THE ARITHMETIC KUZNETSOV FORMULA ON $GL(3)$, I: THE WHITTAKER CASE.

JACK BUTTCANE

Abstract. The original formulae of Kuznetsov for $SL(2, \mathbb{Z})$ allowed one to study either a spectral average via Kloosterman sums or to study an average of Kloosterman sums via a spectral interpretation. In previous papers, we have developed the spectral Kuznetsov formulae at the minimal weights for $SL(3, \mathbb{Z})$, and in these formulae, the big-cell Kloosterman sums occur with weight functions attached to four different integral kernels, according to the choice of signs of the indices. These correspond to the $J$- and $K$-Bessel functions in the case of $GL(2)$. In this paper, we demonstrate a linear combination of the spherical and weight-one $SL(3, \mathbb{Z})$ Kuznetsov formulae that isolates one particular integral kernel, which is the spherical $GL(3)$ Whittaker function. Using the known inversion formula of Wallach, we give the first arithmetic Kuznetsov formula for $SL(3, \mathbb{Z})$ and use it to study smooth averages and the Kloosterman zeta function attached to this particular choice of signs.

1. Introduction

The meromorphic continuation of the Poincaré series for $SL(2, \mathbb{Z})$ was first conducted by Selberg [27] for the purposes of bounding Fourier coefficients of automorphic forms. The Fourier coefficients of the $SL(2, \mathbb{Z})$ Poincaré series can be expressed in terms of the Kloosterman zeta function

$$\tilde{Z}_2(s) := \sum_{c=1}^{\infty} \frac{S(m, n; c)}{c^{1+s}},$$

where

$$S(m, n; c) = \sum_{x \equiv 1 (mod c)} e \left( \frac{mx + nx}{c} \right), \quad e(t) = e^{2\pi it}$$

is the classical Kloosterman sum. Weil’s bound $S(m, n; c) \ll_{m,n} c^{\frac{3}{2}+\epsilon}$ implies the Kloosterman zeta function converges absolutely on $\text{Re}(s) > \frac{1}{2}$, and Selberg’s study of Poincaré series implies it has meromorphic continuation to all $s$ with poles at integral shifts of the spectral parameters of the $SL(2, \mathbb{Z})$ Maass cusp forms, as well as certain poles arising from the continuous spectrum.

Somewhat later, Kuznetsov [20] used their spectral expansion to resolve a conjecture of Linnik [21], itself motivated by problems in additive number theory:

\[ 29 \text{ May } 2018. \]
2010 Mathematics Subject Classification. Primary 11L05, 11F72; Secondary 11F55.
During the time of this research, the author was supported by NSF grant DMS-1601919.
Theorem 1 (Kuznetsov). For $X > 1$ a large parameter and $m, n \in \mathbb{Z}$, $mn > 0$,
\[ \sum_{c \leq X} \frac{S(m, n; c)}{c} \ll_{m, n} X^{\frac{A}{2}} (\log X)^{\frac{3}{2}}, \]

For contrast, Weil’s square-root cancellation bound would give $X^{\frac{A}{2} + \epsilon}$, so we are seeing cancellation between terms in the sum over the moduli $c$. Away from $SL(2, \mathbb{Z})$ where the Selberg eigenvalue conjecture is known, generalizations of Kuznetsov’s result to congruence subgroups typically involve an additional term $X^\theta$ where $0 < \theta < \frac{1}{2}$ comes from bounds on the spectral parameters of the relevant Maass cusp forms.

Kuznetsov’s proof came from the construction of two formulae that strongly tie moduli sums of Kloosterman sums to the study of automorphic forms. The first, which we call the spectral Kuznetsov formula, expresses certain averages of a test function over $SL(2, \mathbb{Z})$ automorphic forms as sums of Kloosterman sums, weighted by integral transforms of the test function. The second, which we call the arithmetic Kuznetsov formula, operates in reverse, expressing averages of Kloosterman sums weighted by a test function as sums over automorphic forms weighted by integral transforms of the test function. The first direction may be thought of as a generalization of Petersson’s trace formula for modular forms to the $SL(2, \mathbb{Z})$ Maass forms.

The first breakthrough in the study of Poincaré series on higher-rank groups came in the paper [4]. Among other things, that paper proves in detail the meromorphic continuation of a spherical $SL(3, \mathbb{Z})$ Poincaré series. The meromorphic continuation of the corresponding Kloosterman zeta function is also claimed [4, eq. (1.20)], but the proof was never provided.

In the papers [7,10,11], the author completed the generalization of the spectral Kuznetsov formula to $SL(3, \mathbb{Z})$ and gave the non-spherical analogs, i.e. Kuznetsov-type trace formulae for automorphic forms which are non-trivial on $SO(3, \mathbb{R})$. In this paper, we turn our attention to building the arithmetic Kuznetsov formulae on $SL(3, \mathbb{Z})$. These will consider sums of the long-element $SL(3, \mathbb{Z})$ Kloosterman sum. We denote that sum by $S_{w}(\psi_{m}, \psi_{n}, c)$, where $m, n \in \mathbb{Z}^2$ are the indices and $c \in \mathbb{N}^2$ are the moduli. The definition of this exponential sum is somewhat unpleasant and not entirely relevant to the discussion, so we postpone that until Section 2, below.

We have square-root cancellation bounds originally due to Stevens [28] of the form
\[ S_{w}(\psi_{m}, \psi_{n}, c) \ll_{m, n} \sqrt{(c_1, c_2)(c_1c_2)^{A/2 + \epsilon}}, \]
though [14] has improved the $\sqrt{(c_1, c_2)}$ term (and much more explicit bounds are given in [6, Theorem 2] and [2, eq. (2.10)]).

The author’s thesis [6] provides a first attempt at an arithmetic Kuznetsov formula using a first-term inversion formula on the spherical Kuznetsov formula.

Theorem 2. Let $X_1, X_2 > 1$ be large parameters. Suppose $m, n \in \mathbb{Z}^2$ such that $m_1n_2m_2n_1 \neq 0$, and $f$ smooth and compactly supported on $(\mathbb{R}^+)^2$, then
\[ \sum_{\epsilon \in \{\pm 1\}^2} \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w}(\psi_{m}, \psi_{n}, c)}{c_1c_2} f\left(\frac{\psi_{m_1n_2m_2n_1}}{c_1^2}, \frac{\psi_{m_2n_1m_1n_2}}{c_2^2}\right) \ll_{m, n, f, \epsilon} (X_1X_2)^{\theta + \epsilon} + X_1^{\frac{A}{2} + \epsilon} + X_2^{\frac{A}{2} + \epsilon}, \]
where \( \theta \) is any bound toward the Ramanujan-Selberg conjecture for spherical \( SL(3, \mathbb{Z}) \) Maass forms.

The Kim-Sarnak bound [18] shows \( \theta = \frac{5}{14} \) is acceptable. The square-root cancellation bound (1) implies the sum is at most \( (X_1 X_2)^{\frac{1}{2} + \varepsilon} \), so we are seeing cancellation between terms of the \( c \) sum, provided the moduli are not too far apart. That is, the terms \( X_1^{\frac{1}{2} + \varepsilon} + X_2^{\frac{1}{2} + \varepsilon} \) become comparable to the square-root cancellation bound when, say, \( c_1 < c_2^{1/2 + \varepsilon} \).

The method of [6] can be regarded, in a much simplified form, as writing a Fourier coefficient of a spherical \( SL(3, \mathbb{Z}) \) Poincaré series in the form

\[
\tilde{Z}(s) + F(s_1 + \frac{1}{3}, s_2 - \frac{1}{6}) + G(s_1 - \frac{1}{6}, s_2 + \frac{1}{3})
\]

where \( \tilde{Z}(s) \), \( F(s) \) and \( G(s) \) are holomorphic on \( \text{Re}(s_1), \text{Re}(s_2) > \frac{1}{6} \), and

\[
\tilde{Z}(s) = \sum_{w \in W} C_w(m, n, s) \tilde{Z}_w(s),
\]

is the full Kloosterman zeta function with \( W \) the Weyl group, \( C_w(m, n, s) \) a quotient of gamma functions (and powers of \( m \) and \( n \)) which varies with \( w \in W \), and \( \tilde{Z}_w(s) \) the sign-independent Kloosterman zeta function at each Weyl element; in particular,

\[
\tilde{Z}_{w_1}(s) = \sum_{\varepsilon \in \{-1\}^2} \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w_1}(\psi_m, \psi_n, c)}{c_1^{1+3s_1} c_2^{1+3s_2}}.
\]

The extra terms \( F \) and \( G \) here are the direct cause of the terms \( X_1^{\frac{1}{2} + \varepsilon} + X_2^{\frac{1}{2} + \varepsilon} \) in Theorem 2; the strange combination of shifts is because we save factors of \( \frac{1}{c_1} \) and \( \frac{1}{c_2} \) directly. The term \( F(s_1 + \frac{1}{3}, s_2 - \frac{1}{6}) \) gives some analytic continuation in the variable \( \tilde{s}_1 := 2s_1 + s_2 \) and \( G(s_1 - \frac{1}{6}, s_2 + \frac{1}{3}) \) gives some analytic continuation in \( \tilde{s}_2 := 2s_2 + s_1 \), but the sum of the two terms gives no analytic continuation in either variable. It is important to note that the functions \( F \) and \( G \) are not sums of shifts of \( \tilde{Z} \); this is because the kernel functions vary with the signs \( \varepsilon \), even though their first-term asymptotics do not. In this way, it is not possible to use the methods of [6] to obtain the meromorphic continuation of the long-element Kloosterman zeta functions. We expect that the methods of [4] would encounter similar difficulties.

Here we will separate out a particular choice of signs and apply a true inversion formula. As we will see, it is necessary to have both the spherical and weight-one Kuznetsov formulae in order to accomplish this.

We treat the long-element \( SL(3, \mathbb{Z}) \) Kloosterman sum \( S_{w_1}(\psi_m, \psi_n, c) \) in the case \( m_1 n_2, m_2 n_1 > 0 \). When the moduli are coprime, the long-element Kloosterman sum factors into classical Kloosterman sums as

\[
S_{w_1}(\psi_m, \psi_n, c) = S(m_1, -n_2 c_2; c_1) S(m_2 c_1, -n_1; c_2)
\]

so one might (counter-intuitively) term this as the \(-,-\) case for the Kloosterman sums.

The arithmetic Kuznetsov formula attaches a spectral interpretation to smooth averages of Kloosterman sums. We state the full formula in Theorem 5, below. A simple application gives the following bound:
Theorem 3. Let \( X_1, X_2 > 0 \) with \( X_1 X_2 > 1 \). Suppose \( m, n \in \mathbb{Z}^2 \) such that \( m_1 n_2, m_2 n_1 > 0 \), and \( f \) smooth and compactly supported on \((\mathbb{R}^+)^2\), then

\[
\sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w_t}(\psi_m, \psi_n, c)}{c_1 c_2} f\left( \frac{X_1 c_1}{c_1^2}, \frac{X_2 c_2}{c_2^2} \right) \ll_{m,n,f,\epsilon} (X_1 X_2)^{\theta+\epsilon} + X_1^{-1-\epsilon} + X_2^{-1-\epsilon},
\]

where \( \theta \) is any bound toward Ramanujan-Selberg for the spherical and weight-one \( SL(3,\mathbb{Z}) \) Maass forms.

