by the supremum (if we do not do this the choices of  $X$  and  $k$  depend on  $t, t_j$ ), and redefine  $W$ .

**Proof of Lemma 4.1.** We use a method similar to that of Section 5 of [25], which we modified from [2].

Begin by writing the approximate functional equation (3.1) in the shorthand form  $L(u_j \times \phi, \frac{1}{2} + it) = \sum_n a_n V(n/X) + \epsilon \sum_n b_n V^*(nX)$ . This is valid for all  $X > 0$ , so we freely integrate against  $X^{-1} dX$  from  $X = e^{-\frac{1}{2}}$  to  $X = e^{\frac{1}{2}}$ . Then by Cauchy’s inequality we deduce

$$\left| L\left(u_j \times \phi, \frac{1}{2} + it\right) \right|^2 \leq 2 \int_{e^{-\frac{1}{2}}}^{e^{\frac{1}{2}}} \left| \sum_n a_n V(n/X) \right|^2 \frac{dX}{X} + 2 \int_{e^{-\frac{1}{2}}}^{e^{\frac{1}{2}}} \left| \sum_n b_n V^*(nX) \right|^2 \frac{dX}{X}. \tag{4.5}$$

Notice that changing variables  $X \rightarrow X^{-1}$  in the latter term makes the two terms more symmetric. Indeed, the latter term then becomes the same as the former term with  $\phi$  replaced by its dual and  $t$  replaced by  $-t$ . We focus on the former term.

We find a simpler expression for  $V(x)$ . By Stirling’s approximation, we have for  $\lambda \in \mathbb{C}$  fixed,  $\text{Re}(s) > 0$ ,  $s$  small compared to  $T$ , and with the shorthand  $Z = \frac{1}{2} + it \pm it_j - \lambda$ , that for certain polynomials  $P_n$ , we have

$$\log \frac{\Gamma(\frac{Z+s}{2})}{\Gamma(\frac{Z}{2})} = \frac{s}{2} \log\left(\frac{Z}{2}\right) + \frac{P_1(s)}{Z} + \frac{P_2(s)}{Z^2} + \dots + O\left(\frac{P_k(s)}{Z^k}\right).$$

With  $Z' = \frac{1}{2} + it \mp it_j - \lambda$ , we have

$$\frac{1}{2} \log\left(\frac{Z}{2}\right) + \frac{1}{2} \log\left(\frac{Z'}{2}\right) = \log |Z/2| + \frac{1}{2} \log\left(1 + \frac{Z' - \bar{Z}}{\bar{Z}}\right).$$

A computation shows that



$$\frac{1}{2} \log \left( 1 + \frac{Z' - \bar{Z}}{\bar{Z}} \right) = \frac{Q_1(t)}{t_j} + \frac{Q_2(t)}{t_j^2} + \dots + O \left( \frac{Q_k(t)}{t_j^k} \right),$$

where each  $Q_l$  is a polynomial of degree  $l$ . Combining these estimates for the gamma factors, we obtain the asymptotic expansion

$$V(x) = \frac{1}{2\pi i} \int_{(\sigma)} (\pi^3 x)^{-s} \frac{G(s)}{s} |q(t, t_j)|^{s/2} \left( 1 + \frac{R_1(s, t)}{t_j} + \dots + O \left( \frac{R_k(s, t)}{t_j^k} \right) \right) ds, \quad (4.6)$$

where each  $R_l(s, t)$  is a polynomial in  $s$  and  $t$  of degree at most  $l$  in terms of  $t$ , and  $q(t, t_j)$  is the product of three  $Z$  and three  $Z'$  terms with  $\lambda$  replaced by  $\alpha, \beta, \gamma$ . Note  $|q(t, t_j)| \asymp T^6$ , uniformly in  $t$  and  $t_j$ . To be clear, we should choose a  $G(s)$  that decays rapidly for  $\text{Im}(s)$  large, such as  $e^{s^2}$ , truncate the  $s$ -integral at say  $\log^2 T$ , apply Stirling’s formula, and then relax the truncation, all with an acceptable error.

The representation (4.6) with  $\sigma > 0$  very large shows that we may truncate the sum over  $n$  at  $T^{3+\varepsilon}$  with an acceptable error term. With this truncation in place, we then fix  $\sigma = 1$  and insert the asymptotic expansion (4.6) into (4.5). According to (4.6), write  $V = V_0 + O(T^{3\sigma} (\frac{U}{T})^k)$  (so that  $V_0$  accounts for all the terms in the expansion except for the error term). For  $k$  sufficiently large in terms of  $U$  and  $T$ , this error term is acceptable. Then change variables  $X \rightarrow X\pi^3 T^3 / |q(t, t_j)|^{1/2}$  and by positivity extend the  $X$ -integral to a fixed interval, say  $[X_0, X_1] \subset (0, \infty)$ . Now

$$\begin{aligned} \left| L \left( u_j \times \phi, \frac{1}{2} + it \right) \right|^2 &\ll \int_{X_0}^{X_1} \left| \sum_n \frac{\lambda_{\phi \times u_j}(n)}{n^{\frac{1}{2}+it}} \int_{(1)} \frac{G(s)}{s} \left( \frac{XT^3}{n} \right)^s \left( 1 + \dots + \frac{R_{k-1}(s, t)}{t_j^{k-1}} \right) ds \right|^2 dX \\ &+ (\text{similar term}) + O(T^{-200}). \end{aligned}$$

Here the sum over  $n$  is truncated at  $T^{3+\varepsilon}$ , but since the contour can be shifted far to the right, we may relax this condition without making a new error term. Next write  $R_l(s, t) = R_{0,l}(s) + R_{1,l}(s)t + \dots + R_{l,l}(s)t^l$ , and use Cauchy’s inequality on the asymptotic expansion to obtain

$$\begin{aligned} \left| L \left( u_j \times \phi, \frac{1}{2} + it \right) \right|^2 &\ll \int_{X_0}^{X_1} \sum_{i \leq l < k} \frac{U^{2i}}{T^{2l}} \left| \sum_n \frac{\lambda_{\phi \times u_j}(n)}{n^{\frac{1}{2}+it}} V_{i,l}(n/(XT^3)) \right|^2 dX + (\text{similar}) \\ &+ O(T^{-200}), \end{aligned}$$

where

$$V_{i,l}(x) = \frac{1}{2\pi i} \int_{(1)} \frac{G(s)}{s} R_{i,l}(s) x^{-s} ds.$$

Clearly each  $V_{i,l}$  satisfies (4.1) with  $Y = 1$ . The “similar” term has the same form as the displayed term, except the coefficients are conjugated and the weight function is different (the Langlands parameters  $\alpha, \beta, \gamma$  are changed) yet it satisfies the same bounds (the proof was for any  $\alpha, \beta, \gamma$

fixed), so again it takes the same form as the right-hand side of (4.2). The total number of terms in the asymptotic expansion is finite.  $\square$

**Theorem 4.3.** *With conditions as in Lemma 4.1, we have*

$$\mathcal{M}(T, U, \Delta) \ll T^{3+\varepsilon}. \tag{4.7}$$

From Theorem 4.3, we deduce Theorem 1.1 by taking  $U = T^{1-\varepsilon}$ , and covering the interval  $[T, 2T]$  with  $O(T^\varepsilon)$  subintervals of the form  $[T', T' + \Delta]$  with  $\Delta = T^{1-\varepsilon}$ .

**Lemma 4.4.** *Let*

$$h(r) = \frac{r^2 + \frac{1}{4}}{T^2} \left[ \exp\left(-\left(\frac{r-T}{\Delta}\right)^2\right) + \exp\left(-\left(\frac{r+T}{\Delta}\right)^2\right) \right]. \tag{4.8}$$

*Then there exists a smooth function  $w$  with support in a dyadic interval  $[P, 2P]$  with*

$$P \ll T^{3+\varepsilon} \tag{4.9}$$

*satisfying (4.1) with  $Y = P$  such that*

$$\mathcal{M}(T, U, \Delta) \ll T^\varepsilon H(U, T, \Delta, w) + O(T^{-100}),$$

*where*

$$H(U, T, \Delta, w) = \int_{-U}^U \sum_{t_j} \alpha_j h(t_j) \left| \sum_n \frac{\lambda_{u_j \times \phi}(n)}{n^{\frac{1}{2}+it}} w(n) \right|^2 dt.$$

**Proof.** We start with (4.4). From the lower bound  $\alpha_j \gg t_j^{-\varepsilon}$  we may attach the weight  $\alpha_j$  at the cost of  $O(T^\varepsilon)$ . By positivity, we freely attach the weight function  $h$  and extend the summation to all  $t_j$ . Finally, we apply a smooth dyadic partition of unity to the inner  $n$ -sum, and apply Cauchy’s inequality to this sum over the partition (only the terms with  $P \ll T^{3+\varepsilon}$  are relevant by a trivial bound), completing the proof.  $\square$

**Lemma 4.5.** *Let  $g$  be a fixed nonnegative, even, Schwartz function satisfying  $g(t) \geq 1$  for  $|t| \leq 2$  whose Fourier transform is compactly supported, and define for any finite sequence of complex numbers  $b_n$ ,*

$$\mathcal{H}(U, T, \Delta; b_n) = \int_{-\infty}^{\infty} g(t/U) \sum_{t_j} \alpha_j h(t_j) \left| \sum_n \lambda_j(n) n^{-it} b_n \right|^2 dt. \tag{4.10}$$

*Then for some  $L$  with*

$$L \ll \sqrt{P}, \tag{4.11}$$

and with coefficients

$$a_{n,l} = n^{-\frac{1}{2}} A(l, n) w_L(n), \tag{4.12}$$

where  $w_L(n)$  has support in  $n \asymp P/L^2$  and satisfies (4.1) with  $Y = P/L^2$ , we have

$$H(U, T, \Delta, w) \ll T^\varepsilon \sum_{l \asymp L} l^{-1} \mathcal{H}(U, T, \Delta; a_{n,l}) + O(T^{-200}). \tag{4.13}$$

**Proof.** We begin by writing  $\lambda_{n_j \times \phi}(n)$  in terms of  $\lambda_j(n)$  and  $A(l, n)$  using (2.3), and using Cauchy’s inequality on the sum over  $l$ . Then we break up the sum over  $l$  into  $O(\log T)$  dyadic segments with  $l \asymp L \ll \sqrt{P}$ .

