Applications of the Kuznetsov formula on $GL(3)$

Valentin Blomer

Abstract We develop a fairly explicit Kuznetsov formula on $GL(3)$ and discuss the analytic behavior of the test functions on both sides. Applications to Weyl's law, exceptional eigenvalues, a large sieve and L-functions are given.

1 Introduction

The Bruggeman–Kuznetsov formula [4, 12, 27] is one of the most powerful tools in the analytic theory of automorphic forms on $GL(2)$ and the cornerstone for the investigation of moments of families of L-functions, including striking applications to subconvexity and non-vanishing. It can be viewed as a relative trace formula for the group $G = GL(2)$ and the abelian subgroup $U_2 \times U_2 \subseteq G \times G$ where U_2 is the group of unipotent upper triangular matrices. The Kuznetsov formula in the simplest case is an equality of the shape

V. Blomer (✉)
Mathematisches Institut, Bunsenstr. 3-5, 37073 Göttingen, Germany
e-mail: blomer@uni-math.gwdg.de

$$2 \sum_j \frac{\lambda_j(n)\overline{\lambda_j(m)}}{L(\mathrm{Ad}^2 u_j, 1)} h(t_j) + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sigma_t(n)\overline{\sigma_t(m)}}{|\zeta(1+2it)|^2} h(t)\, dt$$

$$= \delta_{n,m} \int_{-\infty}^{\infty} h(t)\, d_{\mathrm{spec}} t + \sum_c \frac{1}{c} S(n, m, c) h^{\pm}\left(\frac{|nm|}{c^2}\right) \qquad (1.1)$$

where

- $n, m \in \mathbb{Z} \setminus \{0\}$,
- $\delta_{n,m}$ is the Kronecker symbol,
- the sum on the left-hand side runs over an orthogonal basis of Hecke–Maaß cusp forms u_j for the group $SL_2(\mathbb{Z})$ having spectral parameter t_j and Hecke eigenvalues $\lambda_j(n)$ for $n \in \mathbb{N}$ (and $\lambda_j(-n) := \pm\lambda_j(n)$ depending on whether u_j is even or odd),
- $\sigma_t(n)$ is the Fourier coefficient of an Eisenstein series defined by

$$\sigma_t(n) = \sum_{d_1 d_2 = |n|} d_1^{it} d_2^{-it},$$

- $d_{\mathrm{spec}} t = \pi^{-2} t \tanh(\pi t)\, dt$ is the spectral measure,
- h is some sufficiently nice, even test function, and
- $h^{\pm}$ is a certain integral transform of h, the sign being $\mathrm{sgn}(nm)$, described in (1.2).

There have been many generalizations of the Kuznetsov to other groups of real rank one or products thereof, see e.g. [5, 28, 31], the first of which covers also the groups $SL_2(\mathbb{C})$, $SO(n, 1)$ and $SU(2, 1)$; see also [11, 15] for interesting applications. For the groups $GL(n)$, $n > 2$, Kuznetsov-type formulae are available [17, Theorem 11.6.19], [36], but they are in considerably less explicit form.

The power of the $GL(2)$ Kuznetsov formula lies in the fact that one can choose arbitrary (reasonable) test functions on either side of the formula, and the relevant integral transforms are completely explicit in terms of Bessel functions. In fact, we have

$$h^{\pm}(x) = \int_0^{\infty} \mathcal{J}^{\pm}(t, x) h(t)\, d_{\mathrm{spec}} t \qquad (1.2)$$

where

$$\mathcal{J}^{\pm}(t, x) = 2\pi i \sinh(\pi t)^{-1} \left\{ \begin{array}{c} J_{2it}(4\pi\sqrt{x}) - J_{-2it}(4\pi\sqrt{x}) \\ I_{2it}(4\pi\sqrt{x}) - I_{-2it}(4\pi\sqrt{x}) \end{array} \right\};$$

this is best understood in terms of its Mellin transform:

$$\widehat{\mathcal{J}}^{\pm}(t,s) = \int_0^\infty \mathcal{J}^{\pm}(t,x) x^{s-1}\, dx = \frac{G_t(s)}{G_t(1-s)} \mp \frac{G_t(1+s)}{G_t(2-s)}$$

where

$$G_t(s) = \pi^{-s} \Gamma\left(\frac{s+it}{2}\right) \Gamma\left(\frac{s-it}{2}\right).$$

In addition, this transform can be inverted and is essentially unitary:

$$h(t) \approx \int_0^\infty \mathcal{J}^{\pm}(t,x) h^{\pm}(x)\, \frac{dx}{x}.$$

There is a subtlety, as in the $+$ case the image of the map $h \mapsto h^+$ is not dense, but its complement is well-understood. These formulas together with standard facts about Bessel functions make it possible to apply the Kuznetsov formula in both directions. Unfortunately, such explicit knowledge is not available for $GL(n)$, $n \geq 3$.

The aim of this article to provide a "semi-explicit" version of the Kuznetsov formula for $GL(3)$ together with some careful analysis of the various terms occurring on both sides of the formula, and to give some applications in Theorems 1–4 below. On the way we will prove a number of useful auxiliary results for $GL(3)$ Whittaker functions, Eisenstein series and Kloosterman sums that may be helpful for further investigation of $GL(3)$ automorphic forms. The proof of the Kuznetsov formula proceeds along classical lines: we compute the inner product of two Poincaré series in two ways: by spectral decomposition and by unfolding and computing the Fourier expansion of the Poincaré series. The latter has been worked out in great detail in [8].

The *spectral side* (8.1) of the $GL(3)$ formula consists of three terms:

- the contribution of the cuspidal spectrum,
- the contribution of the minimal parabolic Eisenstein series,
- the contribution of the maximal parabolic Eisenstein series.

The *arithmetic side* (8.2) contains four terms:

- the diagonal contribution corresponding to the identity element in the Weyl group,
- two somewhat degenerate terms[1] corresponding to $\left(\begin{smallmatrix} & & 1 \\ 1 & & \\ & 1 & \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} & 1 & \\ & & 1 \\ 1 & & \end{smallmatrix}\right)$,
- the contribution of the long Weyl element.

---

[1] In all our applications they will be negligible.

Interestingly, the two remaining elements in the Weyl group do not contribute as long as n_1, n_2, m_1, m_2 are non-zero; in fact, these two furnish the $GL(2)$ formula which is hidden in the degenerate terms of the Fourier expansion of the Poincaré series. On the arithmetic side, the various variables $n_1, n_2, m_1, m_2, D_1, D_2$ appearing in the test function are explicitly coupled similarly as on the right hand side of (1.1).

The spectral side (8.1) contains the weight function

$$h(\nu_1, \nu_2) = \left| \int_0^\infty \int_0^\infty \tilde{W}_{\nu_1, \nu_2}(y_1, y_2) F(y_1, y_2) \frac{dy_1 \, dy_2}{(y_1 y_2)^3} \right|^2$$

where $\tilde{W}_{\nu_1, \nu_2}$ is a normalized Whittaker function on $GL(3)$ and F is any compactly supported test function (or with sufficiently rapid decay at 0 and ∞ in both variables). In principle this integral transform is invertible: it has been shown in [18] that a natural generalization of the Kontorovich–Lebedev transform inversion formula holds for Whittaker functions on $GL(n)$, hence we have a recipe to find a suitable F to construct our favorite non-negative function h. Proceeding in this way would however considerably complicate the analysis of the arithmetic side, and hence we take a different route which is somewhat less precise, but more convenient for applications. In Proposition 3 below we show roughly the following: taking

$$F(y_1, y_2) = \left(\tau_1 \tau_2 (\tau_1 + \tau_2) \right)^{1/2} f_1(y_1) f_2(y_2) y_1^{i(\tau_1 + 2\tau_2)} y_2^{i(2\tau_1 + \tau_2)}$$

for some fixed functions f_1, f_2 with compact support in $(0, \infty)$ yields a non-negative smooth bump function h with $h(\nu_1, \nu_2) \asymp 1$ for $\nu_j = i\tau_j + O(1)$ and rapid decay outside this range. In other words, h is a good approximation to the characteristic function of a unit square in the (ν_1, ν_2)-plane. Integration over τ_1, τ_2 can now give a good approximation to the characteristic function of any reasonable shape. Passing to a larger region in this way will in fact improve the performance of sum formula and ease the estimations on the arithmetic side.

The test functions on the arithmetic side are completely explicit in (8.3), (8.4) and given as a multiple integral. At least in principle a careful asymptotic analysis should yield a complete description of the behavior of this function, but this seems very complicated. Nevertheless, we are able to give some nontrivial (and in some cases best possible) bounds in Proposition 5 that suffice for a number of applications that we proceed to describe.

The commutative algebra $\mathcal{D}$ of invariant differential operators of $SL_3(\mathbb{R})$ acting on $L^2(SL_3(\mathbb{R})/SO_3)$ is generated by two elements (see [17, p. 153]), the Laplacian and another operator of degree 3. One class of eigenfunctions of $\mathcal{D}$ is given by the power functions I_{ν_1, ν_2} defined in (2.11) below. A Maaß form ϕ for the group $SL_3(\mathbb{Z})$ with *spectral parameters* ν_1, ν_2 is an element in

$L^2(SL_3(\mathbb{Z}) \backslash SL_3(\mathbb{R})/SO_3)$ that is an eigenfunction of $\mathcal{D}$ with the same eigenvalues as I_{ν_1, ν_2} and vanishes along all parabolics, that is,

$$\int_{(SL_3(\mathbb{Z}) \cup U) \backslash U} \phi(uz) \, du = 0$$

for $U = \left\{ \begin{pmatrix} 1 & & * \\ & 1 & * \\ & & 1 \end{pmatrix} \right\}$ and $\left\{ \begin{pmatrix} 1 & * & * \\ & 1 & \\ & & 1 \end{pmatrix} \right\}$ (and then automatically for the minimal parabolic). We choose an orthonormal basis $\{\phi_j\} = \mathcal{B} \subseteq L^2(SL_3(\mathbb{Z}) \backslash SL_3(\mathbb{R})/SO_3)$ of Hecke–Maaß cusp forms (i.e. Maaß forms that are eigenfunctions of the Hecke algebra as in [17, Sect. 6.4]) with spectral parameters $\nu_1^{(j)}, \nu_2^{(j)}$. If no confusion can arise, we drop the superscripts (j).

This can be re-phrased in more representation theoretic terms. Let $SL_3(\mathbb{R}) = NAK$ be the Iwasawa decomposition where $K = SO_3$, N is the standard unipotent subgroup and A is the group of diagonal matrices with determinant 1 and positive entries, and let $\mathfrak{a}$ be the Lie algebra of A. An infinite-dimensional, irreducible, everywhere unramified cuspidal automorphic representation π of $GL_3(\mathbb{A}_{\mathbb{Q}})$ with trivial central character is generated by a Hecke–Maaß form ϕ_j for $SL_3(\mathbb{Z})$ as above. The local (spherical) representation π_∞ is an induced representation from the parabolic subgroup NA of the extension of a character $\chi : A \to \mathbb{C}^\times$, $\mathrm{diag}(x_1, x_2, x_3) \mapsto x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$ with $\alpha_1 + \alpha_2 + \alpha_3 = 0$. In this way we can identify the spherical cuspidal automorphic spectrum with a discrete subset of the Lie algebra $\mathfrak{a}_{\mathbb{C}}^*/W$ (W the Weyl group), where we associate to each Maaß form $\phi_j \in \mathcal{B}$ the linear form $l = (\alpha_1, \alpha_2, \alpha_3) \in \mathfrak{a}_{\mathbb{C}}^*/W$ that contains the (archimedean) Langlands parameters. A convenient basis in $\mathfrak{a}_{\mathbb{C}}^*$ is given by the fundamental weights $\mathrm{diag}(2/3, -1/3, -1/3)$, $\mathrm{diag}(1/3, 1/3, -2/3)$ of SL_3. The coefficients of $l = (\alpha_1, \alpha_2, \alpha_3)$ with respect to this basis can be obtained by evaluating l at the two co-roots $\mathrm{diag}(1, -1, 0), \mathrm{diag}(0, 1, -1) \in \mathfrak{a}$ and are given by $3\nu_1, 3\nu_2$. We then have $\alpha_1 = 2\nu_1 + \nu_2, \alpha_2 = -\nu_1 + \nu_2, \alpha_3 = -\nu_1 - 2\nu_2$. With this normalization, ϕ is an eigenform of the Laplacian with eigenvalue

$$\lambda = 1 - 3\nu_1^2 - 3\nu_1\nu_2 - 3\nu_2^2 = 1 - \frac{1}{2}(\alpha_1^2 + \alpha_2^2 + \alpha_3^2), \qquad (1.3)$$

and the trivial representation is sitting at $(\nu_1, \nu_2) = (1/3, 1/3)$. The Ramanujan conjecture states that the Langlands parameters $\alpha_1, \alpha_2, \alpha_3$ of Maaß forms are purely imaginary (equivalently, the spectral parameters ν_1, ν_2 are purely imaginary). A Maaß form is called exceptional if it violates the Ramanujan conjecture. Modulo the action of the Weyl group, we can always assume that $\Im\nu_1, \Im\nu_2 \geq 0$ (positive Weyl chamber). Switching to the dual Maaß form if necessary, we can even assume without loss of generality $0 \leq \Im\nu_1 \leq \Im\nu_2$.

A count of the Maaß forms $\phi \in \mathcal{B}$ inside the ellipse $\lambda \leq T^2$ described by (1.3) is referred to as Weyl's law. The number of such forms is known

to be $cT^5 + O(T^3)$ for some constant c, see [24, 29]. As a first test case of the Kuznetsov formula we show a result of comparable strength as [24, Proposition 4.5] that turns out to be a simple corollary of the Kuznetsov formula. A similar upper bound has recently been proved by X. Li [26]. Let $L(\phi \times \tilde{\phi}, s)$ be the Rankin–Selberg L-function (see (4.2) below). Then the following weighted count of the cuspidal spectrum in a small ball of radius $O(1)$ in $\mathfrak{a}^*_{\mathbb{C}}$ holds.

Theorem 1 *There are absolute constants $c_1, c_2 > 0$, T_0, $K \geq 1$ with the following property*: *for all $T_1, T_2 \geq T_0$ we have*

$$c_1 T_1 T_2 (T_1 + T_2) \leq \sum_{\substack{|v_1^{(j)} - iT_1| \leq K \\ |v_2^{(j)} - iT_2| \leq K}} \left(\operatorname*{res}_{s=1} L(\phi_j \times \tilde{\phi}_j, s) \right)^{-1} \leq c_2 T_1 T_2 (T_1 + T_2).$$

It is standard to estimate the residue from above, but due to possible Siegel zeros a good lower bound is not known. If $\phi = \text{sym}^2 u$ for some Hecke–Maaß form $u \in L^2(SL_2(\mathbb{Z}) \backslash \mathfrak{h}^2)$ with spectral parameter $v \in i\mathbb{R}$, then Ramakrishnan and Wong [30] have shown that no Siegel zeros exist:

$$\operatorname*{res}_{s=1} L\left(\text{sym}^2 u \times \text{sym}^2 u, s\right) = \left(1 + |v|\right)^{o(1)}.$$

In general we will only be able to prove the following bounds: if ϕ has spectral parameters v_1, v_2, then setting $C := (1 + |v_1 + v_2|)(1 + |v_1|)(1 + |v_2|)$ we have

$$C^{-1} \ll \operatorname*{res}_{s=1} L(\phi \times \tilde{\phi}, s) \ll C^\varepsilon. \tag{1.4}$$

In particular it follows (after possibly enlarging the constant K in Theorem 1) that in each ball inside $i\mathfrak{a}^*$ of sufficiently large constant radius, there exist cusp forms. We will prove (1.4) in Lemma 2 below.

Miller [29] proved that almost all forms are non-exceptional, that is, the number of exceptional forms $\phi_j \in \mathcal{B}$ with $\lambda_j \leq T^2$ is $o(T^5)$. This was, among other things, strengthened in [24] to $O(T^3)$. By unitarity and the standard Jacquet–Shalika bounds towards the Ramanujan conjecture[2] (cf. (2.4) below) the spectral parameters v_1, v_2 of an exceptional Maaß form are of the form (assuming $0 \leq \Im v_1 \leq \Im v_2$)

$$(v_1, v_2) = (2\rho/3, -\rho/3 + i\gamma), \quad \gamma \geq 0, \ |\rho| \leq 1/2,$$

see (2.8) below. It is an easy corollary of Theorem 1 that there are $O(T^{2+\varepsilon})$ exceptional eigenvalues with $\gamma = T + O(1)$, but more can be shown which

---

[2]Better bounds are available by the work of Luo–Rudnick–Sarnak, but this is not needed here. Even the value of the constant $1/2$ is irrelevant.

can be viewed as a density theorem for exceptional eigenvalues and interpolates nicely between the Jacquet–Shalika bounds and the tempered spectrum.

Theorem 2 *For any $\varepsilon > 0$ we have*

$$\sum_{\substack{\phi_j \ exceptional \\ \gamma_j = T + O(1)}} T^{4|\rho_j|} \ll_\varepsilon T^{2+\varepsilon}.$$

Next we prove a large sieve type estimate for Hecke eigenvalues. Let $A_j(n, 1)$ denote the Hecke eigenvalues of the Hecke–Maaß cusp form ϕ_j.

Theorem 3 *Let $N \geq 1$, $T_1, T_2 \geq T_0$ sufficiently large, and let $\alpha(n)$ be a sequence of complex numbers. Then*

$$\sum_{\substack{T_1 \leq |v_1^{(j)}| \leq 2T_1 \\ T_2 \leq |v_2^{(j)}| \leq 2T_2}} \left| \sum_{n \leq N} \alpha(n) A_j(n, 1) \right|^2 \ll_\varepsilon \left(T_1^2 T_2^2 (T_1 + T_2) + T_1 T_2 N^2 \right)^{1+\varepsilon} \|\alpha\|_2^2$$

(1.5)

for any $\varepsilon > 0$ where $\|\alpha\|_2 = (\sum_n |\alpha(n)|^2)^{1/2}$.

The first term is optimal on the right hand side is optimal. Most optimistically one could hope for an additional term of size N (instead of $T_1 T_2 N^2$), but in any case our result suffices for an essentially optimal bound of the second moment of a family of genuine $GL(3)$ L-functions. This seems to be the first bound of this kind in the literature. For large sieve inequalities in the level aspect (with very different proofs) see [14, Theorem 4] and [35].

Theorem 4 *For $T \geq 1$ and any $\varepsilon > 0$ we have*

$$\sum_{T \leq |v_1^{(j)}|, |v_2^{(j)}| \leq 2T} \left| L(\phi_j, 1/2) \right|^2 \ll_\varepsilon T^{5+\varepsilon}.$$

More applications of the $GL(3)$ Kuznetsov formula to the Sato–Tate distribution of $GL(3)$ Hecke eigenvalues and a version of Theorem 2 for the Langlands parameters at finite places will be given in a forthcoming paper [2, Theorems 1–3].

After the paper was submitted, two other interesting approaches to the $GL(3)$ Kuznetsov formula have been developed independently by Buttcane [10] and Goldfeld–Kontorovich [19]. The present technique, however, gives the strongest bounds for the Kloosterman terms in the Kuznetsov formula which are indispensable for applications to L-functions as in Theorems 3 and 4. One may compare, for instance, with [19] for which the reader

is referred to the appendix which features in Theorem 5 another result of independent interest.

It would be very interesting to generalize the present results and techniques to congruence subgroups of $SL_3(\mathbb{Z})$ of the type

$$\Gamma_0(q) = \left\{ \gamma \in SL_3(\mathbb{Z}) \,\middle|\, \gamma \equiv \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix} \pmod{q} \right\}.$$

The analytic parts of the present argument (in particular the bounds for Whittaker functions and the corresponding integral transforms) work without any change. One needs a more general Bruhat decomposition to calculate the Fourier expansion of the relevant Poincaré series, and it would be useful to have an explicit spectral decomposition for the space $L^2(\Gamma_0(q)\backslash \mathfrak{h}^3)$. This along with further applications will be addressed in [1].

2 Whittaker functions

Let $\nu_1, \nu_2 \in \mathbb{C}$. We introduce the notation

$$\nu_0 := \nu_1 + \nu_2 \tag{2.1}$$

and (as in the introduction)

$$\alpha_1 = 2\nu_1 + \nu_2, \qquad \alpha_2 = -\nu_1 + \nu_2, \qquad \alpha_3 = -\nu_1 - 2\nu_2. \tag{2.2}$$

The transformations

$$(\nu_1, \nu_2) \to (-\nu_1, \nu_0) \to (\nu_2, -\nu_0) \to (-\nu_2, -\nu_1) \to (-\nu_0, \nu_1)$$
$$\to (\nu_0, -\nu_2) \tag{2.3}$$

leave $\{\alpha_1, \alpha_2, \alpha_3\}$ invariant, and they also leave $\{|\Im\nu_0|, |\Im\nu_1|, |\Im\nu_2|\}$ invariant. For convenience we assume the Jacquet–Shalika bounds towards the Ramanujan conjecture

$$\max\left(|\Re\alpha_1|, |\Re\alpha_2|, |\Re\alpha_3|\right) \leq \frac{1}{2}, \tag{2.4}$$

and we always assume unitaricity

$$\{\alpha_1, \alpha_2, \alpha_3\} = \{-\overline{\alpha_1}, -\overline{\alpha_2}, -\overline{\alpha_3}\}. \tag{2.5}$$

It is elementary to deduce from (2.4) that

$$\max\left(|\Re\nu_0|, |\Re\nu_1|, |\Re\nu_2|\right) \leq \frac{1}{3} \tag{2.6}$$

and to deduce from (2.5) that

$$\nu_0, \nu_1, \nu_2, \alpha_1, \alpha_2, \alpha_3 \in i\mathbb{R} \tag{2.7}$$

or

$$\{\alpha_1, \alpha_2, \alpha_3\} \in \{\rho + i\gamma, -\rho + i\gamma, -2i\gamma\},$$
$$\{\nu_1, \nu_2, \nu_0\} \in \{2\rho/3, -\rho/3 + i\gamma, \rho/3 + i\gamma\} \tag{2.8}$$
$$\text{or its translates under (2.3)}$$

with $\rho, \gamma \in \mathbb{R}$ and $|\rho| \leq 1/2$ by (2.4). The choice

$$(\alpha_1, \alpha_2, \alpha_3) = (\rho + i\gamma, -\rho + i\gamma, -2i\gamma),$$
$$(\nu_1, \nu_2, \nu_0) = (2\rho/3, -\rho/3 + i\gamma, \rho/3 + i\gamma) \tag{2.9}$$

is unique if we require $\Im\nu_2 \geq \Im\nu_1 \geq 0$, $\gamma \geq 0$.