We note below that the Kim-Sarnak result \( \theta = \frac{5}{14} \) extends to these forms, as well. Again, the square-root cancellation bound is \( (X_1 X_2)^{\frac{1}{2}+\epsilon} \), so we are seeing cancellation between terms of the \( c \) sum, even when the moduli are very far apart. In fact, the terms \( X_1^{-1-\epsilon} + X_2^{-1-\epsilon} \) only become comparable to the square-root cancellation bound when, say \( c_1 < c_2^{1/5+\epsilon} \). As is typical for smooth averages, the Ramanujan-Selberg conjecture \( \theta = 0 \) implies the bound \( (X_1 X_2)^{\epsilon} + X_1^{-1-\epsilon} + X_2^{-1-\epsilon} \). Aside from the separation of signs, the bound in this theorem is stronger than that of Theorem 2 due to the use of the full inversion formula, Theorem 9, which effectively removes the \( X_1^{\frac{1}{2}+\epsilon} + X_2^{\frac{1}{2}+\epsilon} \) terms. The terms \( X_1^{-1-\epsilon} + X_2^{-1-\epsilon} \) derive from a particularly vexing group of Maass cusp forms in the neighborhood of the self-dual forms, which will haunt us throughout the paper.

We also obtain the meromorphic continuation of the simplest weighted Kloosterman zeta function for this choice of signs:

Theorem 4. Let \( m, n \in \mathbb{Z}^2 \) with \( m_1 n_2, m_2 n_1 > 0 \), and for \( s \in \mathbb{C}^2 \), define

\[
\tilde{s} = (2s_1 + s_2, s_1 + 2s_2),
\]

then the Kloosterman zeta function

\[
Z_{m,n}^*(s) := \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w_t}(\psi_m, \psi_n, c)}{c_1 c_2} f_s\left( \frac{2\sqrt{m_1 n_2 c_2}}{c_1}, \frac{2\sqrt{m_2 n_1 c_1}}{c_2} \right),
\]

\[
f_s(y) = y_1^{2\tilde{s}_1} y_2^{2\tilde{s}_2} \exp \left( -\pi^2 (y_1^2 + y_2^2) \right),
\]

initially convergent on \( \text{Re}(s_1), \text{Re}(s_2) > \frac{1}{2} \), has meromorphic continuation to all of \( s \in \mathbb{C}^2 \) with poles whenever

\[
\tilde{s}_1 = -\mu_i - \ell, \quad \text{or} \quad \tilde{s}_2 = \mu_i - \ell,
\]

with \( \mu = \mu_\varphi \) for some cusp form \( \varphi \) of weight at most one, \( i = 1, 2, 3 \) and \( \ell \in \mathbb{N}_0 = \mathbb{N} \cup \{0\} \).

The above theorem does not list the poles coming from the Eisenstein series, of which there are many, as these are somewhat more complicated. In Section 2.2 we discuss the orders and residues of the cuspidal poles and briefly indicate how one may determine the poles of the Eisenstein series terms.

The weighted Kloosterman zeta function used in the theorem does not lend itself to proving the meromorphic continuation of the unweighted Kloosterman zeta function, due to the same complications described above. It seems that even having the arithmetic Kuznetsov formula is not sufficient to obtain said continuation, and we discuss some reasons why this fails in Section 2.2.
The reason why our arithmetic Kuznetsov formula in the \(m_1n_2, m_2n_1 > 0\) case involves only the spherical and weight-one forms is because that term is zero in the remaining \(d \geq 2\) spectral Kuznetsov formulae [11], for algebraic reasons. This behavior is very similar to the case of \(SL(2, \mathbb{Z})\), where the Whittaker function is the \(K\)-Bessel function. It is not immediately clear how to isolate the remaining signs, but any arithmetic Kuznetsov formulae in those cases will certainly involve all of the spectral Kuznetsov formulae and will require constructing inversion formulae with discrete spectra. The meromorphic continuation of the sign-independent Kloosterman zeta function (2), if it exists, should then also involve the Maass cusp forms of every weight.

An alternative to the current approach to studying sums of Kloosterman sums is to define the Poincaré series directly in terms of the Bruhat decomposition (see [13]); this bypasses the need for the inversion of integral transforms. The construction of such Poincaré series on \(SL(n, \mathbb{Z})\) was conducted by Yangbo Ye [30], but the necessary spectral interpretation was not available to complete the analysis. With the uniform spectral expansion of [8, 9], this type of analysis is now a feasible method to obtain the remaining Kloosterman zeta functions. We leave these to future papers.

2. Results

As in [4], we define the exponential sum

\[
S(m_1, m_2, n_1, n_2; D_1, D_2) = \sum_{B_1, C_1 (\text{mod } D_1)} e \left( \frac{m_1 B_1 + n_1 (Y_1 D_2 - Z_1 B_2)}{D_1} + \frac{m_2 B_2 + n_2 (Y_2 D_1 - Z_2 B_1)}{D_2} \right),
\]

where the sum is restricted to

\[
D_1 C_2 + B_1 B_2 + D_2 C_1 \equiv 0 \pmod{D_1}, \quad (B_1, C_1, D_1) = (B_2, C_2, D_2) = 1,
\]

and the \(Y_i\) and \(Z_i\) are defined by

\[
Y_1 B_1 + Z_1 C_1 \equiv 1 \pmod{D_1}, \quad Y_2 B_2 + Z_2 C_2 \equiv 1 \pmod{D_2}.
\]

Then the long element \(SL(3, \mathbb{Z})\) Kloosterman sum is

\[
S_{\psi}(\psi_m; \psi; c) = S(-n_2, -n_1, m_1, m_2; c_1, c_2).
\]

Note the difference in the definitions between [7] and [10], as explained in [10, Section 4]; we follow [10].

2.1. The arithmetic Kuznetsov formula. Let \(S_2^d\) be bases of even or odd spherical \(SL(2, \mathbb{Z})\) Maass cusp forms as \(d = 0\) or \(d = 1\), respectively. For \(\phi \in S_2^d\), denote its spectral parameters by \((\mu_\phi, -\mu_\phi)\) with \(\mu_\phi \in \mathbb{R}\) (by [26]) and the Hecke eigenvalues of the maximal parabolic Eisenstein series of weight \(d\) attached to \(\phi \in S_2^d\) with the additional spectral parameter \(r \in \mathbb{C}\) by \(\lambda_\phi(m, r)\) (see [8, Section 5.5]).

Let \(S_3^d\) be a basis of vector-valued \(SL(3, \mathbb{Z})\) Maass cusp forms of weight \(d\). For \(\varphi \in S_3^d\) we denote its Hecke eigenvalues by \(\lambda_\varphi(m)\) and its spectral parameters by \(\mu_\varphi\). Denote the Hecke eigenvalues of the minimal parabolic spherical Eisenstein series by \(\lambda_E(m, \mu)\) (see [8, Section 4.2]).
The cuspidal and Eisenstein series parts of both $d = 0, 1$ (minimal-weight) spectral Kuznetsov formulae with $m, n \in \mathbb{Z}^2$, $m_1m_2n_1n_2 \neq 0$ can be written

$$
C^d(F) = \frac{2\pi}{3} \sum_{\varphi \in S_3^d} F(\mu_\varphi) \frac{\lambda_\varphi(n)\overline{\lambda_\varphi(n)}}{L(\text{Ad}^2 \varphi, 1)},
$$

$$
\mathcal{E}_\text{max}^d(F) = \frac{1}{2i} \sum_{\varphi \in S_2^d} \int_{\Re(r)=0} \frac{F(\mu_\varphi + r, -\mu_\varphi + r, -2r)\overline{\lambda_\varphi(m, r)\lambda_\varphi(n, r)}}{L(\phi, 1 + 3r)L(\phi, 1 - 3r)L(\text{Ad}^2 \phi, 1)} dr,
$$

$$
\mathcal{E}_\text{min}^0(F) = -\frac{1}{12\pi} \int_{\Re(\mu)=0} \frac{F(\mu)\lambda_E(n, \mu)\overline{\lambda_E(m, \mu)}}{\prod_{i<j} |\zeta(1 + \mu_i - \mu_j)|^2} d\mu,
$$

where $F$ is a smooth function of $\mu$ satisfying certain regularity and symmetry conditions; see Theorem 10, and

$$
d\mu = d\mu_1 d\mu_2 = d\mu_1 d\mu_3 = d\mu_2 d\mu_3.
$$

For a smooth, compactly supported function $f$ on $(\mathbb{R}^+)^2$, define the integral transforms

$$
F_0(\mu) = F_0(f; \mu) = \frac{16}{\pi^4} \cos(f(\mu)) \int_{(\mathbb{R}^+)^2} f(t) W^{0*}(t, -2\mu) dt_1 dt_2 (t_1 t_2)^2,
$$

$$
F_1(\mu) = F_1(f; \mu) = \frac{1}{2} \tan \frac{\pi}{2}(\mu_1 - \mu_3) \tan \frac{\pi}{2}(\mu_2 - \mu_3) F_0(\mu),
$$

where

$$
\cos^0(\mu) = \frac{2}{\pi} \prod_{i<j} \cos \frac{\pi}{2}(\mu_i - \mu_j),
$$

and $W^{0*}(y, \mu)$ is the completed spherical Whittaker function

$$
W^{0*}(y, \mu) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\pi y_1)^{1-s_1} (\pi y_2)^{1-s_2} G^0(s, \mu) \frac{ds_1 ds_2}{(2\pi i)^2},
$$

$$
G^0(s, \mu) = \prod_{i=1}^{3} \Gamma \left( \frac{s_1 - \mu_i}{2} \right) \Gamma \left( \frac{s_2 + \mu_i}{2} \right) \Gamma \left( \frac{s_1 + s_2}{2} \right),
$$

with spectral parameters $\mu \in \mathbb{C}^3$, $\mu_1 + \mu_2 + \mu_3 = 0$. The contours of (11) should pass to the right of the poles of $G^0(s, \mu)$. Notice that in case $\pi$ is a permutation of $-\mu$ (in particular if $\Re(\mu) = 0$), we have $W^{0*}(y, \mu) = W^{0*}(y, -\mu)$.

The arithmetic Kuznetsov formula is given in terms of these integral transforms.