Considering the value of  $L$  which maximizes the bound, this gives a bound of the form (4.13) except that the weight function is of the form  $w(l^2n)$  which unfortunately depends on  $l$ . We remove the  $l$ -dependence by multiplying  $w(l^2n)$  by  $w_1(n)$ , say, which satisfies (4.1) with  $Y = N/L^2$ , is supported in an interval of the form  $n \asymp Y$ , and is identically one on the union of the supports of  $w(l^2n)$ , for  $l \asymp L$ . Then we separate variables with the Mellin technique, writing

$$w_1(n)w(l^2n) = w_1(n) \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{w}(iy)(l^2n)^{-iy} dy.$$

By the rapid decay of  $\tilde{w}$ , we may truncate the integral at  $|y| \leq T^\varepsilon \leq U$  with an acceptable error. Then we apply Cauchy–Schwarz to take the  $y$ -integral to the outside and change variables  $t \rightarrow t - y$ . By positivity, we extend the  $t$ -integral to  $|t| \leq 2U$ , and integrate trivially over  $y$ . Finally, by positivity we attach the weight function  $g(t/U)$  and extend the integral to  $\mathbb{R}$ .  $\square$

Set

$$N = P/L^2, \tag{4.14}$$

so that  $w_L$  satisfies (4.1) with  $Y = N$ .

With the above reductions, our goal for the rest of this paper is to prove the following.

**Theorem 4.6.** *With  $a_{n,l}$  given by (4.12), we have*

$$\sum_{l \asymp L} l^{-1} \mathcal{H}(U, T, \Delta; a_{n,l}) \ll T^{3+\varepsilon} \sum_{l^2n \ll T^{3+\varepsilon}} \frac{|A(l, n)|^2}{ln}. \tag{4.15}$$

We now briefly explain how Theorem 4.6 implies Theorem 4.3 (and hence Theorem 1.1).

Using the polynomial growth of the Rankin–Selberg convolution  $L(\phi \times \phi, s)$ , one can show that (see Remark 12.1.8 of [6])

$$\sum_{l^2n \leq x} |A(l, n)|^2 \ll x.$$

Actually, if  $\phi$  is a Hecke eigenform then one can include more terms and obtain

$$\sum_{ln \leq x} |A(l, n)|^2 \ll x^{1+\varepsilon}, \tag{4.16}$$

which we explain now. The Hecke relations for  $\phi$  imply

$$A(l, n) = \sum_{d|(l,n)} \mu(d) A(l/d, 1) A(1, n/d).$$

Inserting this into the left-hand side of (4.16), applying Cauchy’s inequality to the sum over  $d$ , and using the standard divisor function bound, we quickly obtain (4.16).

One of the basic techniques used throughout this paper is to apply an asymptotic expansion to a particular quantity and reduce the estimation of the entire quantity to that of the leading-order term, as the lower-order terms have all the essential characteristics of the main term yet are of smaller magnitude.

For the rest of the paper we fix

$$\Delta = T^{1-\varepsilon}. \tag{4.17}$$

### 5. Applying the Kuznetsov formula

**Lemma 5.1.** *Let  $K_0(U, T, \Delta; b_{n,l})$  denote a sum of the form*

$$\begin{aligned} K_0 &= \frac{\Delta T}{\sqrt{N}} \int_{-\infty}^{\infty} g(t/U) \sum_{m,n} b_{m,l} \overline{b_{n,l}} \left(\frac{m}{n}\right)^{-it} \\ &\times \sum_{c \leq \frac{NT^\varepsilon}{\Delta T}} \frac{S(m, n; c)}{\sqrt{c}} e\left(\frac{-2\sqrt{mn}}{c}\right) e^{\frac{iT^2c}{2\pi\sqrt{mn}}} w_1\left(\frac{c\Delta T}{\sqrt{mn}}\right), \end{aligned} \tag{5.1}$$

where  $g$  is as in Lemma 4.5,

$$b_{n,l} = \frac{A(l, n)}{\sqrt{n}} w(n), \tag{5.2}$$

where  $w$  has support in the dyadic interval  $[N, 2N]$  and satisfies (4.1) with  $Y = N$ , and  $w_1$  is a smooth function on  $\mathbb{R}^+$  satisfying (4.1) with  $Y = 1$ . If for all  $L \ll T^{3/2+\varepsilon}$  we have

$$\sum_{l \asymp L} l^{-1} |K_0(U, T, \Delta; b_{n,l})| \ll T^{3+\varepsilon} \sum_{l^2 n \ll T^{3+\varepsilon}} \frac{|A(l, n)|^2}{ln}, \tag{5.3}$$

then Theorem 4.6 holds.

Our strategy of proof for Theorem 4.6 is thus to show (5.3) holds.

**Proof of Lemma 5.1.** Our first step in estimating  $\mathcal{H}$  is to apply Lemma 3.2. It is a somewhat involved task to analyze the integral transform  $\check{h}$ , but Jutila and Motohashi [16] have obtained a precise asymptotic expansion of  $\check{h}$  for the particular choice (4.8). By (3.19) of [16], we obtain an asymptotic expansion for  $\check{h}(x)$  with leading term

$$\frac{4}{\pi} \sqrt{\frac{2}{x}} \Delta T \exp\left(-\left(\frac{2\Delta T}{x}\right)^2\right) \cos\left(x - 2T^2x^{-1} + \frac{\pi}{4}\right). \tag{5.4}$$

Actually they wrote the expansion in terms of a critical point  $u_0$  solving  $\sinh u_0 = \frac{2T}{x}$ , which has the expansion  $u_0 = \frac{2T}{x} + O\left(\left(\frac{T}{x}\right)^3\right)$ ; in fact,  $u_0$  is holomorphic in terms of  $T/x$ . Strictly speaking, the asymptotic is of the form  $\cos\left(x - 2T^2x^{-1} + \frac{\pi}{4} + O(T^4x^{-3})\right)$ , which can be expanded into power series with leading term (5.4) provided  $\Delta \geq T^{1/3+\epsilon}$ .

Applying the Kuznetsov formula gives that

$$\mathcal{H}(U, T, \Delta; a_{n,l}) + (\text{Eisenstein}) = D + K(U, T, \Delta; a_{n,l}), \tag{5.5}$$

say, where  $D$  corresponds to the diagonal term and  $K$  is the sum of Kloosterman sums. The Eisenstein contribution is nonnegative and can be discarded for purposes of estimation of  $\mathcal{H}$ . An easy computation gives

$$D \ll U \Delta T \sum_{n \ll N} \frac{|A(l, n)|^2}{n},$$

which is sufficient for the goal of (4.15).

Inserting the asymptotic expansion for  $\check{h}(x)$ , noting that the exponential decay in (5.4) naturally allows the truncation

$$c \leq \frac{NT^\epsilon}{\Delta T}, \tag{5.6}$$

and writing  $2 \cos(y) = e^{iy} + e^{-iy}$ , we obtain an analogous asymptotic expansion for  $K$  of the form

$$K = K_1 + K_{-1} + \dots + K_r + K_{-r} + O(T^{-200}),$$

for some absolute constant  $r$ , where each  $K_i$  is of the form (5.1), and by a simple symmetry argument each  $|K_{-i}| = |K_i|$ . In fact, the lower-order terms would have a weight function  $f$  that is smaller by a certain power of  $T$ , but it only complicates the notation to include this behavior. Thus Theorem 4.6 follows from (5.3).  $\square$

### 6. Diophantine approximation

The extreme oscillation of the term  $e(\frac{-2\sqrt{mn}}{c})$  is a source of difficulty in exploiting cancellation in the sum over  $m$ . For instance, an application of the  $GL(3)$  Voronoi formula would lead to a sum where the dual variable has size  $\approx T^6$  (for “typical” choices of the parameters), which is a catastrophic loss (though there turns out to be a gain in the simplicity of the arithmetical properties of the new sum). It seems necessary to somehow dampen the oscillations of this exponential. To do so, note that the  $t$ -integral forces  $m$  and  $n$  to be close: essentially  $m = n(1 + O(U^{-1}))$ , so that the identity  $-2\sqrt{mn} = -m - n + (\sqrt{m} - \sqrt{n})^2$  implies a close approximation (note that  $(\sqrt{m} - \sqrt{n})^2 \ll (m - n)^2/N \ll NU^{-2}$ ). Although  $e(\frac{-m-n}{c})$  is just as oscillatory as  $e(\frac{-2\sqrt{mn}}{c})$  (meaning the arguments of the exponential are of the same order of magnitude), it has the property of being periodic in  $m$  and  $n$  modulo  $c$  so that one can treat this term arithmetically and absorb the “remainder”  $e(\frac{(\sqrt{m}-\sqrt{n})^2}{c})$  into the weight function, which is much less oscillatory. This is a key idea in this paper. The papers [1] and [17] also found arithmetical features of this phase but only after applying a summation formula.

