THE WEYL BOUND FOR TRIPLE PRODUCT L-FUNCTIONS

VALENTIN BLOMER, SUBHAJIT JANA, and PAUL D. NELSON

Abstract

Let π_1, π_2, π_3 be three cuspidal automorphic representations for the group $\mathrm{SL}(2,\mathbb{Z})$*, where π_1 and π_2 are fixed and π_3 has large analytic conductor. We prove a subconvex bound for $L(1/2,\pi_1\otimes\pi_2\otimes\pi_3)$ of Weyl-type quality. Allowing π_3 to be an Eisenstein series, we also obtain a Weyl-type subconvex bound for $L(1/2+it,\pi_1\otimes\pi_2)$.*

Contents

1. Introduction

1.1. Weyl-type subconvexity

Subconvexity estimates belong to the core topics in the theory of L-functions and are one of the most challenging testing grounds for the strength of existing technology. If C denotes the analytic conductor of the relevant L-function (restricted to the parameters of interest), then the Phragmén–Lindelöf principle gives the bound $C^{1/4+\varepsilon}$ for the L-function on the central line $\Re s = 1/2$. In the most favorable cases, one can obtain an upper bound $C^{1/6+\varepsilon}$, which we refer to as a *Weyl-type subconvex bound*. For instance, a classical result states that the Riemann zeta function satisfies the bound

$$\zeta(1/2+it) \ll_\varepsilon \left(1+|t|\right)^{1/6+\varepsilon}$$

DUKE MATHEMATICAL JOURNAL
Vol. 172, No. 6, © 2023 DOI 10.1215/00127094-2022-0058
Received 14 February 2021. Revision received 5 April 2022.
First published online 4 April 2023.
2020 *Mathematics Subject Classification*. Primary 11M41; Secondary 11F70, 11F72.

on the critical line. Based on work of Weyl [49], it was proved first by Hardy and Littlewood (cf. [34]), and first written down by Landau [32] in a slightly refined form and generalized to all Dirichlet L-functions. Results of similar strength exist in the Archimedean aspect for automorphic L-functions of degree 2, starting with the work of Good [17] and culminating in the hybrid bound of Jutila and Motohashi [29]. We also have a Weyl-type bound of degree 4 in some limited cases pertaining to Rankin–Selberg L-functions, such as (see [30], [33])

$$L(1/2, f \otimes g) \ll_{g,\varepsilon} C(f)^{1/3+\varepsilon}$$

for two cusp forms f, g for $\mathrm{SL}_2(\mathbb{Z})$, where $C(f)$ denotes the conductor of f as defined in Section 2.2.5 below. Although slightly better bounds are available for GL(1) (see [9], [37]), the Weyl exponent marks a natural barrier that has never been improved, and rarely been reached, beyond GL(1). We note that for some applications (see [16], [35]), the essential input is a Weyl-type subconvex bound (or something approaching it), rather than merely any nontrivial subconvex bound.[1] We note also that Petrow and Young recently established Weyl-type subconvex bounds for GL(1) in the level aspect, improving spectacularly upon the decades-old Burgess-type bounds (see [40], [42], [43]).

A celebrated result of Bernstein and Reznikov established for the first time subconvex bounds for certain L-functions of degree 8. For two fixed spherical cuspidal automorphic representations π_1, π_2 (i.e., generated by Maass forms for $\mathrm{SL}_2(\mathbb{Z})$) and another spherical cuspidal automorphic representation π_3 of large analytic conductor $C(\pi_3)$, they proved (see [3], [4]) that

$$\sum_{T \le C(\pi_3)^{1/2} \le T+T^{1/3}} L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3) \ll_{\pi_1,\pi_2,\varepsilon} T^{5/3+\varepsilon}, \tag{1.1}$$

which implies (by nonnegativity of the central value) in particular $L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3) \ll_{\pi_1,\pi_2,\varepsilon} C(\pi_3)^{5/6+\varepsilon}$. The proof employs a beautiful combination of representation theory, invariant norms, and asymptotic analysis of oscillatory Airy-type integrals. In fact, their result is really an estimate for triple product periods, as L-functions enter only through the period formula of Watson and Ichino [22].

We observe, however, that (1.1) is not optimal. The Lindelöf hypothesis suggests the Weyl-type bound with an exponent $4/3$ instead of $5/3$. That such a result might be within reach was indicated by Suvitie [46]. For a fixed holomorphic cusp form F of weight k and a Maass form h, she showed that

$$\sum_{T \le C(h)^{1/2} \le T+T^{1/3}} e^{\pi r_h} \left|\left\langle y^k |F|^2, h\right\rangle\right|^2 \ll_{F,\varepsilon} T^{2k-\frac{2}{3}+\varepsilon}, \tag{1.2}$$

[1]Indeed, to show that the number of zeros on $i[1,\infty)$ of a holomorphic Hecke eigenform f of weight k tends to infinity as $k \to \infty$, the proof in [16] needs $L(1/2+it, f) \ll k^{0.335}$ with polynomial dependence in t.

which via the Watson–Ichino formula translates into

$$\sum_{T\le C(h)^{1/2}\le T+T^{1/3}} L(1/2, F\otimes F\otimes h) \ll_{F,\varepsilon} T^{4/3+\varepsilon}.$$

This L-function is not primitive, as it factorizes into a degree 6 and a degree 2 L-function, but the same argument would work for FG instead of $|F|^2$ in (1.2). More seriously, however, the proof starts by replacing F with a holomorphic Poincaré series and unfolding the inner product, a route that is not available in general. In fact, an attempt to generalize this to Maass forms remained incomplete [45] and seems not to work. In particular, the work of Bernstein–Reznikov remained unimproved.

In this article, we establish the Weyl-type bound for triple product L-functions in a uniform fashion for all combinations of local types at infinity, that is, any of the three factors can be holomorphic or Maass. As mentioned before, the Weyl-type bound marks the natural limit of all present day approaches to subconvexity. The key novelty in our work is the method. We combine in a substantial way representation theory, local harmonic analysis, and analytic number theory to establish a robust method for the subconvexity problem for triple product L-functions.

THEOREM 1

Let π_1, π_2 be two fixed cuspidal automorphic representations for the group $\mathrm{SL}_2(\mathbb{Z})$. *Let π_3 run over cuspidal automorphic representations for* $\mathrm{SL}_2(\mathbb{Z})$ *with conductor satisfying $T \le C(\pi_3)^{1/2} \le T + T^{1/3}$. Then*

$$\sum_{T\le C(\pi_3)^{1/2}\le T+T^{1/3}} L(1/2, \pi_1\otimes\pi_2\otimes\pi_3) \ll_{\pi_1,\pi_2,\varepsilon} T^{4/3+\varepsilon},$$

in particular,

$$L(1/2, \pi_1\otimes\pi_2\otimes\pi_3) \ll_{\pi_1,\pi_2,\varepsilon} C(\pi_3)^{2/3+\varepsilon}$$

for every $\varepsilon > 0$.

An inspection of the proof shows that the dependence on the analytic conductors of π_1, π_2 is polynomial. Under Watson's formula in [48], the latter estimate translates to bounds for triple product integrals of Maass forms φ_j of eigenvalue $1/4 + t_j^2$ ($j = 1, 2, 3$):

$$\Big|\int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} \varphi_1\varphi_2\varphi_3\Big|^2 \ll_{\varphi_1,\varphi_2,\varepsilon} |t_3|^{-2/3+\varepsilon} \exp\big(-\pi|t_3|\big),$$

giving a further improvement beyond that in [4] on the general exponential decay bounds of [44].

We can allow π_3 to be an Eisenstein series, and our proof yields as a by-product a Weyl-type bound for Rankin–Selberg L-functions.

THEOREM 2

Let π_1, π_2 be two fixed cuspidal automorphic representations for the group $\mathrm{SL}_2(\mathbb{Z})$. *Then*

$$\int_T^{T+T^{1/3}} \left|L(1/2+it,\pi_1\otimes\pi_2)\right|^2 dt \ll_{\pi_1,\pi_2,\varepsilon} T^{4/3+\varepsilon},$$

in particular,

$$L(1/2+it,\pi_1\otimes\pi_2) \ll_{\pi_1,\pi_2,\varepsilon} \left(1+|t|\right)^{2/3+\varepsilon}$$

for every $\varepsilon>0$ and $t\in\mathbb{R}$.

For the rest of the paper, all implied constants may depend on ε, and we suppress it in subsequent formulas. The weaker bound $L(1/2+it,\pi_1\otimes\pi_2)\ll_{\pi_1,\pi_2}(1+|t|)^{5/6+\varepsilon}$ is implicit in [4, Remarks 7.2.2.2] and was the best known result until now. By a method purely based on analytic number theory, the bound $L(1/2+it,\pi_1\otimes\pi_2)\ll_{\pi_1,\pi_2}(1+|t|)^{15/16+\varepsilon}$ was recently shown in [1]. For bounds of triple product L-functions in the level aspect, see [20] and [47].

1.2. Remarks

(1) Our results feature "pure" exponents of Weyl-type quality that are independent of bounds towards the Ramanujan conjecture or the Selberg eigenvalue conjecture. The proof uses at one place that the smallest nonzero Laplace eigenvalue is larger than $3/16$.

(2) In principle, the proof produces an asymptotic formula. If ψ is a sufficiently regular test function with "essential support" in $[T,T+H]$, for example, $\psi(t)=\exp(-(t-T)^2H^{-2})$, and $t_{\pi_3}\asymp\sqrt{C(\pi_3)}$ (cf. (2.20)) denotes the spectral parameter of π_3, then with the same notation and under the same assumptions as in Theorem 1, one can relate

$$\sum_{\pi_3}\psi(t_{\pi_3})\frac{L(1/2,\pi_1\otimes\pi_2\otimes\pi_3)}{L(1,\mathrm{Ad}^2\pi_3)} + \int_{t\in\mathbb{R}}\psi(t)\frac{|L(1/2+it,\pi_1\otimes\pi_2)|}{|\zeta(1+2it)|^2}\frac{dt}{2\pi} \tag{1.3}$$

to

$$cL(1,\mathrm{Ad}^2\pi_1)L(1,\mathrm{Ad}^2\pi_2)TH+O(T^{3/2+\varepsilon}H^{-1/2})$$

for a suitable constant c (depending on ψ).

(3) The proof of Theorem 1 has the shape of a *reciprocity formula* as for instance in [7]. A spectral sum in a window $[T, T+H]$ as in (1.3) is ultimately transformed into a spectral sum of similar shape with spectral parameter up to $\ll T/H$ (see (5.30), (5.28), (5.13)). This is analogous to the discussion after (1.11) in [7], and a new instance of a reciprocity phenomenon. The optimal choice is $H = T^{1/3}$, in which case both spectral sums have length $T^{4/3}$. This yields the Weyl bound, and we see that the Weyl bound is indeed the natural limit from the point of view of spectral analysis. An abstract version of the underlying reciprocity formula is displayed in (1.7) below which features central L-values as well as their "canonical square roots."