Let

$$\mathfrak{h}^2 = \left\{ z = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \begin{pmatrix} y & \\ & 1 \end{pmatrix} \, \middle| \, y > 0, \, x \in \mathbb{R} \right\}$$
$$\cong GL_2(\mathbb{R})/(O_2 Z_2) \cong SL_2(\mathbb{R})/SO_2 \quad \text{and}$$

$$\mathfrak{h}^3 = \left\{ z = \begin{pmatrix} 1 & x_2 & x_3 \\ & 1 & x_1 \\ & & 1 \end{pmatrix} \begin{pmatrix} y_1 y_2 & & \\ & y_1 & \\ & & 1 \end{pmatrix} \, \middle| \, y_1, y_2 > 0, \, x_1, x_2, x_3 \in \mathbb{R} \right\}$$
$$\cong GL_3(\mathbb{R})/(O_3 Z_3) \cong SL_3(\mathbb{R})/SO_3.$$

The group $SL_3(\mathbb{Z})$ acts faithfully on $\mathfrak{h}^3$ by left multiplication.

The Whittaker function $\mathcal{W}^{\pm}_{\nu_1, \nu_2} : \mathfrak{h}^3 \to \mathbb{C}$ is given by[3] (analytic continuation in ν_1, ν_2 of)

$$\mathcal{W}^{\pm}_{\nu_1, \nu_2}(z) = \int_{\mathbb{R}^3} I_{\nu_1, \nu_2}\left(\begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix} \begin{pmatrix} 1 & u_2 & u_3 \\ & 1 & u_1 \\ & & 1 \end{pmatrix} z \right)$$
$$\times e(-u_1 \mp u_2) \, du_1 \, du_2 \, du_3 \tag{2.10}$$

with

$$I_{\nu_1, \nu_2}(z) = y_1^{1+2\nu_1+\nu_2} y_2^{1+\nu_1+2\nu_2}. \tag{2.11}$$

---

[3]Some authors use different signs in the long Weyl element, but since the I_{ν_1, ν_2} function depends only on y_1, y_2, this leads to the same definition.

Compared to [7][4] we have re-normalized the indices $\nu_j \to 1/3 + \nu_j$. By the formula of Takhtadzhyan–Vinogradov we have $\mathcal{W}^{\pm}_{\nu_1,\nu_2}(z) = e(x_1 \pm x_2)W_{\nu_1,\nu_2}(y_1, y_2)$ where[5]

$$
W_{\nu_1,\nu_2}(y_1, y_2) = 8y_1y_2\left(\frac{y_1}{y_2}\right)^{\frac{\nu_1-\nu_2}{2}} \prod_{j=0}^{2} \pi^{\frac{1}{2}+\frac{3}{2}\nu_j}\Gamma\left(\frac{1}{2}+\frac{3}{2}\nu_j\right)^{-1}
$$

$$
\times \int_0^{\infty} K_{\frac{3}{2}\nu_0}\left(2\pi y_2\sqrt{1+1/u^2}\right)K_{\frac{3}{2}\nu_0}\left(2\pi y_1\sqrt{1+u^2}\right)u^{\frac{3}{2}(\nu_1-\nu_2)}\frac{du}{u}.
$$

(2.12)

It is convenient to slightly re-normalize this function: let

$$
c_{\nu_1,\nu_2} := \pi^{-3\nu_0}\prod_{j=0}^{2}\Gamma\left(\frac{1}{2}+\frac{3}{2}\nu_j\right)\left|\Gamma\left(\frac{1}{2}+\frac{3}{2}i\Im\nu_j\right)\right|^{-1}
$$

(2.13)

and

$$
\tilde{W}_{\nu_1,\nu_2}(y_1, y_2) := W_{\nu_1,\nu_2}(y_1, y_2)c_{\nu_1,\nu_2}
$$

$$
= 8\pi^{\frac{3}{2}}\prod_{j=0}^{2}\left|\Gamma\left(\frac{1}{2}+\frac{3}{2}i\Im\nu_j\right)\right|^{-1}y_1y_2\left(\frac{y_1}{y_2}\right)^{\frac{\nu_1-\nu_2}{2}}
$$

$$
\times \int_0^{\infty} K_{\frac{3}{2}\nu_0}(2\pi y_2\sqrt{1+1/u})K_{\frac{3}{2}\nu_0}(2\pi y_1\sqrt{1+u})u^{\frac{3}{4}(\nu_1-\nu_2)}\frac{du}{u}.
$$

If $\nu_1, \nu_2 \in i\mathbb{R}$, this changes the original Whittaker function only by a constant on the unit circle, in the situation (2.9) it changes the order of magnitude by a bounded factor. Often the Whittaker function is defined entirely without the normalizing Gamma-factors in the denominator of (2.12) in which case it is often referred to as the completed Whittaker function. It is convenient not to work with the completed Whittaker function here, see Remark 3 below. (Of course, $\tilde{W}_{\nu_1,\nu_2}$ is not analytic in the indices any more.)

---

[4]In [17, p. 154, third display] the values of ν_1, ν_2 are interchanged in the definition of I-function, but the following formulas are again in accordance with Bump's definition.

[5]The normalization is complicated: the leading constant in [17, (6.1.3)] should be 8 instead of 4, while the definition [33, (1.1)] differs from (2.10), in addition to the Gamma-factors, by a factor 2.

The $GL(2)$-analogue of this function is

$$W_\nu(y) := 2\pi^{\frac{1}{2}} \left| \Gamma\left(\frac{1}{2} + \nu\right) \right|^{-1} \sqrt{y} K_\nu(2\pi y), \tag{2.14}$$

see [17, p. 65].

We proceed to collect analytic information on the $GL(3)$ Whittaker function. We have the double Mellin inversion formula ([17, p. 155], [7, (10.1)])

$$\tilde{W}_{\nu_1,\nu_2}(y_1, y_2) = \frac{y_1 y_2 \pi^{\frac{3}{2}}}{\prod_{j=0}^{2} |\Gamma(\frac{1}{2} + \frac{3}{2} i \Im \nu_j)|}$$

$$\times \frac{1}{(2\pi i)^2} \int_{(c_2)} \int_{(c_1)} \frac{\prod_{j=1}^{3} \Gamma(\frac{1}{2}(s_1 + \alpha_j)) \prod_{j=1}^{3} \Gamma(\frac{1}{2}(s_2 - \alpha_j))}{4\pi^{s_1+s_2} \Gamma(\frac{1}{2}(s_1 + s_2))}$$

$$\times y_1^{-s_1} y_2^{-s_2} \, ds_1 \, ds_2. \tag{2.15}$$

This implies in particular that $\tilde{W}_{\nu_1,\nu_2}$ is invariant under the transformations (2.3) which is the reason for including the normalization constant c_{ν_1,ν_2}. Note that

$$\tilde{W}_{\nu_1,\nu_2}(y_1, y_2) = \tilde{W}_{\nu_2,\nu_1}(y_2, y_1) = \overline{\tilde{W}_{\overline{\nu_1},\overline{\nu_2}}(y_1, y_2)}. \tag{2.16}$$

Uniform bounds for Bessel functions are rare in the literature, but frequently needed in the $GL(2)$ theory. We are not aware of any uniform bound for a $GL(n)$ Whittaker function with $n > 2$. Although the proofs of Theorems 1–5 do not require bounds for individual Whittaker functions, we record here for future reference the following uniform result.

Proposition 1 *Let* $\nu_1, \nu_2 \in \mathbb{C}$ *satisfy* (2.6)–(2.8) *and write* $\theta = \max(|\Re\alpha_1|, |\Re\alpha_2|, |\Re\alpha_3|) \leq 1/2$. *Let* $\theta < \sigma_1 < \sigma_2$ *and* $\varepsilon > 0$. *Then for any* $\sigma_1 \leq c_1, c_2 \leq \sigma_2$ *we have*

$$\tilde{W}_{\nu_1,\nu_2}(y_1, y_2) \ll_{\sigma_1,\sigma_2,\varepsilon} \frac{y_1 y_2}{(1 + |\nu_1| + |\nu_2|)^{1/2-\varepsilon}} \left(\frac{y_1}{1 + |\nu_1| + |\nu_2|}\right)^{-c_1}$$

$$\times \left(\frac{y_2}{1 + |\nu_1| + |\nu_2|}\right)^{-c_2}.$$

Remark 1 This result can be refined somewhat, in particular for small and large y_1, y_2. In the "transitional range" it is not too far from the truth. For instance, if $\nu_1 = \nu_2 = iT$ are large (and purely imaginary) and $y_1 = y_2 = \frac{3}{2\pi}T - \frac{1}{100}T^{1/3}$, then the integral in (2.12) is non-oscillating, and it follows from (2.12) and known properties of the K-Bessel function that $\tilde{W}_{\nu_1,\nu_2}(y_1, y_2) \asymp y_1 y_2 / T^{4/3}$, whereas our bound gives $\tilde{W}_{\nu_1,\nu_2}(y_1, y_2) \ll y_1 y_2 T^{\varepsilon-\frac{1}{2}}$.

Proof Let us first assume that ν_1, ν_2 are purely imaginary; we write $\tilde{\nu}_j := \Im \nu_j$. By (2.16) and the invariance under (2.3) we can assume without loss of generality that

$$0 \leq \tilde{\nu}_1 \leq \tilde{\nu}_2. \tag{2.17}$$

By (2.15) and Stirling's formula, we have

$$\tilde{W}_{\nu_1,\nu_2}(y_1, y_2)$$

$$\ll_{\sigma_1,\sigma_2} y_1 y_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\prod_{j=1}^{3}(1 + |it_1 + \alpha_j|)^{\frac{c_1-1}{2}} \prod_{j=1}^{3}(1 + |it_2 - \alpha_j|)^{\frac{c_2-1}{2}}}{(1 + |t_1 + t_2|)^{\frac{c_1+c_2-1}{2}}}$$

$$\times \exp\left(\frac{3\pi}{4}\sum_{j=0}^{2}|\nu_j| - \frac{\pi}{4}\sum_{j=1}^{3}|it_1 + \alpha_j| - \frac{\pi}{4}\sum_{j=1}^{3}|it_2 - \alpha_j|\right.$$

$$\left. + \frac{\pi}{4}|t_1 + t_2|\right) y_1^{-c_1} y_2^{-c_2} \, dt_1 \, dt_2 \tag{2.18}$$

for $\sigma_1 < c_1, c_2 < \sigma_2$. It is elementary to check that

$$\frac{3\pi}{4}\sum_{j=0}^{2}|\nu_j| - \frac{\pi}{4}\sum_{j=1}^{3}|it_1 + \alpha_j| - \frac{\pi}{4}\sum_{j=1}^{3}|it_2 - \alpha_j| + \frac{\pi}{4}|t_1 + t_2| \leq 0 \tag{2.19}$$

with equality if and only if

$$\tilde{\nu}_1 - \tilde{\nu}_2 \leq t_1 \leq \tilde{\nu}_1 + 2\tilde{\nu}_2, \qquad \tilde{\nu}_2 - \tilde{\nu}_1 \leq t_2 \leq 2\tilde{\nu}_1 + \tilde{\nu}_2, \tag{2.20}$$

or

$$-2\tilde{\nu}_1 - \tilde{\nu}_2 \leq t_1 \leq \tilde{\nu}_1 - \tilde{\nu}_2, \qquad -\tilde{\nu}_1 - 2\tilde{\nu}_2 \leq t_2 \leq \tilde{\nu}_2 - \tilde{\nu}_1. \tag{2.21}$$

For $a < b$ let $w_{a,b}(t)$ be defined by

$$w_{a,b} := \min\left(1, e^{-\frac{\pi}{4}(t-b)}, e^{-\frac{\pi}{4}(a-t)}\right);$$

then the exp-factor in (2.18) is bounded by

$$w_{\tilde{\nu}_1-\tilde{\nu}_2, \tilde{\nu}_1+2\tilde{\nu}_2}(t_1) w_{\tilde{\nu}_2-\tilde{\nu}_1, 2\tilde{\nu}_1+\tilde{\nu}_2}(t_2) + w_{-2\tilde{\nu}_1-\tilde{\nu}_2, \tilde{\nu}_1-\tilde{\nu}_2}(t_1) w_{-\tilde{\nu}_1-2\tilde{\nu}_2, \tilde{\nu}_2-\tilde{\nu}_1}(t_2),$$

and hence we have

$$\tilde{W}_{\nu_1,\nu_2}(y_1,y_2) \ll y_1 y_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(w_{\tilde{\nu}_1-\tilde{\nu}_2,\tilde{\nu}_1+2\tilde{\nu}_2}(t_1) w_{\tilde{\nu}_2-\tilde{\nu}_1,2\tilde{\nu}_1+\tilde{\nu}_2}(t_2) \right.$$

$$\left. + w_{-2\tilde{\nu}_1-\tilde{\nu}_2,\tilde{\nu}_1-\tilde{\nu}_2}(t_1) w_{-\tilde{\nu}_1-2\tilde{\nu}_2,\tilde{\nu}_2-\tilde{\nu}_1}(t_2) \right)$$

$$\times \frac{\prod_{j=1}^3 (1+|it_1+\alpha_j|)^{\frac{c_1-1}{2}} \prod_{j=1}^3 (1+|it_2-\alpha_j|)^{\frac{c_2-1}{2}}}{(1+|t_1+t_2|)^{\frac{c_1+c_2-1}{2}}}$$

$$\times y_1^{-c_1} y_2^{-c_2} \, dt_1 \, dt_2.$$

We consider only the first summand in the first line of the preceding display. The second summand is similar. By a shift of variables, the first term equals

$$y_1^{1-c_1} y_2^{1-c_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_{0,3\tilde{\nu}_2}(t_1) w_{0,3\tilde{\nu}_1}(t_2)$$

$$\times \frac{(1+|t_1-3\tilde{\nu}_2|)^{\frac{c_1-1}{2}} (1+|t_1|)^{\frac{c_1-1}{2}} (1+|t_1+3\tilde{\nu}_1|)^{\frac{c_1-1}{2}}}{(1+|t_1+t_2|)^{\frac{c_1+c_2-1}{2}}}$$

$$\times \left(1+|t_2+3\tilde{\nu}_2|\right)^{\frac{c_2-1}{2}} \left(1+|t_2|\right)^{\frac{c_2-1}{2}} \left(1+|t_2-3\tilde{\nu}_1|\right)^{\frac{c_2-1}{2}} \, dt_1 \, dt_2. \quad (2.22)$$

It is straightforward to estimate this expression. For convenience we provide the details. We recall our assumption (2.17) and split the t_1, t_2 integration into several ranges. Let $\tilde{\nu}_1 \le R \le \tilde{\nu}_2$, and define

$$\mathcal{I}_- := \{t_1 \le \tilde{\nu}_1\}, \qquad \mathcal{I}_R := \{R \le t_1 \le 2R\}, \qquad \mathcal{I}_+ := \{t_1 \ge 2\tilde{\nu}_2\},$$

$$\mathcal{J}_- := \{t_2 \le \tilde{\nu}_1\}, \qquad \mathcal{J}_+ := \{t_2 \ge \tilde{\nu}_1\}.$$

We estimate the double integral in all 6 ranges for $c_1, c_2 > 0$:

$$\int_{\mathcal{I}_-} \int_{\mathcal{J}_-} \ll \int_{-\infty}^{\tilde{\nu}_1} \int_{-\infty}^{\tilde{\nu}_1} \min\left(1, e^{\frac{\pi}{4}t_1}\right) \min\left(1, e^{\frac{\pi}{4}t_2}\right) \left((1+\tilde{\nu}_1)(1+\tilde{\nu}_2)\right)^{\frac{c_1+c_2}{2}-1}$$

$$\times \frac{(1+|t_1|)^{\frac{c_1-1}{2}} (1+|t_2|)^{\frac{c_2-1}{2}}}{(1+|t_1+t_2|)^{\frac{c_1+c_2-1}{2}}} \, dt_2 \, dt_1$$

$$\ll \left((1+\tilde{\nu}_1)(1+\tilde{\nu}_2)\right)^{\frac{c_1+c_2}{2}-1} (1+\tilde{\nu}_1)^{\frac{3}{2}};$$

$$\int_{\mathcal{I}_-} \int_{\mathcal{J}_+} \ll \int_{-\infty}^{\tilde{\nu}_1} \int_{\tilde{\nu}_1}^{\infty} \min\left(1, e^{\frac{\pi}{4}t_1}\right) \min\left(1, e^{\frac{\pi}{4}(3\tilde{\nu}_1-t_2)}\right) (1+\tilde{\nu}_1)^{-\frac{1}{2}}$$

$$\times (1+\tilde{\nu}_2)^{\frac{c_1+c_2}{2}-1} \left(1+|t_1|\right)^{\frac{c_1-1}{2}} \left(1+|t_2-3\tilde{\nu}_1|\right)^{\frac{c_2-1}{2}} \, dt_2 \, dt_1$$

$$\ll \left((1+\tilde{\nu}_1)(1+\tilde{\nu}_2)\right)^{\frac{c_1+c_2}{2}-1} (1+\tilde{\nu}_1)^{\frac{3}{2}};$$

$$\int_{\mathcal{I}_R}\int_{\mathcal{J}_-} \ll \int_R^{2R}\int_{-\infty}^{\tilde{v}_1} \min\left(1, e^{\frac{\pi}{4}t_2}\right)(1+\tilde{v}_1)^{\frac{c_2-1}{2}}(1+\tilde{v}_2)^{\frac{c_1+c_2}{2}-1}$$

$$\times (1+R)^{\frac{c_1-c_2-1}{2}}\left(1+|t_2|\right)^{\frac{c_2-1}{2}} dt_2\, dt_1$$

$$\ll (1+\tilde{v}_1)^{c_2}(1+\tilde{v}_2)^{\frac{c_1+c_2}{2}-1}(1+R)^{\frac{c_1-c_2+1}{2}};$$

$$\int_{\mathcal{I}_R}\int_{\mathcal{J}_+} \ll \int_{-R}^{2R}\int_{\tilde{v}_1}^{\infty} \min\left(1, e^{\frac{\pi}{4}(3\tilde{v}_1-t_2)}\right)(1+\tilde{v}_1)^{\frac{c_2-1}{2}}(1+\tilde{v}_2)^{\frac{c_1+c_2}{2}-1}$$

$$\times (1+R)^{\frac{c_1-c_2-1}{2}}\left(1+|t_2-3\tilde{v}_1|\right)^{\frac{c_2-1}{2}} dt_2\, dt_1$$

$$\ll (1+\tilde{v}_1)^{c_2}(1+\tilde{v}_2)^{\frac{c_1+c_2}{2}-1}(1+R)^{\frac{c_1-c_2+1}{2}};$$

$$\int_{\mathcal{I}_+}\int_{\mathcal{J}_-} \ll \int_{2\tilde{v}_2}^{\infty}\int_{-\infty}^{\tilde{v}_1} \min\left(1, e^{\frac{\pi}{4}(3\tilde{v}_2-t_1)}\right)\min\left(1, e^{\frac{\pi}{4}t_2}\right)(1+\tilde{v}_1)^{\frac{c_2-1}{2}}$$

$$\times (1+\tilde{v}_2)^{\frac{c_1}{2}-1}\left(1+|t_1-3\tilde{v}_2|\right)^{\frac{c_1-1}{2}}\left(1+|t_2|\right)^{\frac{c_2-1}{2}} dt_2\, dt_1$$

$$\ll (1+\tilde{v}_1)^{c_2}(1+\tilde{v}_2)^{c_1-\frac{1}{2}};$$

$$\int_{\mathcal{I}_+}\int_{\mathcal{J}_+} \ll \int_{2\tilde{v}_2}^{\infty}\int_{\tilde{v}_1}^{\infty} \min\left(1, e^{\frac{\pi}{4}(3\tilde{v}_2-t_1)}\right)\min\left(1, e^{\frac{\pi}{4}(3\tilde{v}_1-t_2)}\right)(1+\tilde{v}_1)^{\frac{c_2-1}{2}}$$

$$\times (1+\tilde{v}_2)^{\frac{c_1}{2}-1}\left(1+|t_1-3\tilde{v}_2|\right)^{\frac{c_1-1}{2}}\left(1+|t_2-3\tilde{v}_1|\right)^{\frac{c_2-1}{2}} dt_2\, dt_1$$

$$\ll (1+\tilde{v}_1)^{c_2}(1+\tilde{v}_2)^{c_1-\frac{1}{2}}.$$

Combining all 6 previous bounds, and summing over dyadic numbers R, we obtain the bound of the proposition if $v_1, v_2 \in i\mathbb{R}$.

It remains to consider the situation (2.8). The exponential factor does not change, but the fraction in the first line of (2.18) becomes now

$$\frac{(1+|t_1+\gamma|)^{\frac{c_1+\rho-1}{2}}(1+|t_1+\gamma|)^{\frac{c_1-\rho-1}{2}}(1+|t_1-2\gamma|)^{\frac{c_1-1}{2}}}{(1+|t_1+t_2|)^{\frac{c_1+c_2-1}{2}}}$$

$$\times \left(1+|t_2-\gamma|\right)^{\frac{c_2+\rho-1}{2}}\left(1+|t_2-\gamma|\right)^{\frac{c_2-\rho-1}{2}}\left(1+|t_2+2\gamma|\right)^{\frac{c_2-1}{2}}. \quad (2.23)$$

This is independent of ρ, and the same calculation goes through. □

We recall an important formula of Stade [33] (cf. also (2.16) and observe that Stade's definition [33, (1.1)] of the Whittaker function has v_1 and v_2 interchanged, and his Whittaker function is, up to Gamma-factors, twice our Whittaker function).