**Theorem 5.** Suppose $m, n \in \mathbb{Z}^2$ such that $m_1n_2, m_2n_1 > 0$, and $f$ smooth and compactly supported on $(\mathbb{R}^+)^2$, then

$$
\mathcal{K}_L(f) := \sum_{c_1, c_2 \in \mathbb{N}} S_{w_1}(\psi_m, \psi_n, c) f \left( \frac{2\sqrt{m_1n_2}}{c_1}, \frac{2\sqrt{m_2n_1}}{c_2} \right).
$$

has the spectral interpretation

$$
\mathcal{K}_L(f) = C^0(F_0) + C^1(F_1) + \mathcal{E}_\text{max}^0(F_0) + \mathcal{E}_\text{min}^0(F_0) + 2\mathcal{E}_\text{max}^1(F_1).
$$
Of course, the hypotheses on \( f \) may be reduced to absolute convergence of \( K_L(f) \), absolute convergence of \((8)\), and a bound of the form \( F_0(\mu) \ll (1 + \|\mu\|)^{-\frac{5}{2} - \varepsilon} \), with \( \|\mu\| \) the usual Euclidean norm on \( \mathbb{C}^3 \), see Section 7. An optimal result is well beyond the scope of this paper, but the following proposition gives a sufficient (though certainly not necessary) set of conditions:

**Proposition 6.** Suppose \( f : (\mathbb{R}^+)^2 \rightarrow \mathbb{C} \) and that there exists \( \sigma_1 > \theta, \sigma_2 \geq 0 \), with \( \theta \) as in Theorem 3, such that

1. \( f(y) \) is bounded by \((y_1y_2)^{1+\varepsilon}\) as \( y_1y_2 \to 0 \),
2. for \( j_1, j_2 \leq 2N, j_1 + j_2 \leq 2N \), the derivatives \( \hat{\partial}_{y_1}^{j_1} \hat{\partial}_{y_2}^{j_2} f(y) \) exist and are continuous,
3. for \( j_1, j_2 \leq 2N, j_1 + j_2 \leq 2N - 1 \), the derivatives \( y_1^{j_1} y_2^{j_2} \hat{\partial}_{y_1}^{j_1} \hat{\partial}_{y_2}^{j_2} (y_1y_2 f(y)) \) are bounded by \((y_1y_2)^{1+\varepsilon}\) as \( y_1y_2 \to 0 \) and are otherwise polynomially bounded in the coordinates of \( y \),
4. \( (\Delta_1)^N (y_1y_2 f(y)) \ll (y_1 y_2)^{1 + \sigma_1 (y_1 + y_2)^{2 \sigma_2}} \) for all \( y \in (\mathbb{R}^+)^2 \),

where \( N \) is the integer strictly larger than \( \sigma_1 + \sigma_2 + \frac{9}{4} \), and

\[
(14) \quad \Delta_1 = -y_1^2 \partial_{y_1}^2 - y_2^2 \partial_{y_2}^2 + y_1 y_2 \partial_{y_1} \partial_{y_2} + 4 \pi^2 (y_1^2 + y_2^2).
\]

Then \((13)\) holds for the test function \( f \).

The strength of Theorem 3 relies on non-trivial bounds on the spectral parameters; this was essentially proved in [18]:

**Theorem 7** (Kim, Sarnak). The spectral parameters \( \mu \) of a cusp form of weight at most one satisfy \(|\text{Re}(\mu_j)| < \frac{5}{14}\).

Although the theorem given in [22] is only stated for spherical forms, it trivially extends to the weight-one case; see the discussion in [10, Section 9.1].

Theorem 3 now follows from Theorem 5 and an analysis of the spectral expansion; the proof is given in Section 9, below.

### 2.2. The Kloosterman zeta function.

At this point, one would like to give the meromorphic continuation of the unweighted Kloosterman zeta function. We detour momentarily to analyze the failure of the obvious choice: It seems that one should, with the usual limiting argument, be able to take \( f(y) = (\pi y_1)^{2s_1}(\pi y_2)^{2s_2} \) in Theorem 5 so that the Kloosterman sum side becomes

\[
K_L(f) = (4\pi^2 m_1 n_2)^{s_1} (4\pi^2 m_2 n_1)^{s_2} \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{y_1}(\psi_{m_1}, \psi_{n_1}, c)}{c_1^{2s_1 - s_2} c_2^{2s_2 - s_1}},
\]

which converges on

\[
\{ s \in \mathbb{C}^2 \mid \frac{1}{2} < \text{Re}(s_1) < 2\text{Re}(s_2) - \frac{1}{2}, \frac{1}{2} < \text{Re}(s_2) < 2\text{Re}(s_1) - \frac{1}{2} \}.
\]

If we do this, then \( F_0(\mu) \) becomes

\[
F_0(\mu) = \frac{4}{\pi^4} \cos^0(\mu) G^0(2s, -2\mu),
\]

and here is where our problems begin. The function \( F_0(\mu) \) has exponential decay unless one of the \( \mu_i \) is small (see Section 8), but if, say, \( \mu_2 \) is not too large (compared to \( \text{Im}(s) \)),
and $\mu_1 = -\mu_3 = T$ is large (which rules out the complementary spectrum), we see $F_0(\mu) = T^{2\text{Re}(s_1 + s_2) - 2}$ for $\text{Re}(\mu) = 0$ by Stirling’s formula. Unfortunately, the number of such cusp forms (arithmetically weighted) is within a constant multiple of $T^4$ (see [1, Theorem 1] for the spherical case), so the spectral expansion only converges absolutely for $2\text{Re}(s_1 + s_2) < -2$. So it seems there is no overlap between the two representations, or more accurately, the usual limiting argument doesn’t apply; in particular, it is not possible to choose $\sigma_1$ and $\sigma_2$ for Proposition 6 due to the term $4\pi^2(y_1^2 + y_2^2)$ in (14), which requires increasing $\sigma_1$ and $\sigma_2$ as the number of derivatives $N$ grows at precisely the same rate. This differs from the behavior of the spectral expansions of the Kloosterman zeta functions on $GL(2)$ which have exponential convergence in the spectral expansion [19, 25], and we note that on the generic region where all $|\mu_i| = T$, this naive test function would also produce exponential convergence.

The solution on $GL(2)$ is to use an exponential to control the behavior at infinity, and write the weighted Kloosterman zeta function as a sum of shifts of the unweighted Kloosterman zeta function [27]. The meromorphic continuation of the unweighted function can then be obtained from that of the weighted function in stages. This iterative approach was formalized by Kiral [19], who obtained a sort of Möbius inversion for the power series expansion of the exponential to express the unweighted Kloosterman zeta function as a sum of shifts of the weighted Kloosterman zeta function, in the opposite-sign case.

To that end, one might choose

$$f(y) = (\pi y_1)^{2\alpha_1}(\pi y_2)^{2\alpha_2} \exp \left(-(\pi^9 y_1^5 y_2^4)^2 - (\pi^9 y_1^4 y_2^5)^2\right).$$

Then Proposition 6 applies with

$$\sigma_1 = \theta + \epsilon, \quad \sigma_2 = \epsilon, \quad N = 3$$

(not an optimal choice) on the region

$$10s_1 - 8s_2, 10s_2 - 8s_1 > 24 + \theta.$$ 

Unfortunately, it can be seen that the spectral side for this new weight function does not converge for, say $-\frac{1}{2} < \text{Re}(s_1) < 0$ and $\text{Re}(s_2) > -\frac{1}{2}$, for exactly the same reason as the naive test function in the previous attempt. In fact, studying the failure of this example leads one to believe that any test function of the form

$$f(y) = (\pi y_1)^{2\alpha_1}(\pi y_2)^{2\alpha_2} f_0(y)$$

where $f_0(y)$ satisfies some conditions on its Mellin transform (e.g. existence) and

$$f_0(y) = 1 + O\left(\sum_i y_1^{\alpha_i} y_2^{\beta_i}\right)$$

as $y \to 0$ with $\alpha_i$ and $\beta_i$ strictly positive, will encounter a similar difficulty. This is unfortunate, as we need $2\alpha_i - \beta_i \geq 0, 2\beta_i - \alpha_i \geq 0$ with one inequality strict — that is, the exponents of both $c_1$ and $c_2$ must be non-positive with one exponent strictly negative — in each error term to apply the iterative method.

Thus we consider the simple weighted Kloosterman zeta function of the theorem. We collect the relevant analysis into the following proposition:
Proposition 8. For \( f_s, \tilde{s} \) and \( Z^*_{m,n}(s) \) as in Theorem 4,

1. \((13)\) holds with \( K_L(f_s) = Z^*_{m,n}(s) \) on the region \( 2 \text{Re}(\tilde{s}_1), 2 \text{Re}(\tilde{s}_2) > 1 \),

2. \( F_0(s, \mu) := F_0(f_s, \mu) \) and \( F_1(s, \mu) := F_1(f_s, \mu) \) extend to meromorphic functions of all \( s \) and \( \mu \),

3. when \( |\tilde{s}| < T, \text{Re}(\mu) \) is in some compact set \( M \), and

\[ |\tilde{s}_1 + \mu_j + \ell|, |\tilde{s}_2 - \mu_j + \ell| > \eta > 0, \quad \text{for each } j \text{ and all } \ell \in \mathbb{N}_0, \]

we have \( F_0(s, \mu), F_1(s, \mu) \ll_{T, M, \eta} (1 + ||\mu||)^{-100} \),

4. when the coordinates of \( \mu \) are distinct modulo \( \mathbb{Z} \), \( F_0(s, \mu) \) and \( F_1(s, \mu) \) have simple poles whenever \( \tilde{s}_1 = -\mu_j - \ell \) or \( \tilde{s}_2 = \mu_j - \ell, \ell \in \mathbb{N}_0 \),

5. when \( \mu = (it, it, -2it) \) with \( 0 \neq t \in \mathbb{R} \), \( F_0(s, \mu) \) and \( F_1(s, \mu) \) have double poles whenever \( \tilde{s}_1 = -it - \ell \) or \( \tilde{s}_2 = it - \ell \), and simple poles whenever \( \tilde{s}_1 = 2it - \ell \) or \( \tilde{s}_2 = -2it - \ell, \ell \in \mathbb{N}_0 \),

6. \( F_0(s, 0) \) has triple poles whenever \( \tilde{s}_1 = -\ell \) or \( \tilde{s}_2 = -\ell, \ell \in \mathbb{N}_0 \),

7. \( F_1(s, 0) = 0 \).

These computations are of similar type to [17, Section 9.2] and [4], and we put them off to Sections 8 and 10. Notice the order of each pole is the maximum number of terms of \((4)\) which coincide for any given \( \mu \).

From parts 1, 2 and 3 of the proposition, the right-hand side of \((13)\) gives a meromorphic continuation of \( Z^*_{m,n}(s) \) to all \( s \). The spectral parameters \( \mu \) of spherical and weight-one forms must satisfy the unitarity condition and the bound towards Ramanujan-Selberg \( (\text{e.g. } \theta = \frac{5}{14} \text{ as in Theorem 7}) \):

\[ \max_i |\text{Re}(\mu_i)| < \frac{1}{2}, \quad -\pi \text{ a permutation of } \mu. \]

For such \( \mu \), having \( \mu_j - \mu_k \in \mathbb{Z} \) implies \( \mu \) is a permutation of \((it, it, -2it)\), but \( F_0(s, \mu) \) is permutation-invariant in \( \mu \) and \( F_1(s, \mu) \) is zero on the other permutations, so the proposition covers all relevant cases in \( \mu \). Thus Theorem 4 follows immediately from Proposition 8 and the associated Weyl laws [1, Theorem 1] and [10, Theorem 1].