(4) Implicit in the proof of Theorem 1 is an alternative description of the central triple product L-value $L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3)$ in terms of a certain shifted convolution problem very roughly of the shape

$$L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3) \approx \Big| \frac{1}{t_{\pi_3}^{1/2}} \sum_{\nu \asymp 1} \sum_{m \ll t_{\pi_3}^2} \frac{\lambda_{\pi_3}(m)\lambda_{\pi_2}(m+\nu)\lambda_{\pi_1}(\nu)}{m^{1/4}} \times \exp\Big(\pm 2it_{\pi_3}\sqrt{\frac{\nu}{m}}\Big)\Big|^2. \tag{1.4}$$

The "$\approx$" sign has to be interpreted in a broad sense (see Section 6 for details). In the generic range $m \asymp t_{\pi_3}^2$, the oscillatory factor is flat (see Section 2.4 for definition of flatness).

1.3. Comparison with Bernstein–Reznikov and Michel–Venkatesh

It is instructive to compare our approach for studying the moment (1.1) with those of Bernstein and Reznikov [3], [4] and Michel and Venkatesh [36].

To begin, we briefly sketch the approach to the subconvexity problem introduced by Bernstein and Reznikov. We borrow some presentation features from Michel and Venkatesh (see especially [36, Sections 1.1.1 and 1.1.3]). The starting point of this approach is the triple product formula in [22]: for unit vectors $v_j \in \pi_j$, we have

$$\frac{L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3)}{L(1, \mathrm{Ad}^2\pi_1)L(1, \mathrm{Ad}^2\pi_2)L(1, \mathrm{Ad}^2\pi_3)} \mathscr{L}_\infty(v_1, v_2, v_3) = \Big| \int_{g \in \mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})} v_1 v_2 v_3(g)\, dg \Big|^2 \tag{1.5}$$

for a suitable local factor $\mathscr{L}_\infty(v_1, v_2, v_3)$ (a constant multiple of a matrix coefficient integral).

For unit vectors $v_1 \in \pi_1$ and $v_2 \in \pi_2$, we consider the inner product identity

$$\langle v_1 v_2, v_1 v_2 \rangle = \int_{g \in \mathrm{SL}_2(\mathbb{Z}) \backslash \mathrm{SL}_2(\mathbb{R})} |v_1 v_2|^2(g)\, dg = \big\langle |v_1|^2, |v_2|^2 \big\rangle. \tag{1.6}$$

By expanding each of these inner products over the spectrum of $L^2(\mathrm{SL}_2(\mathbb{Z}) \backslash \mathrm{SL}_2(\mathbb{R}))$ and applying (1.5), we obtain a spectral identity of families of L-functions, roughly of the shape

$$\sum_{\pi_3} h(\pi_3) L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3) \approx 1 + \sum_{\sigma} \tilde{h}(\sigma) \sqrt{L(1/2, \pi_1 \otimes \pi_1 \otimes \sigma) L(1/2, \pi_2 \otimes \pi_2 \otimes \sigma)}. \tag{1.7}$$

Here π_3 and σ run over cuspidal automorphic representations of $\mathrm{SL}_2(\mathbb{Z})$, the square roots of L-values are "canonical square roots" (in the sense of [36, Section 1.1.3]), and the meaning of "$\approx$" is that

- we have suppressed adjoint L-factors and other proportionality constants, and
- we have elided the contribution of the continuous spectrum and all degenerate terms except for the "expected main term" 1, which arises from the inner product $\langle |v_1|^2, 1 \rangle \langle 1, |v_2|^2 \rangle = 1$ (up to volume factors).

The weight functions h and $\tilde{h}$ depend upon the choice of vectors v_1 and v_2.

The weights $h(\pi_3)$ and the L-values $L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3)$ are known to be nonnegative, so if we can bound the right-hand side of (1.7) by $O(1)$ (the natural limit, in view of the expected main term), then we deduce by dropping all but one term the estimate

$$L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3) \ll 1/h(\pi_3). \tag{1.8}$$

Given some π_3 with $\mathrm{cond}(\pi_3)^{1/2} \asymp T$, we now face the optimization problem of choosing unit vectors v_1 and v_2 for which $h(\pi_3)$ is as large as possible, so that the bound (1.8) is as strong as possible. Bernstein and Reznikov [4, (2.6.3), Proposition 9.1] showed (for π_j spherical) that one may choose v_1 and v_2 so that $h(\pi_3)$ is roughly $T^{-5/3}$. This choice and a suitable bound for the global period eventually yields their estimate (1.1). Michel and Venkatesh [36, Sections 3.6 and 3.7] (for π_1 tempered and spherical) employed a simpler choice of vectors for which $h(\pi_3)$ is of size T^{-2}; for this choice, the estimate (1.8) only recovers the convexity bound, but Michel and Venkatesh managed to apply the amplification method to save a further small power of T (in a more general "all aspects" setting).

To approach the Weyl bound $L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3) \ll T^{4/3+\varepsilon}$ using (1.7) would seem to require producing v_1 and v_2 for which $h(\pi_3)$ is at least $T^{-4/3}$, but the analysis of Bernstein and Reznikov strongly suggests that their lower bound $T^{-5/3}$ is best possible. To obtain a stronger lower bound thus requires a more flexible class of weight functions $h(\pi_3)$. Such a class may be obtained from the generalization of (1.6)

to higher-rank tensors $\sum_j v_{1,j} \otimes v_{2,j} \in \pi_1 \otimes \pi_2$, namely,

$$\sum_{j,k} \langle v_{1,j} v_{2,j}, v_{1,k} v_{2,k} \rangle = \sum_{j,k} \langle v_{1,j} \overline{v_{1,k}}, v_{2,k} \overline{v_{2,j}} \rangle.$$

Such tensors yield more flexible forms of (1.7). One could hope to find a tensor $\sum_j v_{1,j} \otimes v_{2,j}$ for which the corresponding weight $h(\pi_3)$ localizes on π_3 satisfying $T \le \mathrm{cond}(\pi_3)^{1/2} \le T + T^{1/3}$ and for which the right-hand side of the corresponding spectral identity as in (1.7) may be effectively bounded. To implement this idea in practice would require a careful study of the spaces of test functions $\{h\}$ and $\{\tilde{h}\}$ as well as the transform relating them. Unfortunately, such a study has not yet been carried out, and does not seem straightforward in the generality of Theorem 1 (which, we should emphasize, imposes no local conditions on the representations π_j).

The method of this paper consists of two stages. We first use a somewhat crude choice of v_1 and v_2 (like in the work of Michel and Venkatesh) and an unfolding technique (see Section 1.5, in particular (1.11)) to express the L-values of interest as bilinear forms in the Hecke eigenvalues of the varying form π_3. We then average over the spectral window $T \le C(\pi_3)^{1/2} \le T + T^{1/3}$ of interest by means of analytic number theory, and in particular the Kuznetsov formula.

Our approach has in common with the works of Bernstein and Reznikov and Michel and Venkatesh that we make use of the well-developed theory of integral representations of L-functions to produce and analyze our test vectors. In this respect, our basic framework owes much to those works. The essential difference is that we analyze sums over much narrower spectral windows, and the more technical difference is that we implement this analysis using the Kuznetsov formula rather than the spectral theory of triple product periods.

The main advantage of our approach is that we can make full use of available technology related to the Kuznetsov formula (abundance of test functions, explicit integral transforms, Bessel function asymptotics, large sieve estimates, and so on), whose avatars are not available at the level of the triple product periods. To the best of our knowledge, this paper is the first to employ such a combination of the theory of integral representations and analytic number theory. We hope that this methodology will be useful more broadly.

1.4. Analytic number theory

Having discussed the representation-theoretic ideas in the previous subsection, we now give a brief sketch of how analytic number theory can handle expressions like (1.4) that can be extracted from the triple product formula. This is a precursor to the analysis in Section 5. The trivial bound in the m-sum recovers convexity, and if we had square-root cancellation in the m-sum we would obtain Lindelöf. We now

consider the sum

$$\sum_{2\pi T\le r_h\le 2\pi(T+H)} L(1/2, f\otimes g\otimes h). \tag{1.9}$$

An application of the Kuznetsov and Voronoi formulas gives a dual shifted convolution problem that can be treated by a δ-symbol method. With a final application of the spectral large sieve, the sum (1.9) can be shown to be $\ll (TH)^{1+\varepsilon}$ for $H = T^{1/3}$, which establishes the Weyl bound. For convenience, we describe a toy version of this argument, restricting each parameter to the generic range. This is somewhat misleading because smaller ranges of m in (1.4) are punished by an additional oscillation which complicates matters considerably, but it nevertheless gives a flavor for what is happening. Restricting (1.4) to $\nu = 1$ (for simplicity) and applying the Kuznetsov formula, we obtain an expression roughly of the shape (cf. (5.14))

$$\frac{H^{3/2}}{T^{5/2}} \sum_{m_1,m_2\asymp T^2} \sum_{c\asymp T/H} \lambda_f(m_1)\lambda_f(m_2) S(m_1-1, m_2-1, c) \times e\Big(\frac{2\sqrt{m_1m_2}}{c} - \frac{T^2c}{\sqrt{m_1m_2}}\Big),$$

provided that $H \ge T^{1/3}$. (For smaller H, more terms in the exponential would be necessary.) The complicated exponential is reminiscent of the uniform asymptotic expansion of the J-Bessel function at imaginary index. The leading term of the Kuznetsov kernel equals the Voronoi kernel, a feature that is now crucially exploited: applying the Voronoi summation formula to m_2, the dual variable will be close to m_1 and a large portion of the oscillation disappears. We obtain roughly (cf. (5.22))

$$\frac{H^{5/2}}{T^{5/2}} \sum_{c\asymp T/H} \sum_{h\asymp T^2/H^2} S(h-1,-1,c) \sum_{m\asymp T^2} \lambda_f(m)\lambda_f(m+h) e\Big(\frac{2Th^{1/2}}{m^{1/2}}\Big).$$

The inner sum is now a shifted convolution problem with a moderately oscillatory factor of size $T(h/m)^{1/2} \asymp T/H$. It can be spectrally decomposed by a delta-symbol method (cf. Section 5.7). Another application of the Voronoi and Kuznetsov formulas (cf. Sections 5.8 and 5.9) leads to a spectral sum of length T/H having $(T/H)^2$ terms by Weyl's law. Note that this is the length of the h-sum, so the large sieve (which is itself an application of the Kuznetsov formula) can show its full power on the h-sum, leading to the desired final bound. As an aside, we see that this analysis employs the Kuznetsov formula three times, in various directions.

We finally remark that a direct approach to (1.9), by an approximate functional equation followed by the Kuznetsov formula, appears to be hopeless: we would obtain

sums over $\lambda_f(n)\lambda_g(n)$ with $n \asymp T^4$ against oscillatory functions of (combined arithmetic and analytic) conductor of size T^2 (regardless of the choice of H), so that a GL(2) × GL(2) Voronoi summation formula would not reduce the length of summation. In other words, after applying the Kuznetsov formula, we run out of moves immediately.