Proposition 2 *Let* $\nu_1, \nu_2, \mu_1, \mu_2, s \in \mathbb{C}$. *We use the notation* (2.1) *and* (2.2), *and define similarly* μ_0 *and* $\beta_1, \beta_2, \beta_3$ *in terms of* μ_1, μ_2. *Then we have an equality of meromorphic functions in* s:

$$\int_0^\infty \int_0^\infty \overline{\tilde{W}_{\nu_1,\nu_2}(y_1, y_2)} \tilde{W}_{\mu_1,\mu_2}(y_1, y_2) \left(y_1^2 y_2\right)^s \frac{dy_1\, dy_2}{(y_1 y_2)^3}$$

$$= \frac{\pi^{3-3s} \prod_{j,k=1}^3 \Gamma(\frac{s+\alpha_j+\overline{\beta}_k}{2})}{4\Gamma(\frac{3}{2}s) \prod_{j=0}^2 |\Gamma(\frac{1}{2} + \frac{3}{2}i\Im\nu_j)| \prod_{j=0}^2 |\Gamma(\frac{1}{2} + \frac{3}{2}i\Im\overline{\mu}_j)|}.$$

3 Integrals over Whittaker functions

For our purposes it is convenient to consider the double Mellin transform of the product of $\tilde{W}_{\nu_1,\nu_2}$ with some rapidly decaying function. We are not aware of any explicit formula in the literature, but the next proposition gives an asymptotic result which is sufficient for our purposes. This is one of the key ingredients in this paper, and therefore we present all details of the lengthy proof.

Proposition 3 *Let* $A, X_1, X_2 \geq 1, \tau_1, \tau_2 \geq 0$, *and assume that* $\tau_1 + \tau_2$ *is sufficiently large in terms of* A. *Let* $\nu_1, \nu_2 \in \mathbb{C}$ *satisfy* (2.6)–(2.8) *and in addition* $\Im\nu_1, \Im\nu_2 \geq 0$. *Let*

$$t_1 = \tau_1 + 2\tau_2, \qquad t_2 = 2\tau_1 + \tau_2.$$

Fix two non-zero smooth functions $f_1, f_2 : (0, \infty) \to [0, 1]$ *with compact support. Let* $\varepsilon > 0$ *and let*

$$I = I(\nu_1, \nu_2, t_1, t_2, X_1, X_2)$$
$$= \left| \int_0^\infty \int_0^\infty \tilde{W}_{\nu_1,\nu_2}(y_1, y_2) f_1(X_1 y_1) f_2(X_2 y_2) y_1^{it_1} y_2^{it_2} \frac{dy_1\, dy_2}{(y_1 y_2)^3} \right|.$$

If $X_1 = X_2 = 1$, *then*

$$I \ll \left(|\Im\nu_1 - \tau_1| + |\Im\nu_2 - \tau_2|\right)^{-A} \prod_{j=0}^2 \left(1 + |\nu_j|\right)^{-\frac{1}{2}} \tag{3.1}$$

Moreover, there is a constant c *depending on* f_1, f_2 *such that the following holds*: *if*

$$X_1 = X_2 = 1, \qquad \tau_1, \tau_2 \gg 1, \qquad |\Im\nu_1 - \tau_1| \leq c, \qquad |\Im\nu_2 - \tau_2| \leq c, \tag{3.2}$$

then

$$I \asymp \prod_{j=0}^{2} \left(1 + |v_j|\right)^{-1/2}; \tag{3.3}$$

and if v_1, v_2 are given by (2.9) and in addition

$$X_2 \gg 1, \qquad |\rho| \geq \varepsilon, \qquad \gamma \gg (X_1 X_2)^{\varepsilon}, \qquad |\gamma - \tau_2| \leq c, \qquad |\tau_1| \leq c, \tag{3.4}$$

then

$$I \asymp X_1 X_2^{1+|\rho|} \prod_{j=0}^{2} \left(1 + |v_j|\right)^{-1/2}. \tag{3.5}$$

All implied constant depend at most on A, ε, f_1, f_2 and the sign $\gg$ should be interpreted as "up to a sufficiently large constant".

Remark 2 This should be roughly interpreted as follows: given t_1, t_2, X_1, X_2 as above, I as a function of $v_1, v_2 \in \{z \in \mathbb{C} : |\Im z| \leq 1/2\}$ is under some technical assumptions a function with a bump at $\Im v_1 = \tau_1$ and $\Im v_2 = \tau_2$ of size $X_1 X_2^{1+\max_j |\Re \alpha_j|} (v_0 v_1 v_2)^{-1/2}$ with rapid decay away from this point. Most of the time we will put $X_1 = X_2 = 1$. Only if we need a test function that blows up at exceptional eigenvalues we will choose X_2 to be large. The asymmetry in X_1, X_2 in (3.4) and (3.5) is due to the special choice (2.9).

Proof By Parseval's formula and (2.15) the double integral in question equals

$$\int_{(1/2)} \int_{(1/2)} \frac{\widehat{f_1}(-1 + it_1 - u_1)\, \widehat{f_2}(-1 + it_2 - u_2)}{X_1^{-1+it_1-u_1} X_2^{-1+it_2-u_2}}$$

$$\times \frac{\prod_{j=1}^{3} \Gamma(\tfrac{1}{2}(u_1 + \alpha_j)) \prod_{j=1}^{3} \Gamma(\tfrac{1}{2}(u_2 - \alpha_j))}{4\pi^{u_1+u_2-\frac{3}{2}} \Gamma(\tfrac{1}{2}(u_1 + u_2)) \prod_{j=0}^{2} |\Gamma(\tfrac{1}{2} + \tfrac{3}{2} i \Im v_j)|} \frac{du_1 du_2}{(2\pi i)^2} \tag{3.6}$$

Let us first assume that

$$|v_1| + |v_2| \leq \frac{1}{100} (\tau_1 + \tau_2). \tag{3.7}$$

In this case the conditions (3.2) and (3.4) are void, so we only need to show (3.1) and take $X_1 = X_2 = 1$. We apply Stirling's formula to the Γ-quotient.

We argue as in the proof of Proposition 1, see (2.19) and the surrounding discussion. The exponential part is given by

$$\exp\left(\frac{3\pi}{4}\sum_{j=0}^{2}|\Im v_j| - \frac{\pi}{4}\sum_{j=1}^{3}\left|\Im(u_1+\alpha_j)\right| - \frac{\pi}{4}\sum_{j=1}^{3}\left|\Im(u_2-\alpha_j)\right|$$
$$+ \frac{\pi}{4}\left|\Im(u_1+u_2)\right|\right).$$

As before let us write $\tilde{v}_j = \Im v_j$ and assume without loss of generality (2.17). Using (2.20) and (2.21) together with the rapid decay of $\widehat{f}_1$ and $\widehat{f}_2$ it is easy to see that by our present assumption (3.7) we can bound I by

$$\ll_{A,f_1,f_2} (t_1+t_2)^{-A}.$$

In the range (3.7) this is acceptable for (3.1).

Let us now assume

$$|v_1| + |v_2| \geq \frac{1}{100}(\tau_1+\tau_2). \tag{3.8}$$

We want to shift the two contours in (3.6) to $-\infty$. To check convergence, we first shift the u_1-integral to $\Re u_1 = -2A - 1$ for some large integer A. We observe that

$$\widehat{f}(s) \ll_{B,f} |s|^{-B} C^{\Re s+B}$$

for $\Re s > 0$, any $B \geq 0$ and some constant $C > 0$ depending only on f (one can take $C := \sup\{x > 0 \mid f(x) \neq 0\}$). We also recall that the reflection formula for the Gamma-function implies the uniform bound

$$\Gamma(-s) \ll e^{-\frac{\pi}{2}|\Im s|}\left(\frac{|s|}{e}\right)^{-\Re s-\frac{1}{2}}$$

for $\Re s > 0$, $\min_{n\in\mathbb{Z}}(\Re s - n) > 1/50$. It is now easy to see that the remaining u_1-integral for $A \to \infty$ vanishes, and we are left with the sum over the residues. Next we shift in the same way the u_2-integral to $-\infty$, and express (3.6) as an absolutely convergent double sum over residues. Let us first assume that $v_1 \neq 0$ so that $\alpha_1, \alpha_2, \alpha_3$ are pairwise distinct. For $j \in \{1, 2, 3\}$ we denote by k, l two integers such that $\{j, k, l\} = \{1, 2, 3\}$. Similarly for $r \in \{1, 2, 3\}$ let s, t be such that $\{r, s, t\} = \{1, 2, 3\}$. Then (3.6) equals

$$\sum_{\substack{1\le j,r\le 3 \\ j\ne r}}\sum \frac{\Gamma(\frac{-\alpha_j+\alpha_k}{2})\Gamma(\frac{-\alpha_j+\alpha_l}{2})\Gamma(\frac{\alpha_r-\alpha_s}{2})\Gamma(\frac{\alpha_r-\alpha_t}{2})}{\prod_{j=0}^{2}|\Gamma(\frac{1}{2}+\frac{3}{2}i\Im v_j)|\,4\pi^{-\alpha_j+\alpha_r-\frac{3}{2}}\Gamma(\frac{-\alpha_j+\alpha_r}{2})}$$

$$\times\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{4(-1)^{n+m}(\frac{-\alpha_j+\alpha_r}{2}-1)_{n+m}}{n!m!(\frac{-\alpha_j+\alpha_k}{2}-1)_n(\frac{-\alpha_j+\alpha_l}{2}-1)_n(\frac{\alpha_r-\alpha_s}{2}-1)_m(\frac{\alpha_r-\alpha_t}{2}-1)_m}$$

$$\times\frac{\widehat{f_1}(-1+\alpha_j+it_1+2n)\,\widehat{f_2}(-1-\alpha_r+it_2+2m)}{X_1^{-1+\alpha_j+it_1+2n}X_2^{-1-\alpha_r+it_2+2m}}. \tag{3.9}$$

We can bound the second and third line of (3.9) by

$$\sum_{n,m}\ll_{A,f_1,f_2} X_1^{1-\Re\alpha_j}X_2^{1+\Re\alpha_r}\left(1+|\alpha_j+it_1|\right)^{-A}\left(1+|\alpha_r-it_2|\right)^{-A}. \tag{3.10}$$

Here we have used that by (2.2) we have the following equality of multisets:

$$\{-\alpha_j+\alpha_k,\,-\alpha_j+\alpha_l,\,\alpha_r-\alpha_s,\,\alpha_r-\alpha_t\}\setminus\{-\alpha_j+\alpha_r\}$$
$$=\{\pm 3v_0,\,\pm 3v_1,\,\pm 3v_2\} \tag{3.11}$$

for a certain choice of signs (depending on j,r) whenever $j\ne r$. In addition we see that we have in the special case $j=3$, $r\in\{1,2\}$ (that is, $\alpha_j=-v_1-2v_2$, $\alpha_r=2v_1+v_2$ or $-v_1+v_2$)

$$\sum_{(n,m)\ne(0,0)}\ll_{A,f_1,f_2} X_1^{1-\Re\alpha_j}X_2^{1+\Re\alpha_r}\frac{(1+|\alpha_j+it_1|)^{-A}(1+|\alpha_r-it_2|)^{-A}}{\min((1+|v_1|)X_2^2,(1+|v_2|)X_1^2)}. \tag{3.12}$$

We will now carefully analyze all 6 terms $1\le\alpha_j,\alpha_r\le 3$, $j\ne r$ in the main term under the assumption $\tau_1,\tau_2\ge 0$, $\Im v_2\ge\Im v_1\ge 0$ and show that they all satisfy the bound (3.1). Moreover, under the assumption (3.2), the term $j=3$, $r=1$ is of order of magnitude (3.3) and dominates all other terms. Similarly we will show (3.5). We will first make the extra assumption

$$|v_1|\ge\varepsilon.$$

This ensures that $\alpha_1,\alpha_2,\alpha_3$ are not too close together (note that $|v_2|$ must be large by (3.8)). By Stirling's formula, (3.11) and (2.9) (in the non-tempered case),

$$\left| \frac{\Gamma(\frac{-\alpha_j + \alpha_k}{2})\Gamma(\frac{-\alpha_j + \alpha_l}{2})\Gamma(\frac{\alpha_r - \alpha_s}{2})\Gamma(\frac{\alpha_r - \alpha_t}{2})}{\prod_{j=0}^{2} |\Gamma(\frac{1}{2} + \frac{3}{2}i\Im \nu_j)| \Gamma(\frac{-\alpha_j + \alpha_r}{2})} \right|$$

$$\begin{cases} \ll_\varepsilon \prod_{n=0}^{2}(1 + |\nu_n|)^{O(1)}, & \{j, r\} = \{1, 2\} \text{ and } \nu_1 \in \mathbb{R}, \\ \asymp_\varepsilon \prod_{n=0}^{2}(1 + |\nu_n|)^{-\frac{1}{2}}, & \text{otherwise.} \end{cases} \tag{3.13}$$

The non-negativity of f_1, f_2 implies that the absolute values of the Mellin transforms $|\widehat{f_1}(-1 + \alpha_3 + it_1)|$, $|\widehat{f_2}(-1 - \alpha_1 + it_2)|$ are bounded from below in the range (3.2) and (3.4) if c is sufficiently small. Combining (3.13) and (3.12), we see that under the assumption (3.2) the term $j = 3$, $r = 1$ satisfies (3.3). Note that (3.2) forces $|\nu_1|$, $|\nu_2|$ to be sufficiently large and excludes (2.8). Similarly, under the assumption (3.4) the term $j = 3$, $r = 1$ (if $\rho > 0$) or the term $j = 3$, $r = 2$ (if $\rho < 0$) satisfies (3.5), whereas the other term is of smaller order of magnitude. This is also consistent with (3.1).

It remains to show that all other terms satisfy (3.1), and are of lesser order of magnitude than (3.3) and (3.5) under the respective conditions. Under the assumption (3.2) all 5 terms $(j, r) \neq (3, 1)$ satisfy $|\alpha_j + it_1| + |\alpha_r - it_2| \geq 3 \min(|\Im \nu_1|, |\Im \nu_2|) + O(1)$ and can therefore be bounded by (recall (2.17))

$$\ll_{\varepsilon, A, f_1, f_2} |\nu_1|^{-A} \prod_{n=0}^{2}\left(1 + |\nu_n|\right)^{-\frac{1}{2}}$$

which is dominated by (3.3). Similarly, under the assumption (3.4) the 4 terms $(\alpha_j, \alpha_r) \notin \{(3, 1), (3, 2)\}$ satisfy $|\alpha_j + it_1| + |\alpha_r - it_2| \geq 3|\Im \nu_2| + O(1)$ and can therefore be bounded by

$$\ll_{\varepsilon, A, f_1, f_2} |\nu_2|^{-A} \prod_{n=0}^{2}\left(1 + |\nu_n|\right)^{-\frac{1}{2}}$$

which is again dominated by (3.5). We proceed now to show (3.1) for $X_1 = X_2 = 1$. It follows from (3.10) and (3.13) that all 6 terms $\alpha_j \neq \alpha_r$ contribute

$$\ll_{\varepsilon, A, f_1, f_2} \left((1 + |\alpha_j + it_1|)(1 + |\alpha_r - it_2|)\right)^{-A}$$

$$\times \begin{cases} \prod_{n=0}^{2}(1 + |\nu_n|)^{O(1)}, & \{j, r\} = \{1, 2\} \text{ and } \nu_1 \in \mathbb{R}, \\ \prod_{n=0}^{2}(1 + |\nu_n|)^{-\frac{1}{2}}, & \text{otherwise} \end{cases}$$

to the main term. This is in agreement with (3.1) if we can show

$$|\Im \alpha_j + t_1| + |\Im \alpha_r - t_2| \geq \frac{1}{2}\left(|\tilde{\nu}_1 - \tau_1| + |\tilde{\nu}_2 - \tau_2|\right).$$

This is clear for $j = 1$ *or* $r = 3$ by the positivity assumption $\tilde{v}_1, \tilde{v}_2, \tau_1, \tau_2 \geq 0$ (recall the notation $\tilde{v}_j = \Im v_j$), even without the factor $1/2$. We check the other 3 cases. In the case $j = 3, r = 1$ we have again the stronger inequality

$$|-\tilde{v}_1 - 2\tilde{v}_2 + \tau_1 + 2\tau_2| + |2\tilde{v}_1 + \tilde{v}_2 - 2\tau_1 - \tau_2| \geq |\tau_1 - \tilde{v}_1| + |\tau_2 - \tilde{v}_2|$$

which follows from the easy to check inequality $|a| + |b| \leq |a + 2b| + |b + 2a|$. In the case $j = 3, r = 2$ we need to show

$$|-\tilde{v}_1 - 2\tilde{v}_2 + \tau_1 + 2\tau_2| + |-\tilde{v}_1 + \tilde{v}_2 - 2\tau_1 - \tau_2| \geq \frac{1}{2}\big(|\tau_1 - \tilde{v}_1| + |\tau_2 - \tilde{v}_2|\big).$$

If $\tau_1 - \tilde{v}_1$ and $\tau_2 - \tilde{v}_2$ are of the same sign, the first term dominates the right hand side; if $\tau_1 - \tilde{v}_1 \leq 0$, $\tau_2 - \tilde{v}_2 \geq 0$, the second term dominates the right hand side; in either case we do not need the factor $1/2$. Finally if $\tau_1 - \tilde{v}_1 \geq 0$, $\tau_2 - \tilde{v}_2 \leq 0$, we distinguish the two cases $\tau_1 - \tilde{v}_1$ greater or smaller than $\tilde{v}_2 - \tau_2$: in the former case the second term dominates the right hand side, because $|-\tilde{v}_1 + \tilde{v}_2 - 2\tau_1 - \tau_2| \geq 2(\tau_1 - \tilde{v}_1) - (\tilde{v}_2 - \tau_2)$, and in the latter the first term dominates the right hand side, because $|-\tilde{v}_1 - 2\tilde{v}_2 + \tau_1 + 2\tau_2| \geq 2(\tilde{v}_2 - \tau_2) - (\tau_1 - \tilde{v}_1)$. Finally the case $j = 2, r = 1$ amounts to showing

$$|-\tilde{v}_1 + \tilde{v}_2 + \tau_1 + 2\tau_2| + |2\tilde{v}_1 + \tilde{v}_2 - 2\tau_1 - \tau_2| \geq \frac{1}{2}\big(|\tau_1 - \tilde{v}_1| + |\tau_2 - \tilde{v}_2|\big)$$

which can be seen as above after interchanging indices.

Finally we need to treat the case $0 \neq |v_1| < \varepsilon$ and $|v_2| \gg |\tau_1| + |\tau_2|$. Here the condition (3.4) is empty, and if τ_1, τ_2 are sufficiently large, the condition (3.2) is also empty, so we only need to show the upper bound (3.1) for $X_1 = X_2 = 1$. We return to (3.9) and partition the 6 terms (j, r) into three pairs

$$\big\{(3, 2), (3, 1)\big\}, \quad \big\{(2, 3), (1, 3)\big\}, \quad \big\{(2, 1), (1, 2)\big\}.$$

The contribution of the first pair is

$$\sum_{n,m=0}^{\infty} \frac{(-1)^{n+m}}{n!m!}$$

$$\times \frac{\Gamma(\frac{3v_2}{2} - n)\Gamma(\frac{3v_0}{2} - n)\Gamma(\frac{3v_2}{2} - m)\widehat{f_1}(-1 - v_1 - 2v_2 + s_1 + 2n)}{\Gamma(\frac{3v_2}{2} - n - m)\prod_{j=0}^{2}|\Gamma(\frac{1}{2} + \frac{3}{2}i\Im v_j)|4\pi^{-\frac{3}{2}+v_1+3v_2}}$$

$$\times \left(\frac{\Gamma(\frac{-3v_1}{2} - m)\widehat{f_2}(-1 + v_1 - v_2 + it_2 + 2m)}{\pi^{-v_1}}\right.$$

$$\left. + \frac{\Gamma(\frac{3v_1}{2} - m)\widehat{f_2}(-1 - 2v_1 - v_2 + it_2 + 2m)}{\pi^{2v_1}}\right).$$

For $|\nu_1| < \varepsilon$ the second line can be bounded by the mean value theorem. Then we use the functional equation $s\Gamma(s) = \Gamma(s+1)$ of the Gamma-function in connection with Stirling's formula as before and bound the preceding display by

$$\ll_{A,f_1,f_2} \left(|\alpha_1 + it_1| + |\nu_2 + it_2|\right)^{-A} \prod_{n=0}^{2} \left(1 + |\nu_n|\right)^{-\frac{1}{2}}$$

and argue as before. The same argument with different indices works for the pair $\{(2,3),(1,3)\}$. The last pair is only a small variation; its contribution is given by

$$\sum_{n,m=0}^{\infty} \frac{(-1)^{n+m}\pi^{3/2}}{n!m!4\prod_{j=0}^{2}|\Gamma(\frac{1}{2} + \frac{3}{2}i\Im\nu_j)|}$$
$$\left(\widehat{f_1}(-1 - \nu_1 + \nu_2 + it_1 + 2n)\widehat{f_2}(-1 - 2\nu_1 - \nu_2 + it_2 + 2m)\right.$$
$$\times \frac{\Gamma(\frac{-3\nu_2}{2} - n)\Gamma(\frac{3\nu_1}{2} - n)\Gamma(\frac{3\nu_0}{2} - m)\Gamma(\frac{3\nu_1}{2} - m)}{\pi^{3\nu_1}\Gamma(\frac{3\nu_1}{2} - n - m)}$$
$$+ \widehat{f_1}(-1 + 2\nu_1 + \nu_2 + it_1 + 2n)\widehat{f_2}(-1 + \nu_1 - \nu_2 + it_2 + 2m)$$
$$\left.\times \frac{\Gamma(\frac{-3(\nu_2+\nu_1)}{2} - n)\Gamma(\frac{-3\nu_1}{2} - n)\Gamma(\frac{3(\nu_0-\nu_1)}{2} - m)\Gamma(\frac{-3\nu_1}{2} - m)}{\pi^{-3\nu_1}\Gamma(\frac{-3\nu_1}{2} - n - m)}\right).$$

For small ν_1, this can again be estimated by the mean value theorem giving the crude bound

$$\ll_{A,f_1,f_2} \left(|\nu_2 + it_1| + |\nu_2 - it_2|\right)^{-A}\left(1 + |\nu_2|\right)^{O(1)}$$

which is admissible for (3.1). This completes the proof of the proposition under the additional assumption that $\alpha_1, \alpha_2, \alpha_3$ are pairwise distinct, that is $\nu_1 \neq 0$. The case $\nu_1 = 0$ follows by continuity. $\qquad\square$

An inspection of the proof, in particular (3.9)–(3.13), shows that for τ_1, τ_2 sufficiently large one has

$$\left| I(v_1, v_2, t_1, t_2, 1, 1) \right|^2$$

$$\sim \frac{(2\pi)^3}{3^3 |v_0 v_1 v_2|} \left| \widehat{f_1}(-1 + i\tau_1 - v_1 + 2i\tau_2 - 2v_2) \right.$$

$$\left. \times \widehat{f_2}(-1 + 2i\tau_1 - 2v_1 + i\tau_2 - v_2) \right|^2 \tag{3.14}$$

for $v_1, v_2 \in i\mathbb{R}$ in a neighborhood of $i\tau_1, i\tau_2$, respectively, and it is very small outside this region.