It is known [12,23,24] that \( \mu = 0 \) does not correspond to a spherical cusp form on \( SL(3, \mathbb{Z}) \), so we only need to worry about \( \mu = 0 \) in the terms \( \mathcal{E}_{\text{max}}^0 \) and \( \mathcal{E}_{\text{min}}^0 \). On the other hand, it is known that \( 0 \) is not the spectral parameter of any cusp form on \( SL(2, \mathbb{Z}) \) [26], i.e. in \((6)\) we have \( \mu_\phi \neq 0 \), so the \( \mu \) argument of \( F_0(s, \mu) \) is never zero in \( \mathcal{E}_{\text{max}} \), and the integrand of \( \mathcal{E}_{\text{min}} \) has a triple zero at \( \mu = 0 \) due to the poles of the six zeta functions in the denominator there. So the point \( \mu = 0 \) and its triple pole need not enter into our computations.

The residues of \( Z^*_{m,n}(s) \) at say \( \tilde{s}_1 = -\mu_k - \ell \) for \( \mu \) the spectral parameters of some cusp form (of either weight) and \( \ell \in \mathbb{N}_0 \) are then

\[ \frac{2\pi}{3} \sum_{d \mid 1,0} \sum_{j \in \{1,2,3\}} \sum_{\varphi \in S^d} \frac{\lambda_\varphi(m)\lambda_\varphi(n)}{L(\text{Ad}^d \varphi, 1)} \text{ res } \tilde{s}_1 = -\mu_k - \ell F_d(s, \mu_\varphi) + \text{ Eisenstein contribution}. \]

The residues of \( F_0(s, \mu) \) (from which follow those of \( F_1(s, \mu) \)), up to symmetry and away from \( \mu = 0 \), are listed in \((42),(34),(36)\).

We have omitted the polar analysis of the minimal and maximal parabolic Eisenstein terms, as they are somewhat more complicated. Essentially, when \( \tilde{s} \) is close to a pole of
$F_d(s, \mu)$, we deform the $\mu$ or $r$ contour past the pole, then ask if $\tilde{s}$ is a pole of the resulting residual term. The essential idea is the same as writing

$$y^s e^{-y} = \begin{cases} \int_{\Re(u) = \epsilon} \Gamma(u + s) y^{-u} \frac{du}{2\pi i} & \Re(s) \geq 0, \\ \int_{\Re(u) = \epsilon} \Gamma(u + s) y^{-u} \frac{du}{2\pi i} + \sum_{\ell=0}^{T} \frac{(-1)^\ell}{\ell!} y^{s+\ell} & -T - 1 \leq \Re(s) < -T \in \mathbb{N}_0. \end{cases}$$

The summands of the $\ell$ sum here are referred to as the residual terms. (Note that the $\epsilon$ in the contour necessarily depends on $\Re(s)$.)

Poles of the residual terms may result from the zeros of the $L$-functions in the denominators of (6) or (7), or from secondary poles of the residue of $F_d(s, \mu)$. This is further complicated by the fact that it is not possible to manipulate the $r$ or $\mu$ contours so that all three coordinates of $\mu$ lie on the same side of $\Re(\mu) = 0$ (since the coordinates sum to zero), so these secondary poles have a different form when one or more of $\tilde{s}_i \in -\mathbb{N}_0 + i\mathbb{R}$.

The possible secondary poles of the residual terms coming from the minimal parabolic Eisenstein series occur whenever $\Re(\tilde{s}_1), \Re(\tilde{s}_2) \leq 0$ and $\tilde{s}_1 + \tilde{s}_2 \in -\mathbb{N}_0$. The possible poles coming from the zeta functions in the denominator occur whenever $\Re(\tilde{s}_1) \leq -\ell_1 \in -\mathbb{N}_0$, $\Re(\tilde{s}_2) \leq -\ell_2 \in -\mathbb{N}_0$, and any of the numbers

$$\tilde{s}_1 + \ell_1 + \tilde{s}_2 + \ell_2, \quad 2(\tilde{s}_1 + \ell_1) - (\tilde{s}_2 + \ell_2), \quad 2(\tilde{s}_2 + \ell_2) - (\tilde{s}_1 + \ell_1)$$

are of the form $\pm(1 - z)$ for $z$ any non-trivial zero of $\zeta(z)$ (the trivial zeros are killed by $\cos^0(\mu)$).

The possible secondary poles of the residual terms coming from the maximal parabolic Eisenstein series attached to $\phi \in S_2^d$ occur whenever $\Re(\tilde{s}_1) \leq -\ell_1 \in -\mathbb{N}_0$, $\Re(\tilde{s}_2) \leq -\ell_2 \in -\mathbb{N}_0$ and either $\tilde{s}_1 + \tilde{s}_2 \in -\mathbb{N}_0$ or any of the numbers (17) is $\pm 2\mu_\phi$. The possible poles coming from the $L$-functions in the denominator occur whenever one of $\Re(\tilde{s}_i) \leq -\ell_i \in -\mathbb{N}_0$ and either of the numbers

$$\frac{1}{2}(\tilde{s}_i + \ell_i), \quad \tilde{s}_i + \ell_i \pm \mu_\phi,$$

are of the form $\pm \frac{1}{3}$ for $z$ any zero of $L(\phi, z)$.

3. Acknowledgements

The author would like to thank Valentin Blomer for many helpful comments. The author would also like to thank the referees for their comments which uncovered a significant mistake with the Kloosterman zeta function in a previous version of this paper.

4. The spherical Whittaker function

We follow the notation of [8, Section 2.1]. In particular, let $U(\mathbb{R})$ be the space of upper-triangular, unipotent matrices in $PSL(3, \mathbb{R})$ and $Y^+$ the space of positive diagonal matrices, in the form $y = \text{diag}(y_1y_2, y_1, 1)$ with measure $dy = \frac{dy_1dy_2}{(y_1y_2)^2}$. We frequently treat $y \in Y^+$ as elements of $(\mathbb{R}^+)^2$ as the multiplication is the same, and the same for the signed diagonal matrices $Y \cong (\mathbb{R}^\times)^2$. 

If we take the usual power function \( p_{\rho+\mu}(y) = |y_1|^{1-\mu_3}|y_2|^{1+\mu_1} \) with \( \rho = (1,0,-1) \), then the incomplete spherical Whittaker function is defined by the Jacquet integral

\[
W^0(g, \mu, \psi_n) = \int_{U(\mathbb{R})} p_{\rho+\mu}(wuyg)\overline{\psi_n(u)}du,
\]

for \( \psi_n \) a character of \( U(\mathbb{R}) \). It can be shown (see [15] or the more complicated cases of [9]) that this has the Mellin-Barnes integral (11) with

\[
\Lambda(\mu) = \pi^{-\frac{3}{2}+\mu_3-\mu_1} \Gamma\left( \frac{1+\mu_1-\mu_2}{2} \right) \Gamma\left( \frac{1+\mu_2-\mu_3}{2} \right) \Gamma\left( \frac{1+\mu_3-\mu_1}{2} \right),
\]

\[W^0(y, \mu) := \Lambda(\mu)W^0(y, \mu, \psi_{1,1}).\]

This fixes the various normalization constants to give us a consistent starting point.

The power function induces an action of the Weyl group \( W \) (as in [8, Section 2.1]) on the coordinates of \( \mu \) by \( p_{\mu w}(y) = p_{\mu}(wy^{-1}) \) for \( w \in W \):

\[
\begin{align*}
\mu^I & = (\mu_1, \mu_2, \mu_3), & \mu^{w_2} & = (\mu_2, \mu_1, \mu_3), & \mu^{w_3} & = (\mu_1, \mu_3, \mu_2), \\
\mu^{w_4} & = (\mu_3, \mu_1, \mu_2), & \mu^{w_5} & = (\mu_2, \mu_3, \mu_1), & \mu^{w_1} & = (\mu_3, \mu_2, \mu_1).
\end{align*}
\]

We say that a function \( F(\mu) \) is Weyl invariant if \( F(\mu) = F(\mu^w) \) for all \( w \in W \).

It is known that the Whittaker functions give a continuous basis of the Schwartz-class functions on \((\mathbb{R}^+)^2\) by the Kontorovich-Lebedev theorem of Wallach [29] (see [16]):

**Theorem 9.** Let \( F(\mu) \) be Schwartz-class, Weyl-invariant and holomorphic for \( \Re(\mu) \) in a neighborhood of \( 0 \), and let \( f(y) \) be Schwartz class on \( Y^+ \), then

\[
F(\mu) = \int_{Y^+} \int_{\Re(\mu')=0} F(\mu')W^{0*}(y, \mu')\sin^0(\mu')d\mu'W^{0*}(y, -\mu)dy,
\]

\[
f(y) = \int_{\Re(\mu)=0} \int_{Y^+} f(t)W^{0*}(t, -\mu)dtW^{0*}(y, \mu)\sin^0(\mu)d\mu,
\]

with the integrals to be taken iteratively, where

\[
(18) \quad \sin^0(\mu) := \frac{1}{6(2\pi i)^2} \prod_{i \neq j} \Gamma\left( \frac{\mu_i-\mu_j}{2} \right) = \frac{1}{192\pi^5} \prod_{i \neq j} (\mu_i - \mu_j) \sin \frac{\pi}{2}(\mu_i - \mu_j).
\]

Note: In comparing to [16, Definitions 1.1, 1.2 and Theorem 1.3], we have the conversion \( 3it = (\mu_1 - \mu_2, \mu_2 - \mu_3) \) and Stade’s formula [16, eq. (2.1)] in that paper is missing a factor \( 2^{n-2} \) from the denominator of the right-hand side. The conjugations may be explained by \( W^{0*}(y, \mu) = W^{0*}(y, -\mu) \) on \( \Re(\mu) = 0 \). The method of proof in [16] requires holomorphy of \( F \), and they technically only prove one direction, but Wallach’s proof culminating in [29, Section 15.10.3 eqs. (1) and (2)] is somewhat daunting. The constants and the spectral measure are verified through the computation of Section 5.1.