1.5. Unfolding

Following [46], we now sketch a beautiful, but completely different approach to the formula (1.4), specific to the case of discrete series representations π_1, π_2. Suppose that π_1, π_2 are generated by holomorphic forms f, g and that π_3 is generated by a Maass form h of spectral parameter $t_h \geq 0$. By Watson's formula, we have

$$L(1/2, f \otimes g \otimes h) \approx e^{\pi t_h} t_h^{2-2k} \left| \int_{z \in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} \bar{f}(z) g(z) h(z) y^k \, d\mu(z) \right|^2.$$

We write g as a linear combination of Poincaré series, and without much loss of generality we assume that

$$g(z) = P_n(z) = \sum_{\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_\infty \backslash \mathrm{SL}_2(\mathbb{Z})} (cz + d)^{-k} e(n\gamma z)$$

is the nth Poincaré series. We insert the Fourier expansions and unfold. This gives a y-integral

$$\begin{aligned} &\int_0^\infty y^k \sqrt{y} \cosh\Big(\frac{\pi t}{2}\Big) K_{it}(2\pi m y) e^{-2\pi(m+n)y} \frac{dy}{y^2} \\ &\quad \approx \frac{e^{-\pi t/2} t^{2k-2}}{m^{k-1/2}} \exp\Big(\pm 2it\sqrt{\frac{n}{m}}\Big) \min\Big(\frac{m^{k/2-1/4}}{t^{k-1/2}}, 1\Big) \end{aligned}$$

for large t and fixed n. This integral was first analyzed by Good [17, Section 4] in terms of hypergeometric ${}_2F_1$ functions; the analysis is long and difficult. At least if the weight k is fixed but relatively large (for small k, one needs to work a little harder), the typical range is $m \ll t^2$, and we obtain that $L(1/2, f \otimes g \otimes h)$ is essentially the absolute-square of a linear combination of

$$\frac{1}{t_h^{1/2}} \sum_{m \ll t_h^2} \frac{\lambda_h(m)\lambda_f(m+n)}{m^{1/4}} \exp\Big(\pm 2it_h\sqrt{\frac{n}{m}}\Big) \tag{1.10}$$

for a fixed number of n's. It is very interesting to note that this resembles closely (1.4). The previous argument is due to [46].

The unfolding step is obviously not applicable in the setup of Bernstein and Reznikov. Following an idea of Zagier [50], the formal identity

$$\begin{aligned}\int_{z\in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} f(z)\,d\mu(z) &= \frac{\pi}{3}\int_{z\in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} f(z)\operatorname*{res}_{s=1} E(z,s)\,d\mu(z)\\ &= \frac{\pi}{3}\oint_s \int_{z\in \Gamma_\infty\backslash\mathbb{H}} f(z)y^s\,d\mu(z)\frac{ds}{2\pi i}\end{aligned} \tag{1.11}$$

for an $\mathrm{SL}_2(\mathbb{Z})$-invariant function f can be used to mimic unfolding in more general situations (see [19, Appendix A] for a related idea). Applied to the triple product of three classical Maass forms it yields a y-integral involving three K-Bessel functions

$$\int_0^\infty K_{it_1}(m_1y)K_{it_2}(m_2y)K_{it_3}(m_3y)y^{s-1/2}\,dy$$

for $m_1+m_2+m_3=0$. This can still be analyzed to some extent, but the resulting highly oscillatory shifted convolution problems become untreatable with the required precision. This is the reason why the attempt on the Maass case in [45] remained incomplete. Nevertheless, the unfolding step (1.11) is also present in our argument (cf. Section 6). It acts as a hinge between the triple product identity and sums over Fourier coefficients, leading eventually to the description (1.4) for the central L-value in terms of shifted convolution sums.

1.6. The analytic test vector problem

The art in using the triple product formula (1.5) to bound L-functions consists of choosing appropriate test vectors. The traditional test vector problem asks for explicit v_1, v_2, v_3 for which the local factor $\mathcal{L}_\infty(v_1,v_2,v_3)$ is nonzero. Spherical vectors are often test vectors in this sense, but are usually *not* the best choice due to the exponential decay of the local factor. For analytic applications, it is useful to work with *analytic* test vectors: vectors for which the local factor is not merely nonvanishing, but enjoys (informally speaking) a reasonable quantitative lower bound. Michel and Venkatesh [36, Section 3.6.1] gave a robust supply of test vectors under local assumptions relevant for Rankin–Selberg subconvexity. We will revisit and extend their approach to the triple product setting, removing all local assumptions in a uniform way.

For our analysis of test vectors, we adopt the language of *analytic newvectors* in [26], which is well suited for keeping track of the essential invariance properties. Analytic newvectors are approximate Archimedean analogues of the classical p-adic newvectors introduced by Casselman [11] (see also [25]). Let $K_0(p^N)$ denote the standard congruence subgroup consisting of matrices in $\mathrm{PGL}_2(\mathbb{Z}_p)$ whose lower left entry is divisible by p^N. Let ξ be a generic irreducible representation of $\mathrm{PGL}_2(\mathbb{Q}_p)$.

Denote by $c(\xi)$ the conductor exponent of ξ, so that $p^{c(\xi)}$ is the usual arithmetic conductor of ξ. The main result of local newvector theory in [11] is that there is a unique (up to scalar) nonzero vector $v \in \xi$ such that $\xi(g)v = v$ for all $g \in K_0(p^{c(\xi)})$. Such vectors v are called *newvectors*.

An Archimedean analogue of the family of congruence subgroups $K_0(p^N) \subseteq \mathrm{PGL}_2(\mathbb{Z}_p)$ is the family of subsets $K_0(X, \tau)$ that is defined by the image in $\mathrm{PGL}_2(\mathbb{R})$ of the set

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}) : \begin{array}{ll} |a-1| < \tau, & |b| < \tau, \\ |c| < \dfrac{\tau}{X}, & |d-1| < \tau \end{array} \right\}.$$

Here X is a large positive parameter, thought of as tending off to infinity, while $\tau \in (0, 1)$ is taken small but fixed. An Archimedean analogue of local newvector theory is given by [26, Theorem 1]: for each fixed $0 \le \vartheta < 1/2$ and arbitrary $\delta > 0$, there is a $\tau > 0$ so that for every generic irreducible unitary ϑ-tempered (see Section 2.3) representation π of $\mathrm{PGL}_2(\mathbb{R})$, there is a unit vector $v \in \pi$ such that

$$\left\| \pi(g)v - v \right\| < \delta \quad \text{for all } g \in K_0(C(\pi), \tau).$$

We refer to such vectors as *analytic newvectors* (suppressing, for terminological brevity, the dependence of this notion upon the parameters δ and τ). Such vectors v may be constructed explicitly as fixed bump functions in the Kirillov model (see [26, Theorem 7]).

Inspired by the construction of [36, Section 3.6.1], we approach the analytic test vector problem for the local triple product periods $L_\infty(v_1, v_2, v_3)$ by choosing v_1 and v_3 to be analytic newvectors. The choice of v_2 is simplest to describe when π_2 is a principal series representation. In that case, we describe v_2 in the induced model by a function on the lower triangular subgroup supported within $O(1/X)$ of the identity. More generally, we make use of the fact that π_2 may be embedded in a (not necessarily unitary) principal series representation.

1.7. Plan for the paper

Having chosen test vectors v_1, v_2, v_3 as indicated above, we need to solve three main problems.

- We need to compute (a lower bound for) the matrix coefficient integral $\mathcal{L}_\infty(v_1, v_2, v_3)$. This will be done in Section 3. The idea of the proof, as in [36, Section 3.7.2], is to write the matrix coefficient integral as the square of a Rankin–Selberg integral and then to estimate the latter by playing the support properties of v_1 in its Kirillov model and v_2 in its induced model against the invariance properties of v_3. One subtlety is that we have not assumed that

any of our representations belongs to the principal series. For this reason, the reduction to Rankin–Selberg integrals is achieved in general only after embedding π_2 into a principal series representation and using the standard intertwining operator to normalize its inner product.

- We use (a refined version of) the formula (1.11) to compute the right-hand side of (1.5). This leads to an integral of three Archimedean Whittaker functions that will be computed asymptotically in Section 4. The proof involves several applications of the local functional equation and stationary phase analysis, but no input concerning special functions beyond Stirling's formula. This yields the expression (1.4). In other words, choosing test vectors as above has the exact same effect as using Poincaré series in the holomorphic case and unfolding (cf. (1.10)). This is a rather remarkable feature.
- We need to bound the shifted convolution problem (1.4). This can be done by analytic number theory, roughly as indicated in Section 1.4. This is the content of Section 5. It is here that we implement the "hard analysis" required by our short spectral summation.

Theorems 1 and 2 are then an easy consequence and will be derived in Section 6.

2. Preliminaries

2.1. Basic notation

Throughout we work with the group $G := \mathrm{PGL}_2(\mathbb{R})$, and its subgroup N of unipotent upper triangular matrices which are equipped with the usual Haar measures. We use the notation

$$n(x) := \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \qquad n'(x) := \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}, \qquad a(y) := \begin{pmatrix} y & 0 \\ 0 & 1 \end{pmatrix},$$

$$k(\theta) := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad w := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We view $k(\theta)$ as a function of $\theta \in \mathbb{R}/\pi\mathbb{Z}$. For convenience, we may assume that θ is taken in the interval $[-\pi/2, \pi/2]$.

We write $d^\times y = dy/|y|$ for the Haar measure on $\mathbb{R}^\times$ and dx for the Lebesgue measure on $\mathbb{R}$. We fix a G-invariant measure dg on $N\backslash G$ given in Iwahori coordinates by

$$N\backslash G \ni g = a(y)n'(x), \quad dg = \frac{d^\times y}{|y|}\, dx,$$

and in Iwasawa coordinates by

$$N\backslash G \ni g = a(y)k(\theta), \quad dg = \frac{d^\times y}{|y|}\, d\theta.$$

We equip $\mathrm{GL}_2(\mathbb{R})$ with the Haar measure compatible with the chosen Haar measures on $\mathbb{R}^\times$ and G via the short exact sequence $1 \to \mathbb{R}^\times \to \mathrm{GL}_2(\mathbb{R}) \to G \to 1$.

Let $\mathfrak{X}(\mathbb{R}^\times)$ denote the character group of $\mathbb{R}^\times$. Each $\chi \in \mathfrak{X}(\mathbb{R}^\times)$ is uniquely of the form $\chi = |\cdot|^s \operatorname{sgn}^a$ for some $s \in \mathbb{C}$ and $a \in \{0,1\}$. We set $\Re(\chi) := \Re(s)$, $\Im(\chi) := \Im(s)$, and $C(\chi) := (1 + |\Im s|)/(2\pi)$ (cf. (2.14)). The group $\mathfrak{X}(\mathbb{R}^\times)$ is a complex manifold with respect to the coordinate charts $\chi \mapsto s$. For a function $f : \mathfrak{X}(\mathbb{R}^\times) \to \mathbb{C}$ of sufficient decay and $\sigma \in \mathbb{R}$, we define the contour integral

$$\int_{\Re(\chi)=\sigma} f(\chi)\, d\chi := \frac{1}{2} \sum_{a\in\{0,1\}} \int_{\Re(s)=\sigma} f\big(|\cdot|^s \operatorname{sgn}^a\big)\, \frac{ds}{2\pi i}.$$

For a smooth function $f : \mathbb{R}^\times \to \mathbb{C}$ of sufficient decay, we then have the Mellin inversion formula

$$f(y) = \int_{\Re(\chi)=\sigma} \chi(y) \Big(\int_{t\in\mathbb{R}^\times} f(t)\chi^{-1}(t)\, \frac{dt}{|t|} \Big)\, d\chi. \tag{2.1}$$

2.2. Local γ-factors, Stirling's formula, and the analytic conductor

Let ρ be a finite-dimensional representation of the Weil group $W_\mathbb{R}$. Let $\psi(x) := e^{2\pi i x}$ be the standard additive character of $\mathbb{R}$. The local γ-factor of ρ is defined as usual by

$$\gamma(s,\rho) := \gamma(s,\rho,\psi) := \epsilon(s,\rho,\psi) \frac{L(1-s,\tilde{\rho})}{L(s,\rho)}, \tag{2.2}$$

where ϵ and L denote the ϵ-factor and L-factor, respectively, a description of which can be found in [48, Section 3.1] for the cases relevant in this paper. We regard ψ as fixed once and for all, and for this reason we drop it from the notation.