Another pleasant feature of  $e(\frac{-m-n}{c})$  is that  $m$  and  $n$  are naturally separated. However, this “twisting” of the Kloosterman sum by the exponential  $e(\frac{-m-n}{c})$  has the side effect of creating terms of the form  $e(\frac{(h-1)m}{c})$  where  $h$  is coprime to  $c$ , so that it does not always hold that  $h - 1$  is coprime to  $c$ , thus making the use of the large sieve inequalities or Voronoi summation problematic. Naturally one can factor out the greatest common divisor of  $h - 1$  and  $c$  to proceed further; that is the content of the following

**Lemma 6.1.** *For all integers  $m, n$ , and positive integers  $c$ , we have*

$$S(m, n; c)e\left(\frac{-m - n}{c}\right) = \sum_{ab=c} \sum_{\substack{x \pmod{b} \\ (x(x+a), b)=1}} e\left(\frac{\bar{x}m - (\bar{x} + a)n}{b}\right). \tag{6.1}$$

**Remark.** This calculation seems to have been first performed by Luo [19], but see also [14] for some curious connections.

**Proof of Lemma 6.1.** By opening the Kloosterman sum, the left-hand side of (6.1) is

$$\sum_{\substack{h \pmod{c} \\ (h,c)=1}} e\left(\frac{(h - 1)m + (\bar{h} - 1)n}{c}\right).$$

Write  $(h - 1, c) = a, c = ab$ , and change variables  $h \equiv 1 + ax \pmod{c}$  where now  $x$  runs modulo  $b$  and satisfies  $(x(1 + ax), b) = 1$ . Note that  $1 + ax - 1 \equiv -ax(1 + \bar{a}x) \pmod{b}$ . Replacing  $x$  by  $\bar{x}$  gives (6.1).  $\square$

Inserting (6.1) into (5.1) gives

$$K_0 = \frac{\Delta T}{\sqrt{N}} \sum_{ab \leq \frac{NT\epsilon}{\Delta T}} \frac{1}{\sqrt{ab}} \sum_{\substack{x \pmod{b} \\ (x(x+a), b)=1}} \sum_{m,n} b_m \bar{b}_n e\left(\frac{\bar{x}m - (\bar{x} + a)n}{b}\right) Z(m, n), \tag{6.2}$$

where

$$Z(m, n) = e^{\frac{iT^2 ab}{2\pi\sqrt{mn}}} w_1\left(\frac{ab\Delta T}{\sqrt{mn}}\right) e\left(\frac{(\sqrt{m} - \sqrt{n})^2}{ab}\right) \int_{-\infty}^{\infty} g(t/U) \left(\frac{m}{n}\right)^{-it} dt. \tag{6.3}$$

Let

$$J = \sum_{m,n} b_m \bar{b}_n e\left(\frac{\bar{x}m - (\bar{x} + a)n}{b}\right) Z(m, n), \tag{6.4}$$

so that

$$K_0 \ll \frac{\Delta T}{\sqrt{N}} \sum_{ab \ll \frac{N}{\Delta T} T^\varepsilon} \frac{1}{\sqrt{ab}} \sum_{\substack{x \pmod{b} \\ (x(x+a), b)=1}} |J|. \tag{6.5}$$

### 7. Separation of variables

Next we want to separate the variables  $m$  and  $n$  in  $Z(m, n)$ , for which we shall use oscillatory integral transforms. This will allow us to express  $J$  in terms of a bilinear form so that we can apply the powerful technology of the large sieve inequality.

**Lemma 7.1.** *Let  $K_{00}(U, T, \Delta; b_{n,l})$  denote an expression of the form*

$$K_{00} = \Delta T^{1+\varepsilon} \sum_{ab \leq \frac{NT^\varepsilon}{\Delta T}} \frac{1}{ab} \int_{v \ll 1} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_m b_m e\left(\frac{\bar{x}m}{b}\right) m^{iy_0} e\left(\frac{v\sqrt{mN}}{Uab}\right) \right|^2 dv, \tag{7.1}$$

where  $b_m = b_{m,l}$  satisfies the condition (5.2) as in Lemma 5.1, and  $y_0 \ll T^\varepsilon$  is fixed. If

$$\sum_{l \geq L} l^{-1} K_{00}(U, T, \Delta; b_{n,l}) \ll T^{3+\varepsilon} \sum_{l^2 n \ll T^{3+\varepsilon}} \frac{|A(l, n)|^2}{ln}, \tag{7.2}$$

then Theorem 4.6 holds.

**Proof.** We find an asymptotic expansion  $J = \sum_{1 \leq i \leq r} J_i + O(T^{-200})$  such that when inserted into (6.5), gives a bound of the form, say  $|K_0| \ll \sum_i |K_{0,i}| + O(T^{-150})$ , and such that each  $|K_{0,i}|$  is of the form (7.1). In this way we reduce the estimation of  $K_0$  to that of  $K_{00}$ .

First attach a smooth, compactly-supported weight function  $w_2(\sqrt{mn}/N)$  to  $Z(m, n)$  that takes the value 1 for all  $m$  and  $n$  in the support of the implicit weight function appearing in  $b_m \bar{b}_n$ , and write  $Z(m, n) = Z_1(m, n) Z_2(m, n)$ , where

$$Z_1(m, n) = e^{\frac{iT^2 ab}{2\pi\sqrt{mn}}} w_1\left(\frac{ab\Delta T}{\sqrt{mn}}\right) w_2\left(\frac{\sqrt{mn}}{N}\right), \tag{7.3}$$

and

$$Z_2(m, n) = e\left(\frac{(\sqrt{m} - \sqrt{n})^2}{ab}\right) \int_{-\infty}^{\infty} g(t/U) \left(\frac{m}{n}\right)^{-it} dt. \tag{7.4}$$

We separate variables in each of  $Z_1$  and  $Z_2$  in turn. By Mellin inversion, we obtain

$$Z_1(m, n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_{a,b,\Delta,T,N}(iy) \left(\frac{N}{\sqrt{mn}}\right)^{iy} dy, \tag{7.5}$$

where

$$F_{a,b,\Delta,T,N}(iv) = \int_0^{\infty} w_1\left(\frac{ab\Delta T}{Nx}\right) x^{-1} w_2(x) e^{\frac{iT^2 ab}{2\pi N x}} x^{iv} dx.$$

If  $|y| \gg T^\varepsilon$  then repeated integration by parts shows  $F(iy)$  is smaller than any negative power of  $T$ . For  $|y| \ll T^\varepsilon$ , a trivial bound shows  $F(iy) \ll 1$ . This separates the variables  $m$  and  $n$  by a very short integral (essentially no cost).

Next we separate variables in

$$Z_2(m, n) = e\left(\frac{(\sqrt{m} - \sqrt{n})^2}{ab}\right) U \hat{g}\left(\frac{U}{2\pi} \log \frac{m}{n}\right). \tag{7.6}$$

Recalling that  $\hat{g}$  has compact support, which restricts  $m$  and  $n$  so that  $|m - n| \ll U^{-1}N$ , we can then use a Taylor expansion to write

$$\log \frac{m}{n} = 2 \log\left(1 + \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}\right) = 2 \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}} - \left(\frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}\right)^2 + \dots, \tag{7.7}$$

noting  $|\frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}| \ll U^{-1}$ . Then we get an asymptotic expansion for  $\hat{g}$  in the form

$$\hat{g}\left(\frac{U}{2\pi} \log \frac{m}{n}\right) = \hat{g}\left(\frac{U}{\pi} \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}\right) - \frac{1}{2} U^{-1} \left(U \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}\right)^2 \hat{g}'\left(\frac{U}{\pi} \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}\right) + \dots \tag{7.8}$$

Note that each successive term has the same form as the leading term, but is smaller by a power of  $U$ , so we shall treat the generic term in what follows.