5. **Spectral Kuznetsov Formulae**

We refer to [10, Section 4.5] for the definitions of the power-series solutions \( J_w(y, \mu) \) to the differential equations satisfied by the Kuznetsov kernel functions. Their precise definition is not needed here, as we will only use them in an algebraic sense.
For $F(\mu)$ Schwartz-class and holomorphic for $\Re(\mu)$ in a neighborhood of 0, we define the integral transformation

$$H^d_w(F; y) = \frac{1}{|y_1 y_2|} \int_{\Re(\mu) = 0} F(\mu) K^d_w(y, \mu) \text{spec}^d(\mu) d\mu$$

using the spectral weights

$$\text{spec}^0(\mu) := \frac{\sin^0(\mu)}{\cos^0(\mu)} = \frac{1}{384\pi^4} \prod_{i<j} (\mu_i - \mu_j) \tan \frac{\pi}{2}(\mu_i - \mu_j),$$

$$\text{spec}^1(\mu) = \frac{1}{64\pi^4} \left( \prod_{i<j} (\mu_i - \mu_j) \right) \cot \frac{\pi}{2}(\mu_1 - \mu_3) \cot \frac{\pi}{2}(\mu_2 - \mu_3) \tan \frac{\pi}{2}(\mu_1 - \mu_2),$$

and the kernel functions

$$K^0_1(y, \mu) = 1,$$

$$K^0_w(y, \mu) = \frac{1}{8\pi} \sum_{w \in W_3} \frac{J_w(y, \mu^w)}{\sin \frac{\pi}{2}(\mu_1^w - \mu_3) \sin \frac{\pi}{2}(\mu_2^w - \mu_3)},$$

$$J^1_w(y, \mu) = -\sin \frac{\pi}{2} (\mu_1 - \mu_2) J_w(y, \mu) - i \varepsilon_1 \cos \frac{\pi}{2} (\mu_1 - \mu_3) J_{w^4}(y, \mu^w),$$

$$+ i \varepsilon_1 \cos \frac{\pi}{2} (\mu_2 - \mu_3) J_{w^4}(y, \mu^w),$$

$$K^1_w(y, \mu) = \frac{1}{8\pi} \sum_{w \in W_3} \frac{J_{w^4}(y, \mu^w)}{\cos \frac{\pi}{2}(\mu_1 - \mu_3) \sin \frac{\pi}{2}(\mu_2 - \mu_3)},$$

$$K^d_w(y, \mu) = K^d_{w^4}((-y_2, y_1), -\mu),$$

$$K^0_w(y, \mu) = \frac{1}{16\pi} \sum_{w \in W_3} \frac{J_w(y, \mu^w) - J_{w^4}(y, \mu^w)}{\prod_{i<j} \sin \frac{\pi}{2}(\mu_i - \mu_j)},$$

$$J^1_w(y, \mu) = \varepsilon_2 J_w(y, \mu) + \varepsilon_1 J_{w^4}(y, \mu^w) + \varepsilon_1 \varepsilon_2 J_{w^4}(y, \mu^w),$$

$$K^1_w(y, \mu) = -\frac{1}{16\pi} \sum_{w \in W_3} \frac{J_{w^4}(y, \mu^w) - J_{w^4}(y, \mu^w)}{\cos \frac{\pi}{2}(\mu_1 - \mu_3) \sin \frac{\pi}{2}(\mu_2 - \mu_3)},$$

with $\varepsilon = \text{sgn}(y)$, and $W_3 = \{1, w_4, w_5\}$ the subgroup of order 3 permutations in the Weyl group. The weight-one formulae are [10, (4)-(7)] and the surrounding displays; the spherical formulae are derived in Section 5.1, below.

With these integral transforms, we consider the following averages of Kloosterman sums at each Weyl cell with $m, n \in \mathbb{Z}^2$, $m_1 m_2 n_1 n_2 \neq 0$:

$$K^d_{w^4}(F) = \delta_{|m_1|=|m_1|} H^d_{1}(F; (1, 1))$$

$$K^d_{w^4}(F) = \sum_{\varepsilon \in \{\pm 1\}^2} \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w^4}(\psi_{c_1}, \psi_{c_2}, c)}{c_1 c_2} H^d_{w^4}(F; (\varepsilon_1 \varepsilon_2 m_1 m_2 n_1 n_2, 1)), $$
\[ \mathcal{K}_s^d(F) = \sum_{\varepsilon \in \{\pm 1\}^2} \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w_5}(\psi_{m_1} \psi_{m_2}, c)}{c_1 c_2} H_{w_5}^d \left( F; \left( 1, \varepsilon_1 \varepsilon_2 \frac{m_1^2 m_2 n_1}{c_1^2 n_2^2} \right) \right), \]

\[ \mathcal{K}_L^d(F) = \sum_{\varepsilon \in \{\pm 1\}^2} \sum_{c_1, c_2 \in \mathbb{N}} \frac{S_{w_1}(\psi_{m_1} \psi_{m_2}, c)}{c_1 c_2} H_{w_1}^d \left( F; \left( \varepsilon_1 \varepsilon_2 \frac{m_1 n_2 c_2}{c_1}, \varepsilon_1 \frac{m_2 n_1 c_1}{c_2} \right) \right), \]

and at last we are ready for the spectral Kuznetsov formulæ:

**Theorem 10.** Let \( F(\mu) \) be Schwartz-class and holomorphic on \( \{ \mu | |\text{Re}(\mu)| < \frac{1}{2} + \delta \} \) for some \( \delta > 0 \). Suppose \( m, n \in \mathbb{Z}^2 \) with \( m_1 m_2 n_1 n_2 \neq 0 \) and recall (5)-(7).

0. If \( F(\mu) \) is invariant under all \( \mu \mapsto \mu^w \) and \( F(\mu) = 0 \) whenever \( \mu_i - \mu_j = \pm 1, i \neq j \), then we have the spherical Kuznetsov formula,

\[ C^0(F) + C_{\text{max}}^0(F) + C_{\text{min}}^0(F) = \mathcal{K}_s^0(F) + \mathcal{K}_s^0(F) + \mathcal{K}_L^0(F). \]

1. If \( F(\mu) \) is invariant under the transposition \( \mu_1 \leftrightarrow \mu_2 \) and \( F(\mu) = 0 \) whenever \( \mu_1 - \mu_2 = \pm 1 \), then we have the weight-one Kuznetsov formula,

\[ C_1(F) + 2C_{\text{max}}^1(F) = \mathcal{K}_s^1(F) + \mathcal{K}_s^1(F) + \mathcal{K}_L^1(F). \]

In the weight-one case, this is [10, Theorem 6]; the spherical case is [7, Theorems 1 and 4], with the corrected constants which we now compute.

5.1. **The constants in the spherical formula.** In what follows, the leading constants in the terms of the arithmetic and spectral sides of both spectral Kuznetsov formulæ will be essential to the construction of the arithmetic Kuznetsov formula. Unfortunately, these constants in [7] are generally incorrect. The key difficulties are that the paper [7] relied on outside formulæ such as spherical inversion (section 1.8 there; the constant there is extremely difficult to verify), the Mellin transform of the Whittaker function (section 1.4 there), Li’s Kuznetsov formula (section 1.9 there) as computed in [15] and [5], especially Stade’s formula (eq. (17) there), that [7] has the extra step of computing the Fourier transform of the spherical function (eq. (22) there), and that [7] attempts to evaluate the final \( x' \) and \( t \) integrals simultaneously (see the comment at the bottom of page 5 in [10]).

The papers [10] and [11] avoid these difficulties by computing all of the formulæ directly, and there are far fewer steps. Most of these computational errors in the constants of [7] were eventually corrected, to the point that the ratio of the constants on the trival term of the arithmetic side and cuspidal term of the spectral side is known to be correct by heuristic verification (see [3]), but the remaining terms have no such means of verification, and it is difficult to check that the various corrections were properly carried through to those terms.

The solution is to repeat the computation from first principles using the methods of [10] and [11]. We briefly summarize the computation here as it is essentially identical to those papers.

Stade’s formula is computed (from (11) as in [10, Section 5]) to be

\[ \Psi^0(\mu, \mu', t) := \int_{Y'} W^{0*}(y, \mu) W^{0*}(y, \mu')(y_1^2 y_2)' dy = \frac{1}{4\pi^3 \Gamma \left( \frac{3t}{2} \right)} \prod_{i,j} \Gamma \left( \frac{t + \mu_i + \mu_j'}{2} \right), \]
and at $t = 1$ and $t = 0$, this is
\[
\frac{1}{\cos^{0}(\mu)} := \Psi^{0}(\mu, -\mu, 1), \quad \frac{1}{\sin^{0}(\mu)} = 2(2\pi i)^{2} \lim_{t \to 0} t^{2} \Psi^{0}(\mu, -\mu, t),
\]
which produce the formulae (10), (18) and (19).

The generalization of Kontorovich-Lebedev inversion was already stated in Theorem 9.

As in [11, Section 7.2] (see also the last paragraph of the introduction to [10, Section 7.1]),
the kernel functions are defined by the Riemann integral
\[
K_{w}^{0}(y, \mu)W^{0*}(t, \mu) = \int_{U_{w}(\mathbb{R})} W^{0*}(ywxt, \mu) \overline{\psi_{1,1}(x)}dx, \quad y \in Y, t \in Y^{+},
\]
and we determine the combinations of the power-series solutions by the asymptotics (via (11))
\[
W^{0*}(y, \mu) \sim \frac{1}{2\pi} \sum_{w \in W} \pi^{-\mu_{w}} |y| |1 - \mu_{w}| \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \int_{-\infty}^{\infty} (\pi y_{2})^{1-s_{2}} \Gamma \left( \frac{s_{2} + \mu_{w}}{2} \right) \frac{ds_{2}}{2\pi}
\]
\[
\sim \sum_{w \in W} \pi^{\mu_{w} - \mu_{w}} \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) W^{0*}(t, \mu),
\]
as $y_{1} \to 0$ in the first case or $y \to 0$ in general in the second, for $\mu$ with $\text{Re}(\mu) = 0$ and
$\mu_{i} \neq \mu_{j}, i \neq j$. For the long-element term inserting the $y \to 0$ asymptotic into the definition
gives
\[
K_{w_{1}}^{0}(y, \mu)W^{0*}(t, \mu) \sim \sum_{w \in W} \pi^{\mu_{w} \mu_{w}} \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \frac{\Lambda(\mu^{w})}{\Lambda(\mu^{w})} W^{0*}(t, \mu),
\]
and inserting the $y_{1} \to 0$ asymptotic gives
\[
K_{w_{1}}^{0}(y, \mu)W^{0*}(t, \mu) \sim \sum_{w \in W} \pi^{1-3\mu_{w}} |y| |1 - \mu_{w}| \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \Gamma \left( \frac{\mu_{w}^{\mu} - \mu_{w}}{2} \right) \frac{\Lambda(\mu^{w})}{\Lambda(\mu^{w})} W^{0*}(t, \mu),
\]
along $y_{2} = 1$. Then comparing with the first-term asymptotics of each $J_{w}(y, \mu)$ determines
the kernel functions completely, giving (22) and (26).

The formulas (21) and (25) for $K_{I}^{0}(y, \mu)$ and $K_{w_{1}}^{0}(y, \mu)$ follow from $U_{I}(\mathbb{R}) = \{I\}$ and
$W^{0*}(y, \mu) = W^{0*}((y_{2}, y_{1})$), $-\mu$) with the reasoning of [10, Section 7.1.4]. This completes the
corrections for the constants on the transforms $H_{w}$ in [7, Theorem 1] as well as the formulae
for the functions $K_{w}$ on page 6686 of [7], keeping in mind the notational differences discussed
in the introduction to [10, Section 4].