The analytic conductor $C(\rho)$ has been defined in various slightly different ways (see, e.g., [24, Section 2], [23, Section 5], [36, Section 3.1.8]). For many applications, it is unimportant precisely which definition is used: what matters is just that $C(\rho)$ controls the local γ-factor in the sense that for small enough s, and under favorable conditions, one has at least the rough approximation

$$\gamma(s,\rho) \approx C(\rho)^{1/2-s}. \tag{2.3}$$

For the purposes of this article, it will be convenient to normalize the definition of $C(\rho)$ somewhat more precisely, so that a correspondingly more precise form of (2.3) holds. While we could work with ad hoc definitions, it is useful to present this in a slightly more general context. The purpose of the following computation is to give a uniform asymptotic formula for the local gamma factors in the cases relevant for our application. This is achieved in (2.20) and (2.21) below and used in Section 4.5.

2.2.1. Stirling's formula

With the principal branch of the logarithm, we have

$$\Gamma(z) = z^{-1/2}\Big(\frac{z}{e}\Big)^z \big(\mathcal{G}_N(z) + O_{N,\varepsilon}\big(|z|^{-N}\big)\big), \quad \big|\arg(z)\big| \le \pi - \varepsilon, |z| \ge \varepsilon$$

for some smooth function $\mathcal{G}_N$ satisfying

$$|z|^j \frac{d^j}{dz^j} \mathcal{G}_N(z) \ll_{j,N} 1$$

for all $N, j \in \mathbb{Z}_{\ge 0}$.

2.2.2. Characters of $\mathbb{R}^\times$

Set $\Gamma_{\mathbb{R}}(s) := \pi^{-s/2}\Gamma(s/2)$. The basic Archimedean local γ-factors over $\mathbb{R}$ are given (with respect to the standard character ψ of $\mathbb{R}$, as above) by

$$\gamma(s, \operatorname{sgn}^a) = i^a \frac{\Gamma_{\mathbb{R}}(1 - s + a)}{\Gamma_{\mathbb{R}}(s + a)} \quad \big(s \in \mathbb{C}, a \in \{0, 1\}\big),$$

corresponding to the characters $|\cdot|^s \operatorname{sgn}^a$ of $\mathbb{R}^\times$. For $s = \sigma + i\tau$, we define $g_{\sigma,a}(\tau)$ by writing

$$\gamma(s, \operatorname{sgn}^a) = \Big(\frac{|\tau|}{2\pi e}\Big)^{1/2-s} g_{\sigma,a}(\tau).$$

The factor $g_{\sigma,a}(\tau)$ is "mild" in the sense that whenever σ is restricted to a fixed interval and $\min_{n\in\mathbb{N}} |s - n| \ge \varepsilon$ for some fixed $\varepsilon > 0$, we have

$$\partial_\tau^j g_{\sigma,a}(\tau) \ll \big(1 + |\tau|\big)^{-j} \tag{2.4}$$

for all fixed $j \in \mathbb{Z}_{\ge 0}$; this estimate follows from Stirling's formula for $|\tau| \ge 1$ and is otherwise trivial.

From this estimate, we derive a useful approximation for the local variation of $\gamma(s, \operatorname{sgn}^a)$, as follows. For $w = u + iv$, we may write

$$\gamma(s + w, \operatorname{sgn}^a) = \Big(\frac{|v|}{2\pi}\Big)^{1/2-s-w} \exp\big(i\phi_v(\tau)\big) g_{\sigma,u,a}(\tau, v), \tag{2.5}$$

where

$$\phi_v(\tau) := -(v + \tau)\log\Big(\frac{|1 + \tau/v|}{e}\Big) \tag{2.6}$$

and $g_{\sigma,u,a}(\tau, v) = (e^{-1}|1 + \tau/v|)^{1/2-\sigma-u} g_{\sigma+u,a}(\tau + v)$. For $|v| \ge \max(1, 2|\tau|)$ and $\sigma, u \ll 1$, we conclude from (2.4) that

$$\partial_\tau^{j_1} \partial_v^{j_2} g_{\sigma,u,a}(\tau, v) \ll |v|^{-j_1-j_2}. \tag{2.7}$$

2.2.3. Characters of $\mathbb{C}^\times$

We now record the analogous discussion over $\mathbb{C}$. Set $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$. The basic local γ-factors over $\mathbb{C}$ are given with respect to the standard additive character $\psi_{\mathbb{C}}(x) := e^{2\pi i(x+\bar{x})}$ of $\mathbb{C}$ by

$$\gamma_{\mathbb{C}}(s, \mathrm{sgn}_{\mathbb{C}}^a) = i^{a+1}\frac{\Gamma_{\mathbb{C}}(1-s+|a|/2)}{\Gamma_{\mathbb{C}}(s+|a|/2)} \quad (s \in \mathbb{C}, a \in \mathbb{Z}), \tag{2.8}$$

corresponding to the character $|\cdot|_{\mathbb{C}}^s \mathrm{sgn}_{\mathbb{C}}^a$ of $\mathbb{C}^\times$; here $|z|_{\mathbb{C}} := z\bar{z}$, $\mathrm{sgn}_{\mathbb{C}}(z) := z/|z|$. We extend the definition (2.8) to arbitrary $a \in \mathbb{R}$ by taking $i^{a+1} := \exp(\frac{i\pi}{2}(a+1))$. We suppose henceforth that $a \geq 0$. For $s = \sigma + i\tau$, we define $g_{\mathbb{C},\sigma}(\frac{a}{2} + i\tau)$ by writing

$$\gamma_{\mathbb{C}}(s, \mathrm{sgn}_{\mathbb{C}}^a) = \Big(\frac{|a/2+i\tau|}{2\pi e}\Big)^{1-2s}\Big(\frac{a/2+i\tau}{|a/2+i\tau|}\Big)^{-a} g_{\mathbb{C},\sigma}\Big(\frac{a}{2}+i\tau\Big).$$

Again, for σ restricted to a fixed interval and (s,a) a fixed distance away from poles of $\gamma_{\mathbb{C}}(s, \mathrm{sgn}_{\mathbb{C}}^a)$, Stirling's formula implies that

$$\partial_a^{j_1}\partial_\tau^{j_2} g_{\mathbb{C},\sigma}(z) \ll_{\mathcal{D},j_1,j_2} \big(1+|\tau|+|a|\big)^{-j_1-j_2} \tag{2.9}$$

for $j_1, j_2 \geq 0$. We write

$$\gamma_{\mathbb{C}}(s+w, \mathrm{sgn}_{\mathbb{C}}^a) = \Big(\frac{|a/2+iv|}{2\pi}\Big)^{1-2s-2w} \times \exp\big(i\phi_{\mathbb{C},v,a}(\tau)\big) g_{\mathbb{C},\sigma,u}\Big(\tau, \frac{a}{2}+iv\Big), \tag{2.10}$$

where

$$\phi_{\mathbb{C},v,a}(\tau) := -2(v+\tau)\log\Big(\frac{1}{e}\Big|1+\frac{i\tau}{a/2+iv}\Big|\Big) - a\arg\Big(\frac{a}{2}+i(\tau+v)\Big) \tag{2.11}$$

and $g_{\mathbb{C},\sigma,u}(\tau, \frac{a}{2}+iv) = (e^{-1}|1+i\tau/(\frac{a}{2}+iv)|)^{1-2\sigma-2u} g_{\mathbb{C},\sigma+u}(\frac{a}{2}+i(v+\tau))$. For $|a/2+iv| \geq \max(1, 2|\tau|)$ and $\sigma, u \ll 1$, we obtain from (2.9) that

$$\partial_\tau^{j_1}\partial_v^{j_2}\partial_a^{j_3} g_{\mathbb{C},\sigma,u}\Big(\tau, \frac{a}{2}+iv\Big) \ll \big(|v|+|a|\big)^{-j_1-j_2-j_3}. \tag{2.12}$$

2.2.4. The general definition

Any n-dimensional representation of the Weil group $W_{\mathbb{R}}$ is isomorphic to a direct sum

$$\rho = \big(\oplus_{j=1}^{n_1} |\cdot|^{w_j}\,\mathrm{sgn}^{a_1}\big) \oplus \big(\oplus_{j=1}^{n_2} |\cdot|_{\mathbb{C}}^{z_j}\,\mathrm{sgn}_{\mathbb{C}}^{b_j}\big), \tag{2.13}$$

where $n = n_1 + 2n_2$, $w_j, z_j \in \mathbb{C}$, $a_j \in \{0,1\}$, and $b_j \in \mathbb{Z}_{\geq 1}$. Here we identify the indicated characters of $\mathbb{C}^\times \cong W_{\mathbb{C}}$ with the corresponding 2-dimensional induced rep-

resentations of $W_{\mathbb{R}}$. The local γ-factor of ρ is now given by

$$\gamma(s,\rho) = \prod_{j=1}^{n_1} \gamma(s+w_j, \mathrm{sgn}^{a_j}) \prod_{j=1}^{n_2} \gamma_{\mathbb{C}}(s+z_j, \mathrm{sgn}_{\mathbb{C}}^{b_j}).$$

Write $w_j = u_j + iv_j$ and $z_j = x_j + iy_j$. We define the analytic conductor

$$C(\rho) := \prod_{j=1}^{n_1} \frac{\max(1,|v_j|)}{2\pi} \prod_{j=1}^{n_2} \frac{\max(1,b_j/2+iy_j)^2}{(2\pi)^2} \tag{2.14}$$

and, using (2.6) and (2.11), the phase function

$$\phi_\rho(\tau) := \sum_{j=1}^{n_1} \phi_{v_j}(\tau) + \sum_{j=1}^{n_2} \phi_{\mathbb{C},y_j,b_j}(\tau)$$

and the factors

$$g_\sigma(\tau,\rho) := \prod_{j=1}^{n_1} g_{\sigma,u_j,a_j}(\tau,v_j) \prod_{j=1}^{n_2} g_{\mathbb{C},\sigma,x_j}\Big(\tau, \frac{b_j}{2}+iy_j\Big),$$

$$e_\rho := \prod_{j=1}^{n_1} \Big(\frac{|v_j|}{2\pi}\Big)^{-iv_j} \prod_{j=1}^{n_2} \Big(\frac{b_j/2+iy_j}{2\pi}\Big)^{-2iy_j}.$$

By the *dual* (resp., *conjugate*) of ρ, we mean the representation obtained by negating (resp., by conjugating) the parameters w_j, z_j in (2.13). We summarize the previous discussion in the following lemma.