Thus it suffices to consider

$$Z_3(m, n) = e\left(\frac{(\sqrt{m} - \sqrt{n})^2}{ab}\right) U w_3\left(U \frac{\sqrt{m} - \sqrt{n}}{\sqrt{n}}\right),$$

where  $w_3$  is compactly supported and satisfies (4.1) with  $Y = 1$ . Let  $z = \frac{\sqrt{m} - \sqrt{n}}{\sqrt{ab}}$ , and  $Z = \frac{1}{U} \sqrt{\frac{n}{ab}}$ , so that



$$Z_3(m, n) = U e(z^2) w_3\left(\frac{z}{Z}\right).$$

Suppose  $Z \gg T^\varepsilon$  (in our application this is always satisfied since  $ab \ll (\Delta T)^{-1+\varepsilon} N$ ,  $\Delta = T^{1-\varepsilon}$ , and  $U \leq \Delta$ ). Then by Fourier inversion,

$$Z_3(m, n) = \int_{-\infty}^{\infty} Y(v) e\left(\frac{v(\sqrt{m} - \sqrt{n})}{\sqrt{ab}}\right) dv, \quad Y(v) = U \int_{-\infty}^{\infty} e(z^2 - vz) w_3\left(\frac{z}{Z}\right) dz.$$

With inspiration here from pp. 431–432 of [23], we want to apply the Plancherel formula. The function  $e(z^2)$  is not  $L^2$  so instead we argue directly by using  $e(z^2) = \frac{e^{\pi i/4}}{\sqrt{2}} \int_{-\infty}^{\infty} e(tz - \frac{t^2}{4}) dt$  (which can be checked directly, the integral seen to converge uniformly on integration by parts) and reversing the order of integration. Thus

$$Y(v) = \frac{e^{\pi i/4}}{\sqrt{2}} U Z \int_{-\infty}^{\infty} e\left(-\left(\frac{v+y}{2}\right)^2\right) \widehat{w}_3(-yZ) dy.$$

Expanding the square, we get

$$Y(v) = \sqrt{2} e^{\pi i/4} U Z e\left(\frac{-v^2}{4}\right) \int_{-\infty}^{\infty} e(-y^2 - vy) \widehat{w}_3(-2yZ) dy.$$

Next truncate the integral at  $T^{-\varepsilon}$  (with negligible error), expand  $e(-y^2)$  into a Taylor series taking  $O(1/\varepsilon)$  terms so that the remainder is  $O(T^{-2009})$ , and then extend the integral back to  $\mathbb{R}$ . This gives an asymptotic expansion for  $Y(v)$ , with leading-order term say  $Y_1$  given by

$$Y_1(v) = \sqrt{2} e^{\pi i/4} U Z e\left(\frac{-v^2}{4}\right) \int_{-\infty}^{\infty} e(-vy) \widehat{w}_4(-2yZ) dy = U \frac{e^{\pi i/4}}{\sqrt{2}} e\left(\frac{-v^2}{4}\right) w_3\left(\frac{v}{2Z}\right).$$

The lower-order terms are similar but multiplied by powers of  $Z^{-1}$ , and with derivatives of  $w_3$  replacing  $w_3$ , which of course are also compactly supported and satisfy (4.1) with  $Y = 1$ . Thus we conclude that

$$Z_2(m, n) \sim \frac{e^{\pi i/4}}{\sqrt{2}} \int_{-\infty}^{\infty} e\left(\frac{-v^2}{4U^2}\right) w_3\left(\frac{v}{2UZ}\right) e\left(\frac{v(\sqrt{m} - \sqrt{n})}{U\sqrt{ab}}\right) dv. \tag{7.9}$$

Combining (7.5) and (7.9), we get an asymptotic expansion for  $J$  with leading-order term  $J_1$  of the form

$$\begin{aligned}
 |J_1| &\ll \int_{-T^\varepsilon}^{T^\varepsilon} \int_{|v| \ll \sqrt{\frac{N}{ab}}} \left| \sum_m b_m e\left(\frac{\bar{x}m}{b}\right) m^{-iy/2} e\left(\frac{v\sqrt{m}}{U\sqrt{ab}}\right) \right| \\
 &\times \left| \sum_n b_n e\left(\frac{(\bar{x}+a)n}{b}\right) n^{iy/2} e\left(\frac{v\sqrt{n}}{U\sqrt{ab}}\right) w_3\left(\frac{v}{\sqrt{\frac{n}{ab}}}\right) \right| dv dy. \tag{7.10}
 \end{aligned}$$

Using the simple inequality  $|A|B| \leq \frac{1}{2}|A|^2 + \frac{1}{2}|B|^2$ , we obtain a bound for  $|J_1|$  with two similar terms, one involving  $w_3$ , the other without  $w_3$ . Consider the term with  $w_3$ , the other term being similar yet easier. Inserting this bound into (6.5), we get a term of the form

$$\begin{aligned}
 &\frac{\Delta T}{\sqrt{N}} \sum_{ab \leq \frac{NT^\varepsilon}{\Delta T}} \frac{1}{\sqrt{ab}} \int_{-T^\varepsilon}^{T^\varepsilon} \int_{-\infty}^{\infty} \sum_{\substack{x \pmod{b} \\ (x(x+a), b)=1}} \left| \sum_n b_n e\left(\frac{(\bar{x}+a)n}{b}\right) n^{iy/2} \right. \\
 &\times \left. e\left(\frac{v\sqrt{n}}{U\sqrt{ab}}\right) w_3\left(\frac{v}{\sqrt{\frac{n}{ab}}}\right) \right|^2 dv dy.
 \end{aligned}$$

Now we simply drop the condition  $(x, b) = 1$  by positivity, change variables  $x \rightarrow x - a$  and change variables  $v \rightarrow \frac{\sqrt{N}}{\sqrt{ab}}v$ . In this way we arrive at a term of the form

$$\Delta T \sum_{ab \leq \frac{NT^\varepsilon}{\Delta T}} \frac{1}{ab} \int_{-T^\varepsilon}^{T^\varepsilon} \int_{|v| \ll 1} \sum_{\substack{x \pmod{b} \\ (x, b)=1}} \left| \sum_n b_n e\left(\frac{\bar{x}n}{b}\right) n^{iy/2} e\left(\frac{v\sqrt{nN}}{Uab}\right) w_3(v\sqrt{N/n}) \right|^2 dv dy,$$

which after summing over  $l$  and bounding the  $y$ -integral by its length times the supremum over  $y$  would give the desired form for Lemma 7.1, except for the presence of  $w_3(v\sqrt{\frac{N}{n}})$ . However, we can remove this dependence by a simple Mellin inversion similarly to how we handled  $Z_1(m, n)$ , so we omit the details (recall that the support on  $b_n$  implicitly has  $n \asymp N$ ). The lower-order terms involving powers of  $Z^{-1}$  can also be seen to have the same form.  $\square$

We shall use different methods of estimation depending on the sizes of  $a$  and  $b$ . Let  $K_{A,B} = K_{A,B}(U, T, \Delta, a_{n,l})$  denote the same sum as  $K_{00}$  given by (7.1) but with  $A < a \leq 2A$  and  $B < b \leq 2B$ , so that

$$K_{0,0}(U, T, \Delta; a_{n,l}) \ll \log^2 T \sup_{AB \leq \frac{NT^\varepsilon}{\Delta T}} K_{A,B}(U, T, \Delta, a_{n,l}), \tag{7.11}$$

where (5.6) translates to give the condition

$$AB \ll \frac{NT^\varepsilon}{\Delta T}. \tag{7.12}$$

By changing variables  $v \rightarrow \frac{ab}{AB}v$  and summing trivially over  $a$ , note that

$$K_{A,B} \ll T^\varepsilon \frac{\Delta T}{B} \int_{v \ll 1} \sum_{b > B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_m b_m e\left(\frac{\bar{x}m}{b}\right) m^{iy_0} e\left(\frac{v\sqrt{mN}}{UAB}\right) \right|^2 dv. \tag{7.13}$$

We can immediately apply Lemma 3.1, noting as well that  $|b_m|^2 \ll |a_m|^2$ , to get

**Lemma 7.2.** *We have*

$$K_{A,B}(U, T, \Delta; a_n) \ll T^\varepsilon \Delta T (B + UA) \sum_{n \leq N} |a_n|^2 \ll T^\varepsilon \left( \frac{T^3}{A} + T^2 UA \right) \sum_{n \leq N} |a_n|^2. \tag{7.14}$$

This estimate is sufficient for (7.2) for  $A$  small ( $A \ll U^{-1} T^{1+\varepsilon}$ ), so we henceforth assume

$$A \geq T^\varepsilon, \tag{7.15}$$

which simplifies some later work. Also, notice that we used no special properties of the coefficients  $a_n$  so far.

**8. Voronoi summation**

In order to improve on Lemma 7.2 we resort to use special properties of the coefficients  $a_n$ . Our tool is the  $GL(3)$  Voronoi summation formula proved by [20]. We will state this important formula in a form developed by X. Li [17].

**Theorem 8.1** (Miller–Schmid). *Let  $\psi$  be a smooth function with compact support on the positive reals. Then*

$$\begin{aligned} \sum_n A(l, n) e\left(\frac{n\bar{x}}{b}\right) \psi(n) &= \frac{b\pi^{-\frac{5}{2}}}{4i} \sum_{n_1|bl} \sum_{n_2>0} \frac{A(n_2, n_1)}{n_1 n_2} S\left(lx, n_2; \frac{bl}{n_1}\right) \Psi_1\left(\frac{n_2 n_1^2}{b^3 l}\right) \\ &+ \frac{b\pi^{-\frac{5}{2}}}{4i} \sum_{n_1|bl} \sum_{n_2>0} \frac{A(n_2, n_1)}{n_1 n_2} S\left(lx, -n_2; \frac{bl}{n_1}\right) \Psi_2\left(\frac{n_2 n_1^2}{b^3 l}\right), \end{aligned} \tag{8.1}$$

for certain integral transforms  $\Psi_1$  and  $\Psi_2$ .

We need an explicit asymptotic expansion of  $\Psi_1$  and  $\Psi_2$ , which is provided by Lemma 6.1 of [18] (generalizing Lemma 3 of Ivić [10]). Each of  $\Psi_1$  and  $\Psi_2$  is a linear combination of two other functions  $\Psi_0(x)$  and  $x^{-1}\Psi_{0,0}(x)$ , say, where each has similar asymptotic behavior, so it suffices to treat  $\Psi_0(x)$ .