4 Maass forms

Let $\Gamma = SL_3(\mathbb{Z})$. We denote by

$$P = \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & * \end{pmatrix} \subseteq \Gamma, \qquad P_1 = \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 1 \end{pmatrix} \subseteq \Gamma$$

the maximal parabolic subgroup, and by

$$U_3 = \begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix} \subseteq \Gamma$$

the standard unipotent group. Analogously, let $U_2 := \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \subseteq SL_2(\mathbb{Z})$.

A Maaß cusp form $\phi : \Gamma \backslash \mathfrak{h}^3 \to \mathbb{C}$ with spectral parameters v_1, v_2 (that is, of type $(1/3 + v_1, 1/3 + v_2)$ in the notation of [17]) for the group Γ has a Fourier expansion of the type

$$\phi(z) = \sum_{m_1=1}^{\infty} \sum_{m_2 \neq 0} \frac{A_\phi(m_1, m_2)}{|m_1 m_2|}$$

$$\times \sum_{\gamma \in U_2 \backslash SL_2(\mathbb{Z})} \mathcal{W}_{v_1, v_2}^{\mathrm{sgn}(m_2)} \left(\left(\begin{pmatrix} |m_1 m_2| & & \\ & m_1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} \gamma & \\ & 1 \end{pmatrix} z \right) c_{v_1, v_2} \right)$$

$$\tag{4.1}$$

with $\mathcal{W}_{v_1, v_2}^{\pm}$ as in (2.10) and c_{v_1, v_2} as in (2.13). The Fourier coefficients are given by

$$\int_0^1 \int_0^1 \int_0^1 \phi(z) e(-m_1 x_1 - m_2 x_2) \, dx_1 \, dx_2 \, dx_3$$

$$= \frac{A_\phi(m_1, m_2)}{|m_1 m_2|} \tilde{W}_{v_1, v_2}(m_1 y_1, |m_2| y_2).$$

We have

$$A_\phi(m_1, m_2) = A_\phi(m_1, -m_2),$$

see [17, Proposition 6.3.5]. Hence one can alternatively write the Fourier expansion as a sum over $m_1, m_2 \geq 1$, $\gamma \in U_2 \backslash GL_2(\mathbb{Z})$. We will use this observation in the proof of Lemma 1.

It is expected that ν_1, ν_2 are imaginary, but we certainly know that (2.6)–(2.8) hold. If ϕ is an eigenfunction of the Hecke algebra (see [17, Sect. 6.4]), we define its L-function by $L(\phi, s) := \sum_m A_\phi(1, m) m^{-s}$, and the Rankin–Selberg L-function by

$$L(\phi \times \tilde{\phi}, s) := \zeta(3s) \sum_{m_1, m_2} \frac{|A_\phi(m_1, m_2)|^2}{m_1^{2s} m_2^s}. \tag{4.2}$$

It follows from [25, Theorem 2] or [6, Corollary 2] that the coefficients are essentially bounded on average, uniformly in ν:

$$\sum_{m \leq x} |A_\phi(m, 1)|^2 \ll x \left(x(1 + |\nu_1| + |\nu_2|)\right)^\varepsilon. \tag{4.3}$$

The space of cusp forms is equipped with an inner product

$$\langle \phi_1, \phi_2 \rangle := \int_{\Gamma \backslash \mathfrak{h}^3} \phi_1(z) \overline{\phi_2(z)} \, dx_1 \, dx_2 \, dx_3 \, \frac{dy_1 \, dy_2}{(y_1 y_2)^3}.$$

It is known that $L(\phi \times \tilde{\phi}, s)$ can be continued holomorphically to $\mathbb{C}$ with the exception of a simple pole at $s = 1$ whose residue is proportional to $\|\phi\|^2$ [17, Theorem 7.4.9]. The proportionality constant is given in the next lemma.

Lemma 1 *For a Hecke eigenform ϕ as in* (4.1) *with $A_\phi(1, 1) = 1$ we have* $\|\phi\|^2 \asymp \mathrm{res}_{s=1} L(\phi \times \tilde{\phi}, s)$.

Remark 3 This lemma shows that the normalization of the Whittaker functions $\tilde{W}_{\nu_1, \nu_2}$ is well chosen in the sense that an arithmetically normalized cusp form $A_\phi(1, 1) = 1$ should roughly have norm 1. The main point is that $\tilde{W}_{\nu_1, \nu_2}$ has roughly norm 1 with respect to the inner product

$$(f, g) := \int_0^\infty \int_0^\infty f(y_1, y_2) \overline{g(y_1, y_2)} \det \begin{pmatrix} y_1 y_2 & & \\ & y_1 & \\ & & 1 \end{pmatrix} \frac{dy_1 \, dy_2}{(y_1 y_2)^3}.$$

Proof This is standard Rankin–Selberg theory. We use the maximal parabolic Eisenstein series

$$E(z, s; \mathbf{1}) := \sum_{\gamma \in P \backslash \Gamma} \det(\gamma z)^s = \frac{1}{2} \sum_{\gamma \in P_1 \backslash \Gamma} \det(\gamma z)^s, \quad \Re s > 1.$$

It follows from (5.7) below that

$$\|\phi\|^2 = \frac{3\zeta(3)}{2\pi} \operatorname*{res}_{s=1} \langle \phi, \phi E(., \bar{s}, \mathbf{1}) \rangle.$$

We follow the unfolding argument of [17, pp. 227–229] and [16, Sect. 3]. Unfolding the Eisenstein series, we see

$$\langle \phi, \phi E(., \bar{s}, \mathbf{1}) \rangle = \frac{1}{2} \int_{P_1 \backslash \mathfrak{h}^3} |\phi(z)|^2 (y_1^2 y_2)^s \, dx_1 \, dx_2 \, dx_3 \, \frac{dy_1 \, dy_2}{(y_1 y_2)^3}.$$

Let $\mathcal{F}$ denote a fundamental domain for $\left\{ \begin{pmatrix} 1 & * \\ & 1 & * \\ & & 1 \end{pmatrix} \right\} \backslash \mathfrak{h}^3$, and let $\widetilde{GL}_2(\mathbb{Z}) := \left\{ \begin{pmatrix} \gamma & \\ & 1 \end{pmatrix} \mid \gamma \in GL_2(\mathbb{Z}) \right\} \subseteq GL_3(\mathbb{Z})$. Then $P_1 \backslash \mathfrak{h}^3$ is in 2-to-1 correspondence with $\widetilde{GL}_2(\mathbb{Z}) \backslash \mathcal{F}$. Inserting the Fourier expansion of one factor and unfolding once again, we obtain

$$\langle \phi, \phi E(., \bar{s}, \mathbf{1}) \rangle$$
$$= \sum_{m_1=1}^{\infty} \sum_{m_2=1}^{\infty} \frac{|A_\phi(m_1, m_2)|^2}{|m_1 m_2|^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\tilde{W}_{\nu_1, \nu_2}(m_1 y_1, |m_2| y_2)|^2$$
$$\times (y_1^2 y_2)^s \frac{dy_1 \, dy_2}{(y_1 y_2)^3}$$
$$= \frac{L(\phi \times \tilde{\phi}, s)}{\zeta(3s)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\tilde{W}_{\nu_1, \nu_2}(y_1, y_2)|^2 (y_1^2 y_2)^s \frac{dy_1 \, dy_2}{(y_1 y_2)^3}.$$

The lemma follows now easily from Stade's formula.   □

We are now ready to prove (1.4).

Lemma 2 *For an arithmetically normalized Hecke–Maaß cusp form ϕ with spectral parameters ν_1, ν_2 as above we have*

$$\left((1+|\nu_0|)(1+|\nu_1|)(1+|\nu_2|) \right)^{-1} \ll \|\phi\|^2 \ll_\varepsilon \left((1+|\nu_0|)(1+|\nu_1|)(1+|\nu_2|) \right)^\varepsilon$$

for any $\varepsilon > 0$.

Proof We conclude from Lemma 1 that as in (4.3) the upper bound follows directly from [25, Theorem 2] or [6, Corollary 2]. We proceed to prove the lower bound. The idea is taken from [13, Lemma 4]. We can assume that one of v_1, v_2 is sufficiently large. Since $A_\phi(1, 1) = 1$, we have

$$\tilde{W}_{v_1,v_2}(y_1, y_2) = \int_0^1 \int_0^1 \int_0^1 \phi(z)e(-x_1 - x_2)\, dx_1\, dx_2\, dx_3.$$

for any $y_1, y_2 > 0$. By Cauchy–Schwarz we get

$$\left|\tilde{W}_{v_1,v_2}(y_1, y_2)\right|^2 \leq \left(\int_0^1 \int_0^1 \int_0^1 \left|\phi(z)\right|^2 dx_1\, dx_2\, dx_3\right)^{1/2} \left|\tilde{W}_{v_1,v_2}(y_1, y_2)\right|.$$

Integrating this inequality and using Cauchy–Schwarz again, we find

$$\int_1^\infty \int_1^\infty \left|\tilde{W}_{v_1,v_2}(y_1, y_2)\right|^2 \left(y_1^2 y_2\right)^{1/2} \frac{dy_1\, dy_2}{y_1^3 y_2^3}$$

$$\leq \left(\int_1^\infty \int_1^\infty \int_0^1 \int_0^1 \int_0^1 \left|\phi(z)\right|^2 \frac{dx_1\, dx_2\, dx_3\, dy_1\, dy_2}{(y_1 y_2)^3}\right)^{1/2}$$

$$\times \left(\int_0^\infty \int_0^\infty \left|\tilde{W}_{v_1,v_2}(y_1, y_2)\right|^2 y_1^2 y_2 \frac{dy_1\, dy_2}{(y_1 y_2)^3}\right)^{1/2}.$$

Since $[1, \infty)^2 \times [0, 1]^3$ is contained in a fundamental domain for $SL_3(\mathbb{Z})\backslash\mathfrak{h}^3$ (see e.g. [21]), we obtain together with Proposition 2 that the right hand side is

$$\ll \|\phi\| \left(\frac{\prod_{j,k=1}^3 \Gamma(\frac{1+\alpha_j+\overline{\alpha}_k}{2})}{\prod_{j=0}^2 |\Gamma(\frac{1}{2} + \frac{3}{2}i\Im v_j)|^2}\right)^{1/2} \asymp \|\phi\|.$$

The left hand side is

$$\geq \int_0^\infty \int_0^\infty \left|\tilde{W}_{v_1,v_2}(y_1, y_2)\right|^2 \left(y_1^2 y_2\right)^{1/2} \frac{dy_1\, dy_2}{y_1^3 y_2^3}$$

$$- \int_0^\infty \int_0^\infty \left|\tilde{W}_{v_1,v_2}(y_1, y_2)\right|^2 \left(y_1^2 y_2\right)^{1/4} \frac{dy_1\, dy_2}{y_1^3 y_2^3}.$$

Again by Proposition 2, this is

$$\asymp \left((1 + |v_0|)(1 + |v_1|)(1 + |v_2|)\right)^{-1/2}$$
$$+ O\left(\left((1 + |v_0|)(1 + |v_1|)(1 + |v_2|)\right)^{-3/4}\right)$$
$$\gg \left((1 + |v_0|)(1 + |v_1|)(1 + |v_2|)\right)^{-1/2}$$

if one of v_1, v_2 is sufficiently large. $\qquad\square$

We briefly discuss cusp forms $u : SL_2(\mathbb{Z}) \backslash \mathfrak{h}^2 \to \mathbb{C}$ for the group $SL_2(\mathbb{Z})$ and spectral parameter $v \in i\mathbb{R}$ (Selberg's eigenvalue conjecture is known for $SL_2(\mathbb{Z})$). A cusp form u has a Fourier expansion

$$u(z) = \sum_{m \neq 0} \frac{\rho_u(m)}{\sqrt{m}} W_v(|m|y) e(mx)$$

where W_v was defined in (2.14). Similarly as in Lemma 1 we see that an arithmetically normalized newform u has norm

$$\|u\|^2 = \int_{SL_2(\mathbb{Z}) \backslash \mathfrak{h}^2} |u(z)|^2 \frac{dx\,dy}{y^2} = 2L(\mathrm{Ad}^2 u, 1). \tag{4.4}$$

Indeed, the Eisenstein series $E(z, s) = \frac{1}{2} \sum_{\gamma \in U_2 \backslash SL_2(\mathbb{Z})} \Im(\gamma z)^s$ has residue $3/\pi$ at $s = 1$, hence by (2.14)

$$\|u\|^2 = \frac{\pi}{3} \mathop{\mathrm{res}}_{s=1} \sum_{m \neq 0} \frac{|\rho(m)|^2}{|m|} \frac{4\pi}{|\Gamma(1/2 + v)|^2}$$

$$\times \int_{-\infty}^{\infty} |m|y K_v(2\pi |m|y) K_{\bar{v}}(2\pi |m|y) y^s \frac{dy}{y^2}$$

$$= \frac{2\pi}{3\zeta(2)} L(\mathrm{Ad}^2 u, 1) \frac{4\pi}{|\Gamma(1/2 + v)|^2}$$

$$\times \frac{\Gamma(\frac{1+v+\bar{v}}{2}) \Gamma(\frac{1-v+\bar{v}}{2}) \Gamma(\frac{1+v-\bar{v}}{2}) \Gamma(\frac{1-v-\bar{v}}{2})}{8}$$

$$= 2L(\mathrm{Ad}^2 u, 1);$$

the evaluation of the integral follows from [20, 6.576.4] or Stade's formula for $GL(2)$. Again we see that an arithmetically normalized cusp form u is essentially L^2-normalized, and W_v has roughly norm one with respect to the inner product

$$(f, g) := \int_0^{\infty} f(y) \overline{g(y)} \det \begin{pmatrix} y & \\ & 1 \end{pmatrix} \frac{dy}{y^2}.$$

5 Eisenstein series

There are three types of Eisenstein series on the space $L^2(\Gamma \backslash \mathfrak{h}^3)$ according to the decomposition

$$L^2(SL_2(\mathbb{Z}) \backslash \mathfrak{h}^2) = L^2_{\mathrm{Eis}} \oplus L^2_{\mathrm{cusp}} \oplus \mathbb{C} \cdot \mathbf{1}.$$

The first term gives rise to *minimal parabolic Eisenstein series*: for $z \in \mathfrak{h}^3$ and $\Re\nu_1, \Re\nu_2$ sufficiently large we define the minimal parabolic Eisenstein series

$$E(z, \nu_1, \nu_2) := \sum_{\gamma \in U_3 \backslash \Gamma} I_{\nu_1, \nu_2}(\gamma z)$$

where I_{ν_1, ν_2} was defined in (2.11). It has meromorphic continuation in ν_1 and ν_2, and its non-zero Fourier coefficients are given by

$$\int_0^1 \int_0^1 \int_0^1 E(z, \nu_1, \nu_2) e(-m_1 x_1 - m_2 x_2) \, dx_1 \, dx_2 \, dx_3$$

$$= \frac{A_{(\nu_1, \nu_2)}(m_1, m_2)}{|m_1 m_2|} \frac{W_{\nu_1, \nu_2}(m_1 y_1, |m_2| y_2)}{\zeta(1 + 3\nu_0)\zeta(1 + 3\nu_1)\zeta(1 + 3\nu_2)}$$

$$= \frac{A_{(\nu_1, \nu_2)}(m_1, m_2)}{|m_1 m_2|} \frac{\tilde{W}_{\nu_1, \nu_2}(m_1 y_1, |m_2| y_2) c_{\nu_1, \nu_2}^{-1}}{\zeta(1 + 3\nu_0)\zeta(1 + 3\nu_1)\zeta(1 + 3\nu_2)} \tag{5.1}$$

(cf. (2.12) and (2.13) for the notation) where

$$A_{(\nu_1, \nu_2)}(m_1, m_2) = |m_1|^{\nu_1 + 2\nu_2} |m_2|^{2\nu_1 + \nu_2} \sigma_{-3\nu_2, -3\nu_1}\big(|m_1|, |m_2|\big)$$

and $\sigma_{\nu_1, \nu_2}(m_1, m_2)$ is the multiplicative function defined by

$$\sigma_{\nu_1, \nu_2}\big(p^{k_1}, p^{k_2}\big) = p^{-\nu_2 k_1} \frac{\left| \begin{pmatrix} 1 & p^{\nu_2(k_1 + k_2 + 2)} & p^{(\nu_1 + \nu_2)(k_1 + k_2 + 2)} \\ 1 & p^{\nu_2(k_1 + 1)} & p^{(\nu_1 + \nu_2)(k_1 + 1)} \\ 1 & 1 & 1 \end{pmatrix} \right|}{\left| \begin{pmatrix} 1 & p^{2\nu_2} & p^{2(\nu_1 + \nu_2)} \\ 1 & p^{\nu_2} & p^{\nu_1 + \nu_2} \\ 1 & 1 & 1 \end{pmatrix} \right|}.$$

This is a combination of [7, (6.5), (6.7), (6.8), (7.3), Theorem 7.2]. An alternative description is given as follows: $A_{(\nu_1, \nu_2)}(m_1, m_2)$ is defined by

$$A_{(\nu_1, \nu_2)}(1, m) = \sum_{d_1 d_2 d_3 = m} d_1^{\alpha_1} d_2^{\alpha_2} d_3^{\alpha_3}$$

and the symmetry and Hecke relation

$$A_{(\nu_1, \nu_2)}(m, 1) = \overline{A_{(\nu_1, \nu_2)}(1, m)} = A_{(\nu_2, \nu_1)}(1, m),$$

$$A_{(\nu_1, \nu_2)}(n, m) = \sum_{d | (n, m)} \mu(d) A_{(\nu_1, \nu_2)}\left(\frac{n}{d}, 1\right) A_{(\nu_1, \nu_2)}\left(1, \frac{m}{d}\right), \tag{5.2}$$

cf. [17, Theorems 6.4.11 and 10.8.6] and note his different choice of the I-function.

Next we define *maximal parabolic Eisenstein series*. Let $s \in \mathbb{C}$ have sufficiently large real part and let $u : SL_2(\mathbb{Z})\backslash\mathfrak{h}^2 \to \mathbb{C}$ be a Hecke–Maaß cusp form with $\|u\| = 1$, spectral parameter $\nu \in i\mathbb{R}$ and Hecke eigenvalues $\lambda_u(m)$. Then we define

$$E(z, s; u) := \sum_{\gamma \in P\backslash\Gamma} \det(\gamma z)^s u\big(\pi(\gamma z)\big) \qquad (5.3)$$

where

$$\pi : \mathfrak{h}^3 \to \mathfrak{h}^2, \qquad \begin{pmatrix} y_1 y_2 & x_2 y_1 & x_3 \\ & y_1 & x_1 \\ & & 1 \end{pmatrix} \mapsto \begin{pmatrix} y_2 & x_2 \\ & 1 \end{pmatrix}$$

is the restriction to the upper left corner. It has a meromorphic continuation in s, and as the minimal parabolic Eisenstein series it is an eigenform of all Hecke operators; in particular for $s = 1/2 + \mu$ it is an eigenform of $T(1, m)$ with eigenvalue

$$B_{(\mu,u)}(1, m) = \sum_{d_1 d_2 = |m|} \lambda_u(d_1) d_1^{-\mu} d_2^{2\mu},$$

see [17, Proposition 10.9.3]. We extend this definition to all pairs of integers by the Hecke relations (5.2). Coupling this with [17, Proposition 10.9.1], we conclude that the non-zero Fourier coefficients

$$\int_0^1 \int_0^1 \int_0^1 E(z, 1/2 + \mu; u)e(-m_1 x_1 - m_2 x_2)\, dx_1\, dx_2\, dx_3$$

are proportional to

$$\frac{B_{(\mu,u)}(m_1, m_2)}{|m_1 m_2|} \tilde{W}_{\mu - \frac{1}{3}\nu, \frac{2}{3}\nu}(m_1 y_1, |m_2| y_2), \qquad (5.4)$$

and the proportionality constant is

$$\frac{c}{L(u, 1 + 3\mu)L(\mathrm{Ad}^2 u, 1)^{1/2}} \qquad (5.5)$$

for some absolute non-zero constant c. This can be seen by setting $m_1 = m_2 = 1$ and comparing with [32, Theorem 7.1.2] in the special case $G = GL(3)$, $M = GL(2) \times GL(1)$, $m = 1$, $s = 3\mu$ and observing (4.4).

A degenerate case of (5.3) occurs if we choose ϕ to be the constant function (of course, this is not a cusp form). For $\Re s$ sufficiently large and $z \in \mathfrak{h}^3$ let

$$E(z, s, \mathbf{1}) := \sum_{\gamma \in P\backslash\Gamma} \det(\gamma z)^s. \qquad (5.6)$$

This function has a meromorphic continuation to all $s \in \mathbb{C}$, and it has a simple pole at $s = 1$ with constant residue

$$\operatorname*{res}_{s=1} E(z, s, \mathbf{1}) = \frac{1}{3}\left(\Gamma(3/2)\zeta(3)\pi^{-3/2}\right)^{-1} = \frac{2\pi}{3\zeta(3)}, \qquad (5.7)$$

see [16, Corollary 2.5].[6] As the constant function on $SL_2(\mathbb{Z})\backslash\mathfrak{h}^2$ is the residue of an Eisenstein series on $SL_2(\mathbb{Z})\backslash\mathfrak{h}^2$, the Eisenstein series (5.6) is a residue of a minimal parabolic Eisenstein series and has only degenerate terms in its Fourier expansion.