LEMMA 1

Suppose that ρ is isomorphic to its conjugate dual. Then

$$\gamma(s,\rho) = e_\rho C(\rho)^{1/2-s} \exp\big(i\phi_\rho(\tau)\big) g_\sigma(\tau,\rho). \tag{2.15}$$

If moreover ρ is self-dual (equivalently, self-conjugate), then

$$e_\rho = 1.$$

Proof

The content of our hypothesis is that we have the equalities of multisets

$$\big\{(w_1,a_1),\dots,(w_{n_1},a_{n_1})\big\} = \big\{(-\overline{w_1},a_1),\dots,(-\overline{w_{n_1}},a_{n_1})\big\}, \tag{2.16}$$

$$\big\{(z_1,b_1),\dots,(z_{n_2},b_{n_2})\big\} = \big\{(-\overline{z_1},b_1),\dots,(-\overline{z_{n_2}},b_{n_2})\big\}. \tag{2.17}$$

It follows that

$$\prod_{j=1}^{n_1}\Big(\frac{|v_j|}{2\pi}\Big)^{-u_j} = \prod_{j=1}^{n_2}\Big(\frac{|b_j/2+iy_j|}{2\pi}\Big)^{-x_j} = 1.$$

We deduce (2.15) by multiplying together the identities (2.5) and (2.10). Assuming moreover that ρ is self-dual, we obtain the additional equalities of multisets as in (2.16) and (2.17), but without the conjugations, which in turn give that $e_\rho = 1$. □

The primary hypothesis of Lemma 1 is satisfied if, for instance, ρ corresponds to a *unitary* representation π of $\mathrm{GL}_n(\mathbb{R})$, while the full hypotheses are satisfied if π is self-dual.

2.2.5. Examples of interest

We consider in this paper cuspidal automorphic representations π for $\mathrm{SL}_2(\mathbb{Z})$. Each such π defines a generic irreducible unitary representation of $\mathrm{PGL}_2(\mathbb{R})$, hence a 2-dimensional representation ρ_π of $W_\mathbb{R}$. We set

$$\gamma(s,\pi) := \gamma(s,\rho_\pi), \tag{2.18}$$

and similarly define $C(\pi)$, $\phi_\pi(\tau)$ and $g_\sigma(\tau,\pi)$. The possibilities for ρ_π are as follows:

(1) π is a principal series representation $\pi = \pi(r,a)$ obtained by normalized induction of the character $|\cdot|^{ir}\operatorname{sgn}^a$ for some $r \in \mathbb{R}\cup(-1/2,1/2)i$ and $a\in\{0,1\}$, in which case $\rho_\pi = |\cdot|^{ir}\operatorname{sgn}^a \oplus |\cdot|^{-ir}\operatorname{sgn}^a$, or

(2) π is a discrete series representation $\pi=\pi(k)$ of lowest weight $k\in 2\mathbb{Z}_{\geq 1}$, in which case $\rho_\pi = \operatorname{sgn}_\mathbb{C}^{k-1}$.

We note that any such π is self-dual, hence any such ρ is both self-dual and self-conjugate; this property is evident in each example. Thus for $s=\sigma+i\tau$, we have

$$\gamma(s,\pi) = C(\pi)^{1/2-s}\exp\big(i\phi_\pi(\tau)\big)g_\sigma(\tau,\pi), \tag{2.19}$$

where:

- for $\pi=\pi(r,a)$, we have

$$\begin{aligned} C(\pi) &= \Big(\frac{\max(1,|\Re(r)|)}{2\pi}\Big)^2,\\ \phi_\pi(\tau) &= -(r+\tau)\log\Big(\Big|1+\frac{\tau}{r}\Big|\Big) - (-r+\tau)\log\Big(\Big|1-\frac{\tau}{r}\Big|\Big) + 2\tau; \end{aligned} \tag{2.20}$$

- for $\pi=\pi(k)$,

$$\begin{aligned} C(\pi) &= \Big(\frac{\max(1,(k-1)/2)}{2\pi}\Big)^2, \\ \phi_\pi(\tau) &= -2\tau\log\Big(\frac{1}{e}\Big|1+\frac{i\tau}{(k-1)/2}\Big|\Big) - (k-1)\arg\Big(\frac{k-1}{2}+i\tau\Big); \end{aligned} \tag{2.21}$$

and $g_\sigma(\tau,\pi)$ varies mildly in the sense given by the estimates (2.7) and (2.12); namely:

- for $|r| \geq \max(1, 2|\tau|)$ and $\sigma \ll 1$,

$$\partial_\tau^{j_1}\partial_r^{j_2} g_\sigma\big(\tau,\pi(r,a)\big) \ll |r|^{-j_1-j_2};$$

- for $|k-1| \geq 4|\tau|$ and $\sigma \ll 1$,

$$\partial_\tau^{j_1}\partial_k^{j_2} g_\sigma\big(\tau,\pi(k)\big) \ll |k|^{-j_1-j_2}. \tag{2.22}$$

We note that, while $\pi(k)$ is not defined as a representation for nonintegral k, each of the factors $\gamma(s,\pi(k))$, $C(\pi(k))$, $\phi_{\pi(k)}$ and hence also $g_\sigma(\tau,\pi(k))$ is defined for any $k \in \mathbb{R}_{\geq 1}$ (see after (2.8)). For this reason, it makes sense to differentiate with respect to k in (2.22).

On one occasion, we will apply Lemma 1 to a Rankin–Selberg convolution $\pi_1 \otimes \pi_2 \otimes \chi$ of a pair of generic irreducible unitary representations of $\mathrm{PGL}_2(\mathbb{R})$, twisted further by a character χ of $\mathrm{GL}_1(\mathbb{R})$. Writing $\chi = \chi_0|\cdot|^{\Re(\chi)}$ with χ_0 unitary, we have

$$\begin{aligned} \gamma(1/2, \pi_1\otimes\pi_2\otimes\chi) &= \gamma\big(1/2+\Re(\chi), \pi_1\otimes\pi_2\otimes\chi_0\big) \\ &\ll C(\pi_1\otimes\pi_2\otimes\chi)^{-\Re(\chi)}, \end{aligned} \tag{2.23}$$

where in the second step we invoke Lemma 1 and the accompanying Stirling estimates, using the unitarity of π_1, π_2 and χ_0 to verify its hypotheses.

2.3. *General bounds for Whittaker functions*

Let π be a generic irreducible unitary representation of $G := \mathrm{PGL}_2(\mathbb{R})$. We recall that "generic" means that there is a G-equivariant embedding, necessarily unique,

$$\begin{aligned} &\pi \hookrightarrow \big\{W : G \to \mathbb{C} \text{ smooth } \mid W\big(n(x)g\big) = e(x)W(g)\big\}, \\ &v \mapsto W_v, \end{aligned}$$

where G acts on the space on the right-hand side by right translation. The image of π under such an embedding is called the *Whittaker model* of π with respect to ψ. An invariant inner product on π is given by

$$\langle v_1, v_2\rangle_\pi := \int_{\mathbb{R}^\times} W_{v_1}\big(a(y)\big)\overline{W_{v_2}\big(a(y)\big)}\, d^\times y. \tag{2.24}$$

When we speak below of π being realized in its Whittaker model, we mean that we identify π with its image under such an embedding, with inner product normalized as in (2.24).

Fix $\vartheta \in [0, 1/2)$. We say that π is *ϑ-tempered* if it lies in the discrete series or if, writing π as a Langlands quotient of an isobaric sum $\sigma_1 \otimes |\det|^{s_1} \boxplus \sigma_2 \otimes |\det|^{s_2}$, we have that each $|\Re(s_i)| \le \vartheta$. Then π is 0-tempered in the above sense if and only if it is tempered in the usual sense, that is, its matrix coefficients lie in $L^{2+\varepsilon}(G)$ for each $\varepsilon > 0$.

In what follows, we work exclusively with smooth vectors in such representations. Thus "let $v \in \pi$" is shorthand for "let v be a smooth vector in π."

We denote by $\mathcal{S}_d$ the Sobolev norm on π defined in [36, (2.6)]. It takes finite values on smooth vectors.

LEMMA 2

Let π be a ϑ-tempered generic irreducible unitary representation of G, realized in its Whittaker model. For each $W \in \pi$ and all $y \in \mathbb{R}^\times$ and $z \in \mathbb{R}$ with $|z| \le 1000$, we have

$$(y\partial_y)^{j_2}\partial_z^{j_1} W\big(a(y)wn(z)\big) \ll \mathcal{S}_d(W)\min\big(|y|^{1/2-\vartheta}, |y|^{-N}\big) \tag{2.25}$$

and

$$(y\partial_y)^{j_2}\partial_z^{j_1} W\big(a(y)n'(z)\big) \ll \mathcal{S}_d(W)\min\big(|y|^{1/2-\vartheta}, |y|^{-N}\big) \tag{2.26}$$

for all $j_1, j_2, N \in \mathbb{Z}_{\ge 0}$, where $d \in \mathbb{Z}_{\ge 0}$ and the implied constants depend at most upon j_1, j_2, N.

Proof

See [36, Sections 2.4.1 and 3.2.3]. □

2.4. Smooth weight functions

Let $\mathcal{X}$ be a large parameter which will be clear from the context. We adopt the convention that ε denotes a fixed (i.e., independent of $\mathcal{X}$) positive quantity, whose precise meaning may change from line to line. As usual, the notation $A \ll B$ means that $|A| \le C|B|$ for some fixed C; we introduce subscripts as in $A \ll_j B$ to signify that C may depend upon j. We use the notation $A \asymp B$ to denote that A and B are nonzero real numbers for which A/B lies in some fixed compact subset of $(0, \infty)$; we then have $A \ll B \ll A$. We introduce the abbreviation

$$A \preccurlyeq B \iff A \ll_\varepsilon \mathcal{X}^\varepsilon B.$$

We call an expression *negligible* if it is $\ll_N \mathcal{X}^{-N}$ for any $N > 0$. We call a smooth function $V : \mathbb{R}^n \to \mathbb{C}$ *flat* if

$$x_1^{j_1} \cdots x_n^{j_n} V^{(j_1,\dots,j_n)}(x_1,\dots,x_n) \preccurlyeq_{\mathbf{j}} 1 \tag{2.27}$$

for all $\mathbf{j} \in \mathbb{Z}^n_{\geq 0}$. Clearly if V is flat, then so is $\exp(iV)$. If in addition V has fixed compact support in $(0,\infty)^n$, then we call it *nice*. We generally let V denote a nice function in one or more variables, *not necessarily the same at every occurrence*. In practice, V may depend on some additional parameters having certain prescribed sizes; it will always be clear from the context with respect to which variables "flatness" is applied (in which case all implied constants are uniform in these parameters).

For a nice function V, we may separate variables in $V(x_1,\dots,x_n)$ by first inserting a redundant function $V(x_1)\cdots V(x_n)$ that is 1 on the support of V and then applying Mellin inversion

$$\begin{aligned} V(x_1,\dots,x_n) &= V(x_1,\dots,x_n)V(x_1)\cdots V(x_n) \\ &= \int_{\Re(s_1)=0} \cdots \int_{\Re(s_n)=0} \widehat{V}(s_1,\dots,s_n) \\ &\quad \times \big(V(x_1)\cdots V(x_n)x_1^{-s_1}\cdots x_n^{-s_n}\big)\frac{ds_1\cdots ds_n}{(2\pi i)^n}. \end{aligned}$$

Here we can truncate the vertical integrals at height $|\Im s| \preccurlyeq 1$ at the cost of a negligible error. We will often separate variables in this way without explicit mention.

2.5. *Integration by parts and stationary phase*

We quote the following lemmas from [6, Section 8] and its extension in [31, Section 3].