**Lemma 8.2** (Ivić, Li). *Suppose  $\psi(r)$  is supported on  $[N, 2N]$ . Then there exist constants  $c_{j,\pm}$  such that*

$$\Psi_0(x) = \sum_{j=1}^L \sum_{\pm} c_{j,\pm} x \int_0^\infty \psi(r) e(\pm 3(xr)^{1/3}) \frac{dr}{(xr)^{j/3}} + O((xN)^{-\frac{L+2}{3}}). \tag{8.2}$$

An easy contour shift argument shows that  $\Psi_0(x)$  has rapid decay for  $xN \rightarrow \infty$ , and is bounded for  $xN \ll 1$  (see the original expression (6.12) of [18]).

Write (7.13) in the form

$$K_{A,B} \ll \frac{\Delta T^{1+\varepsilon}}{B} \int \sum_{v \ll 1} \sum_{\substack{b \asymp B \\ x \pmod{b} \\ (x,b)=1}} |V(v; x, b)|^2 dv,$$

where

$$V(v; x, b) = N^{-\frac{1}{2}} \sum_m A(l, m) e\left(\frac{xm}{b}\right) m^{iy_0} e\left(\frac{v\sqrt{mN}}{UAB}\right) w(m), \tag{8.3}$$

where  $w$  is a function as in Lemma 4.4; notice that we wrote  $n^{-\frac{1}{2}} = N^{-\frac{1}{2}} \left(\frac{N}{n}\right)^{\frac{1}{2}}$ , and absorbed the latter term into the weight function  $w$ .

Applying the Voronoi summation formula to  $V(v; x, b)$  with  $\psi(r) = r^{iy_0} e\left(\frac{v\sqrt{rN}}{UAB}\right) w(r)$ , we obtain  $V(v; x, b) = V_1(v; x, b) + V_2(v; x, b)$ , say, corresponding to the two terms on the right-hand side of (8.1). We accordingly write  $K_{A,B} \ll K_{A,B}^+ + K_{A,B}^-$ . Changing variables by  $x \rightarrow -x$  shows that  $K_{A,B}^-$  is of a form similar to that of  $K_{A,B}^+$ , so we shall henceforth only treat  $K_{A,B}^+$ .

First we claim we may assume  $xN \gg T^\varepsilon$ . Otherwise, using  $b \ll \frac{NT^\varepsilon}{AT^2}$ , then

$$n_1^2 n_2 \ll \frac{b^3 l}{N} \ll T^\varepsilon \frac{N^2 l}{A^3 T^6}.$$

Recalling that  $N \ll T^{3+\varepsilon}/L^2$ , we get that  $n_1^2 n_2 \ll \frac{T^\varepsilon}{L^3 A^3}$ . By (7.15), this condition is never satisfied.

Since  $xN \gg T^\varepsilon$ , Lemma 8.2 gives an asymptotic expansion of  $V_1$ . As usual, we treat the leading-order term, say  $K_{A,B}^0$ , which takes the form

$$K_{A,B}^0 = \frac{\Delta T^{1+\varepsilon}}{B} \int \sum_{v \ll 1} \sum_{\substack{b \asymp B \\ x \pmod{b} \\ (x,b)=1}} \left| \sum_{n_1 | bl} \sum_{n_2 > 0} \frac{A(n_2, n_1)}{n_1 n_2} S\left(lx, n_2; \frac{bl}{n_1}\right) \Phi\left(\frac{n_2 n_1^2}{b^3 l}\right) \right|^2 dv, \tag{8.4}$$

where  $\Phi$  is a function of the form

$$\Phi(\lambda) = \frac{b}{\sqrt{N}} \lambda \int_0^\infty w(r) r^{iy_0} e\left(-3(\lambda r)^{1/3} + \frac{u\sqrt{rN}}{UAB}\right) \frac{dr}{(\lambda r)^{1/3}},$$

where  $w$  is smooth, supported on  $[N, 2N]$ , satisfying (4.1) with  $Y = N$ , and  $y_0 \ll T^\varepsilon$ , possibly after changing variables  $v \rightarrow -v$  or  $y_0 \rightarrow -y_0$ . The change of variables  $r \rightarrow Nr$  gives, with  $w_N(r) = w(Nr)$ ,

$$\Phi(\lambda) = \frac{b}{\sqrt{N}} (\lambda N)^{2/3} N^{iy_0} \int_0^\infty w_N(r) r^{-\frac{1}{3} + iy_0} e\left(-3(\lambda r N)^{1/3} + \frac{v\sqrt{rN}}{UAB}\right) dr. \tag{8.5}$$

Our plan now is to express  $K_{A,B}^0$  into a form where we can apply Lemma 3.1; the  $v$ -integral is critical for an extra saving effect, and as such it is important to understand the phase of the integral transform  $\Phi$  (not just its magnitude).

**9. Asymptotic behavior of  $\Phi(x)$**

Under the assumption  $xN \gg T^\epsilon$  (with  $y_0 \ll T^{\epsilon/100}$ , say), the first term in the exponential in (8.5) dominates over the phase of  $y^{iv_0}$ . Unless the two terms in the exponential are of the same order of magnitude and of opposite signs (in particular,  $v$  must be positive), then an easy integration by parts argument shows that  $\Phi(\lambda)$  is negligible (smaller than  $T^{-2009}$ ). That is,  $\Phi(\lambda)$  is small unless

$$\lambda \asymp \frac{v^3 N^2}{(UAB)^3}, \tag{9.1}$$

with certain absolute implied constants. Now suppose (9.1) holds.

We shall treat general integrals of the form

$$I = \int_0^\infty f(y)e(\alpha y^{1/2} - \beta y^{1/3}) dy,$$

where  $\alpha, \beta > 0$ ,  $\alpha \asymp \beta$ , and  $f$  satisfies

$$f \text{ is smooth of compact support on } \mathbb{R}^+, \text{ satisfying } f^{(j)}(y) \ll T_0^j, \tag{9.2}$$

for some parameter  $1 \leq T_0 \ll |\alpha|^{1/100}$ . The stationary phase method easily gives the main term for  $I$ , but a search of the literature did not find an adequate asymptotic expansion. In this section we show that  $I$  has an asymptotic expansion (as  $\alpha \rightarrow \infty$ ) with leading term equal to

$$I \sim \frac{6(\frac{2\beta}{3\alpha})^5}{(2\beta)^{\frac{1}{2}}} e\left(\frac{-4\beta^3}{27\alpha^2} + \frac{1}{8}\right) f\left(\left(\frac{2\beta}{3\alpha}\right)^6\right), \tag{9.3}$$

and where the lower-order terms have the same phase, but are smaller by powers of  $\alpha$ .

Applying (9.3) to  $\Phi$ , we find that

$$\Phi(\lambda) \sim b\sqrt{\lambda}h\left(\frac{(UAB)^3\lambda}{v^3N^2}\right)e\left(\frac{-4\lambda(UAB)^2}{v^2N}\right)\left(\frac{\lambda^2}{v^6}\right)^{iy_0} z(N, U, A, B),$$

where  $h$  is a smooth function of compact support on  $\mathbb{R}^+$ , satisfying (4.1) with  $Y = 1$ , and  $z(N, U, A, B)$  is some bounded function (not depending on either  $v$  or  $\lambda$ ). Noting  $b\sqrt{\lambda} = \sqrt{n_1 n_2} \frac{1}{\sqrt{\frac{bl}{n_1}}}$  and inserting this expression into (8.4), we obtain

$$\begin{aligned}
 K_{A,B}^0 &\ll \frac{\Delta T^{1+\varepsilon}}{B} \int_{v \ll 1} \sum_{b \succ B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_{n_1|bl} \sum_{n_2} \frac{A(n_2, n_1)}{\sqrt{n_1 n_2}} \frac{S(lx, n_2; \frac{bl}{n_1})}{\sqrt{bl/n_1}} \right. \\
 &\quad \left. \times h\left(\frac{(UAB)^3 n_2 n_1^2}{v^3 b^3 l N^2}\right) e\left(\frac{-4n_2 n_1^2 (UAB)^2}{v^2 b^3 l N}\right) (n_1 n_2^2)^{2iy_0} \right|^2 dv. \tag{9.4}
 \end{aligned}$$

We will continue with this expression in the following section.

**Proof of (9.3).** Our goal is to use known properties of the Airy function, using ideas similar to those appearing in Section 7. First apply the change of variables  $y = t^6$  to get

$$I = \int_0^\infty h(t) e(\alpha t^3 - \beta t^2) dt,$$

where  $h(t) = f(t^6)(6t^5)$  satisfies (9.2). We will show

$$I \sim \frac{1}{(2\beta)^{\frac{1}{2}}} e\left(\frac{-4\beta^3}{27\alpha^2} + \frac{1}{8}\right) h\left(\frac{2\beta}{3\alpha}\right), \tag{9.5}$$

which immediately implies (9.3).