6 Kloosterman sums

As usual we write

$$S(m, n, c) := \sum_{d \,(\mathrm{mod}\, c)}^* e\left(\frac{md + n\bar{d}}{c}\right)$$

for the standard Kloosterman sum. We introduce now $GL(3)$ Kloosterman sums; the following account is taken from [8].

For $n_1, n_2, m_1, m_2 \in \mathbb{Z}$, $D_1, D_2 \in \mathbb{N}$ we define

$$\begin{aligned} &S(m_1, m_2, n_1, n_2, D_1, D_2) \\ &:= \sum_{\substack{B_1, C_1 \,(\mathrm{mod}\, D_1) \\ B_2, C_2 \,(\mathrm{mod}\, D_2) \\ (D_1, B_1, C_1)=(D_2, B_2, C_2)=1 \\ D_1 C_2 + B_1 B_2 + C_1 D_2 \equiv 0 \,(D_1 D_2)}} e\left(\frac{m_1 B_1 + n_1(Y_1 D_2 - Z_1 B_2)}{D_1}\right) \\ &\qquad \times e\left(\frac{m_2 B_2 + n_2(Y_2 D_1 - Z_2 B_1)}{D_2}\right) \end{aligned} \qquad (6.1)$$

where Y_1, Y_2, Z_1, Z_2 are chosen such that

$$Y_1 B_1 + Z_1 C_1 \equiv 1 \;(\mathrm{mod}\; D_1), \qquad Y_2 B_2 + Z_2 C_2 \equiv 1 \;(\mathrm{mod}\; D_2).$$

It can be shown that this expression is well-defined [8, Lemmas 4.1, 4.2]. Clearly it depends only on $m_1, n_1 \,(\mathrm{mod}\, D_1)$ and $m_2, n_2 \,(\mathrm{mod}\, D_2)$, and satisfies [8, Properties 4.4, 4.5]

$$\begin{aligned} S(m_1, m_2, n_1, n_2, D_1, D_2) &= S(m_2, m_1, n_2, n_1, D_2, D_1) \\ &= S(n_1, n_2, m_1, m_2, D_1, D_2). \end{aligned} \qquad (6.2)$$

---

[6]Note that the Eisenstein series in [16, p. 164] differs by a factor two from our definition. In [17, Theorem 7.4.4] our definition is used, but the factor $1/2$ seems to have got lost in the last display of p. 224 and the following argument.

Moreover, if $p_1 p_2 \equiv q_1 q_2 \equiv 1 \pmod{D_1 D_2}$, then [8, Property 4.3]

$$S(p_1 m_1, p_2 m_2, q_1 m_1, q_2 m_2, D_1, D_2) = S(m_1, m_2, m_1, m_2, D_1, D_2).$$

Finally we have the factorization rule [8, Property 4.7]

$$S\big(m_1, m_2, n_1, n_2, D_1 D_1', D_2 D_2'\big)$$
$$= S\big(\overline{D_1'}^2 D_2' m_1, \overline{D_2'}^2 D_1' m_2, n_1, n_2, D_1, D_2\big)$$
$$\times S\big(\overline{D_1}^2 D_2 m_1, \overline{D_2}^2 D_1 m_2, n_1, n_2, D_1', D_2'\big) \tag{6.3}$$

whenever $(D_1 D_2, D_1' D_2') = 1$ and inverses are taken with respect to the product of the respective moduli, that is,

$$\overline{D_1} D_1 \equiv \overline{D_2} D_2 \equiv 1 \pmod{D_1' D_2'}, \qquad \overline{D_1'} D_1' \equiv \overline{D_2'} D_2' \equiv 1 \pmod{D_1 D_2}.$$

This implies in particular

$$S(m_1, m_2, n_1, n_2, D_1, D_2) = S(D_2 m_1, n_1, D_1) S(D_1 m_2, n_2, D_2),$$
$$(D_1, D_2) = 1. \tag{6.4}$$

For a prime p and $l \geq 1$ we have [8, Property 4.10]

$$S\big(m_1, m_2, n_1, n_2, p, p^l\big) = S(n_1, 0, p) S\big(m_2, n_2 p, p^l\big)$$
$$+ S(m_1, 0, p) S\big(n_2, m_2 p, p^l\big)$$
$$+ \delta_{l=1}(p - 1). \tag{6.5}$$

Essentially best possible ("Weil-type") upper bounds for $S(m_1, m_2, n_1, n_2, D_1, D_2)$ have been given by Stevens [34, Theorem 5.1]. The dependence on m_1, m_2, n_1, n_2 has been worked out in [9, p. 39].

Lemma 3 *For any integers* $n_1, n_2, m_1, m_2 \in \mathbb{Z} \setminus \{0\}$, $D_1, D_2 \in \mathbb{N}$ *and any* $\varepsilon > 0$ *we have*

$$S(m_1, m_2, n_1, n_2, D_1, D_2)$$
$$\ll (D_1 D_2)^{1/2+\varepsilon} \big((D_1, D_2)(m_1 n_2, [D_1, D_2])(m_2 n_1, [D_1, D_2])\big)^{1/2}$$

where $[., .]$ *denotes the least common multiple. In particular,*

$$\sum_{D_1 \leq X_1} \sum_{D_2 \leq X_2} \big|S(m, \pm 1, n, 1, D_1, D_2)\big| \ll (X_1 X_2)^{3/2+\varepsilon} (n, m)^\varepsilon \tag{6.6}$$

if $mn \neq 0$. *All implied constants depend only on* ε.

Proof It remains to show (6.6) which is straightforward:

$$\sum_{D_1 \leq X_1} \sum_{D_2 \leq X_2} \left| S(m, \pm 1, n, 1, D_1, D_2) \right|$$

$$\ll (X_1 X_2)^{1/2+\varepsilon} \sum_{d_1 | m} \sum_{d_2 | n} (d_1 d_2)^{1/2} \sum_{D} \sum_{\substack{D_1 \leq X_1/D \\ [d_1, d_2] | DD_1 D_2}} \sum_{D_2 \leq X_2/D_2} D^{1/2}$$

$$\ll (X_1 X_2)^{3/2+\varepsilon} \sum_{d_1 | m} \sum_{d_2 | n} \frac{(d_1 d_2)^{1/2}}{[d_1, d_2]} \ll (X_1 X_2)^{3/2+\varepsilon} (n, m)^{\varepsilon}. \qquad \square$$

Next we define a different class of Kloosterman sums: If $D_1 \mid D_2$, we put

$$\tilde{S}(m_1, n_1, n_2, D_1, D_2)$$

$$:= \sum_{\substack{C_1(D_1), C_2(D_2) \\ (C_1, D_1) = (C_2, D_2/D_1) = 1}} e\left(\frac{m_1 C_1 + n_1 \overline{C_1} C_2}{D_1}\right) e\left(\frac{n_2 \overline{C_2}}{D_2/D_1}\right).$$

Again this sum depends only on $m_1, n_1 \pmod{D_1}$ and $n_2 \pmod{D_2/D_1}$, and for $p_1 q_1 \equiv 1 \pmod{D_1}$, $p_2 q_2 \equiv 1 \pmod{D_2}$ we have [8, Property 4.13]

$$\tilde{S}(m_1 p_1, n_1 q_1 p_2, n_2 q_2, D_1, D_2) = \tilde{S}(m_1, n_1, n_2, D_1, D_2).$$

We have the factorization rule [8, Property 4.15]

$$\tilde{S}(m_1, n_1, n_2, D_1 D_1', D_2 D_2')$$

$$= \tilde{S}(m_1 \overline{D_1'}, n_1 D_2', n_2 \overline{D_2'}^2, D_1, D_2) \tilde{S}(m_1 \overline{D_1}, n_1 D_2, n_2 \overline{D_2}^2, D_1', D_2')$$

whenever $(D_2, D_2') = 1$ and all terms are defined. Finally we have for a prime number p and $1 \leq l < k$ [8, Properties 4.16, 4.17]

$$\tilde{S}(m_1, n_1, n_2, p^l, p^l) = \begin{cases} p^{2l} - p^{2l-1}, & p^l \mid m_1, \ p^l \mid n_1 \\ -p^{2l-1}, & p^{l-1} \| m_1, \ p^l \mid n_1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$\tilde{S}(m_1, n_1, n_2, p^l, p^k) = 0$$

unless

- $k < 2l$ and $p^{2l-k} \mid n_1$, or
- $k = 2l$, or
- $k > 2l$ and $p^{k-2l} \mid n_2$.

In particular

$$\tilde{S}(m_1, n_1, n_2, D_1, D_2) = 0 \quad \text{unless} \quad D_1^2 \mid n_1 D_2.$$

A sharp bound was proved by Larsen [8, Appendix]:

$$\tilde{S}(m_1, n_1, n_2, D_1, D_2) \ll \min\big((n_2, D_2/D_1)D_1^2, (m_1, n_1, D_1)D_2\big)(D_1 D_2)^\varepsilon. \tag{6.7}$$

7 Poincaré series

Let $F : (0, \infty)^2 \to \mathbb{C}$ be a smooth compactly supported function (or sufficiently rapidly decaying at 0 and ∞ in both variables). Let

$$F^*(y_1, y_2) := F(y_2, y_1).$$

For two positive integers m_1, m_2 and $z \in \mathfrak{h}^3$ let $\mathcal{F}_{m_1,m_2}(z) := e(m_1 x_2 + m_2 x_2)F(m_1 y_1, m_2 y_2)$. Then we consider the following Poincaré series:

$$P_{m_1,m_2}(z) := \sum_{\gamma \in U_3 \backslash \Gamma} \mathcal{F}_{m_1,m_2}(\gamma z).$$

Unfolding shows

$$\langle \phi, P_{m_1,m_2} \rangle = \int_{U_3 \backslash \mathfrak{h}^3} \phi(z)\overline{\mathcal{F}_{m_1,m_2}(z)}\, dx_1\, dx_2\, dx_3 \frac{dy_1\, dy_2}{(y_1 y_2)^3}$$

$$= \int_0^\infty \int_0^\infty \int_0^1 \int_0^1 \int_0^1 \phi(z)e(-m_1 x_1 - m_2 x_2)\, dx_1\, dx_2\, dx_3$$

$$\times \overline{F(m_1 y_1, m_2 y_2)} \frac{dy_1\, dy_2}{(y_1 y_2)^3} \tag{7.1}$$

for an arbitrary automorphic form ϕ. In particular, if ϕ is given as in (4.1), we find

$$\langle \phi, P_{m_1,m_2} \rangle$$

$$= \int_0^\infty \int_0^\infty \frac{A_\phi(m_1, m_2)}{m_1 m_2} \tilde{W}_{\nu_1, \nu_2}(m_1 y_1, m_2 y_2)\overline{F(m_1 y_1, m_2 y_2)} \frac{dy_1\, dy_2}{(y_1 y_2)^3}$$

$$= m_1 m_2 A_\phi(m_1, m_2) \int_0^\infty \int_0^\infty \tilde{W}_{\nu_1, \nu_2}(y_1, y_2)\overline{F(y_1, y_2)} \frac{dy_1\, dy_2}{(y_1 y_2)^3}. \tag{7.2}$$

We want to apply (7.1) also with $\phi = P_{n_1, n_2}$ where n_1, n_2 is another pair of positive integers. The Fourier expansion of P_{n_1, n_2} has been computed explicitly in [8, Theorem 5.1]: For $m_1, m_2 > 0$ we have

$$\int_0^1 \int_0^1 \int_0^1 P_{n_1, n_2}(z) e(-m_1 x_1 - m_2 x_2) \, dx_1 \, dx_2 \, dx_3 = S_1 + S_{2a} + S_{2b} + S_3, \tag{7.3}$$

where

$$S_1 = \delta_{m_1, n_1} \delta_{m_2, n_2} F(n_1 y_1, n_2 y_2),$$

$$S_{2a} = \sum_{\epsilon = \pm 1} \sum_{\substack{D_1 | D_2 \\ m_2 D_1^2 = n_1 D_2}} \tilde{S}(\epsilon m_1, n_1, n_2, D_1, D_2) \tilde{J}_F(y_1, y_2, \epsilon m_1, n_1, n_2, D_1, D_2),$$

$$S_{2b} = \sum_{\epsilon = \pm 1} \sum_{\substack{D_2 | D_1 \\ m_1 D_2^2 = n_2 D_1}} \tilde{S}(\epsilon m_2, n_2, n_1, D_2, D_1)$$

$$\times \tilde{J}_{F^*}(y_2, y_1, \epsilon m_2, n_2, n_1, D_2, D_1),$$

$$S_3 = \sum_{\epsilon_1, \epsilon_2 = \pm 1} \sum_{D_1, D_2} S(\epsilon_1 m_1, \epsilon_2 m_2, n_1, n_2, D_1, D_2)$$

$$\times J(y_1, y_2, \epsilon_1 m_1, \epsilon_2 m_2, n_1, n_2, D_1, D_2).$$

The Kloosterman sums have been defined in Sect. 6 and the weight functions are given as follows:

$$\tilde{J}_F(y_1, y_2, m_1, n_1, n_2, D_1, D_2)$$

$$= y_1^2 y_2 \int_{\mathbb{R}^2} e(-m_1 x_1 y_1) e\left(\frac{n_1 D_2 y_2}{D_1^2} \cdot \frac{x_1 x_2}{x_1^2 + 1}\right)$$

$$\times e\left(\frac{n_2 D_1}{y_1 y_2 D_2^2} \cdot \frac{x_2}{x_1^2 + x_2^2 + 1}\right)$$

$$\times F\left(\frac{n_1 D_2 y_2}{D_1^2} \cdot \frac{\sqrt{x_1^2 + x_2^2 + 1}}{x_1^2 + 1}, \frac{n_2 D_1}{y_1 y_2 D_2^2} \cdot \frac{\sqrt{x_1^2 + 1}}{x_1^2 + x_2^2 + 1}\right) dx_1 \, dx_2, \tag{7.4}$$

$$J(y_1, y_2, m_1, m_2, n_1, n_2, D_1, D_2)$$

$$= (y_1 y_2)^2 \int_{\mathbb{R}^3} e(-m_1 x_1 y_1 - m_2 x_2 y_2) e\left(-\frac{n_1 D_2}{D_1^2 y_2} \cdot \frac{x_1 x_3 + x_2}{x_3^2 + x_2^2 + 1}\right)$$

$$\times e\left(-\frac{n_2 D_1}{D_2^2 y_1} \cdot \frac{x_2(x_1 x_2 - x_3) + x_1}{(x_1 x_2 - x_3)^2 + x_1^2 + 1}\right)$$

$$\times F\left(\frac{n_1 D_2}{D_1^2 y_2} \cdot \frac{\sqrt{(x_1 x_2 - x_3)^2 + x_1^2 + 1}}{x_3^2 + x_2^2 + 1},\right.$$

$$\left.\frac{n_2 D_1}{D_2^2 y_1} \cdot \frac{\sqrt{x_3^2 + x_2^2 + 1}}{(x_1 x_2 - x_3)^2 + x_1^2 + 1}\right) dx_1 \, dx_2 \, dx_3. \tag{7.5}$$

8 Spectral decomposition

We have the following spectral decomposition theorem [17, Proposition 10.13.1]: If $\phi \in L^2(\Gamma \backslash \mathfrak{h}^3)$ is orthogonal to all residues of Eisenstein series, then

$$\phi = \sum_j \langle \phi, \phi_j \rangle \phi_j + \int_{(0)} \int_{(0)} \langle \phi, E(., \nu_1, \nu_2) \rangle E(., \nu_1, \nu_2) \frac{d\nu_1 \, d\nu_2}{(4\pi i)^2}$$

$$+ \sum_j \int_{(0)} \langle \phi, E(., 1/2 + \mu; u_j) \rangle E(., 1/2 + \mu; u_j) \, d\mu$$

$$+ \frac{1}{2\pi i} \int_{(0)} \langle \phi, E(., 1/2 + \mu; \mathbf{1}) \rangle E(., 1/2 + \mu; \mathbf{1}) \frac{d\mu}{2\pi i}$$

where the first j-sum runs over an orthonormal basis of cusp forms ϕ_j for $SL_3(\mathbb{Z})$ and the second j-sum runs over an orthonormal basis of cusp forms u_j for $SL_2(\mathbb{Z})$.

Therefore we have for 4 positive integers n_1, n_2, m_1, m_2 an equality of the type

$$\frac{\langle P_{n_1,n_2}, P_{m_1,m_2} \rangle}{n_1 n_2 m_1 m_2}$$

$$= \sum_j \frac{\langle P_{n_1,n_2}, \phi_j \rangle \langle \phi_j, P_{m_1,m_2} \rangle}{n_1 n_2 m_1 m_2} + \cdots \quad \text{(continuous spectrum)}.$$

We refer to the right hand side as the spectral side and to the left hand side as the arithmetic side.

We proceed to describe the spectral side and the arithmetic side more precisely. We define an inner product on $L^2((0, \infty)^2, dy_1 \, dy_2/(y_1 y_2)^3)$ by

$$\langle f, g \rangle := \int_0^\infty \int_0^\infty f(y_1, y_2) \overline{g(y_1, y_2)} \frac{dy_1 \, dy_2}{(y_1 y_2)^3}.$$

Let $\{\phi_j\}$ denote an arithmetically normalized ortho*gonal* basis of the space of cusp forms on $L^2(SL_3(\mathbb{Z}) \backslash \mathfrak{h}^3)$ that we assume to be eigenfunctions of the

Hecke algebra with eigenvalues $A_j(m_1, m_2)$. Let $\{u_j\}$ be an arithmetically normalized orthogonal basis of the space of cusp forms on $L^2(SL_2(\mathbb{Z})\backslash\mathfrak{h}^2)$ that we assume to be eigenfunctions of the Hecke algebra with eigenvalues $\lambda_j(m)$ and spectral parameter $\nu_j \in i\mathbb{R}$.

Proposition 4 *Keep the notation developed so far. Let $F : (0, \infty)^2 \to \mathbb{C}$ be a smooth compactly supported function, and let $m_1, m_2, n_1, n_2 \in \mathbb{N}$. Then for some absolute constant $c > 0$ the following equality holds:*

$$
\sum_j \frac{\overline{A_j(n_1, n_2)}A_j(m_1, m_2)}{\|\phi_j\|^2} \left|\langle \tilde{W}_{\nu_1,\nu_2}, F\rangle\right|^2
$$

$$
+ \frac{1}{(4\pi i)^2} \int_{(0)} \int_{(0)} \frac{\overline{A_{(\nu_1,\nu_2)}(n_1, n_2)}A_{(\nu_1,\nu_2)}(m_1, m_2)}{|\zeta(1 + 3\nu_0)\zeta(1 + 3\nu_1)\zeta(1 + 3\nu_2)|^2}
$$

$$
\times \left|\langle \tilde{W}_{\nu_1,\nu_2}, F\rangle\right|^2 d\nu_1 \, d\nu_2
$$

$$
+ \frac{c}{2\pi i} \sum_j \int_{(0)} \frac{\overline{B_{(\mu, u_j)}(n_1, n_2)}B_{(\mu, u_j)}(m_1, m_2)}{|L(u_j, 1 + 3\mu)|^2 L(\mathrm{Ad}^2 u_j, 1)} \left|\langle \tilde{W}_{\mu - \frac{1}{3}\nu_j, \frac{2}{3}\nu_j}, F\rangle\right|^2 d\mu
$$

$$
= \Sigma_1 + \Sigma_{2a} + \Sigma_{2b} + \Sigma_3, \tag{8.1}
$$

where

$$
\Sigma_1 = \delta_{m_1, n_1}\delta_{m_2, n_2}\|F\|^2,
$$

$$
\Sigma_{2a} = \sum_{\epsilon = \pm 1} \sum_{\substack{D_1|D_2 \\ m_2 D_1^2 = n_1 D_2}} \frac{\tilde{S}(\epsilon m_1, n_1, n_2, D_1, D_2)}{D_1 D_2} \tilde{\mathcal{J}}_{\epsilon; F}\left(\sqrt{\frac{n_1 n_2 m_1}{D_1 D_2}}\right),
$$

$$
\Sigma_{2b} = \sum_{\epsilon = \pm 1} \sum_{\substack{D_2|D_1 \\ m_1 D_2^2 = n_2 D_1}} \frac{\tilde{S}(\epsilon m_2, n_2, n_1, D_2, D_1)}{D_1 D_2} \tilde{\mathcal{J}}_{\epsilon; F^*}\left(\sqrt{\frac{n_1 n_2 m_2}{D_1 D_2}}\right), \tag{8.2}
$$

$$
\Sigma_3 = \sum_{\epsilon_1, \epsilon_2 = \pm 1} \sum_{D_1, D_2} \frac{S(\epsilon_1 m_1, \epsilon_2 m_2, n_1, n_2, D_1, D_2)}{D_1 D_2}
$$

$$
\times \mathcal{J}_{\epsilon_1, \epsilon_2}\left(\frac{\sqrt{m_1 n_2 D_1}}{D_2}, \frac{\sqrt{m_2 n_1 D_2}}{D_1}\right).
$$

The weight functions $\tilde{\mathcal{J}}$ and $\mathcal{J}$ are given by

$$
\tilde{\mathcal{J}}_{\epsilon;F}(A) = A^{-2} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty e(-\epsilon A x_1 y_1) e\left(y_2 \cdot \frac{x_1 x_2}{x_1^2 + 1}\right)
$$

$$
\times e\left(\frac{A}{y_1 y_2} \cdot \frac{x_2}{x_1^2 + x_2^2 + 1}\right)
$$

$$
\times F\left(y_2 \cdot \frac{\sqrt{x_1^2 + x_2^2 + 1}}{x_1^2 + 1}, \frac{A}{y_1 y_2} \cdot \frac{\sqrt{x_1^2 + 1}}{x_1^2 + x_2^2 + 1}\right)
$$

$$
\times \overline{F(A y_1, y_2)} \, dx_1 \, dx_2 \frac{dy_1 \, dy_2}{y_1 y_2^2}, \tag{8.3}
$$

$$
\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2)
$$
$$
= (A_1 A_2)^{-2} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty e(-\epsilon_1 A_1 x_1 y_1 - \epsilon_2 A_2 x_2 y_2)
$$

$$
\times e\left(-\frac{A_2}{y_2} \cdot \frac{x_1 x_3 + x_2}{x_3^2 + x_2^2 + 1}\right) e\left(-\frac{A_1}{y_1} \cdot \frac{x_2(x_1 x_2 - x_3) + x_1}{(x_1 x_2 - x_3)^2 + x_1^2 + 1}\right)
$$

$$
\times F\left(\frac{A_2}{y_2} \cdot \frac{\sqrt{(x_1 x_2 - x_3)^2 + x_1^2 + 1}}{x_3^2 + x_2^2 + 1}, \frac{A_1}{y_1} \cdot \frac{\sqrt{x_3^2 + x_2^2 + 1}}{(x_1 x_2 - x_3)^2 + x_1^2 + 1}\right)
$$

$$
\times \overline{F(A_1 y_1, A_2 y_2)} \, dx_1 \, dx_2 \, dx_3 \frac{dy_1 \, dy_2}{y_1 y_2}. \tag{8.4}
$$

Proof The spectral side (8.1) follows from[7] (7.2) in combination with (5.1) and (5.4), (5.5). Note that $E(z, 1/2 + \mu, \mathbf{1})$ does not contribute because it has only degenerate terms in its Fourier expansion.