LEMMA 3

Let $Y \geq 1$, $X, P, U, S > 0$, and suppose that w is a smooth function with support on $[\alpha,\beta]$, satisfying

$$w^{(j)}(t) \ll_j XU^{-j}.$$

Suppose that h is a smooth function on $[\alpha,\beta]$ such that

$$\big|h'(t)\big| \gg S$$

for some $S > 0$, and

$$h^{(j)}(t) \ll_j YP^{-j}, \quad \text{for } j = 2,3,\dots.$$

Then

$$\int_{t\in\mathbb{R}} w(t)e^{ih(t)}\,dt \ll_A (\beta-\alpha)X\big[(PS/\sqrt{Y})^{-A}+(SU)^{-A}\big].$$

LEMMA 4

Let $\mathcal{X}$ be a large parameter. Let V be a flat function in the sense of Section 2.4 with support in $\times_{j=1}^{d}[c_{1j},c_{2j}]$ for some fixed intervals $[c_{1j},c_{2j}]\subseteq\mathbb{R}$ not containing 0. Let $X_1,\dots,X_d>0$, $Y\geq\mathcal{X}^{\varepsilon}$. Write $\mathscr{S}=\times_{j=1}^{d}[c_{1j}X_j,c_{2j}X_j]\subseteq\mathbb{R}^d$. Let $\phi:\mathbb{R}^d\to\mathbb{R}$ be a smooth function satisfying the derivative upper bounds[2]

$$\phi^{(j_1,\dots,j_d)}(t_1;t_2,\dots,t_d)\preccurlyeq Y\prod_{i=1}^{d}X_i^{-j_i}$$

for $\boldsymbol{j}\in\mathbb{N}_0^d$ and $(t_1,\dots t_d)\in\mathscr{S}$, as well as the following second derivative lower bound in the first variable:

$$\phi^{(2,0,\dots,0)}(t_1;t_2,\dots,t_d)\gg YX_1^{-2}.$$

Suppose that there exists $t^*=t^*(t_2,\dots,t_d)$ such that $\phi^{(1,0,\dots,0)}(t^*,t_2,\dots,t_d)=0$. Then for any $N>0$, we have

$$\begin{aligned}\int_{\mathbb{R}} V\Big(\frac{t_1}{X_1},\dots,\frac{t_d}{X_d}\Big)e^{i\phi(t_1,\dots,t_d)}\,dt_1&\\ =\frac{X_1}{Y^{1/2}}e^{i\phi(t^*,t_2,\dots,t_d)}W\Big(\frac{t_2}{X_2},\dots,\frac{t_d}{X_d}\Big)&+O_N(\mathcal{X}^{-N})\end{aligned}$$

for a flat function $W=W_N$ with support in $\times_{j=2}^{d}[c_{1j},c_{2j}]$.

3. The local triple product factor

Let π_i for $i=1,2,3$ be generic irreducible unitary representations of G such that:

- π_1 and π_2 are ϑ-tempered, while
- π_3 is tempered.

We regard π_1 and π_2 as fixed. We write $Q=C(\pi_3)$ for the conductor of π_3 and think of Q as a large parameter. The aim of this section is to obtain a lower bound for the local triple product integral $\mathcal{L}_\infty(v_1,v_2,v_3)$ in (1.5) for a certain choice of vectors $v_j\in\pi_j$. The choice will be made at the beginning of Section 3.3 and the result will be stated in Theorem 3 at the end of this section.

[2]The main result in [31, Section 3] states this with $\ll$ instead of $\preccurlyeq$, but our conclusion on W is insensitive to Q^{ε}-powers.

Let ψ denote the additive character of N given by $n(x) \mapsto e(x)$. We realize π_1 (resp., π_3) in its Whittaker model with respect to ψ (resp., $\bar{\psi}$), with inner products normalized as in (2.24).

In this section, we abbreviate $\chi_s := \chi \otimes |\cdot|^s$ for $\chi \in \mathfrak{X}(\mathbb{R}^\times)$ and $s \in \mathbb{C}$.

3.1. Embedding via intertwiners

Let $\chi \in \mathfrak{X}(\mathbb{R}^\times)$. Let $\mathcal{I}(\chi)$ denote the unitarily normalized induction of χ from the standard upper triangular Borel subgroup in G, consisting of smooth $f : G \to \mathbb{C}$ satisfying $f(n(x)a(y)g) = |y|^{1/2}\chi(y)f(g)$. Let $M(\chi)$ denote the standard intertwining operator from the principal series $\mathcal{I}(\chi)$ to $\mathcal{I}(\chi^{-1})$, defined by the integral

$$f \mapsto \int_{x\in\mathbb{R}} f\big(wn(x)\big)\,dx \tag{3.1}$$

for $\Re(\chi) > 0$ and then meromorphically continued to all of $\mathfrak{X}(\mathbb{R}^\times)$.

Let $\chi = |\cdot|^{1/2+k}$ for $k \in \mathbb{Z}_{\geq 0}$, and consider a K-type basis $\{f_{2l}\}_{l\in\mathbb{Z}}$ on $\mathcal{I}(\chi)$. From the computation of [10, Proposition 2.6.3] we see that $M(\chi)f_l = 0$ for $|l| \geq 1 + k$. Thus $M(\chi)$ has a unique infinite-dimensional kernel which is isomorphic to the discrete series D_k of weight k. We normalize $M(\chi)$ as

$$M^*(\chi) := \gamma(0, \chi^2)M(\chi), \tag{3.2}$$

where γ is the local Tate gamma factor as in Section 2.2. Then $M^*(\chi)$ is nonzero for all $\Re(\chi) > 0$ and is meromorphic for all χ. In other words, D_k can be embedded into the principal series representation $\mathcal{I}(\chi)$ with $\chi = |\cdot|^{k+1/2}$ via the normalized intertwining operator $M^*(\chi)$. A similar embedding can be done for the complementary series representation as well (see [10, Section 2.6]).

Let χ with $\Re(\chi) \geq 0$ not be a pole of $M^*(\chi)$. From now on we will only consider χ for which either $\Re(\chi) = 0$ or $\Im(\chi) = 0$. Note that if $\mathcal{I}(\chi)$ is unitary, then χ satisfies this property. We can define a G-invariant sesquilinear pairing on $\mathcal{I}(\chi)$ by

$$(f_1, f_2)_0 := \begin{cases} \int_{x\in\mathbb{R}} f_1(n'(x))\overline{f_2(n'(x))}\,dx & \text{if } \Re(\chi) = 0, \\ \int_{x\in\mathbb{R}} f_1(n'(x))\overline{M^*(\chi)f_2(n'(x))}\,dx & \text{if } \Im(\chi) = 0, \end{cases}$$

for $f_1, f_2 \in \mathcal{I}(\chi)$.

There is a principal series representation $\pi_2^p = \mathcal{I}(\chi)$, with χ of nonnegative real part, into which π_2 embeds. Explicitly:

- If π_2 is a tempered principal series $\mathcal{I}(\chi_0)$, that is, if χ_0 is unitary, then we choose $\chi = \chi_0$.
- If π_2 is the weight-k discrete series D_k, then we choose $\chi = |\cdot|^{1/2+k}$.
- If π_2 is the complementary series attached to $0 < \sigma < 1/2$, then we choose $\chi = |\cdot|^\sigma$.

In each case, we have a G-invariant embedding $\pi_2 \hookrightarrow \pi_2^p = \mathcal{I}(\chi)$ and $c(\chi) \in \mathbb{C}^\times$ such that

$$\langle v_1, v_2 \rangle_{\pi_2} = c(\chi)(f_{v_1}, f_{v_2})_0 =: (f_{v_1}, f_{v_2}),$$

where f_{v_i} are the images of v_i under the above embedding and $\langle , \rangle_{\pi_2}$ is as defined in (2.24). We refer to [10, Section 2.6] for details.

3.2. Local Rankin–Selberg zeta integral

Let $W_1 \in \pi_1$ and $W_3 \in \pi_3$, and let $f_2 \in \mathcal{I}(\chi)$ for some $\chi \in \mathfrak{X}(\mathbb{R}^\times)$. We may parameterize f_2 in terms of a Schwartz function, as follows. Let $e_2 := (0, 1) \in \mathbb{R}^2$, and let $\Phi \in \mathcal{S}(\mathbb{R}^2)$ be a Schwartz function. We define

$$f_2(g) := \int_{t \in \mathbb{R}^\times} \Phi(e_2 t g) \chi_{1/2}\big(\det(tg)\big)\, d^\times t.$$

The above integral converges absolutely for $\Re(\chi) > -1/2$ and continues meromorphically to all $\chi \in \mathfrak{X}(\mathbb{R}^\times)$.

The local $\mathrm{GL}(2) \times \mathrm{GL}(2)$ Rankin–Selberg zeta integral of π_1 and π_3 is defined by

$$\begin{aligned}\Psi(W_1, f_2, W_3) &:= \int_{g \in N \backslash G} W_1(g) f_2(g) W_3(g)\, dg \\ &= \int_{g \in N \backslash \mathrm{GL}_2(\mathbb{R})} W_1(g) W_3(g) \Phi(e_2 g) \chi_{1/2}\big(\det(g)\big)\, dg,\end{aligned}$$

for $\Re(\chi)$ sufficiently large and in general by meromorphic continuation. The $\mathrm{GL}(2) \times \mathrm{GL}(2)$ local functional equation (see [12, Theorem 3.2]) asserts, using the notation (2.18) and (2.2), that

$$\begin{aligned}&\gamma(1/2, \pi_1 \otimes \pi_3 \otimes \chi) \int_{g \in N \backslash \mathrm{GL}_2(\mathbb{R})} W_1(g) W_3(g) \Phi(e_2 g) \chi_{1/2}\big(\det(g)\big)\, dg \\ &\quad = \int_{g \in N \backslash \mathrm{GL}_2(\mathbb{R})} \tilde{W}_1(g) \tilde{W}_3(g) \hat{\Phi}(e_2 g) \chi_{1/2}^{-1}\big(\det(g)\big)\, dg, \end{aligned} \tag{3.3}$$

where $\tilde{W}_i \in \tilde{\pi}_i$ is the contragredient of W_i defined by $\tilde{W}_i(g) = W(wg^{-\top})$ and $\hat{\Phi}$ is the Fourier transform of Φ defined by

$$\hat{\Phi}(y) := \int_{x \in \mathbb{R}^2} \Phi(x) e(y^\top x)\, dx.$$

For χ a fixed distance away from a pole or zero of $\gamma(1/2, \pi_1 \otimes \pi_3 \otimes \chi)$, we have

$$\begin{aligned}\gamma(1/2, \pi_1 \otimes \pi_3 \otimes \chi)^{-1} &\asymp \gamma(1/2, \tilde{\pi}_1 \otimes \tilde{\pi}_3 \otimes \chi^{-1}) \\ &\ll_{\Re(\chi)} C(\pi_1 \otimes \pi_3 \otimes \chi)^{\Re(\chi)} \ll_{\pi_1, \chi} C(\pi_3)^{2\Re(\chi)}, \end{aligned} \tag{3.4}$$

when χ is fixed with $\Re(\chi) \geq 0$. The first estimate above follows from the definition of the gamma factor. The second estimate follows from (2.23). The third estimate follows from repeated application of [21, Lemma A.2].