Next change variables  $t \rightarrow t + \beta/(3\alpha)$  to get

$$I = e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\infty}^\infty h\left(t + \frac{\beta}{3\alpha}\right) e\left(\alpha t^3 - \frac{\beta^2}{3\alpha} t\right) dt.$$

Now use Fourier inversion on  $h_{\alpha,\beta}(t) = h(t + \frac{\beta}{3\alpha})$  (again,  $h_{\alpha,\beta}$  satisfies (9.2) except its support may include negative reals) and reverse the orders of integration (justified by uniform convergence following from integration by parts) to get

$$I = e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\infty}^\infty \widehat{h_{\alpha,\beta}}(-t) \int_{-\infty}^\infty e\left(\alpha y^3 - \left(\frac{\beta^2}{3\alpha} + t\right)y\right) dy dt.$$

From a change of variables and some simple symmetry arguments, note that

$$\int_{-\infty}^\infty e\left(\alpha y^3 - \left(\frac{\beta^2}{3\alpha} + t\right)y\right) dy = \frac{2}{(6\pi\alpha)^{1/3}} \int_0^\infty \cos\left(\frac{1}{3}y^3 - 2\pi \frac{\frac{\beta^2}{3\alpha} + t}{(6\pi\alpha)^{1/3}} y\right) dy,$$

which using the definition of the Airy function  $\text{Ai}(x)$  can be expressed as

$$\frac{2\pi}{(6\pi\alpha)^{1/3}} \text{Ai}\left(-2\pi \frac{\frac{\beta^2}{3\alpha} + t}{(6\pi\alpha)^{1/3}}\right).$$

In terms of  $I$ , we thus have

$$I = \frac{2\pi}{(6\pi\alpha)^{1/3}} e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\infty}^{\infty} \widehat{h_{\alpha,\beta}}(-t) \operatorname{Ai}\left(-2\pi \frac{\beta^2}{(6\pi\alpha)^{1/3}} + t\right) dt.$$

Next change variables by  $t \rightarrow \frac{\beta^2}{3\alpha}t$  to get

$$I = \frac{2\pi}{(6\pi\alpha)^{1/3}} e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\infty}^{\infty} \frac{\beta^2}{3\alpha} \widehat{h_{\alpha,\beta}}\left(-\frac{\beta^2}{3\alpha}t\right) \operatorname{Ai}\left(-2\pi \frac{\beta^2}{(6\pi\alpha)^{1/3}}(1+t)\right) dt.$$

We insert the asymptotic expansion for the Airy function at large negative argument (see (4.07) of [21]), namely

$$\operatorname{Ai}(-x) \sim \frac{1}{\sqrt{\pi}x^{1/4}} \left[ \cos\left(\frac{2}{3}x^{3/2} - \frac{\pi}{4}\right) \sum_k \frac{c_{2k}}{x^{3k}} + \sin\left(\frac{2}{3}x^{3/2} - \frac{\pi}{4}\right) \sum_k \frac{c_{2k+1}}{x^{\frac{3}{2}(2k+1)}} \right],$$

for certain explicit constants  $c_k$  (in particular,  $c_0 = 1$ ). To justify this, we note that integration by parts shows  $\widehat{h_{\alpha,\beta}}(y) \ll (y^{-1}T_0)^j$  where the implied constant depends on the support of  $h$ . Thus we may truncate the integral at  $t \ll \alpha^{-2/3}$ , say, with a negligible error of size any power of  $\alpha^{-1}$ . We then have

$$I \sim \frac{2\sqrt{\pi}}{(6\pi\alpha)^{1/3}} e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\alpha^{-2/3}}^{\alpha^{-2/3}} \frac{\beta^2}{3\alpha} \widehat{h_{\alpha,\beta}}\left(-\frac{\beta^2}{3\alpha}t\right) \frac{\cos\left(\frac{2}{3}\left(\frac{2\pi\beta^2}{(6\pi\alpha)^{1/3}}(1+t)\right)^{3/2} - \frac{\pi}{4}\right)}{\left(\frac{2\pi\beta^2}{(6\pi\alpha)^{1/3}}(1+t)\right)^{1/4}} dt,$$

where the lower-order terms have a similar shape but are multiplied by powers of  $\frac{\alpha^2}{\beta^3}(1+t)^{-3/2} \asymp \alpha^{-1}(1+t)^{-3/2}$ . The terms with  $\cos$  replaced by  $\sin$  are treated similarly, so we work with  $\cos$  only. This expression simplifies as

$$I \sim \frac{2^{1/2}}{\beta^{1/2}} e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\infty}^{\infty} \frac{\beta^2}{3\alpha} \widehat{h_{\alpha,\beta}}\left(-\frac{\beta^2}{3\alpha}t\right) (1+t)^{-1/4} \cos\left(\frac{2}{3}\left(\frac{2\pi\beta^2}{(6\pi\alpha)^{1/3}}(1+t)\right)^{3/2} - \frac{\pi}{4}\right) dt.$$

We expand  $(1+t)^{-1/4}$  into a Taylor series, developing the expansion further.

Now write  $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ , and write  $I \sim I_+ + I_-$  correspondingly. Thus

$$I_{\pm} \sim \frac{1}{(2\beta)^{1/2}} e\left(\frac{-2\beta^3}{27\alpha^2}\right) \int_{-\alpha^{-2/3}}^{\alpha^{-2/3}} \frac{\beta^2}{3\alpha} \widehat{h_{\alpha,\beta}}\left(-\frac{\beta^2}{3\alpha}t\right) e\left(\pm\left(\frac{2\beta^3}{27\alpha^2}(1+t)^{\frac{3}{2}} - \frac{1}{8}\right)\right) dt.$$

Next we take a Taylor series for  $(1+t)^{3/2} = 1 + \frac{3}{2}t + \dots$  in the exponential. The quadratic and higher terms are small (much less than 1) so we take a Taylor series expansion for the exponential of these terms, giving another asymptotic expansion with leading-order term

$$I_{\pm} \sim \frac{e(\frac{\mp 1}{8})}{(2\beta)^{\frac{1}{2}}} e\left(\frac{-2\beta^3(1 \mp 1)}{27\alpha^2}\right) \int_{-\alpha^{-2/3}}^{\alpha^{-2/3}} \frac{\beta^2}{3\alpha} \widehat{h_{\alpha,\beta}}\left(-\frac{\beta^2}{3\alpha}t\right) e\left(\pm\left(\frac{\beta^3}{9\alpha^2}t\right)\right) dt.$$

Extending the integral back to  $\mathbb{R}$  and changing variables back via  $t \rightarrow \frac{3\alpha}{\beta^2}t$  gives

$$I_{\pm} \sim \frac{e(\frac{\mp 1}{8})}{(2\beta)^{\frac{1}{2}}} e\left(\frac{-2\beta^3(1 \mp 1)}{27\alpha^2}\right) \int_{-\infty}^{\infty} \widehat{h_{\alpha,\beta}}(-t) e\left(\pm\left(\frac{\beta}{3\alpha}t\right)\right) dt.$$

Calculating the integral in terms of  $h$ , we get

$$I_{\pm} \sim \frac{e(\frac{\mp 1}{8})}{(2\beta)^{\frac{1}{2}}} e\left(\frac{-2\beta^3(1 \mp 1)}{27\alpha^2}\right) h_{\alpha,\beta}\left(\mp \frac{\beta}{3\alpha}\right).$$

Note  $h_{\alpha,\beta}(\mp \frac{\beta}{3\alpha}) = h(\frac{\beta}{3\alpha}(1 \mp 1))$ . Since  $h$  has support on the positive reals, we have  $h(0) = 0$ , in which case only  $I_-$  contributes to  $I$ , so then

$$I \sim \frac{1}{(2\beta)^{\frac{1}{2}}} e\left(\frac{-4\beta^3}{27\alpha^2} + \frac{1}{8}\right) h\left(\frac{2\beta}{3\alpha}\right).$$

One easily checks that this is the expected main term one obtains from stationary phase.  $\square$

### 10. Cleaning

We have reduced the problem of estimating  $K_{A,B}$  (originally given by (7.13)) to estimating the rather messy expression  $K_{A,B}^0$  given by (9.4), in the sense that any bound for  $K_{A,B}^0$  with a general compactly-supported weight function  $h$  satisfying (4.1) with  $Y = 1$ , is also a bound on  $K_{A,B}$  (plus a negligible error term of size, say  $O(T^{-100})$  which shall be dwarfed by our upper bound on  $K_{A,B}$ ).

We shall make some preliminary transformations to clean up this expression for  $K_{A,B}^0$ . The reader interested in the essential details should consider the crucial case  $v \asymp 1, l = n_1 = 1$ , which greatly simplifies the forthcoming calculations.