Upon combining (7.1) and (7.3), we obtain the arithmetic side (8.2) after applying a linear change of variables

$$
y_1 \mapsto \sqrt{\frac{n_1 n_2}{D_1 D_2 m_1}} y_1, \qquad y_2 \mapsto \frac{y_2}{m_2}
$$

---

[7]Even though $\tilde{W}_{\nu_1,\nu_2}$ just fails to be in $L^2((0,\infty)^2, dy_1 dy_2/(y_1 y_2)^3)$, the inner products exist by the decay properties of F.

in (7.4) and observing $m_2 D_1^2 = n_1 D_2$ (and with interchanged indices for Σ_{2b}), and

$$y_1 \mapsto \sqrt{\frac{n_2 D_1}{m_1 D_2^2}} y_1, \qquad y_2 \mapsto \sqrt{\frac{n_1 D_2}{m_2 D_1^2}} y_2$$

in (7.5). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Formally (8.1) and (8.2) resemble the $GL(2)$ Kuznetsov formula, but in its present form it is relatively useless as long as we do not understand the transforms $|\langle \tilde{W}_{\nu_1,\nu_2}, F \rangle|^2$ and $\tilde{\mathcal{J}}, \mathcal{J}$ for a given test function F. The present formulation has the important advantage that the weight functions on the arithmetic side (8.2) do not depend on $n_1, n_2, m_1, m_2, D_1, D_2$ individually, but only in a coupled fashion. This is, of course, a well-known phenomenon in the $GL(2)$ world.

We choose now

$$F(y_1, y_2) = F_{f,X_1,X_2,R_1,R_2,\tau_1,\tau_2}(y_1, y_2)$$
$$:= \left(R_1 R_2 (R_1 + R_2)\right)^{1/2} f(X_1 y_1) f(X_2 y_2) y_1^{i(\tau_1 + 2\tau_2)} y_2^{i(\tau_2 + 2\tau_1)} \quad (8.5)$$

for $X_1, X_2, R_1, R_2 \geq 1$, $\tau_1, \tau_2 \geq 0$, $\tau_1 + \tau_2 \geq 1$ and f a fixed smooth, nonzero, non-negative function with support in $[1, 2]$. Analytic properties of $\langle \tilde{W}_{\nu_1,\nu_2}, F \rangle$ have been obtained in Proposition 3. We summarize some bounds for the weight functions occurring on the arithmetic side in the following proposition.

Proposition 5 *With the notation developed so far, we have*

$$\|F\|^2 \asymp (X_1 X_2)^2 R_1 R_2 (R_1 + R_2). \qquad (8.6)$$

Let $C_1, C_2 \geq 0$, $\varepsilon > 0$. Then

$$\tilde{\mathcal{J}}_{\epsilon;F}(A) \begin{cases} \ll X_1^2 X_2 R_1 R_2 (R_1 + R_2) \left(\frac{1+A^{2/3}}{\tau_1+\tau_2}\right)^{C_1}, \\ = 0, \quad \text{if } A \leq (100 X_1)^{-3/2} + (100 X_1 X_2)^{-3/4}, \end{cases} \quad (8.7)$$

and

$$\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2) \ll (X_1 X_2)^2 R_1 R_2 (R_1 + R_2)$$
$$\times \left(\frac{1 + A_1^{4/3} A_2^{2/3}(X_1 + X_2)}{\tau_1 + \tau_2}\right)^{C_1} \left(\frac{1 + A_2^{4/3} A_1^{2/3}(X_1 + X_2)}{\tau_1 + \tau_2}\right)^{C_2}$$
$$\times \left((X_1 + A_1)(X_2 + A_2)\right)^{\varepsilon}, \qquad (8.8)$$
$$\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2) = 0, \quad \text{if } \min(A_1 A_2^2, A_2 A_1^2) \leq (100 X_1 X_2)^{-3/2}.$$

In the special case when $A_1, A_2 \leq 1$, $X_1 = 1$, $X_2 = X \geq 1$, $R_1 + R_2 \asymp \tau_1 + \tau_2$ this can be improved to

$$
\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2) \ll X^2 R_1 R_2 \left(\frac{1 + A_1^{4/3} A_2^{2/3} X}{\tau_1 + \tau_2} \right)^{C_1}
$$

$$
\times \left(\frac{1 + A_2^{4/3} A_1^{2/3} X}{\tau_1 + \tau_2} \right)^{C_2} \left((R_1 + R_2) X \right)^{\varepsilon}. \qquad (8.9)
$$

Let g be a fixed smooth function with compact support in $(0, \infty)$. Then for $R_1, R_2 \gg 1$ sufficiently large we have

$$
\int_0^\infty \int_0^\infty g\left(\frac{\tau_1}{R_1} \right) g\left(\frac{\tau_2}{R_2} \right) \mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2) \, d\tau_1 \, d\tau_2
$$

$$
\ll (X_1 X_2)^2 R_1 R_2 (R_1 + R_2) \left(\frac{1 + A_1^{4/3} A_2^{2/3} (X_1 + X_2)}{R_1 + R_2} \right)^{C_1}
$$

$$
\times \left(\frac{1 + A_2^{4/3} A_1^{2/3} (X_1 + X_2)}{R_1 + R_2} \right)^{C_2}
$$

$$
\times \left(R_1 R_2 (X_1 + A_1)(X_2 + A_2) \right)^{\varepsilon}. \qquad (8.10)
$$

On the left hand side we have suppressed the dependence of $\mathcal{J}_{\epsilon_1,\epsilon_2}$ on τ_1, τ_2.

Remark 4 The bounds (8.7), (8.8), (8.10) are not best possible, but (8.9) is likely to be best possible. The important feature is that (8.7) and (8.8) effectively bound A_1, A_2 from below, and therefore D_1, D_2 in (8.2) from above. For example, for the contribution of the long Weyl element, we can essentially assume

$$
D_1 \leq \frac{(m_1 n_2)^{1/3} (m_2 n_1)^{2/3}}{\tau_1 + \tau_2}, \qquad D_2 \leq \frac{(m_1 n_2)^{2/3} (m_2 n_1)^{1/3}}{\tau_1 + \tau_2}
$$

if $X_1 = X_2 = 1$. It is instructive to compare this with the $GL(2)$ situation: one can construct a sufficiently nice test function h on the spectral side with essential support on $[T, T + 1]$ such that the integral transforms $h^{\pm}$ in (1.1) are negligible unless $c \leq \frac{(nm)^{1/2}}{T}$.

The bound (8.10) shows that integration over τ_1, τ_2 can be performed at almost no cost, in other words, we save a factor $(R_1 R_2)^{1-\varepsilon}$ compared to trivial integration.

Remark 5 Choosing $f(y_1, y_2) = e^{-2\pi(y_1+y_2)}(y_1 y_2)^{100}$ (say), the two y-integrals in (8.4) can be computed explicitly using [20, 3.471.9], giving two

Bessel-K-functions with general complex arguments. It is not clear how to take advantage of this fact.

Proof Equation (8.6) is clear. We proceed to prove (8.7). Let us write

$$\xi_1 := x_1^2 + 1, \qquad \xi_2 = x_1^2 + x_2^2 + 1$$

in (8.3). The support of f restricts the variables to

$$(X_1 A)^{-1} \le y_1 \le 2(X_1 A)^{-1}, \qquad X_2^{-1} \le y_2 \le 2 X_2^{-1},$$

$$X_2/(2X_1) \le \xi_2^{1/2} \xi_1^{-1} \le 2 X_2/X_1, \qquad (8.11)$$

$$\left(A^2 X_1 X_2^2\right)^{-1} \le \xi_1^{1/2} \xi_2^{-1} \le 8\left(A^2 X_1 X_2^2\right)^{-1}.$$

The second set of conditions in (8.11) implies $\xi_1 \asymp A^{4/3} X_1^2$ and $\xi_2 \asymp A^{8/3}(X_1 X_2)^2$. Hence the second part of (8.7) is clear and a trivial estimation shows

$$\tilde{\mathcal{J}}_{\epsilon;F}(A) \ll R_1 R_2 (R_1 + R_2) X_1^2 X_2.$$

In certain ranges this can be improved by partial integration. We have

$$\tilde{\mathcal{J}}_{\epsilon;F}(A) = \frac{R_1 R_2 (R_1 + R_2)}{A^{2-i(\tau_1-\tau_2)}} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty (x_1^2 + 1)^{-\frac{3}{2} i \tau_2}$$

$$\times \left(x_1^2 + x_2^2 + 1\right)^{-\frac{3}{2} i \tau_1} e(-A y_1 \epsilon x_1)$$

$$\times e\left(\frac{y_2 x_1 x_2}{x_1^2 + 1}\right) e\left(\frac{A x_2}{y_1 y_2 (x_1^2 + x_2^2 + 1)}\right) \overline{f}(X_1 A y_1) \overline{f}(X_2 y_2)$$

$$\times f\left(X_1 y_2 \frac{\sqrt{x_1^2 + x_2^2 + 1}}{x_1^2 + 1}\right)$$

$$\times f\left(\frac{A X_2}{y_1 y_2} \cdot \frac{x_2 + i\sqrt{x_1^2 + 1}}{x_1^2 + x_2^2 + 1}\right) y_1^{-3i(\tau_1+\tau_2)} y_2^{-3i\tau_1} dx_1 \, dx_2 \frac{dy_1 \, dy_2}{y_1 y_2^2}.$$

We can assume that C_1 is an integer. Then C_1 successive integrations by parts with respect to y_1 yield an additional factor

$$\ll_{C_1} \left(\left(\frac{y_1}{\tau_1 + \tau_2}\right)\left(A|x_1| + \frac{A|x_2|}{\xi_2 y_1^2 y_2} + \frac{1}{y_1}\right)\right)^{C_1} \ll_{C_1} \left(\frac{1 + A^{2/3}}{\tau_1 + \tau_2}\right)^{C_1} \quad (8.12)$$

in the support of f.

The bound (8.8) can be shown similarly, but the estimations are a little more involved. Here we write

$$\xi_1 := (x_1 x_2 - x_3)^2 + x_1^2 + 1, \qquad \xi_2 = x_3^2 + x_2^2 + 1 \qquad (8.13)$$

and truncate

$$(X_1 A_1)^{-1} \le y_1 \le 2(X_1 A_1)^{-1}, \qquad (X_2 A_2)^{-1} \le y_2 \le 2(X_2 A_2)^{-1},$$

$$(2 X_1 X_2 A_2^2)^{-1} \le \xi_1^{1/2} \xi_2^{-1} \le 2(X_1 X_2 A_2^2)^{-1}, \qquad (8.14)$$

$$(2 X_1 X_2 A_1^2)^{-1} \le \xi_2^{1/2} \xi_1^{-1} \le 2(X_1 X_2 A_1^2)^{-1}.$$

This implies

$$\xi_2 \asymp \Xi_2 := A_1^{4/3} A_2^{8/3} (X_1 X_2)^2, \qquad \xi_1 \asymp \Xi_1 := A_2^{4/3} A_1^{8/3} (X_1 X_2)^2 \quad (8.15)$$

which yields in particular the second part of (8.8) as well as

$$x_1 \ll A_2^{2/3} A_1^{4/3} X_1 X_2, \qquad x_2 \ll A_1^{2/3} A_2^{4/3} X_1 X_2,$$

$$x_3 \ll A_1^{2/3} A_2^{4/3} X_1 X_2, \qquad x_1 x_2 - x_3 \ll A_2^{2/3} A_1^{4/3} X_1 X_2. \qquad (8.16)$$

For future purposes we study the volume of the set of (x_1, x_2, x_3) defined by (8.15) or by

$$\xi_2 = \Xi_2 (1 + O(1/R_2)), \qquad \xi_1 = \Xi_1 (1 + O(1/R_1)). \qquad (8.17)$$

Lemma 4 *For $\Xi_1, \Xi_2 \ge 1$ and any $\varepsilon > 0$ we have*

$$\iiint_{\substack{x_1, x_2, x_3 \\ \text{satisfying (8.15)}}} dx_1 \, dx_2 \, dx_3 \ll (\Xi_1 \Xi_2)^{1/2+\varepsilon}.$$

Moreover,

$$\iiint_{\substack{x_1, x_2, x_3 \\ \text{satisfying (8.17)}}} dx_1 \, dx_2 \, dx_3 \ll (\Xi_1 \Xi_2 R_1 R_2)^\varepsilon \frac{(\Xi_1 \Xi_2)^{1/2}}{R_1 R_2}.$$

We postpone the proof to the end of this section. A trivial estimation now implies

$$\mathcal{J}_{\epsilon_1, \epsilon_2}(A_1, A_2) \ll \frac{R_1 R_2 (R_1 + R_2)}{(A_1 A_2)^2} \iiint_{\substack{x_1, x_2, x_3 \\ \text{satisfying (8.15)}}} dx_1 \, dx_2 \, dx_3$$

$$\ll R_1 R_2 (R_1 + R_2)(X_1 X_2)^2 \big((X_1 + A_1)(X_2 + A_2)\big)^\varepsilon.$$

Alternatively we write

$$\mathcal{J}_{\epsilon_1,\epsilon_2}(A_1, A_2) = \frac{R_1 R_2 (R_1 + R_2)}{(A_1 A_2)^2 A_2^{i(\tau_1 - \tau_2)} A_1^{i(\tau_2 - \tau_1)}} \int_0^\infty \int_0^\infty (y_1 y_2)^{-3i(\tau_1 + \tau_2)}$$

$$\times \int_{-\infty}^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \xi_1^{-\frac{3}{2}i\tau_1} \xi_2^{-\frac{3}{2}i\tau_2}$$

$$\times e(-\epsilon_1 A_1 x_1 y_1 - \epsilon_2 A_2 x_2 y_2) e\left(-\frac{A_2}{y_2} \cdot \frac{x_1 x_3 + x_2}{x_3^2 + x_2^2 + 1}\right)$$

$$\times e\left(-\frac{A_1}{y_1} \cdot \frac{x_2(x_1 x_2 - x_3) + x_1}{(x_1 x_2 - x_3)^2 + x_1^2 + 1}\right) \overline{f}(X_1 A_1 y_1)$$

$$\times f\left(\frac{X_1 A_2}{y_2} \cdot \frac{\sqrt{(x_1 x_2 - x_3)^2 + x_1^2 + 1}}{x_3^2 + x_2^2 + 1}\right) \overline{f}(X_2 A_2 y_2)$$

$$\times f\left(\frac{X_2 A_1}{y_1} \cdot \frac{\sqrt{x_3^2 + x_2^2 + 1}}{(x_1 x_2 - x_3)^2 + x_1^2 + 1}\right) dx_1 \, dx_2 \, dx_3 \frac{dy_1 \, dy_2}{y_1 y_2} \tag{8.18}$$

using the notation (8.13). We can assume that C_1, C_2 are integers. Integrating by parts C_1 times with respect to y_1 and C_2 times with respect to y_2 introduces an additional factor

$$\ll_{C_1,C_2} \left(\left(\frac{y_1}{\tau_1 + \tau_2}\right)\left(A_1 |x_1| + \frac{A_1 |x_2(x_1 x_2 - x_3) + x_1|}{y_1^2 \xi_1} + \frac{1}{y_1}\right)\right)^{C_1}$$

$$\times \left(\left(\frac{y_2}{\tau_1 + \tau_2}\right)\left(A_2 |x_2| + \frac{A_2 |x_1 x_3 + x_2|}{y_2^2 \xi_2} + \frac{1}{y_2}\right)\right)^{C_2}$$

and (8.8) follows by (8.16) and the same argument that led to (8.7).

The proof of (8.9) is a small variant of the preceding argument. We need to save an additional power of $R_1 + R_2$ which comes from a more careful treatment of the y_1, y_2-integral. Let $\eta > 0$ be small. If $1 + A_1^{4/3} A_2^{2/3} X \le (R_1 + R_2)^{1-\eta}$, we replace C_1 in (8.8) by $C_1 + \eta^{-1}$ saving a factor

$$\left(\frac{1 + A_1^{4/3} A_2^{2/3} X}{R_1 + R_2}\right)^{-1/\eta} \ge R_1 + R_2.$$

The same argument works if $1 + A_1^{2/3} A_2^{4/3} X \le (R_1 + R_2)^{1-\eta}$. In the remaining case

$$1 + A_1^{2/3} A_2^{4/3} X \ge (R_1 + R_2)^{1-\eta} \quad \text{and} \quad 1 + A_2^{2/3} A_1^{4/3} X \ge (R_1 + R_2)^{1-\eta}$$

it is enough to show

$$\mathcal{J}_{\epsilon_1, \epsilon_2}(A_1, A_2) \, d\tau_1 \, d\tau_2 \ll R_1 R_2 X^2 \big((R_1 + R_2) X \big)^\varepsilon; \tag{8.19}$$

then the bound (8.9) follows with $\varepsilon + \eta(C_1 + C_2)$ instead of ε. To this end, we combine as before (8.18) and the first part of Lemma 4, and need to show that the y_1 and y_2 integral in (8.18) are both $\ll (\tau_1 + \tau_2)^{-1/2} \asymp (R_1 + R_2)^{-1/2}$. Our present assumption $X_1 = 1$, $X_2 = X \ge 1$, $A_1, A_2 \le 1$ together with the size constraints (8.14)–(8.16) imply that the y_1 integral is of the form

$$\int_0^\infty y_1^{-3i(\tau_1 + \tau_2)} e(-\epsilon_1 A_1 x_1 y_1) w(A_1 y_1) \frac{dy_1}{y_1} \tag{8.20}$$

where w is a smooth function with support in $[1, 2]$ and $w^{(j)}(y) \ll_j 1$ uniformly in all other variables. We can assume that $\epsilon_1 = \operatorname{sgn}(x_1)$ and $|x_1| \asymp \tau_1 + \tau_2$, otherwise we can save as many powers of $\tau_1 + \tau_2$ as we wish by repeated partial integration. In that case we make another change of variables and re-write (8.20) as

$$\int_0^\infty e\left(\frac{3}{2\pi} (\tau_1 + \tau_2)(y_1 - \log y_1) \right) w\left(\frac{3(\tau_1 + \tau_2)}{2\pi |x_1|} y_1 \right) \frac{dy_1}{y_1}.$$

A standard stationary phase argument bounds this integral by $(\tau_1 + \tau_2)^{-1/2}$: we cut out smoothly the region $y_1 = 1 + O((\tau_1 + \tau_2)^{-1/2})$ which we estimate trivially. For the rest we apply integration by parts. The treatment of the y_2 integral is very similar. Here our assumptions imply that the integral is of the form

$$\int_0^\infty y_2^{-3i(\tau_1 + \tau_2)} e\left(-\frac{A_2}{y_2} \cdot \frac{x_1 x_3 + x_2}{x_3^2 + x_2^2 + 1} \right) \tilde{w}(X A_2 w) \frac{dy_2}{y_2},$$

and the same stationary phase-type argument gives a saving of $(\tau_1 + \tau_2)^{-1/2}$.