We record a variant of the local functional equation.

LEMMA 5
We have

$$\Psi(W_1, f_2, W_3)\gamma(1/2, \pi_1 \otimes \pi_3 \otimes \chi) = \Psi\big(W_1, M^*(\chi) f_2, W_3\big),$$

where $M^*(\chi)$ is as in (3.2).

Proof
Let $\Re(\chi)$ be sufficiently large. By expanding the definition (3.1) of the intertwining operator, we see that

$$M(\chi) f_2(g) = \chi_{1/2}\big(\det(g)\big) \int_{x\in\mathbb{R}} \int_{t\in\mathbb{R}^\times} \Phi\big((t,x)g\big)\chi^2(t)\, d^\times t\, dx.$$

We use the local Tate functional equation to evaluate the above as (cf. [15, p. 225])

$$\gamma(0, \chi^2)^{-1}\chi\big(\det(g)\big)\big|\det(g)\big|^{-1/2} \int_{t\in\mathbb{R}^\times} \hat{\Phi}\big((t,0)g^{-\top}\big)\chi^{-2}(t)|t|\, d^\times t.$$

Recalling (3.2), we may thus write

$$M^*(\chi) f_2(g) = \int_{t\in\mathbb{R}^\times} \hat{\Phi}(e_2 t w g^{-\top})\chi_{1/2}^{-1}\big(\det(tg^{-\top})\big)\, d^\times t.$$

We use the definition of $\tilde{W}_i$ and change variables $g \mapsto wg^{-\top}$ on the right-hand side of the local functional equation (3.3) to write

$$\begin{aligned}\gamma(1/2, \pi_1 \otimes \pi_3 \otimes \chi) &\int_{g\in N\backslash \mathrm{GL}_2(\mathbb{R})} W_1(g) W_3(g) \Phi(e_2 g)\chi_{1/2}\big(\det(g)\big)\, dg \\ &= \int_{g\in N\backslash \mathrm{GL}_2(\mathbb{R})} W_1(g) W_3(g) \hat{\Phi}(e_2 w g^{-\top})\chi_{-1/2}\big(\det(g)\big)\, dg.\end{aligned}$$

Folding the above integrals over $\mathbb{R}^\times$, the identity follows for $\Re(\chi)$ large. We conclude the proof by meromorphic continuation of the zeta integrals and the intertwiner. □

LEMMA 6
Let π_1 and π_2 be ϑ-tempered with $\vartheta < 1/4$, and let π_3 be tempered. Let $\pi_i \ni v_i \mapsto W_i$ for $i = 1, 3$ be realized in their respective Whittaker models equipped with the inner

products as defined in (2.24). Also let $v_2 \mapsto f_2$ under $\pi_2 \hookrightarrow \pi_2^p = \mathcal{I}(\chi)$ as described in Section 3.1. Then

$$\int_{g\in G} \prod_{i=1}^{3} \langle \pi_i(g)v_i, v_i\rangle_{\pi_i}\, dg = c(\chi)\Psi(W_1, f_2, W_3)\overline{\Psi(W_1, \tilde{f}_2, W_3)},$$

where

$$\tilde{f}_2 = \begin{cases} f_2 & \text{if } \pi_2^p \text{ is a tempered principal series,} \\ M^*(\chi) f_2 & \text{otherwise.} \end{cases}$$

From [36, Section 2.5.1] we have the bound of the matrix coefficients

$$\langle \pi_i(g)v_i, v_i\rangle \ll_{\pi_i} \Xi(g)^{1-2\vartheta} \quad \text{for } i = 1, 2, \qquad \langle \pi_3(g)v_3, v_3\rangle \ll_{\pi_3} \Xi(g).$$

Here Ξ is the Harish-Chandra Ξ-function, which satisfies $\int_{g\in G} \Xi(g)^{2+\epsilon}\, dg < \infty$. Thus from the assumption that $\vartheta < 1/4$, we see that the local triple product integral is absolutely convergent.

Proof

The proof is essentially given in [39, Lemma 2.14.3], but in an analogous metaplectic setting. We modify the relevant part of the proof. Note that the left-hand side of the equation in the lemma is

$$\int_{g\in G} \langle \pi_1(g)W_1, W_1\rangle \langle \pi_2^p(g) f_2, f_2\rangle \langle \pi_3(g)W_3, W_3\rangle\, dg.$$

We define

$$\xi_1 := W_1 f_2, \qquad \xi_2 := W_1 \tilde{f}_2,$$

and note that

$$\xi_i(ng) = \psi(n)\xi_i(g), \quad n \in N, g \in G.$$

Using Iwahori coordinates $g = a(y)n'(x) \in N\backslash G$ and the transformation of f_2 under the Borel subgroup, we compute the absolutely convergent integral

$$\begin{aligned} &\int_{h\in N\backslash G} \xi_1(hg)\overline{\xi_2(h)}\, dh \\ &\quad = \int_{x\in\mathbb{R}} \pi_2^p(g) f_2\big(n'(x)\big)\overline{\tilde{f}_2\big(n'(x)\big)} \\ &\qquad \times \int_{y\in\mathbb{R}^\times} \pi_1(g)W_1\big(a(y)n'(x)\big)\overline{W_1\big(a(y)n'(x)\big)}\, d^\times y\, dx. \end{aligned}$$

The inner integral evaluates to $\langle \pi_1(g)W_1, W_1\rangle$ and consequently, we have

$$\int_{h\in N\backslash G} \xi_1(hg)\overline{\xi_2(h)}\,dh = \langle \pi_1(g)W_1, W_1\rangle \big(\pi_2^p(g)f_2, \tilde{f}_2\big)_0.$$

Hence, the left-hand side of the equation in the lemma equals

$$c(\chi)\int_{g\in G}\int_{h\in N\backslash G} \xi_1(hg)\overline{\xi_2(h)}\langle \pi_3(g)W_3, W_3\rangle\,dh\,dg.$$

The above double integral is only conditionally convergent. We proceed exactly as in the proof of the identity (2.29) in [39] to evaluate the above integral as

$$c(\chi)\int_{h\in N\backslash G} \xi_1(h)W_3(h)\,dh \int_{h\in N\backslash G} \overline{\xi_2(h)W_3(h)}\,dh.$$

The proof is now complete. □

3.3. Choice of vectors

We choose $f_2 \in \pi_2^p$ as before

$$f_2(g) := \int_{t\in\mathbb{R}^\times} \Phi(e_2 tg)\chi_{1/2}\big(\det(tg)\big)\,d^\times t,$$

where Φ is a smooth nonnegative bump function on $\mathbb{R}^2$ sufficiently concentrated around the point $e_2 = (0,1)$ in terms of π_1 and π_2 only. Such a vector has a nonzero preimage $v_2' \in \pi_2$. We choose

$$v_2 := a(Q)v_2',$$

where, as we recall, $Q = C(\pi_3)$ is the conductor of π_3 as in Section 2.2. We choose $v_i \in \pi_i$ for $i = 1, 3$ such that v_i are analytic newvectors, in the sense of Section 1.6; that is, v_i in their Kirillov models (with conjugate additive characters of N) are given by fixed bump functions in $C_c^\infty(\mathbb{R}^\times)$ sufficiently concentrated around 1. We denote by W_i the images of v_i in their Whittaker models for $i = 1, 3$.

We note that

$$\int_{x\in\mathbb{R}} f_2\big(n'(x)\big)\,dx \asymp 1, \qquad \int_{y\in\mathbb{R}^\times} W_1\big(a(y)\big)W_3\big(a(y)\big)\chi_{-1/2}(y)\,d^\times y \asymp 1.$$

We normalize v_1, v_2', v_3 so that both of the above integrals are 1.

LEMMA 7

Let π_1 be ϑ-tempered with $\vartheta < 1/2$, and let π_3 be tempered. Let χ with $\Re(\chi) \geq 0$ and $f_2 \in \mathcal{I}(\chi)$ be as chosen above. Then for $C(\pi_3) = Q$ sufficiently large, we have

$$\Psi\big(W_1, f_2\big(a(Q)\big), W_3\big) \gg_{\pi_1,\pi_2} Q^{-1/2+\Re(\chi)}.$$

Proof

We write the zeta integral with the Iwahori coordinates and change variables to obtain

$$\begin{aligned}
&\chi_{1/2}^{-1}(Q)\Psi\big(W_1, f_2\big(a(Q)\big), W_3\big)\\
&\quad = \chi_{-1/2}^{-1}(Q)\int_{y\in\mathbb{R}^\times}\int_{x\in\mathbb{R}} W_1\Big(a(y)n'\Big(\frac{x}{Q}\Big)\Big)W_3\Big(a(y)n'\Big(\frac{x}{Q}\Big)\Big)\\
&\qquad \times f_2\big(a(yQ)n'(x)\big)\,dx\frac{d^\times y}{|y|}. \qquad (3.5)
\end{aligned}$$

Note that

$$f_2\big(a(yQ)n'(x)\big) = \chi_{1/2}(yQ)\,f_2\big(n'(x)\big),$$

and the support condition of Φ confirms that $f_2(n'(x))$ is supported in a sufficiently small neighborhood of 0. We rewrite the right-hand side of (3.5) as

$$\begin{aligned}
&\int_{y\in\mathbb{R}^\times}\int_{x\in\mathbb{R}}\Big(W_3\Big(a(y)n'\Big(\frac{x}{Q}\Big)\Big) - W_3\big(a(y)\big)\Big)\\
&\qquad \times W_1\Big(a(y)n'\Big(\frac{x}{Q}\Big)\Big)f_2\big(n'(x)\big)\chi_{-1/2}(y)\,dx\,d^\times y\\
&\quad + \int_{y\in\mathbb{R}^\times}\int_{x\in\mathbb{R}}\Big(W_1\Big(a(y)n'\Big(\frac{x}{Q}\Big)\Big) - W_1\big(a(y)\big)\Big)\\
&\qquad \times W_3\big(a(y)\big)f_2\big(n'(x)\big)\chi_{-1/2}(y)\,dx\,d^\times y\\
&\quad + \int_{y\in\mathbb{R}^\times} W_1\big(a(y)\big)W_3\big(a(y)\big)\chi_{-1/2}(y)\,d^\times y\int_{x\in\mathbb{R}} f_2\big(n'(x)\big)\,dx. \qquad (3.6)
\end{aligned}$$

Note that the third integral equals 1 by the choice of normalizations of the vectors.

As π_1 is fixed, we may apply (2.26) to conclude that

$$W_1\big(a(y)n'(x/Q)\big) \ll_N \min\big(|y|^{1/2-\vartheta}, |y|^{-N}\big)$$

for $x \ll 1$. Moreover, given some sufficiently small constant $c > 0$, then for x sufficiently small (in terms of c only) and $C(\pi_i) \le Q$ we have

$$W_i\big(a(y)n'(x/Q)\big) - W_i\big(a(y)\big) \ll \frac{C(\pi_i)}{Q}c|y|^{1/2-\theta-\eta}, \qquad (3.7)$$

for any $\eta > 0$ and $\theta = \vartheta, 0$ if $i = 1, 3$, respectively. This estimate is essentially contained in [26, Section 2.1] and can be seen as follows: using Mellin inversion for $W_i(a(y)n'(x/Q))|y|^{-\sigma}$ for $\sigma = 1/2 - \theta - \eta$ and the PGL(2) × GL(1) local func-

tional equation, we write the above difference (as in [26, Section 2.1][3])

$$\int_{\Re(\chi')=0} \chi'_\sigma(y) \frac{\chi'_\sigma(C(\pi_i))}{\gamma(1/2-\sigma, \pi_i \otimes \chi'^{-1})} \times \int_{t\in\mathbb{R}^\times} \Big(e\Big(-\frac{txC(\pi_i)}{Q}\Big) - 1\Big) W_i\big(a\big(C(\pi_i)t\big)w\big)\chi'_\sigma(t)\, d^\times t\, d\chi',$$

and (3.7) follows as in [26, Section 2.1].