The constants on the spectral side of the spherical formula come from the correction
\[
\int_{\mathbb{R}^{2}} \text{Im}(z_{2})^{\frac{1}{2} + \mu_{w}} (u_{3}^{2} + u_{2}^{2} + 1)^{-\frac{3}{2} - \frac{3}{2}\mu_{1}} du_{2} du_{3} = B \left( \frac{1}{2}, \frac{3\mu_{1} \pm \mu_{1}}{2} \right) B \left( \frac{1}{2}, \frac{3\mu_{1} \pm \mu_{1}}{2} \right),
\]
which brings [7, Appendix A] in line with the computations of [8, Section 5.5]. In terms of \( \cos^0(\mu, \nu) \), the conversion to Hecke eigenvalues given in [7, Appendix A] becomes

\[
\cos^0(\mu, \nu) |\rho^0_{\nu}(1)|^{-2} = \frac{3}{2\pi} L(\Ad^2 \phi, 1),
\]

\[
\cos^0(\mu) |\rho(1, \mu)|^{-2} = \frac{1}{8\pi} |\zeta(1 + \mu_1 - \mu_2)\zeta(1 + \mu_2 - \mu_3)\zeta(1 + \mu_1 - \mu_3)|^2, \quad \text{Re}(\mu) = 0,
\]

\[
\cos^0(\mu, \phi + r, -\mu_\phi + r, -2r) |\rho_{\phi}(1, r)|^{-2} = \frac{1}{\pi} L(\Ad^2 \phi, 1) |L(\phi, 1 + 3r)|^2, \quad \text{Re}(r) = 0.
\]

6. The Arithmetic Kuznetsov Formula

We rely fundamentally on the fact [7, Theorem 2] that

\[
K^0_{w_l}(y, \mu) = \frac{\pi^5}{2} \cos^0(\nu) \sqrt{y_1 y_2 W^0)((2\sqrt{y_1}, 2\sqrt{y_2}), 2\mu),
\]

when \( \text{sgn}(y) = (1, 1) \).

Suppose \( F_0(\mu) \) is given by (8) for some smooth, compactly supported \( f : (\mathbb{R}^+)^2 \to \mathbb{C} \), and define \( F_1(\mu) \) by (9). Notice that \( F_0(\mu) \) satisfies the conditions of the spherical spectral Kuznetsov formula, Theorem 10.0 since \( W^0((t, -2\mu) \) is entire in \( \nu \) and an eigenfunction of the restricted Laplacian (14) (using the positive Laplacian as in [9, eqs. (42),(43)]):

\[
(\lambda_{\nu}(t, \mu)) W^0(t, \mu) = \Delta_1 W^0(t, \mu) = 1 - \frac{\mu^2 + \nu^2 + \mu^2}{2}.
\]

Then \( F_1(\mu) \) also satisfies the conditions of the weight-one spectral Kuznetsov formula, Theorem 10.1. The zeros hypotheses are covered by the \( \cos^0(\nu) \) factor and the symmetries by the Weyl-invariance of \( W^0((t, -2\mu) \).

Note that, in general, for functions \( F_0 \) and \( F_1 \) related by (9), even if \( F_0 \) is positive on the spherical spectrum, \( F_1 \) is not necessarily positive on the weight-one spectrum, and visa versa.

We add the two Kuznetsov formulæ.

6.1. The long-element term. With \( \varepsilon = \text{sgn}(y) \), the long element weight functions can be written

\[
H^0_{w_l}(F_0; y) = -\frac{6}{|y_1 y_2|} \int_{\text{Re}(\mu) = 0} F_0(\mu) J_{w_l}(y, \mu) \frac{\text{spec}^0(\nu)}{16\pi \prod_{1<j} \sin \frac{\pi}{2} (\mu_i - \mu_j) \, d\mu},
\]

\[
H^1_{w_l}(F_1; y) = -\frac{6}{|y_1 y_2|} \int_{\text{Re}(\mu) = 0} (\varepsilon_2 + \varepsilon_1 + \varepsilon_1 \varepsilon_2) F_0(\mu) J_{w_l}(y, \mu) \frac{\text{spec}^0(\nu)}{16\pi \prod_{1<j} \sin \frac{\pi}{2} (\mu_i - \mu_j) \, d\mu},
\]

which follows directly from (26) and (28) using \( F_0(\mu^w) = F_0(\mu) \) for all \( w \in W \). Combining these expressions into \( H^*_{w_l}(F_0; y) := H^0_{w_l}(F_0; y) + H^1_{w_l}(F_1; y) \) with \( 1 + \varepsilon_2 + \varepsilon_1 + \varepsilon_1 \varepsilon_2 = 4\delta_{\varepsilon = (1,1)} \) gives

\[
H^*_{w_l}(F_0; y) = 4\delta_{\varepsilon = (1,1)} H^0_{w_l}(F_0; y) = 4\delta_{\varepsilon = (1,1)} \frac{1}{|y_1 y_2|} \int_{\text{Re}(\mu) = 0} F_0(\mu) K^0_{w_l}(y, \mu) \text{spec}^0(\nu) \, d\mu,
\]
to which we may apply (29). Thus for \( y_1, y_2 > 0 \), we have

\[
H_{w_1}^*(F_0; (y_1^2/4, y_2^2/4)) = \frac{8\pi^5}{y_1 y_2} \int_{\text{Re}(\mu) = 0} F_0(\mu)W^{0*}(y, 2\mu)\sin^0(\mu)\,d\mu,
\]

recalling (19).

The double-angle formula gives

\[
32\pi \cos^0(\mu)\sin^0(\mu) = \sin^0(2\mu),
\]

so applying (8) gives

\[
H_{w_2}^*(F_0; (y_1^2/4, y_2^2/4)) = \frac{4}{y_1 y_2} \int_{\text{Re}(\mu) = 0} \int_{Y^+} t_1 t_2 f(t)W^{0*}(t, -2\mu)dtW^{0*}(y, 2\mu)\sin^0(2\mu)\,d\mu.
\]

Then Kontorovich-Lebedev inversion, Theorem 9, gives

\[
H_{w_1}^*(F_0; (y_1^2/4, y_2^2/4)) = f(y).
\]

6.2. The other terms. The trivial terms satisfy

\[
H_1^1(F_1; I) = \int_{\text{Re}(\mu) = 0} F_0(\mu) \frac{1}{3} \sum_{\alpha \in \omega_3} \frac{1}{2} \tan \frac{\pi}{2} (\mu_1^w - \mu_3^w) \tan \frac{\pi}{2} (\mu_2^w - \mu_3^w) \spec^1(\mu^w)\,d\mu
\]

\[
= - H_1^1(F_0; I),
\]

which follows from the Weyl invariance of \( F_0 \) and the triple tangent identity

\[
\sum_{\alpha \in \omega_3} \tan \frac{\pi}{2} (\mu_1^w - \mu_2^w) = - \prod_{i < j} \tan \frac{\pi}{2} (\mu_i - \mu_j).
\]

Directly from the Weyl invariance of \( F_0 \) and equations (22), (23), (24), (19) and (20) we see the \( w_4 \) weight functions satisfy

\[
H_{w_4}^0(F_0; y) = \frac{3}{|y_1|} \int_{\text{Re}(\mu) = 0} F_0(\mu) J_{w_4}(y, \mu) \sin \frac{\pi}{2} (\mu_1 - \mu_2) \frac{\spec^0(\mu)}{8\pi} \frac{1}{\prod_{i < j} \sin \frac{\pi}{2} (\mu_i - \mu_j)}\,d\mu
\]

\[
= - H_{w_4}^1(F_1; y),
\]

and (25) implies also

\[
H_{w_3}^0(F_0; y) + H_{w_3}^1(F_1; y) = 0.
\]

This completes the proof of Theorem 5.

7. The proof of Proposition 6

Suppose \( f, \sigma_1, \sigma_2, \theta \) and \( N \) are as in the statement of the proposition and \( \{f_\ell\} \) is any sequence of smooth, compactly supported functions on \((\mathbb{R}^+)^2\) tending to \( f \) pointwise such that the derivatives \( \partial_{y_1}^\ell \partial_{y_2}^\ell f_\ell \) also tend to \( \partial_{y_1}^\ell \partial_{y_2}^\ell f \) pointwise. We may assume \( f \) and all \( f_\ell \) satisfy conditions 1, 3 and 4 of the proposition with the same implied constants. (The conditions may be expressed in terms of a collection of seminorms on an \( L^1 \)-space to which we apply the usual density theory.) The proof then consists of three applications of dominated convergence.
On the Kloosterman sum side, condition 1 of the proposition is sufficient since on \( y_1 = \left( \frac{2\sqrt{m_1n_2c_2}}{c_1}, \frac{2\sqrt{m_2n_1c_1}}{c_2} \right) \), we have
\[
(y_1y_2)^{1+\epsilon} = \left( \frac{16m_1m_2n_1n_2}{c_1c_2} \right)^{\frac{1+\epsilon}{2}},
\]
which is sufficient for convergence of the sum of Kloosterman sums by (1). (Note that \( y_1y_2 \to 0 \) as \( c_1, c_2 \to \infty \) so only the behavior as \( y_1y_2 \to 0 \) is relevant.)

On the spectral side, we proceed in two steps: Let \( \tilde{f}(y) = (y_1y_2)^{-1}(\tilde{\Delta}_1)^N(y_1y_2f) \) and similarly for \( \tilde{f}_c \), then with \( \lambda_1(\mu) \) as in (30),
\[
\lim_{\ell \to \infty} F_0(f_\ell; \mu) = \lambda_1(-2\mu)^{-N} \lim_{\ell \to \infty} F_0(\tilde{f_\ell}; \mu) = \lambda_1(-2\mu)^{-N} F_0(\tilde{f}; \mu) = F_0(f_\ell; \mu),
\]
as we now explain.

The first and last equalities of (31) hold because \( \tilde{\Delta}_1 \) is self-adjoint with respect to the inner product
\[
\langle h_1, h_2 \rangle = \int_{\mathbb{R}^+} h_1(y)\overline{h_2(y)}\,dy,
\]
which can be seen directly through integration by parts or by noting that \( \tilde{\Delta}_1 \) is a certain restriction of the Laplacian as in [15, Theorem 6.1.6]. (The elements of the operator algebra considered there are not typically self-adjoint, but the Laplacian certainly is.) Condition 3 of the proposition assures that the boundary terms in the integration by parts are all zero since the Whittaker function has super-polynomial (in fact, exponential) decay as \( y_i \to \infty \). (Note that if some \( y_i \to 0 \) while \( y_1y_2 \to 1 \), then necessarily \( y_3 \to \infty \), so condition 3 is sufficient along the boundaries \( y_i = 0 \) as well as \( y_i = \infty \).)