Finally we prove (8.10). Let

$$Z := R_1 R_2 (X_1 + A_1)(X_2 + A_2).$$

As before we see that we can assume

$$\begin{aligned}
&1 + A_1^{2/3} A_2^{4/3} (X_1 + X_2) \ge (R_1 + R_2)^{1-\eta} \quad \text{and} \\
&1 + A_2^{2/3} A_1^{4/3} (X_1 + X_2) \ge (R_1 + R_2)^{1-\eta};
\end{aligned} \tag{8.21}$$

otherwise we integrate trivially over τ_1, τ_2. In the situation (8.21) it is enough to show[8]

$$\int_{-\infty}^{\infty} \int_{\infty}^{\infty} g\left(\frac{\tau_1}{R_1}\right) g\left(\frac{\tau_2}{R_2}\right) \mathcal{J}_{\epsilon_1, \epsilon_2}(A_1, A_2) \, d\tau_1 \, d\tau_2$$

$$\ll R_1 R_2 (R_1 + R_2)(X_1 X_2)^2 Z^{\varepsilon}; \tag{8.22}$$

then the bound (8.10) follows with $\varepsilon + \eta(C_1 + C_2)$ instead of ε. In order to show (8.22), we integrate (8.18) explicitly over τ_1, τ_2 and observe that

$$\int_{-\infty}^{\infty} \int_{\infty}^{\infty} g\left(\frac{\tau_1}{R_1}\right) g\left(\frac{\tau_2}{R_2}\right) \frac{A_1^{i(\tau_1 - \tau_2)} A_2^{i(\tau_2 - \tau_1)}}{(y_1 y_2)^{3i(\tau_1 + \tau_2)} \xi_1^{\frac{3}{2}i\tau_1} \xi_2^{\frac{3}{2}i\tau_2}} \, d\tau_1 \, d\tau_2$$

$$= R_1 R_2 \tilde{g}\left(\frac{R_1}{2\pi} \log \frac{A_1}{A_2 (y_1 y_2)^3 \xi_1^{3/2}}\right) \tilde{g}\left(\frac{R_2}{2\pi} \log \frac{A_2}{A_1 (y_1 y_2)^3 \xi_2^{3/2}}\right).$$

Since g is smooth, $\tilde{g}$ is rapidly decaying, and up to a negligible error of Z^{-A} we can restrict ξ_1, ξ_2 to

$$\begin{aligned}
\xi_1 &= \frac{A_1^{2/3}}{A_2^{2/3}(y_1 y_2)^2}\left(1 + O\left(\frac{1}{R_1 Z^{\varepsilon}}\right)\right), \\
\xi_2 &= \frac{A_2^{2/3}}{A_1^{2/3}(y_1 y_2)^2}\left(1 + O\left(\frac{1}{R_2 Z^{\varepsilon}}\right)\right).
\end{aligned} \tag{8.23}$$

We note that $A_1^{2/3}(A_2^{2/3}(y_1 y_2)^2)^{-1} \asymp \Xi_1$, $A_2^{2/3}(A_1^{2/3}(y_1 y_2)^2)^{-1} \asymp \Xi_2$ in the notation of (8.15). Hence a trivial estimate bounds the left hand side of (8.22) by

$$\ll \frac{(R_1 R_2)^2 (R_1 + R_2)}{(A_1 A_2)^2} \underset{\substack{x_1, x_2, x_3 \\ \text{satisfying (8.23)}}}{\int \int \int} dx_1 \, dx_2 \, dx_3$$

and the desired bound follows from the second part of Lemma 4. $\qquad \square$

It remains to prove Lemma 4: the conditions $\xi_1 \le \Xi_1$, $\xi_2 \le \Xi_2$ are equivalent to

$$x_2^2 \le \Xi_2 - 1, \qquad x_3^2 \le \Xi_2 - 1 - x_2^2,$$

$$\left(x_1 - \frac{x_2 x_3}{x_2^2 + 1}\right)^2 \le \frac{\Xi_1 - 1}{x_2^2 + 1} - \frac{x_3^2}{(x_2^2 + 1)^2}.$$

---

[8] Recall that $\mathcal{J}_{\epsilon_1, \epsilon_2}(A_1, A_2)$ depends on τ_1, τ_2 although this is not displayed in the notation.

Hence

$$\underset{\xi_1 \leq \Xi_1, \xi_2 \leq \Xi_2}{\iiint} dx_1\, dx_2\, dx_3 \leq 8 \int_0^{\Xi_2^{1/2}} \int_0^{\Xi_2^{1/2}} \frac{\Xi_1^{1/2}}{(x_2^2+1)^{1/2}} dx_3\, dx_2$$

$$\ll (\Xi_1 \Xi_2)^{1/2} \log(1 + \Xi_2).$$

This proves the first part of the lemma. The second part is more technical. The conditions (8.17) imply

$$
\begin{aligned}
&x_2^2 \leq \Xi_2\left(1 + \frac{c}{R_2}\right) - 1, \\
&\Xi_2\left(1 - \frac{c}{R_2}\right) - 1 - x_2^2 \leq x_3^2 \leq \Xi_2\left(1 + \frac{c}{R_2}\right) - 1 - x_2^2, \\
&\frac{\Xi_1(1 - c/R_1) - 1}{x_2^2 + 1} - \frac{x_3^2}{(x_2^2 + 1)^2} \\
&\leq \left(x_1 - \frac{x_2 x_3}{x_2^2 + 1}\right)^2 \leq \frac{\Xi_1(1 + c/R_1) - 1}{x_2^2 + 1} - \frac{x_3^2}{(x_2^2 + 1)^2}
\end{aligned}
\tag{8.24}
$$

for some constant $c > 0$. We separate four cases for the range of x_3.

Case 1. If $(\Xi_1(1 + c/R_1) - 1)(x_2^2 + 1) < x_3^2$, then the condition on x_1 is empty.

Case 2. Let us assume

$$\left(\Xi_1(1 - c/R_1) - 1\right)\left(x_2^2 + 1\right) \leq x_3^2 \leq \left(\Xi_1(1 + c/R_1) - 1\right)\left(x_2^2 + 1\right). \tag{8.25}$$

Then the volume of the x_1-region is

$$\leq 2\left(\frac{\Xi_1(1 + c/R_1) - 1}{x_2^2 + 1} - \frac{x_3^2}{(x_2^2 + 1)^2}\right)^{1/2} \leq 2\left(\frac{2c\,\Xi_1}{R_1(x_2^2 + 1)}\right)^{1/2}.$$

The region (8.25) and the second inequality in (8.24) have a non-empty intersection only if

$$\frac{\Xi_2(1 - c/R_2)}{\Xi_1(1 + c/R_1)} - 1 \leq x_2^2 \leq \frac{\Xi_2(1 + c/R_2)}{\Xi_1(1 - c/R_1)} - 1. \tag{8.26}$$

If $R_1, R_2 > 2c$ (which we may assume), this condition is empty unless $\Xi_2 \geq \frac{1}{3}\Xi_1$. This implies $x_2^2 \leq 3\Xi_2/\Xi_1$ and $x_2^2 + 1 \asymp \Xi_2/\Xi_1$. In the following we will frequently use the inequality $\sqrt{A} - \sqrt{B} \leq (A - B)A^{-1/2}$ for $A \geq B \geq 0$.

Since we are assuming that $\varXi_1$, $\varXi_2$ are sufficiently large, we can deduce from (8.25) and the second inequality in (8.24) that the volume of the x_3-region is

$$\ll \min\left(\frac{\varXi_2^{1/2}}{R_2}, \frac{\varXi_2^{1/2}}{R_1} \right) \tag{8.27}$$

(uniformly in x_2) and hence the total contribution under the assumption (8.25) is

$$\int_{x_2 \in (8.26)} \frac{(\varXi_1 \varXi_2)^{1/2}}{R_1^{1/2}(R_1 + R_2)(x_2^2 + 1)^{1/2}} \, dx_2 \asymp \int_{x_2 \in (8.26)} \frac{\varXi_1}{R_1^{1/2}(R_1 + R_2)} \, dx_2.$$

The region (8.26) describes an interval of length $O((R_1^{-1} + R_2^{-1})\varXi_2/\varXi_1)$ for x_2^2, hence the total contribution is

$$\ll \frac{\varXi_1}{R_1^{1/2}(R_1 + R_2)} \frac{\varXi_2^{1/2}}{\varXi_1^{1/2}} \left(\frac{1}{R_1^{1/2}} + \frac{1}{R_2^{1/2}} \right) \ll \frac{(\varXi_1 \varXi_2)^{1/2}}{R_1 R_2}$$

as claimed.

Case 3. For a parameter $1/3 \leq \alpha \leq c/R_1$ consider the region

$$\big(\varXi_1(1 - 2\alpha) - 1\big)\big(x_2^2 + 1\big) \leq x_3^2 \leq \big(\varXi_1(1 - \alpha) - 1\big)\big(x_2^2 + 1\big). \tag{8.28}$$

The procedure here is very similar to case 2. The x_1-volume is at most

$$\leq \frac{4\varXi_1 c/R_1}{(x_2^2 + 1)^{1/2}} \left(\varXi_1\left(1 + \frac{c}{R_1}\right) - 1 - \frac{x_3^2}{x_2^2 + 1} \right)^{-1/2} \ll \frac{\varXi_1^{1/2}}{R_1(x_2^2 + 1)^{1/2}\alpha^{1/2}}.$$

The region (8.28) and the second inequality in (8.24) have a non-empty intersection only if

$$\frac{\varXi_2(1 - c/R_2)}{\varXi_1(1 - \alpha)} - 1 \leq x_2^2 \leq \frac{\varXi_2(1 + c/R_2)}{\varXi_1(1 - 2\alpha)} - 1. \tag{8.29}$$

In particular this implies $x_2^2 + 1 \asymp \varXi_2/\varXi_1$. As in (8.27) we see that the x_3-volume is $\ll \varXi_2^{1/2} \min(R_2^{-1}, \alpha)$, hence the total contribution in the present subcase is

$$\ll \int_{x_2 \in (8.29)} \frac{\varXi_1 \min(R_2^{-1}, \alpha)}{R_1 \alpha^{1/2}} \, dx_2 \ll \frac{\varXi_1 \min(R_2^{-1}, \alpha)}{R_1 \alpha^{1/2}} \frac{\varXi_2^{1/2}}{\varXi_1^{1/2}} \left(\frac{1}{R_2^{1/2}} + \alpha^{1/2} \right)$$

$$\ll \frac{(\varXi_1 \varXi_2)^{1/2}}{R_1 R_2}.$$

Case 4. Finally we consider the region $x_3^2 \leq (\Xi_1/3 - 1)(x_2^2 + 1)$. In this case the x_1-volume is

$$\leq \frac{4\Xi_1 c/R_1}{(x_2^2+1)^{1/2}} \left(\Xi_1 \left(1 + \frac{c}{R_1}\right) - 1 - \frac{x_3^2}{x_2^2+1} \right)^{-1/2} \ll \frac{\Xi_1^{1/2}}{R_1(x_2^2+1)^{1/2}}.$$

The length of the x_3 interval is at most

$$\leq \frac{4c\Xi_2}{R_2(\Xi_2(1+c/R_2)-1-x_2^2)^{1/2}}.$$

Hence the total contribution is at most

$$\ll \int_{x_2^2 \leq \Xi_2(1+c/R_2)-1} \frac{\Xi_1^{1/2}\Xi_2 \, dx_2}{R_1 R_2 (x_2+1)^{1/2}(\Xi_2(1+c/R_j)-1-x_2^2)^{1/2}}$$

$$\ll \frac{(\Xi_1 \Xi_2)^{1/2}}{R_1 R_2} + \int_{1 \leq x_2^2 \leq \Xi_2(1+c/R_2)-1} \frac{\Xi_1^{1/2}\Xi_2 \, dx_2}{R_1 R_2 x (\Xi_2(1+c/R_j)-1-x_2^2)^{1/2}}.$$

This last integral can be computed explicitly:

$$\int \frac{dx}{x(Z-x^2)^{1/2}} = \frac{\log(x) - \log(Z + \sqrt{Z(Z-x^2)})}{\sqrt{Z}},$$

and the desired bound follows. This completes the proof of the lemma. □

9 Proofs of the theorems

For the *proof of Theorem 1* we choose $n_1 = n_2 = m_1 = m_2 = 1$ and combine (8.1), (8.2), Lemma 1 and Propositions 3 and 5. We choose $\tau_1 = R_1 = T_1$, $\tau_2 = R_2 = T_2$ and $X_1 = X_2 = 1$ in (8.5), fix a function f and drop all these parameters from the notation of F. By the second part of (8.7) and (8.8), the Kloosterman terms $\Sigma_{2a}, \Sigma_{2b}, \Sigma_3$ are finite sums over D_1, D_2 and hence are $O((T_1 T_2)^{-100})$ by the first part of (8.7) and (8.8). The diagonal term (8.6) is $\asymp T_1 T_2(T_1 + T_2)$. On the spectral side, we drop the Eisenstein spectrum and large parts of the cuspidal spectrum to conclude by (3.3) and Lemma 1 the upper bound

$$\sum_{\substack{|v_1^{(j)} - iT_1| \leq c \\ |v_2^{(j)} - iT_2| \leq c}} \left(\operatorname*{res}_{s=1} L(s, \phi_j \times \tilde{\phi}_j) \right)^{-1} \ll T_1 T_2(T_1 + T_2)$$

for some sufficiently small c and $T_1, T_2 \geq T_0$, and hence

$$\sum_{\substack{|\nu_1^{(j)} - iT_1| \leq K \\ |\nu_2^{(j)} - iT_2| \leq K}} \left(\operatorname*{res}_{s=1} L(s, \phi_j \times \tilde{\phi}_j)\right)^{-1} \ll_K T_1 T_2 (T_1 + T_2) \qquad (9.1)$$

for any $K \geq 1$ by adding the contribution of $O_K(1)$ balls. To prove the lower bound, we choose (once and for all) K so large that

$$\sum_{\max(|\nu_1^{(j)} - iT_1|, |\nu_2^{(j)} - iT_2|) \geq K} \|\phi_j\|^{-2} \big|\langle \tilde{W}_{\nu_1, \nu_2}, F\rangle\big|^2 \leq \frac{1}{2}\|F\|^2$$

which is possible by (9.1) and (3.1). We bound the Eisenstein spectrum trivially: the second line of (8.1) contributes $O((T_1 + T_2)^\varepsilon)$ by known bounds for the zeta function on the line $\Re s = 1$, the third line contributes similarly $O((T_1 + T_2)^{1+\varepsilon})$ by Weyl's law for $SL_2(\mathbb{Z})$ and lower bounds for the L-functions in the denominator [22, 23]. Hence we obtain

$$\sum_{\substack{|\nu_1^{(j)} - iT_1| \leq K \\ |\nu_2^{(j)} - iT_2| \leq K}} \|\phi_j\|^{-2} \big|\langle \tilde{W}_{\nu_1, \nu_2}, F\rangle\big|^2 \geq \frac{1}{2}\|F\|^2 + O\big((T_1 + T_2)^{1+\varepsilon}\big),$$

and the lower bound in Theorem 1 follows from (3.1) and (8.6) for T_1, T_2 sufficiently large. $\qquad\square$

The *proof of Theorem 2* proceeds similarly. As mentioned in the introduction, as a direct corollary of Theorem 1 we find that the number of exceptional Maaß forms ϕ_j with $\gamma_j = T + O(1)$ is $O(T^2)$. In order to prove Theorem 2, it is therefore enough to consider those Maaß forms with $|\rho_j| \geq \varepsilon$. Moreover, by symmetry it is enough to bound only Maaß forms satisfying (2.9). In (8.5) we take $\tau_2 = R_2 = T$, $R_1 = 1$, $\tau_1 = 0$, $X_1 = 1$, $X_2 = X = T^\delta$ for some $\delta > 0$ to be chosen later. With this data, the spectral side, after dropping

- the tempered spectrum,
- the Eisenstein spectrum, and
- those parts of the non-tempered spectrum not of the form (2.9) with $|\rho_j| \geq \varepsilon$,

is by (3.5) (note that (3.4) is satisfied) and the upper bound of (1.4) at least

$$\gg T^{-\varepsilon} X^2 \sum_{\substack{\phi_j \text{ as in } (2.9) \\ \gamma_j = T + O(1) \\ |\rho_j| \geq \varepsilon}} X^{2|\rho_j|}.$$

On the arithmetic side, the diagonal term is $\asymp T^2 X^2$ by (8.6). Next by (8.7) we have

$$\Sigma_{2a} \ll T^2 X \sum_{D_1 \ll 1} \frac{|\tilde{S}(\pm 1, 1, 1, D_1, D_1^2)|}{D_1^3} T^{-102} \ll X T^{-100}$$

and

$$\Sigma_{2b} \ll T^2 X^2 \sum_{D_2 \ll X^{1/2}} \frac{|\tilde{S}(\pm 1, 1, 1, D_2, D_2^2)|}{D_2^3} T^{-102} \ll X^{2+\varepsilon} T^{-100}$$

by (6.7). (Note that we are exchanging X_1 and X_2 for Σ_{2b}.) The long Weyl element contributes at most

$$(T X^2)^{1+\varepsilon} \sum_{D_1, D_2 \ll X} \frac{|S(\pm 1, 1, 1, 1, D_1, D_2)|}{D_1 D_2} \left(\frac{1 + X/D_2}{T} \right)^{-C_1}$$

$$\times \left(\frac{1 + X/D_1}{T} \right)^{-C_2} \ll (T X^2)^{1+2\varepsilon} \frac{X}{T} \tag{9.2}$$

which follows by combining (8.9) and (6.6). Choosing $X = T^2$ completes the proof of Theorem 2. $\qquad\square$

We proceed to prove *Theorem 3*. Again we choose $X_1 = X_2 = 1$, $R_1 = T_1$, $R_2 = T_2$ in (8.5), fix a function f and then drop R_1, R_2, X_1, X_2, f from the notation of F and keep only τ_1, τ_2. We also fix a suitable non-negative smooth function g with support in $[1/2, 3]$ as in Proposition 5. Let $T := \max(T_1, T_2)$. The left hand side of (1.5) is, by (3.3) and the upper bound of (1.4),

$$\ll T^\varepsilon \sum_j \frac{1}{\|\phi_j\|^2} \int_{T_1}^{2T_1} \int_{T_2}^{2T_2} \left| \langle \tilde{W}_{\nu_1, \nu_2}, F_{\tau_1, \tau_2} \rangle \right|^2 d\tau_1 \, d\tau_2 \left| \sum_{n \leq N} \alpha(n) A_j(n, 1) \right|^2.$$

We cut the n-sum into dyadic intervals, insert artificially the function g and bound the preceding display by

$$\ll (NT)^\varepsilon \max_{M \leq N} \sum_j \frac{1}{\|\phi_j\|^2}$$

$$\times \int_{-\infty}^{-\infty} \int_{-\infty}^{\infty} g\left(\frac{\tau_1}{T_1} \right) g\left(\frac{\tau_2}{T_2} \right) \left| \langle \tilde{W}_{\nu_1, \nu_2}, F_{\tau_1, \tau_2} \rangle \right|^2 d\tau_1 \, d\tau_2$$

$$\times \left| \sum_{M \leq n \leq 2M} \alpha(n) A_j(n, 1) \right|^2.$$

Next we add artificially the continuous spectrum getting the upper bound

$$(NT)^\varepsilon \max_{M \leq N} \left(\sum_j \int_{-\infty}^\infty \int_{-\infty}^\infty g\left(\frac{\tau_1}{T_1}\right) g\left(\frac{\tau_2}{T_2}\right) \frac{|\langle \tilde{W}_{\nu_1, \nu_2}, F_{\tau_1, \tau_2}\rangle|^2}{\|\phi_j\|^2} \, d\tau_1 \, d\tau_2 \right.$$

$$\times \left| \sum_{M \leq n \leq 2M} \alpha(n) A_j(n, 1) \right|^2$$

$$+ \frac{1}{(4\pi i)^2} \int_{(0)} \int_{(0)} \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{g\left(\frac{\tau_1}{T_1}\right) g\left(\frac{\tau_2}{T_2}\right) |\langle \tilde{W}_{\nu_1, \nu_2}, F_{\tau_1, \tau_2}\rangle|^2}{|\zeta(1 + 3\nu_0)\zeta(1 + 3\nu_1)\zeta(1 + 3\nu_2)|^2} \, d\tau_1 \, d\tau_2$$

$$\times \left| \sum_{M \leq n \leq 2M} \alpha(n) A_{\nu_1, \nu_2}(n, 1) \right|^2 d\nu_1 \, d\nu_2$$

$$+ \frac{c}{2\pi i} \sum_j \int_{(0)} \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{g\left(\frac{\tau_1}{T_1}\right) g\left(\frac{\tau_2}{T_2}\right) |\langle \tilde{W}_{\frac{2}{3}\nu_j, \mu - \frac{1}{3}\nu_j}, F_{\tau_1, \tau_2}\rangle|^2}{|L(1 + 3\mu, u_j)|^2 L(1, \mathrm{Ad}^2 u_j)} \, d\tau_1 \, d\tau_2$$

$$\times \left. \left| \sum_{M \leq n \leq 2M} \alpha(n) B_{\mu, u_j}(n, 1) \right|^2 d\mu \right).$$

We open the squares and apply the Kuznetsov formula, that is, we replace the three terms of the shape (8.1) with the four terms (8.2). We estimate each of them individually. The diagonal term contributes by (8.6)

$$\ll (NT)^\varepsilon \max_{M \leq N} \sum_{M \leq n \leq 2M} |\alpha(m)|^2$$

$$\times \int_{-\infty}^\infty \int_{-\infty}^\infty g\left(\frac{\tau_1}{T_1}\right) g\left(\frac{\tau_2}{T_2}\right) T_1 T_2 (T_1 + T_2) \, d\tau_1 \, d\tau_2$$

$$\ll (NT)^\varepsilon T_1^2 T_2^2 (T_1 + T_2) \|\alpha\|_2^2.$$

This is the first term on the right hand side of (1.5). In the term Σ_{2a} in (8.2) the condition $D_1 \mid D_2$, $D_1^2 = n D_2$ is equivalent to $D_1 = nd$, $D_2 = nd^2$ for some $d \in \mathbb{N}$; hence its contribution is at most

$$\ll (NT)^\varepsilon \max_{M \leq N} \sum_{M \leq n, m \leq 2M} |\alpha(n)\alpha(m)| \sum_{\epsilon = \pm 1} \sum_d \frac{|S(\epsilon m, n, 1, nd, nd^2)|}{n^2 d^3}$$

$$\times \left| \int_{-\infty}^\infty \int_{-\infty}^\infty g\left(\frac{\tau_1}{T_1}\right) g\left(\frac{\tau_2}{T_2}\right) \tilde{\mathcal{J}}_{\epsilon, F}\left(\frac{m^{1/2}}{nd^{3/2}}\right) d\tau_1 \, d\tau_2 \right|.$$

By (8.7), the d-sum is finite, hence in combination with (6.7) this is bounded by

$$\ll (NT)^\varepsilon \max_{M \leq N} \sum_{M \leq n, m \leq 2M} \left|\alpha(n)\alpha(m)\right| T^{-101} \ll N^\varepsilon T^{-100} \|\alpha\|^2.$$

In the term Σ_{2b} in (8.2) the condition $D_2 \mid D_1$ is redundant, and the argument of $\tilde{\mathcal{J}}_{\epsilon, F^*}$ equals $(n/(mD_2^3))^{1/2}$. As before we see that this contributes at most $N^\varepsilon T^{-100} \|\alpha\|^2$.