**Lemma 10.1.** *Suppose that  $K_{A,B}^0$  is any expression of the form (9.4) with  $h$  satisfying (4.1) with  $Y = 1$ , and let  $\mathcal{L} = \mathcal{L}(A, B, L, T, U, \Delta, N_1, N_2)$  be an expression of the form*

$$\mathcal{L} = \frac{\Delta T}{B} \frac{W'}{L} \sum_{l \asymp L} \sum_{b \asymp B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \sum_{\substack{n_1 | bl \\ n_1 \asymp N_1}} \int_{v \asymp 1} \left| \sum_{n_2 \asymp N_2} \frac{b(n_1, n_2)}{\sqrt{n_1 n_2}} \frac{S(lx, n_2; \frac{bl}{n_1})}{\sqrt{bl/n_1}} e\left(\frac{vn_2}{W}\right) \right|^2 dv, \tag{10.1}$$

where  $b(n_1, n_2)$  are complex numbers satisfying  $|b(n_1, n_2)| \ll |A(n_1, n_2)|$ , and where

$$W = \frac{N_2^{2/3} B L^{1/3}}{N_1^{2/3} N^{1/3}}, \quad W' = \frac{U A N_2^{1/3} N_1^{2/3}}{L^{1/3} N^{2/3}}. \tag{10.2}$$



Then for some such choice of  $b(n_1, n_2)$ , we have

$$\sum_{l \asymp L} l^{-1} K_{A,B}^0 \ll T^\varepsilon \sup_{A,B,N_1,N_2} \mathcal{L}, \tag{10.3}$$

where recall (7.12), and where

$$N_1^2 N_2 \ll \frac{LN^2}{(UA)^3}, \tag{10.4}$$

**Proof.** We begin with the representation (9.4) and apply Cauchy’s inequality to take the sum over  $n_1|bl$  outside the absolute values, and divide up the sum over  $n_2$  into dyadic intervals  $n_2 \asymp N_2$ , where in view of (9.1),

$$N_2 \ll \frac{lN^2}{n_1^2(UA)^3}. \tag{10.5}$$

With the shorthand notation

$$K' = \int_{v \ll 1} \left| \sum_{n_2 \asymp N_2} \frac{a(n_1, n_2)}{\sqrt{n_1 n_2}} \frac{S(lx, n_2; \frac{bl}{n_1})}{\sqrt{bl/n_1}} h\left(\frac{(UAB)^3 n_2 n_1^2}{v^3 b^3 l N^2}\right) e\left(\frac{4n_2 n_1^2 (UAB)^2}{v^2 b^3 l N}\right) \right|^2 dv, \tag{10.6}$$

where  $a(n_1, n_2)$  are certain complex numbers with the same absolute value as  $A(n_2, n_1)$ , we have

$$K_{A,B}^0 \ll T^\varepsilon \frac{\Delta T}{B} \sum_{b \asymp B} \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \sum_{n_1|bl} K'. \tag{10.7}$$

Next, locate  $n_1 \asymp N_1$ , so that (10.4) holds, and recall  $l \asymp L$ . For such  $l$  and  $n_1$ , change variables in  $v$  via

$$v \rightarrow \frac{UAB^{3/2} N_2^{1/3} n_1 L^{1/6}}{b^{3/2} l^{1/2} N^{2/3} N_1^{1/3}} v^{-1/2},$$

to get

$$K' \ll W' \int_{v \asymp 1} \left| \sum_{n_2 \asymp N_2} \frac{a(n_1, n_2)}{\sqrt{n_1 n_2}} \frac{S(lx, n_2; \frac{bl}{n_1})}{\sqrt{bl/n_1}} h\left(\left(\frac{v^3 b^3 l n_2^2 N_1^2}{B^3 L N_2^2 n_1^2}\right)^{1/2}\right) e\left(\frac{vn_2}{W}\right) \right|^2 dv,$$

where  $W$  and  $W'$  are given by (10.2), the restriction to  $v \asymp 1$  is redundant to the support of  $h$ , and where we use positivity and the location of the variables to write  $W'$  in terms of capital letters rather than lowercase letters.

Next separate the variables  $v, b, l, n_1$ , and  $n_2$  in  $h\left(\left(\frac{v^3 b^3 l n_2^2 N_1^2}{B^3 L N_2^2 n_1^2}\right)^{1/2}\right)$  by using the Mellin inversion formula; since  $v$  is already located to be  $\asymp 1$ , and  $b \asymp B, l \asymp L$ , and  $n_1 \asymp N_1, n_2 \asymp N_2$ .

This can be done with a Mellin integral of length  $O(T^\varepsilon)$  which can be taken to the outside of  $\sum_l l^{-l} K_{A,B}^0$  (after an application of Cauchy–Schwarz) and then bounded by its length times the supremum. Effectively this simply changes the coefficients to say  $b(n_1, n_2)$  having the same absolute values as  $a(n_1, n_2)$ .  $\square$

**11. Final step: applying the large sieve**

**Lemma 11.1.** *Suppose that  $\mathcal{L}$  is given by (10.1) as in Lemma 10.1. Then*

$$\mathcal{L} \ll T^{3+\varepsilon} \sum_{ln \ll T^{3+\varepsilon}} \frac{|b(l, n)|^2}{ln}. \tag{11.1}$$

Lemma 11.1 shall complete the proof of Theorem 4.6 by the reductions of Lemmas 5.1, 7.1, and 10.1, combined with the bound of Lemma 7.2 for  $A \leq T^\varepsilon$ .

**Proof of Lemma 11.1.** Let  $d = (b, n_1)$  and change variables  $b \rightarrow db, n_1 \rightarrow dn_1$  to get

$$\begin{aligned} \mathcal{L} &\ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \\ &\times \sum_{l \geq L} \int_{v \geq 1} \sum_{d \ll N_1} \sum_{b \geq B/d} \sum_{\substack{x \pmod{bd} \\ (x, bd) = 1}} \sum_{\substack{n_1 | l \\ (n_1, b) = 1 \\ dn_1 \geq N_1}} \left| \sum_{n_2} \frac{b(dn_1, n_2)}{\sqrt{dn_1 n_2}} \frac{S(lx, n_2; \frac{bl}{n_1})}{\sqrt{bl/n_1}} e\left(\frac{vn_2}{W}\right) \right|^2 du. \end{aligned} \tag{11.2}$$

Write  $\frac{l}{n_1} = rs$  where  $r | b^\infty$  (meaning all primes dividing  $r$  also divide  $b$ ) and  $(s, b) = 1$ . Next change variables to eliminate  $l$  and note that the sum over  $x$  only depends modulo  $b$  to give

$$\mathcal{L} \ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \int_{v \geq 1} \sum_{d \ll N_1} d \sum_{b \geq \frac{B}{d}} \sum_{\substack{n_1 r s \leq L \\ (n_1 s, b) = 1 \\ dn_1 \geq N_1}} \sum_{r | b^\infty} \frac{1}{brs} \mathcal{L}_1, \tag{11.3}$$

where as shorthand

$$\mathcal{L}_1 = \sum_{\substack{x \pmod{b} \\ (x, b) = 1}} \left| \sum_{n_2} \frac{b(dn_1, n_2)}{\sqrt{dn_1 n_2}} S(n_1 r s x, n_2; brs) e\left(\frac{vn_2}{W}\right) \right|^2. \tag{11.4}$$

Next we simplify  $\mathcal{L}_1$ , by showing

$$\mathcal{L}_1 \leq br^2 s \sum_{h \pmod{bs}}^* \left| \sum_{n_2} \frac{b(dn_1, rn_2)}{\sqrt{dn_1 rn_2}} e\left(\frac{hn_2}{bs}\right) e\left(\frac{vrn_2}{W}\right) \right|^2. \tag{11.5}$$

**Proof of (11.5).** From the multiplicativity relation for Kloosterman sums, we obtain

$$S(n_1 r s x, n_2; brs) = S(n_1 r s \bar{s} x, n_2 \bar{s}, br) S(n_1 r s \bar{b} r x, n_2 \bar{b} r; s) = S(n_1 r x, n_2 \bar{s}, br) S(0, n_2; s),$$

which becomes  $S(rx, n_2; br)S(0, n_2; s)$  after the change of variables  $x \rightarrow s\bar{n}_1x$  (recall  $n_1$  is coprime to  $br$ ). Thus

$$\mathcal{L}_1 = \sum_{\substack{x \pmod{b} \\ (x,b)=1}} \left| \sum_m b_m S(rx, m; br) \right|^2, \quad \text{with } b_m = \frac{b(dn_1, m)}{\sqrt{dn_1m}} S(0, m; s) e\left(\frac{vm}{W}\right).$$

Next we compute for arbitrary complex numbers  $c_m$ ,

$$\sum_{x \pmod{b}} \left| \sum_m c_m S(rx, m; br) \right|^2 = b \sum_{m_1, m_2} c_{m_1} \bar{c}_{m_2} \sum_{\substack{h_1, h_2 \pmod{br} \\ h_1 \equiv h_2 \pmod{b}}}^* e\left(\frac{h_1 m_1 - h_2 m_2}{br}\right). \tag{11.6}$$

Change variables via  $h_i = y + bz_i, i = 1, 2$ , where  $y$  runs modulo  $b$  and  $z_i$  runs modulo  $r$ . Since  $r|b^\infty$ , the condition that  $(h_i, br) = 1$  is equivalent to  $(y, b) = 1$ . The sum over  $z_i$  vanishes unless  $r|m_i$ , in which case the sum is  $r$ . Thus (11.6) equals

$$br^2 \sum_{r|m_1, m_2} c_{m_1} \bar{c}_{m_2} \sum_{y \pmod{b}}^* e\left(\frac{y(\frac{m_1}{r} - \frac{m_2}{r})}{b}\right) = br^2 \sum_{y \pmod{b}}^* \left| \sum_{r|m} c_m e\left(\frac{y\frac{m}{r}}{b}\right) \right|^2,$$

and hence

$$\mathcal{L}_1 \leq br^2 \sum_{y \pmod{b}}^* \left| \sum_{r|m} c_m e\left(\frac{y\frac{m}{r}}{b}\right) S(0, m; s) \right|^2, \quad \text{with } c_m = \frac{a(dn_1, m)}{\sqrt{dn_1m}} e\left(\frac{vm}{W}\right).$$

By Cauchy’s inequality, we have for any complex coefficients  $b_l$  that

$$\left| \sum_l b_l S(0, l; s) \right|^2 \leq s \sum_{h \pmod{s}}^* \left| \sum_l b_l e\left(\frac{hl}{s}\right) \right|^2. \tag{11.7}$$