The central equality of (31) holds by dominated convergence using the bound of condition 4. Here we need some bound on the Whittaker function, and we use [1, Proposition 1]:

**Theorem 11** (Blomer). Suppose \( \mu \) satisfies the unitaricity condition and the bound towards Ramanujan-Selberg (15). Then for any \( A > a_1 > |a_2| + \theta, a_2 \in \mathbb{R} \), we have the bound
\[
W^{0*}(y, \mu) \ll_A (y_1y_2)^{1-a_1}(y_1/y_2)^{a_2}(1 + \|\mu\|)^{2a_1-\frac{3}{2}+\epsilon}.
\]

We split the integral defining \( F_0(\tilde{f}; \mu) \) or \( F_0(\tilde{f}_\ell; \mu) \) into four pieces along the curves \( y_1y_2 = 1 \) and \( y_1 = y_2 \). Then we apply Theorem 11 on each piece with \( a_1 = \sigma_1 + \sigma_2 + \epsilon \) according to \( y_1y_2 > 1 \) or \( y_1y_2 < 1 \) and \( a_2 = \pm(\sigma_2 - \epsilon) \) according to \( y_1 < y_2 \) or \( y_1 > y_2 \). By condition 4 of the proposition, the resulting integrals converge absolutely, and we have the bound
\[
\lambda_1(-2\mu)^{-N} F_0(\tilde{f}; \mu) \ll \lambda_1(-2\mu)^{-\sigma_1-\sigma_2-\frac{1}{2}}(1 + \|\mu\|)^{2\sigma_1+2\sigma_2-\frac{1}{2}+\epsilon} \ll (1 + \|\mu\|)^{-5-\epsilon},
\]
and the same for \( \lambda_1(-2\mu)^{-N} F_0(\tilde{f}_\ell; \mu) \), with the same implied constants. From the Weyl law [10, Theorem 1] and its spherical counterpart [1, Theorem 1] (which also dominate the Weyl laws for the continuous spectra), a bound of \((1 + \|\mu\|)^{-5-\epsilon}\) is sufficient for dominated convergence on the spectral expansion, so we are done.
8. The poles, residues and decay of $G^0$

Define

$$\tilde{F}_0(u, \mu) = \frac{4}{\pi^4} \cos^0(\mu) G^0(2u, -2\mu),$$

and by analogy,

$$\tilde{F}_1(u, \mu) = \frac{1}{2} \tan \frac{\pi}{2}(\mu_1 - \mu_3) \tan \frac{\pi}{2}(\mu_2 - \mu_3) \tilde{F}_0(u, \mu).$$

These functions are connected to the integral transforms $F_d(f; \mu)$ through Plancherel’s theorem for the Mellin transform.

Since the tangents are particularly non-threatening, we analyze $\tilde{F}_0$ and leave the other to the reader. First, we note that Stirling’s formula implies $\tilde{F}_0(u, \mu)$, for $\text{Re}(\mu)$ and $\text{Re}(u)$ in fixed compact sets, has no exponential growth and in fact decays exponentially in $\text{Im}(\mu)$ or $\text{Im}(u)$ unless

$$\text{Im}(\mu_1) \leq \text{Im}(u_2) \leq \text{Im}(\mu_2) \leq \text{Im}(-u_1) \leq \text{Im}(\mu_3),$$

up to permutation of the coordinates of $\mu$ or the coordinates of $(-u_1, u_2)$. This is somewhat unpleasant to check, so we refer the reader to [1, eqs. (2.19)-(2.21)]. In particular, $\tilde{F}_0(u, (it, it, -2it))$ decays exponentially as $t \to \pm \infty$. Furthermore, $\mu_1 + \mu_2 + \mu_3 = 0$ implies

$$\prod_{i=1}^{3} (1 + |\mu_i|) \gg (1 + \|\mu\|)^2;$$

so for, say, $\|\text{Im}(u)\| \ll (1 + \|\mu\|)^\epsilon$, $|\text{Re}(\mu)| < \frac{1}{2}$ and $\text{Re}(u)$ in some fixed compact set with $\text{Re}(u_1 + u_2) < 0$, we have

$$\tilde{F}_0(u, \mu) \ll (1 + \|\mu\|)^{2\text{Re}(u_1 + u_2) + \epsilon},$$

even when one of the coordinates of $\mu$ is small (say $\text{Im}(\mu_1) \ll (1 + \|\mu\|)^\epsilon$).

Next we compute the residues of $\tilde{F}_0(s, \mu)$: When the coordinates of $\mu$ are distinct modulo $\mathbb{Z}$, we have

$$\text{res}_{u_1 = -\mu_1 - \ell} \tilde{F}_0(u, \mu) = \frac{4}{\pi^4} \cos^0(\mu) \frac{(-1)^\ell \Gamma(\mu_2 - \mu_1 - \ell) \Gamma(\mu_3 - \mu_1 - \ell)}{\Gamma(u_2 - \mu_1 - \ell)} \prod_{i=1}^{3} \Gamma(u_2 - \mu_i),$$

$$\text{res}_{u_1 = -\mu_1 - \ell_1} \text{res}_{u_2 = -\mu_3 - \ell_2} \tilde{F}_0(u, \mu) = \frac{4}{\pi^4} \cos^0(\mu) \frac{(-1)^{\ell_1 + \ell_2} \Gamma(\mu_3 - \mu_1 - \ell_1) \Gamma(\mu_3 - \mu_1 - \ell_2)}{\ell_1! \ell_2!} \frac{\Gamma(\mu_3 - \mu_1 - \ell_1) \Gamma(\mu_3 - \mu_1 - \ell_2)}{\Gamma(u_2 - \mu_1 - \ell_1) \Gamma(u_2 - \mu_1 - \ell_2)} \times \Gamma(\mu_2 - \mu_1 - \ell_1) \Gamma(\mu_2 - \mu_1 - \ell_2).$$

Notice that the exponential parts of the second-order residue (35) cancel.

When $\mu = (it, it, -2it)$ with $0 \neq t \in \mathbb{R}$, the residue at $u_1 = 2it - \ell$ follows the form (34), and the residue at the double pole becomes

$$\text{res}_{u_1 = -it - \ell} \tilde{F}_0(u, \mu) = \frac{4}{\pi^4} \cos^0(\mu) \left( \prod_{i=1}^{3} \Gamma(u_2 - \mu_i) \right) \frac{1}{(\ell!) \Gamma(u_2 - it - \ell)} \Gamma(-3it - \ell) \times (2H_{\ell} - 2\gamma + \psi(-3it - \ell) - \psi(u_2 - it - \ell)), $$

where $H_n$ denotes the $n$-th harmonic number.
where \( H_\ell = \sum_{j=1}^{\ell} \frac{1}{j} \) is the \( \ell \)-th harmonic number, \( \gamma \) is the Euler-Mascheroni constant and 
\[ \psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} \] is the digamma function. Since computing residues at double poles requires differentiation, we do not include the second-order residues here.

Lastly, \( \tilde{F}_0(u, 0) \) has a triple pole at \( u_1 = -\ell, \ell \in \mathbb{N}_0 \); the residue is easily computed with a computer algebra package and is somewhat complicated, so we do not include it here. As pointed out in Section 2.2, the case \( \mu = 0 \) is not relevant to the spectral expansion.

9. The proof of Theorem 3

We apply Theorem 5 to the function
\[
 f_X(y) = f \left( \frac{X_1 y_1^2}{4m_1 n_2}, \frac{X_2 y_2^2}{4m_2 n_1} \right).
\]
Using the Mellin-Barnes integral (11) (i.e. by Plancherel’s theorem for the Mellin transform), we write \( F_0(f_X; \mu) \) in the form
\[
 F_0(f_X; \mu) = \int_{\text{Re}(s)=(\theta+i,\theta+i)} \left( \frac{X_1}{4\pi^2 m_1 n_2} \right)^{s_1} \left( \frac{X_2}{4\pi^2 m_2 n_1} \right)^{s_2} \hat{f}(-s) \tilde{F}_0(s, \mu) \frac{ds_1 ds_2}{(2\pi i)^2},
\]
where \( \hat{f}(s) \) is the usual Mellin transform of \( f \). Now \( \hat{f}(s) \) has super-polynomial decay in \( \text{Im}(s) \), and, as described in Section 8, \( \tilde{F}_0(s, \mu) \) has exponential decay in \( \text{Im}(\mu) \) unless some \( |\text{Im}(\mu)| \ll |\text{Im}(s)| \). Therefore,
\[
 F_0(\mu) \ll_{m,n,f} (X_1 X_2 (1 + \|\mu\|))^{-100}
\]
unless some \( |\text{Im}(\mu_i)| \ll (X_1 X_2)^c \).

From the Weyl law [10, Theorem 1] and its spherical counterpart [1, Theorem 1], the tempered cusp forms and continuous spectrum of Theorem 5 with \( \|\mu\| \ll (X_1 X_2)^c \) contribute at most \( (X_1 X_2)^c \) to the bound in Theorem 3, which is observable from the trivial bound on (37) with the contours at \( \text{Re}(s) = (\epsilon, \epsilon) \). If there exist any non-tempered (i.e. complementary series) cusp forms with spectral parameters, say, \( \mu \) a permutation of \((-x + it, -2it, x + it)\), \( 0 < x < \theta \) and \( t \ll (X_1 X_2)^c \) their contribution is at most \( (X_1 X_2)^{\theta+c} \).

For the troublesome forms which are near self-dual forms, say with one spectral parameter \( |\text{Im}(\mu_i)| \ll (X_1 X_2)^c \) and \( \|\mu\| \gg (X_1 X_2)^{2c} \) (which rules out the non-tempered forms), we suppose WLOG \( i = 1 \) and \( X_1 > X_2 \) (so, in particular \( X_1 > 1 \)) and shift \( \text{Re}(s_1) \) highly negative (saving arbitrarily many powers of \( X_1 (1 + \|\mu\|^2) \)), giving
\[
 F_0(f_X; \mu) \approx \sum_{w \in \mathbb{W}_3} \left( \frac{X_1}{4\pi^2 m_1 n_2} \right)^{-\mu_1^w} \int_{\text{Re}(s_2)=\epsilon} \left( \frac{X_2}{4\pi^2 m_2 n_1} \right)^{s_2} \hat{f}(\mu_1^w, s_2) \res_{s_1=-\mu_1^w} \tilde{F}_0(s, \mu) \frac{ds_1 ds_2}{(2\pi i)^2}
\]
+ lower-order terms.

All of the terms have super-polynomial decay in \( \mu_2 = -\mu_3 \) (coming from \( \hat{f} \)) except \( w = I \), and on that term, we shift \( \text{Re}(s_2) \mapsto -1 - \epsilon \). The residues again have super-polynomial decay in \( \mu_2 \), and the integral on the shifted contour is \( \ll X_2^{-1-\epsilon} \|\mu\|^{-4-\epsilon} \). This is sufficient for this part of the spectral expansion to converge, as the Weyl laws [1, Theorem 1] and [10, Theorem 1] imply there are at most \( T^{4+\epsilon}(X_1 X_2)^c \) forms with \( \mu_1 \ll (X_1 X_2)^c \) and \( \mu_2 \ll -\mu_3 = T \).
10. The proof of Proposition 8

Proposition 6 applies for \( f = f_s \) with \( \sigma_1 = \theta + \epsilon, \sigma_2 = \epsilon, N = 3 \) on the region \( 2\text{Re}(\tilde{s}_1), 2\text{Re}(\tilde{s}_2) > 1 \), which proves part 1. The remaining parts concern

\[
F_0(s, \mu) = \int_{\text{Re}(u_1, u_2) = (\epsilon, \epsilon)} \Gamma(u_1) \Gamma(u_2) \tilde{F}_0(\tilde{s} - u, \mu) \frac{du_1 du_2}{(2\pi i)^2},
\]

after applying (11) as in the previous section.