Finally the long Weyl element finally contributes by (8.10)

$$
\begin{aligned}
\ll \quad & (NT)^\varepsilon \max_{M \leq N} \sum_{M \leq n, m \leq 2M} \left|\alpha(n)\alpha(m)\right| \\
& \times \sum_{\epsilon_1, \epsilon_2 = \pm 1} \sum_{D_1, D_2} \frac{|S(\epsilon_1 m, \epsilon_2, n, 1, D_1, D_2)|}{D_1 D_2} \\
& \times \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g\left(\frac{\tau_1}{T_1}\right) g\left(\frac{\tau_2}{T_2}\right) \mathcal{J}_{\epsilon_1, \epsilon_2}\left(\frac{\sqrt{mD_1}}{D_2}, \frac{\sqrt{nD_2}}{D_1}\right) d\tau_1 \, d\tau_2 \right| \\
\ll_{C_1, C_2} & (NT)^\varepsilon \max_{M \leq N} T_1 T_2 (T_1 + T_2) \sum_{M \leq n, m \leq 2M} \left|\alpha(n)\alpha(m)\right| \\
& \times \sum_{\epsilon_1, \epsilon_2 = \pm 1} \sum_{D_1, D_2} \frac{|S(\epsilon_1 m, \epsilon_2, n, 1, D_1, D_2)|}{D_1 D_2} \left(\frac{1 + M/D_2}{T_1 + T_2}\right)^{C_1} \\
& \times \left(\frac{1 + M/D_1}{T_1 + T_2}\right)^{C_2}
\end{aligned}
\tag{9.3}
$$

for any $C_1, C_2 \geq 0$. Recalling the notation $T = \max(T_1, T_2)$ and using (6.6), it is straightforward to see that the previous display is

$$\ll (NT)^\varepsilon \left(T_1 T_2 N^2\right) \|\alpha\|_2^2.$$

This is the second term on the right hand side of (1.5).   □

Finally we prove *Theorem 4*. To this end, we express $L(\phi_j, 1/2)$ by an approximate functional equation. As we are summing over the archimedean parameters of the L-functions, we need an approximate functional equation whose weight function is essentially independent of the underlying family. This has been obtained in [3, Proposition 1], and we quote the following spe-

cial case. For a Maaß form ϕ_j as in Theorem 4 put

$$(\eta_j)_1 = \frac{1}{4} + \frac{2v_1^{(j)} + v_2^{(j)}}{2}, \qquad (\eta_j)_2 = \frac{1}{4} + \frac{-v_1^{(j)} + v_2^{(j)}}{2},$$

$$(\eta_j)_3 = \frac{1}{4} + \frac{-v_1^{(j)} - 2v_2^{(j)}}{2} \qquad \eta_j = \min_{1 \le l \le 3} \left| (\eta_j)_l \right| \asymp \left| (\eta_j)_2 \right|,$$

$$C_j = \prod_{l=1}^{3} \left| (\eta_j)_l \right|.$$

Moreover, for a multi-index $\mathbf{n} \in \mathbb{N}_0^6$ we write $|\mathbf{n}| = \mathbf{n}(1) + \cdots + \mathbf{n}(6)$ and

$$\eta_j^{-\mathbf{n}} := \prod_{l=1}^{3} (\eta_j)_l^{-\mathbf{n}(2l-1)} (\overline{\eta}_j)_l^{-\mathbf{n}(2l)}.$$

Proposition 6 *Let $G_0 : (0, \infty) \to \mathbb{R}$ be a smooth function with functional equation $G_0(x) + G_0(1/x) = 1$ and derivatives decaying faster than any negative power of x as $x \to \infty$. Let $M \in \mathbb{N}$ and fix a Maaß form ϕ as above. There are explicitly computable rational constants $c_{\mathbf{n},\ell} \in \mathbb{Q}$ depending only on $\mathbf{n}$, ℓ, M such that the following holds for*

$$G(x) := G_0(x) + \sum_{\substack{0 < |\mathbf{n}| < M \\ 0 < \ell < |\mathbf{n}| + M}} c_{\mathbf{n},\ell} \eta_j^{-\mathbf{n}} \left(x \frac{\partial}{\partial x} \right)^{\ell} G_0(x).$$

For any $\varepsilon > 0$ one has

$$L(\phi_j, 1/2) = \sum_{n=1}^{\infty} \frac{A_j(1, n)}{\sqrt{n}} G\left(\frac{n}{\sqrt{C_j}} \right) + \kappa_j \sum_{n=1}^{\infty} \overline{\frac{A_j(1, n)}{\sqrt{n}} G\left(\frac{n}{\sqrt{C_j}} \right)}$$

$$+ O\left(\eta_j^{-M} C_j^{1/4 + \varepsilon} \right), \tag{9.4}$$

where $|\kappa_j| = 1$ and the implied constant depends at most on ε, M, and the function G_0.

It is now a simple matter to prove Theorem 4. We can assume that T is sufficiently large. Let

$$G_\ell(x) := \left(x \frac{\partial}{\partial x} \right)^{\ell} G_0(x).$$

Then the Mellin transform $\widehat{G}_\ell(s)$ is rapidly decaying on vertical lines $\Re s = \sigma > 0$. By (9.4) and Mellin inversion we have

$$\sum_{T \leq |v_1^{(j)}|, |v_2^{(j)}| \leq 2T} \left| L(\phi_j, 1/2) \right|^2$$

$$\ll_{M,\varepsilon} \sum_{\ell \leq 2M} \sum_{T \leq |v_1^{(j)}|, |v_2^{(j)}| \leq 2T} \left(\left| \sum_{n \leq T^{3/2+\varepsilon}} \frac{A_j(n,1)}{\sqrt{n}} \overline{G}_\ell \left(\frac{n}{\sqrt{C_j}} \right) \right| \right.$$

$$\left. + O\left(\eta_j^{-M} C_j^{1/4+\varepsilon} + T^{-100} \right) \right)^2$$

$$\ll_{M,\varepsilon} T^\varepsilon \sum_{T \leq |v_1^{(j)}|, |v_2^{(j)}| \leq 2T} \left(\int_{-T^\varepsilon}^{T^\varepsilon} \left| \sum_{n \leq T^{3/2+\varepsilon}} \frac{A_j(n,1)}{n^{1/2+\varepsilon+it}} \right| dt \right.$$

$$\left. + O\left(\eta_j^{-M+1/4} T^{1/2+\varepsilon} + T^{-100} \right) \right)^2,$$

noting that $C_j \ll \eta_j T^2 \ll T^3$. This is at most

$$\ll T^\varepsilon \left(\max_{|t| \leq T^\varepsilon} \sum_{T \leq |v_1^{(j)}|, |v_2^{(j)}| \leq 2T} \left| \sum_{n \leq T^{3/2+\varepsilon}} \frac{A_j(n,1)}{n^{1/2+\varepsilon+it}} \right|^2 \right.$$

$$\left. + T \sum_{T \leq |v_1^{(j)}|, |v_2^{(j)}| \leq 2T} \left(1 + \left| v_1^{(j)} - v_2^{(j)} \right| \right)^{-2M+1/2} \right).$$

By Theorem 3 and (4.3) the first term is $O(T^{5+\varepsilon})$. By Theorem 1 or (9.1) it is easy to see that the second term is also $O(T^{5+\varepsilon})$. This completes the proof of Theorem 4. $\qquad\square$

Appendix: A theorem of Goldfeld–Kontorovich

A very nice application of the *GL*(3) Kuznetsov formula has been given recently in [19]. The purpose of this appendix is to illustrate how the methods of this paper directly yield a version of [19, Theorem 1.3] with considerably better error terms and without assuming the Ramanujan conjectures. We keep the notation developed so far.

Theorem 5 *For $n_1, n_2, m_1, m_2 \in \mathbb{N}$, $P = n_1 n_2 m_1 m_2$, T sufficiently large, one has*

$$\sum_j \overline{A_j(n_1, n_2)} A_j(m_1, m_2) \frac{h_T(\nu_1, \nu_2)}{\|\phi_j\|^2}$$

$$= \delta_{\substack{m_1=n_1 \\ m_2=n_2}} \sum_j \frac{h_T(\nu_1, \nu_2)}{\|\phi_j\|^2} + O\left((T^2 P^{1/2} + T^3 P^\theta + P^{5/3})(TP)^\varepsilon\right)$$

where $\theta \leq 7/64$ is a bound towards the Ramanujan conjecture on GL(2). Here h_T is non-negative, uniformly bounded on $\{|\Re\nu_1| \leq 1/2\} \times \{|\Re\nu_2| \leq 1/2\}$, $h_T \asymp 1$ on $\{(\nu_1, \nu_2) \mid c \leq \Im\nu_1, \Im\nu_2 \leq T, |\Re\nu_1|, |\Re\nu_2| \leq 1/2\}$ for some absolute constant $c > 0$, and $h_T(\nu_1, \nu_2) \ll_A ((1 + |\nu_1|/T)(1 + |\nu_2|/T))^{-A}$.

For comparison, the error term in [19, Theorem 1.3] (scaled down by T^{-3R}) is $O(T^{3+\varepsilon} P^2)$, but see also [19, Remarks 1.8, 1.19] where possible improvements are mentioned. A more precise discussion on the asymptotic behavior of the test function h_T can be found in Remark 6 below.

Injecting Theorem 5 into the estimates of [19, Sect. 9] and using only $\theta < 1/3$, we obtain the following corollary. For a Hecke–Maaß form ϕ for $SL_3(\mathbb{Z})$ let $\rho(\phi)$ be one of ϕ, $sym^2\phi$ or $Ad\phi$. Let ψ be a smooth test function whose Fourier transform has support in $(-\delta, \delta)$ for some $\delta > 0$. Define $D(\rho(\phi), \psi)$ as in [19, Sect. 1.4].

Corollary 6 (Goldfeld–Kontorovich) *Assume the generalized Riemann hypothesis and the Ramanujan conjectures. Suppose*

$$\delta < 5/23, \qquad \rho(\phi) = sym^2\phi \text{ or } Ad\phi,$$
$$\delta < 10/13, \qquad \rho(\phi) = \phi.$$

Then one has

$$\left(\sum_j \frac{h_T(v_1, v_2)}{\|\phi_j\|^2}\right)^{-1} \sum_j D\big(\rho(\phi_j), \psi\big) \frac{h_T(v_1, v_2)}{\|\phi_j\|^2}$$

$$= \int_{\mathbb{R}} \psi(x) W_\rho(x)\, dx + O\left(\frac{\log\log T}{\log T}\right),$$

where

$$W_\rho(x) = 1, \quad \rho(\phi) = \phi \text{ or } sym^2\phi,$$

$$W_\rho(x) = 1 - \frac{\sin(2\pi x)}{2\pi x}, \quad \rho(\phi) = \mathrm{Ad}\,\phi.$$

In particular, the symmetry types are unitary or symplectic, respectively.

This improves the range of the support of $\hat{\psi}$ by about a factor 3 compared to [19, Theorem 1.13] (see also [19, Remarks 1.18, 1.19]).

Proof of Theorem 5 Let g be a fixed, smooth, non-negative, compactly supported test function. Let R_1, R_2 be sufficiently large, and write $R = R_1 + R_2$. We choose F as in (8.5) with $X_1 = X_2 = 1$ and integrate the equality in Proposition 4 against

$$\int_0^\infty \int_0^\infty g\left(\frac{\tau_1}{R_1}\right) g\left(\frac{\tau_2}{R_2}\right) d\tau_1\, d\tau_2$$

as in (8.10). From Proposition 3, the above mentioned lower bounds for L-functions [22, 23] on the line $\Re s = 1$ and Weyl's law for $GL(2)$ we conclude that the Eisenstein contribution in (8.1) is $O(R^{3+\varepsilon} P^{\theta+\varepsilon})$. From Proposition 3 and Proposition 5 we conclude by the same calculation as in (9.3) that the long Kloosterman sum Σ_3 in (8.2) contributes $O(R^{2+\varepsilon} P^{1/2+\varepsilon})$. Similarly, if $P < R^{3-\varepsilon}$, the other two Kloosterman contributions $\Sigma_{2a} + \Sigma_{2b}$ are $O(R^{-100})$, and are otherwise $O(R^{5+\varepsilon})$ which follows after a straightforward estimate using Proposition 5 and (6.7). Hence in either case their contribution is $O(P^{5/3+\varepsilon})$. We conclude

$$\sum_j \overline{A_j(n_1, n_2)} A_j(m_1, m_2) \frac{h_{R_1, R_2}(v_1, v_2)}{\|\phi_j\|^2}$$

$$= \delta_{\substack{m_1=n_1 \\ m_2=n_2}} H_{R_1, R_2} + O\big((R^2 P^{1/2} + R^3 P^\theta + P^{5/3})(RP)^\varepsilon\big) \quad (10.1)$$

where

$$h_{R_1, R_2}(v_1, v_2) = \int_0^\infty \int_0^\infty g\left(\frac{\tau_1}{R_1}\right) g\left(\frac{\tau_2}{R_2}\right) \big|\langle \tilde{W}_{v_1, v_2}, F\rangle\big|^2 d\tau_1\, d\tau_2$$

and $H_{R_1,R_2} = \int_0^\infty \int_0^\infty g(\frac{\tau_1}{R_1}) g(\frac{\tau_2}{R_2}) \|F\|^2 \, d\tau_1 \, d\tau_2$, but we only need to know that this quantity is independent of n_1, n_2, m_1, m_2.

The weight function h_{R_1,R_2} is uniformly bounded and non-negative. It follows directly from Proposition 3 that

$$h_{R_1,R_2}(\nu_1, \nu_2) \asymp 1 \quad \text{for } |\Re \nu_1|, |\Re \nu_2| \le 1/2, \ \frac{\Im \nu_1}{R_1}, \frac{\Im \nu_2}{R_2} \in \text{supp}(g), \quad (10.2)$$

and rapidly decaying outside the region $\frac{|\nu_1|}{R_1}, \frac{|\nu_2|}{R_2} \in \text{supp}(g)$. In other words, h_{R_1,R_2} is a good approximation of the characteristic function on the square $\Im \nu_1 \asymp R_1, \Im \nu_2 \asymp R_2$.

Applying (10.1) with $n_1 = n_2 = m_1 = m_2 = 1$, we see that

$$H_{R_1,R_2} = \sum_j \frac{h_{R_1,R_2}(\nu_1, \nu_2)}{\|\phi_j\|^2} + O(R^{3+\varepsilon}).$$

Hence we obtain

$$\sum_j \overline{A_j(n_1, n_2)} A_j(m_1, m_2) \frac{h_{R_1,R_2}(\nu_1, \nu_2)}{\|\phi_j\|^2}$$

$$= \delta_{\substack{m_1=n_1 \\ m_2=n_2}} \sum_j \frac{h_{R_1,R_2}(\nu_1, \nu_2)}{\|\phi_j\|^2} + O\big((R^2 P^{1/2} + R^3 P^\theta + P^{5/3})(RP)^\varepsilon\big)$$

whenever R_1, R_2 are sufficiently large. Piecing together dyadic squares, we obtain Theorem 5. $\qquad\square$

Remark 6 The proof of Proposition 3 gives much more precise information on the weight function h_T in Theorem 5. By (3.14), we see that h_{R_1,R_2} described in (10.2) satisfies the more precise asymptotic relation

$$h_{R_1,R_2}(\nu_1, \nu_2) \sim c \frac{R_1 R_2(R_1 + R_2)}{|\nu_1 \nu_2(\nu_1 + \nu_2)|} g\left(\frac{|\nu_1|}{R_1}\right) g\left(\frac{|\nu_2|}{R_2}\right), \quad \nu_1, \nu_2 \in i\mathbb{R},$$

for $R_1, R_2 \to \infty$, where the constant $c > 0$ is given by

$$c = \frac{(2\pi)^3}{3^3} \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \widehat{f}(-1 - ix - 2iy) \widehat{f}(-1 - 2ix - iy) \right|^2 dx \, dy$$

$$= \frac{(2\pi)^3}{3^4} \left(\int_{\mathbb{R}} \left| \widehat{f}(-1 - ix) \right|^2 dx \right)^2$$

for the weight function f in the Poincaré series (8.5). In particular, by varying g one has the flexibility to prescribe asymptotically any reasonable bump function on the tempered spectrum.

References

1. Balakci, D.: Automorphic forms for congruence subgroups of $SL(3, \mathbb{Z})$. PhD thesis, Göttingen, in progress
2. Blomer, V., Buttcane, J., Raulf, N.: A Sato-Tate law for $GL(3)$. Comment. Math. Helv. to appear
3. Blomer, V., Harcos, G.: A hybrid asymptotic formula for the second moment of Rankin–Selberg L-functions. Proc. Lond. Math. Soc. **105**, 473–505 (2012)
4. Bruggeman, R.W.: Fourier coefficients of cusp forms. Invent. Math. **45**, 1–18 (1978)
5. Bruggeman, R.W., Miatello, R.J.: Sum formula for SL_2 over a totally real number field. Mem. Am. Math. Soc. **197**, 919 (2009). vi+81 pp.
6. Brumley, F.: Second order average estimates on local data of cusp forms. Arch. Math. **87**, 19–32 (2006)
7. Bump, D.: Automorphic Forms on $GL(3, \mathbb{R})$. Lecture Notes in Mathematics, vol. 1083. Springer, Berlin (1984)
8. Bump, D., Friedberg, S., Goldfeld, D.: Poincaré series and Kloosterman sums for $SL(3, \mathbb{Z})$. Acta Arith. **50**, 31–89 (1988)
9. Buttcane, J.: Sums of $SL(3, \mathbb{Z})$ Kloosterman sums. PhD thesis, UCLA (2012)
10. Buttcane, J.: On sums of $SL(3, \mathbb{Z})$ Kloosterman sums. Ramanujan J., to appear
11. Cogdell, J., Li, J., Piatetski-Shapiro, I., Sarnak, P.: Poincaré series for $SO(n, 1)$. Acta Math. **167**, 229–285 (1991)
12. Deshouillers, J.-M., Iwaniec, H.: Kloosterman sums and Fourier coefficients of cusp forms. Invent. Math. **70**, 219–288 (1982/83)
13. Deshouillers, J.-M., Iwaniec, H.: The non-vanishing of Rankin–Selberg zeta-functions at special points. In: Hejhal, D., Sarnak, P., Terras, A. (eds.) The Selberg Trace Formula and Related Topics. Contemporary Mathematics, vol. 53. AMS, Providence (1986)
14. Duke, W., Kowalski, E.: Large sieve inequalities for $GL(n)$-forms in the conductor aspect (with appendix by D. Ramakrishnan). Invent. Math. **139**, 1–39 (2000)
15. Elstrodt, J., Grunewald, F., Mennicke, J.: Kloosterman sums for Clifford algebras and a lower bound for the positive eigenvalues of the Laplacian for congruence subgroups acting on hyperbolic spaces. Invent. Math. **101**, 641–685 (1990)
16. Friedberg, S.: A global approach to the Rankin–Selberg convolution for $GL(3, \mathbb{Z})$. Trans. Am. Math. Soc. **300**, 159–174 (1987)
17. Goldfeld, D.: Automorphic Forms and L-Functions for the Group $GL(n, \mathbb{R})$. Cambridge Studies in Advanced Mathematics, vol. 99 (2006)
18. Goldfeld, D., Kontorovich, A.: On the determination of the Plancherel measure for Lebedev–Whittaker transforms on $GL(n)$. Acta Arith. **155**, 15–26 (2012)
19. Goldfeld, D., Kontorovich, A.: On the $GL(3)$ Kuznetsov formula with applications to symmetry types of families of L-functions. In: Proceedings of the International Colloquium on Automorphic Representations and L-Functions, Tata Institute (2012, to appear)
20. Gradshteyn, I.S., Ryzhik, I.M.: Tables of Integrals, Series, and Products, 7th edn. Academic Press, New York (2007)
21. Grenier, D.: On the shape of fundamental domains in $GL(n, \mathbb{R})/O(n)$. Pac. J. Math. **160**, 53–66 (1993)
22. Hoffstein, J., Lockhart, P.: Coefficients of Maass forms and the Siegel zero. Ann. Math. **140**, 161–176 (1994)
23. Hoffstein, J., Ramakrishnan, D.: Siegel zeros and cusp forms. Int. Math. Res. Not. **1995**(6), 279–308 (1995)
24. Lapid, E., Müller, W.: Spectral asymptotics for arithmetic quotients of $SL(n, \mathbb{R})/SO(n)$. Duke Math. J. **149**, 117–154 (2009)
25. Li, Xinnan: Upper bounds on L-functions at the edge of the critical strip. Int. Math. Res. Not. **2010**, 727–755 (2010)

26. Li, Xiaoqing: A spectral mean value theorem for $GL(3)$. J. Number Theory **130**, 2395–2403 (2010)
27. Kuznetsov, N.V.: The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums. Math. USSR Sb. **39**, 299–342 (1981)
28. Miatello, R., Wallach, N.: Kuznetsov formulas for real rank one groups. J. Funct. Anal. **93**, 171–206 (1990)
29. Miller, S.: On the existence and temperedness of cusp forms for $SL_3(\mathbb{Z})$. J. Reine Angew. Math. **533**, 127–169 (2001)
30. Ramakrishnan, D., Wong, S.: On the exceptional zeros of Rankin–Selberg L-functions. Compos. Math. **135**, 211–244 (2003)
31. Reznikov, A.: On the asymptotic formula of Kuznetsov and existence of generic cusp forms. Math. Z. **213**, 557–573 (1993)
32. Shahidi, F.: Eisenstein Series and Automorphic L-Functions. American Mathematical Society Colloquium Publications, vol. 58. Amer. Math. Soc., Providence (2010)
33. Stade, E.: Hypergeometric series and Euler factors at infinity for L-functions on $GL(3, \mathbb{R}) \times GL(3, \mathbb{R})$. Am. J. Math. **115**, 371–387 (1993)
34. Stevens, G.: Poincarè series on $GL(r)$ and Kloostermann sums. Math. Ann. **277**, 25–51 (1987)
35. Venkatesh, A.: Large sieve inequalities for $GL(n)$-forms in the conductor aspect. Adv. Math. **200**, 336–356 (2006)
36. Ye, Y.: A Kuznetsov formula for Kloosterman sums on GL_n. Ramanujan J. **4**, 385–395 (2000)