Note that  $S(0, l; s) = S(0, \frac{l}{r}; s)$  since  $(r, s) = 1$ . Hence

$$\mathcal{L}_1 \leq br^2 s \sum_{h \pmod{s}}^* \sum_{y \pmod{b}}^* \left| \sum_{r|m} c_m e\left(\frac{y\frac{m}{r}}{b}\right) e\left(\frac{hm}{s}\right) \right|^2,$$

which gives (11.5) using the Chinese remainder theorem and changing variables  $m \rightarrow mr$ .  $\square$

Picking back up the chain of reasoning, we insert (11.5) into (11.3), getting

$$\begin{aligned} \mathcal{L} &\ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \int_{v \geq 1} \\ &\times \sum_{d \ll N_1} d \sum_{b \geq \frac{B}{d}} \sum_{\substack{n_1 r s \leq L \\ (n_1 s, b) = 1 \\ dn_1 \asymp N_1}} \sum_{r|b^\infty} r \sum_{y \pmod{bs}}^* \left| \sum_{n_2} \frac{b(dn_1, rn_2)}{\sqrt{dn_1 rn_2}} e\left(\frac{yn_2}{bs}\right) e\left(\frac{vrn_2}{W}\right) \right|^2 dv. \tag{11.8} \end{aligned}$$

Next relax the condition that  $r|b^\infty$  and for convenience locate the variable  $s \asymp S$ , where

$$N_1 r S \asymp dL. \tag{11.9}$$

We need to bound the supremum over  $S$ , where note that we may assume  $S \ll L$ . In addition change variables  $bs \rightarrow c$ , getting

$$\begin{aligned} \mathcal{L} &\ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \\ &\times \sum_{dn_1 \asymp N_1} d \sum_{r \ll \frac{L}{n_1 S}} r \int_{v > 1} \sum_{c \asymp \frac{BS}{d} h} \sum_{h \pmod{c}}^* \left| \sum_{rn_2 \asymp N_2} \frac{b(dn_1, rn_2)}{\sqrt{dn_1 rn_2}} e\left(\frac{yn_2}{c}\right) e\left(\frac{vn_2}{W/r}\right) \right|^2 dv, \end{aligned} \tag{11.10}$$

for some value of  $S$ . Finally, an application of Lemma 3.1 gives

$$\mathcal{L} \ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \sum_{dn_1 \asymp N_1} d \sum_{r \ll \frac{L}{n_1 S}} r \left( \left(\frac{BS}{d}\right)^2 + \frac{W}{r} \right) \sum_{rn_2 \asymp N_2} \frac{|b(dn_1, rn_2)|^2}{dn_1 rn_2}.$$

We simplify this by first letting  $rn_2 \rightarrow n_2$  be a new variable, getting

$$\mathcal{L} \ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \sum_{dn_1 \asymp N_1} d \left( \frac{L}{n_1 S} \left(\frac{BS}{d}\right)^2 + W \right) \sum_{n \asymp N_2} \frac{|b(dn_1, n_2)|^2}{dn_1 n_2}.$$

Similarly, let  $dn_1 \rightarrow n_1$  be a new variable to get

$$\mathcal{L} \ll T^\varepsilon \frac{\Delta T}{B} \frac{W'}{L} \left( \frac{LB^2 S}{N_1} + WN_1 \right) \sum_{l^2 n \ll N_1^2 N_2} \frac{|b(n_1, n_2)|^2}{n_1 n_2}. \tag{11.11}$$

Note that from (10.4) we have  $W' \ll 1$  and

$$W'W = \frac{UABN_2}{N} \ll \frac{BLN}{(UA)^2 N_1^2}.$$

Thus using  $S \ll L$ , the expression (11.11) simplifies as

$$\mathcal{L} \ll T^\varepsilon \Delta T \left( \frac{BL}{N_1} + \frac{N}{N_1(UA)^2} \right) \sum_{l^2 n \ll N_1^2 N_2} \frac{|b(l, n)|^2}{ln}. \tag{11.12}$$

Recalling that  $B \ll \frac{NT^\varepsilon}{A\Delta T}$ ,  $L^2 N \ll T^{3+\varepsilon}$ , and  $\Delta = U = T^{1-\varepsilon}$ , finishes the proof.  $\square$

**Acknowledgments**

I would like to thank Brian Conrey, Adrian Diaconu, Xiaoqing Li, Peter Sarnak, and Akshay Venkatesh for interesting discussions, and the referee for a very careful reading which greatly improved the paper.

## References

- [1] J.B. Conrey, H. Iwaniec, The cubic moment of central values of automorphic  $L$ -functions, *Ann. of Math.* (2) 151 (3) (2000) 1175–1216.
- [2] J.-M. Deshouillers, H. Iwaniec, The nonvanishing of Rankin–Selberg zeta-functions at special points, in: *The Selberg Trace Formula and Related Topics*, Brunswick, Maine, 1984, in: *Contemp. Math.*, vol. 53, Amer. Math. Soc., Providence, RI, 1986, pp. 51–95.
- [3] W. Duke, H. Iwaniec, Estimates for coefficients of  $L$ -functions. II, in: *Proceedings of the Amalfi Conference on Analytic Number Theory*, Maiori, 1989, Univ. Salerno, Salerno, 1992, pp. 71–82.
- [4] P.X. Gallagher, A large sieve density estimate near  $\sigma = 1$ , *Invent. Math.* 11 (1970) 329–339.
- [5] A. Diaconu, P. Garrett, D. Goldfeld, Moments for  $L$ -functions for  $GL_r \times GL_{r-1}$ , preprint, 2009.
- [6] D. Goldfeld, *Automorphic Forms and  $L$ -Functions for the Group  $GL(n, \mathbb{R})$* , Cambridge Stud. Adv. Math., vol. 99, Cambridge Univ. Press, Cambridge, 2006, with an appendix by Kevin A. Broughan.
- [7] A. Good, The square mean of Dirichlet series associated with cusp forms, *Mathematika* 29 (2) (1982) 278–295.
- [8] D.R. Heath-Brown, The twelfth power moment of the Riemann-function, *Quart. J. Math. Oxford Ser. (2)* 29 (116) (1978) 443–462.
- [9] J. Hoffstein, P. Lockhart, Coefficients of Maass forms and the Siegel zero, *Ann. of Math.* (2) 140 (1) (1994) 161–181, with an appendix by Dorian Goldfeld, Jeffrey Hoffstein and Daniel Lieman.
- [10] A. Ivić, On the ternary additive divisor problem and the sixth moment of the zeta-function, in: *Sieve Methods, Exponential Sums, and Their Applications in Number Theory*, Cardiff, 1995, in: *London Math. Soc. Lecture Note Ser.*, vol. 237, Cambridge Univ. Press, Cambridge, 1997, pp. 205–243.
- [11] H. Iwaniec, Fourier coefficients of cusp forms and the Riemann zeta-function, in: *Seminar on Number Theory 1979–1980*, Exp. No. 18, Univ. Bordeaux I, Talence, 1980, 36 pp.
- [12] H. Iwaniec, Small eigenvalues of Laplacian for  $T_0(N)$ , *Acta Arith.* 56 (1) (1990) 65–82.
- [13] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ., vol. 53, Amer. Math. Soc., Providence, RI, 2004.
- [14] H. Iwaniec, X. Li, The orthogonality of Hecke eigenvalues, *Compos. Math.* 143 (3) (2007) 541–565.
- [15] H. Iwaniec, W. Luo, P. Sarnak, Low lying zeros of families of  $L$ -functions, *Publ. Math. Inst. Hautes Études Sci.* 91 (2000) 55–131.
- [16] M. Jutila, Y. Motohashi, Uniform bound for Hecke  $L$ -functions, *Acta Math.* 195 (2005) 61–115.
- [17] X. Li, Bounds for  $GL(3) \times GL(2)$   $L$ -functions and  $GL(3)$   $L$ -functions, *Ann. of Math.*, <http://annals.math.princeton.edu/annals/accepted/090218-Li/090218-Li-v1.pdf>, in press.
- [18] X. Li, The central value of the Rankin–Selberg  $L$ -functions, *Geom. Funct. Anal.* 18 (5) (2009) 1660–1695, doi:10.1007/s00039-008-0692-5.
- [19] W. Luo, The spectral mean value for linear forms in twisted coefficients of cusp forms, *Acta Arith.* 70 (4) (1995) 377–391.
- [20] S.D. Miller, W. Schmid, Automorphic distributions,  $L$ -functions, and Voronoi summation for  $GL(3)$ , *Ann. of Math.* (2) 164 (2) (2006) 423–488.
- [21] F. Olver, *Asymptotics and Special Functions*, Comp. Sci. Appl. Math., Academic Press, New York, London, 1974.
- [22] P. Sarnak, Fourth moments of Grössencharakteren zeta functions, *Comm. Pure Appl. Math.* 38 (2) (1985) 167–178.
- [23] P. Sarnak, Estimates for Rankin–Selberg  $L$ -functions and quantum unique ergodicity, *J. Funct. Anal.* 184 (2) (2001) 419–453.
- [24] J. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc. (N.S.)* 12 (2) (1985) 183–216.
- [25] M. Young, The second moment of  $GL(3) \times GL(2)$   $L$ -functions at special points, preprint, 2009.