When the components of \( \mu \) are distinct modulo \( \mathbb{Z} \), we have

\[
F_0(s, \mu) = \sum_{w \in W} \sum_{j_1, j_2 \geq 0} \Gamma(j_1 + \tilde{s}_1 + \mu w) \Gamma(j_2 + \tilde{s}_2 - \mu w) \text{res}_{u_1 = -\mu w - j_1} \text{res}_{u_2 = \mu w - j_2} \tilde{F}_0(u, \mu),
\]

by shifting the \( u \) contours to \( \infty \), as we may. From (35), this series representation converges rapidly away from the poles and clearly defines a meromorphic function of all \( \tilde{s} \) and \( \mu \), proving part 2, and part 4 is clear. Note that the product of \( \cos^0(\mu) \) and the gamma factors of the summand has exponential decay for \( \text{Im}(\mu) \) large compared to \( \text{Im}(\tilde{s}) \) and away from the poles, which is the content of part 3, so this is proved for \( \mu \) whose components are distinct modulo \( \mathbb{Z} \). There is a minor detail that when, say, \( \mu_1 \equiv \mu_2 \) is large compared to \( \mu_1 - \mu_2 \), the polynomial part of the summand is increasing in \( j_1 \) like \( |\mu_1|^{j_1} \), but to defeat the exponential decay, we would need \( j \sim \frac{|\mu_1|}{\log|\mu_1|} \), at which point the \( j_1! \) in the denominator (as well as the factor \( \Gamma(\mu_2 - \mu_1 - j_1) \)) overwhelms any possible growth.

When \( \mu = (it, it, -2it) \) with \( t \neq 0 \), we have

\[
F_0(s, \mu) = \sum_{i=1}^{3} F_{0,i}(s, t),
\]

\[
F_{0,1}(s, t) = \sum_{j_1, j_2 \geq 0} \frac{(-1)^{j_1 + j_2}}{(j_1! j_2!)(3it - j_1)(3it - j_2)} (\tilde{s}_1 + it + j_1) (\tilde{s}_2 - it + j_2) \left( 2H_{j_2} - 2H_{j_1 + j_2} + 2\psi(j_1 + 1) + \psi(3it - j_1) + \psi(3it - j_2) - \psi(\tilde{s}_1 - it + j_1) - \psi(\tilde{s}_2 - it + j_2) \right)
\]

\[
F_{0,2}(s, t) = \sum_{j_1, j_2 \geq 0} \frac{(-1)^{j_1}}{(j_1!)(3it - j_1)^2} (\tilde{s}_1 + it + j_1) (\tilde{s}_2 + 2it + j_2) \left( \frac{\Gamma(3it - j_1)}{\Gamma(3it - j_1 - j_2)} + \frac{\Gamma(3it - j_1)}{\Gamma(3it - j_1 - j_2)} \right)
\]

\[
F_{0,3}(s, t) = F_{0,2}((s_2, s_1), -t),
\]

and it is easy to see that \( F_0(s, \mu) \) decays exponentially as \( t \to \pm \infty \), which completes the proof of part 3.

The equations (40) and (41) were necessary to show the rapid decay in \( \mu \) for part 3 (for all \( s \in \mathbb{C}^2 \) away from the poles), but to see the poles with their locations, orders and residues,
it is perhaps easier to write
\[
F_0(s, \mu) = \sum_{\substack{j_1 = 0 \atop j_2 = 0}}^{T_1} \sum_{\substack{j_2 = 0 \atop j_1 = 0}}^{T_2} \frac{(-1)^{j_1+j_2}}{j_1! j_2!} \tilde{F}_0(\tilde{s} + j, \mu)
+ \int_{\text{Re}(u_2) = -T_2 - \frac{1}{2}} \Gamma(u_2) \tilde{F}_0(\tilde{s} + (j_1, -u_2), \mu) \frac{du_2}{2\pi i},
\]
assuming \(-\text{Re}(\tilde{s}_i) < T_i \in \mathbb{Z}\). Now the poles are easy to compute from the poles of \(\tilde{F}_0(\tilde{s}, \mu)\); for instance, if \(\ell \in \mathbb{N}_0\), then
\[
\text{res}_{\tilde{s}_1 = -\mu_1 - \ell} F_0(s, \mu) = \sum_{\substack{j_1 = 0 \atop j_2 = 0}}^{\ell} \sum_{\substack{j_2 = 0 \atop j_1 = 0}}^{T_2} \frac{(-1)^{j_1+j_2}}{j_1! j_2!} \text{res}_{\tilde{s}_1 = -\mu_1 - \ell} \tilde{F}_0(\tilde{s} + j, \mu)
+ \int_{\text{Re}(u_2) = -T_2 - \frac{1}{2}} \Gamma(u_2) \text{res}_{\tilde{s}_1 = -\mu_1 - \ell} \tilde{F}_0(\tilde{s} + (j_1, -u_2), \mu) \frac{du_2}{2\pi i},
\]
to which we apply (34) or (36).

REFERENCES

[1] Valentin Blomer, Applications of the Kuznetsov formula on \(GL(3)\), Invent. Math. 194 (2013), no. 3, 673–729.
[2] Valentin Blomer and Jack Buttcane, Global decomposition of \(GL(3)\) Kloosterman sums and the spectral large sieve, Crelles Journal (2018). doi:10.1515/crelle-2017-0034.
[3] Valentin Blomer, Jack Buttcane, and Nicole Raulf, A Sato-Tate law for \(GL(3)\), Comment. Math. Helv. 89 (2014), no. 4, 895–919.
[4] Daniel Bump, Solomon Friedberg, and Dorian Goldfeld, Poincaré series and Kloosterman sums for \(SL(3, \mathbb{Z})\), Acta Arith. 50 (1988), no. 1, 31–89.
[5] Jack Buttcane, Sums of \(SL(3, \mathbb{Z})\) Kloosterman Sums, ProQuest LLC, Ann Arbor, MI, 2012. Thesis (Ph.D.)–University of California, Los Angeles.
[6] Valentin Blomer, On sums of \(SL(3, \mathbb{Z})\) Kloosterman sums, Ramanujan J. 32 (2013), no. 3, 371–419.
[7] Valentin Blomer, The spectral Kuznetsov formula on \(SL(3)\), Trans. Amer. Math. Soc. 368 (2016), no. 9, 6683–6714.
[8] Valentin Blomer, Higher weight on \(GL(3)\), I: The Eisenstein series, Forum Math. (2018). doi:10.1515/forum-2017-0060.
[9] Valentin Blomer, Higher weight on \(GL(3)\), II: The cusp forms (arXiv:1701.04380).
[10] Valentin Blomer, Kuznetsov, Petersson and Weyl on \(GL(3)\), I: The principal series forms (arXiv:1703.09837).
[11] Valentin Blomer, Kuznetsov, Petersson and Weyl on \(GL(3)\), II: The generalized principal series forms (arXiv:1706.08816).
[12] Sultan Catto, Jonathan Huntley, Jay Jorgenson, and David Tepper, On an analogue of Selberg’s eigenvalue conjecture for \(SL_3(\mathbb{Z})\), Proc. Amer. Math. Soc. 126 (1998), no. 12, 3455–3459.
[13] James W. Cogdell and Ilya I. Piatetski-Shapiro, The arithmetic and spectral analysis of Poincaré series, Perspectives in Mathematics, vol. 13, Academic Press, Inc., Boston, MA, 1990.
[14] Romuald Dabrowski and Benji Fisher, *A stationary phase formula for exponential sums over $\mathbb{Z}/p^m\mathbb{Z}$ and applications to $GL(3)$-Kloosterman sums*, Acta Arith. **80** (1997), no. 1, 1–48.

[15] Dorian Goldfeld, *Automorphic forms and $L$-functions for the group $GL(n,\mathbb{R})$*, Cambridge Studies in Advanced Mathematics, vol. 99, Cambridge University Press, Cambridge, 2006. With an appendix by Kevin A. Broughan.

[16] Dorian Goldfeld and Alex Kontorovich, *On the determination of the Plancherel measure for Lebedev-Whittaker transforms on $GL(n)$*, Acta Arith. **155** (2012), no. 1, 15–26.

[17] Henryk Iwaniec, *Spectral methods of automorphic forms*, Second, Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002.

[18] Henry H. Kim, *Functoriality for the exterior square of $GL_4$ and the symmetric fourth of $GL_2$*, J. Amer. Math. Soc. **16** (2003), no. 1, 139–183 (electronic). With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak.

[19] Eren Mehmet Kıral, *Opposite-sign Kloosterman sum zeta function*, Mathematika **62** (2016), no. 2, 406–429.

[20] Nikolai Vasil’evich Kuznetsov, *The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture*. *Sums of Kloosterman sums*, Mat. Sb. (N.S.) **111**(153) (1980), no. 3, 334–383, 479.

[21] Yuri Vladimirovich Linnik, *Additive problems and eigenvalues of the modular operators*, Proc. Internat. Congr. Mathematicians (Stockholm, 1962), 1963, pp. 270–284.

[22] Wenzhi Luo, Zeev Rudnick, and Peter Sarnak, *On Selberg’s eigenvalue conjecture*, Geom. Funct. Anal. **5** (1995), no. 2, 387–401.

[23] Stephen David Miller, *Cusp forms on $SL(3,\mathbb{Z})\backslash SL(3,\mathbb{R})/SO(3,\mathbb{R})$*, ProQuest LLC, Ann Arbor, MI, 1997. Thesis (Ph.D.)–Princeton University.

[24] , *The highest lowest zero and other applications of positivity*, Duke Math. J. **112** (2002), no. 1, 83–116.

[25] Yuichi Motohashi, *On the Kloosterman-sum zeta-function*, Proc. Japan Acad. Ser. A Math. Sci. **71** (1995), no. 4, 69–71.

[26] Walter Roelcke, *über die Wellengleichung bei Grenzkreisgruppen erster Art*, S.-B. Heidelberger Akad. Wiss. Math.-Nat. Kl. **1953/1955** (1953/1955), 159–267 (1956).

[27] Atle Selberg, *On the estimation of Fourier coefficients of modular forms*, Proc. Sympos. Pure Math., Vol. VIII, 1965, pp. 1–15.

[28] Glenn Stevens, *Poincaré series on $GL(r)$ and Kloosterman sums*, Math. Ann. **277** (1987), no. 1, 25–51.

[29] Nolan Russell Wallach, *Real reductive groups. II*, Pure and Applied Mathematics, vol. 132, Academic Press, Inc., Boston, MA, 1992.

[30] Yangbo Ye, *A Kuznetsov formula for Kloosterman sums on $GL_n$*, Ramanujan J. **4** (2000), no. 4, 385–395.

Mathematics Department, 244 Mathematics Building, Buffalo, NY 14260, USA

E-mail address: buttcane@buffalo.edu