Now we define

$$I_2(\chi) := \int_{x\in\mathbb{R}} \big|f_2\big(n'(x)\big)\big|\, dx \le \int_{(t,x)\in\mathbb{R}^2} \Phi(tx,t)|t|^{2\Re(\chi)}\, dt\, dx \ll 1.$$

The first integral in (3.6) is therefore

$$\ll c\, I_2(\chi) \int_{y\in\mathbb{R}^\times} |y|^{1/2-\eta} \min\big(|y|^{1/2-\vartheta-\eta}, |y|^{-N}\big)|y|^{\Re(\chi)-1/2}\, d^\times y \ll c,$$

as $\Re(\chi) \ge 0$ and $\vartheta < 1/2$. The second integral in (3.6) is similarly

$$\ll Q^{-1} I_2(\chi) \int_{y\in\mathbb{R}^\times} W_3\big(a(y)\big)|y|^{1/2-\vartheta}|y|^{\Re(\chi)-1/2}\, d^\times y \ll Q^{-1}.$$

Thus we estimate

$$\Psi\big(W_1, f_2\big(a(Q)\big), W_3\big) \gg Q^{-1/2+\Re(\chi)}\big(1 + O(c) + O(1/Q)\big),$$

which concludes the proof upon choosing c sufficiently small in terms of the implied constants. □

THEOREM 3

Let π_i for $i = 1,2,3$ be generic irreducible unitary representations of G such that π_1 and π_2 are ϑ-tempered with $\vartheta < 1/4$ and π_3 is tempered with sufficiently large conductor Q. The (smooth) vectors $v_i \in \pi_i$ specified at the beginning of Section 3.3 have the following properties.

(i) *We have $\|v_i\| \asymp 1$ for $i = 1,2,3$.*

(ii) *We have $v_1 = v_1^0$ and $v_2 = a(Q)v_2^0$, where v_1^0, v_2^0 are fixed (independent of Q).*

(iii) *For any fixed nontrivial unitary character ψ of N, the vector v_3 is given in the ψ-Kirillov model by a fixed bump function.*

(iv) *We have*

$$\int_{g\in G} \prod_{i=1}^{3} \big\langle \pi_i(g)v_i, v_i\big\rangle dg \ll_{\pi_1,\pi_2} Q^{-1}.$$

[3] In that paper, the authors took $y = 1$.

Proof

Assertions (i), (ii), and (iii) are clear from the construction. (The description of v_3 in the Kirillov model is independent of the choice of ψ: different choices give rise to models that are isomorphic to one another via left translation by a suitable diagonal matrix.)

To verify (iv), we embed $\pi_2 \hookrightarrow \pi_2^p = \mathcal{I}(\chi)$ for some χ with $\Re(\chi) \geq 0$ such that either $\Re(\chi) = 0$ or $\Im(\chi) = 0$. For $\Re(\chi) = 0$, the integral in question evaluates to

$$c(\chi)\left|\Psi\big(W_1, f_2\big(a(Q)\big), W_3\big)\right|^2$$

by Lemma 6, and Lemma 7 implies the required bound. For $\Im(\chi) = 0$ and $\Re(\chi) > 0$, we apply Lemmas 5 and 6 to see that the integral in question is

$$c(\chi)\overline{\gamma(1/2, \pi_1 \otimes \pi_3 \otimes \chi)}\left|\Psi\big(W_1, f_2\big(a(Q)\big), W_3\big)\right|^2.$$

An appeal to (3.4) and Lemma 7 then completes the proof. □

4. A triple Whittaker integral

4.1. Setting and statement of results

In this section, we evaluate asymptotically an integral containing three Whittaker functions which is the crucial ingredient for an understanding of the right-hand side of (1.5). We will not need any knowledge on special functions, but we do use extensively Stirling's formula and stationary phase analysis as described in Sections 2.2 and 2.5.

We retain the basic notation of Section 2.1. We continue to adopt the following setting (as in the previous section):

- π_1 and π_2 are fixed ϑ-tempered generic irreducible unitary representations of G, with $0 \leq \vartheta < 1/2$ fixed.
- π_3 is a varying tempered generic irreducible unitary representation of G, whose analytic conductor (normalized as in Section 2.2) we denote by $Q := C(\pi_3)$.
- Recall that we realize each π_j in its Whittaker model as a space of functions W satisfying $W(n(x)g) = e(x)W(g)$, where $e(x) := e^{2\pi i x}$. Also, recall from (2.24) that we normalize this realization so that the inner product on π_j is given in the Kirillov model by integration over the diagonal subgroup: $\|W\|^2 = \int_{y\in\mathbb{R}^\times} |W(a(y))|^2\, d^\times y$.
- We let $W_j \in \pi_j$ be the image of the vector v_j as in Theorem 3. Thus $W_1 = W_1^0$ and $W_2 = a(Q)W_2^0$ with W_1^0, W_2^0 fixed (independent of Q).

The basic bounds from Lemma 2 can be used for $W_1 = W_1^0$ and W_2^0. We will derive useful bounds for W_3 in Section 4.3 below.

We recall the notation and conventions of smooth weight functions in Section 2.4. Our basic large parameter here is Q, so $A \preccurlyeq B$ means $A \ll_\varepsilon Q^\varepsilon B$. As usual, the value of ε may change from line to line.

For $y_1, y_2 \in \mathbb{R}^\times$ with $y_1 + y_2 \neq 0$, we define $y_3 \in \mathbb{R}^\times$ by requiring that

$$y_1 + y_2 + y_3 = 0. \tag{4.1}$$

We set

$$F(y_1, y_2) := \int_{\theta \in \mathbb{R}/\pi\mathbb{Z}} \prod_{j=1,2,3} W_j\big(a(y_j)k(\theta)\big)\, d\theta. \tag{4.2}$$

The main result of this section is the following estimate for F and its derivatives.

THEOREM 4
We have

$$F(y_1, y_2) = \sum_{\pm} e\Big(\pm 2\sqrt{Q}\Psi\Big(\frac{y_1}{y_2}\Big)\Big)\mathcal{N}_\pm(y_1, y_2) + \mathcal{E}(y_1, y_2),$$

where Ψ is a smooth function satisfying the estimates

$$\Psi(y) = |y|^{1/2} + O\big(|y|^{3/2}\big), \quad \Psi^{(j)}(y) \ll |y|^{1/2-j} \quad (j \in \mathbb{N}) \tag{4.3}$$

and for fixed $j_1, j_2, N \geq 0$ we have

$$(y_1\partial_{y_1})^{j_1}(y_2\partial_{y_2})^{j_2}\mathcal{N}_\pm(y_1, y_2) \preccurlyeq \Big(\frac{|y_2|}{Q}\Big)^{3/4}\big(1 + |y_1|\big)^{-N}\Big(1 + \frac{|y_2|}{Q}\Big)^{-N}$$

and $\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3$, where for some absolute constant $c > 0$ we have

$$\mathcal{E}_1(y_1, y_2) = Q^{-1/2}\big(1 + |y_1| + |y_2|\big)^{-N},$$
$$\mathcal{E}_2(y_1, y_2) = Q^c\big(1 + |y_1| + |Qy_2|\big)^{-N},$$
$$\mathcal{E}_3(y_1, y_2) = \big(1 + |y_1| + |y_2/Q|\big)^{-N} Q^{-N}.$$

Here and in the following, all implied constants may depend on N and j, with or without subscript. The proof gives an exact formula for Ψ which however is irrelevant for our application. It depends mildly on whether π_3 belongs to the principal series or to the discrete series. The key point of Theorem 4 is that it produces the desired and expected oscillatory factor (cf. (1.4)).

4.2. Preliminary decomposition

It will be convenient first to switch to Iwahori coordinates. We may find an even function $\phi_0 \in C_c^\infty(\mathbb{R}/\pi\mathbb{Z})$, supported on $(-\pi/3, \pi/3) + \pi\mathbb{Z}$, so that $\phi_0(\theta) + \phi_0(\pi/2 + \theta) = 1$ for all θ. Setting $z := \tan\theta$ and using the matrix identities

$$k(\theta) = n(-z)a(z^2+1)n'(z),$$
$$k(\pi/2-\theta) = n(z)a(z^2+1)wn(z)$$

and the relation $d\theta = (1+z^2)^{-1}\,dz$, we see that $F(y_1, y_2) = F_1(y_1, y_2) + F_w(y_1, y_2)$, where, with the abbreviation $y_j' := y_j(z^2+1)$, we have

$$F_1(y_1, y_2) := \int_{z\in\mathbb{R}} \Big(\prod_{j=1,2,3} W_j\big(a(y_j')n'(z)\big)\Big)\phi_0\big(\arctan(z)\big)\,\frac{dz}{z^2+1},$$

$$F_w(y_1, y_2) := \int_{z\in\mathbb{R}} \Big(\prod_{j=1,2,3} W_j\big(a(y_j')wn(z)\big)\Big)\phi_0\big(\arctan(z)\big)\,\frac{dz}{z^2+1}.$$

Note that $z \mapsto \phi_0(\arctan(z))$ defines a smooth compactly-supported function on $\mathbb{R}$.

We now further decompose F_1. We write 1 as a sum $\phi^\sharp + \phi^\flat$ of smooth functions on $\mathbb{R}^\times$ with

- $\phi^\flat(z)$ supported on $z \ll Q^{-1}$,
- $\phi^\sharp(z)$ supported on $Q^{-1} \ll z$,

and with each function ϕ in this decomposition satisfying $(z\partial_z)^j\phi(z) \ll 1$. We accordingly decompose $F_1 = F_1^\flat + F_1^\sharp$ by weighting the z-integral. In summary, we have decomposed

$$F = F_w + F_1^\flat + F_1^\sharp. \tag{4.4}$$

We will verify that each of the three terms on the right-hand side of (4.4) satisfies the conclusions of Theorem 4. The first two terms are fairly straightforward to analyze. We treat them in Section 4.4. Indeed, we will see that the contribution of $F_1^\flat$ can be absorbed into the error term $\mathcal{E}$, while F_w is nonnegligible only if $y_1 \preccurlyeq 1$ and y_2 is roughly of size Q, in which case the oscillatory factor $e(\pm Q^{1/2}\Psi(y_1/y_2))$ is flat in the sense of Section 2.4. The somewhat more intricate analysis of $F_1^\sharp$ is carried out in Section 4.5.

4.3. Interlude: Bounds for newvectors

For the relevant asymptotic analysis we will need certain uniform bounds of the test vectors. We record them here.

LEMMA 8

Let $W \in \pi$, $y \in \mathbb{R}^\times$, $z \in \mathbb{R}$, and let $\sigma > -1/2 + \vartheta$. Then $W(a(y)wn(z))$ admits the absolutely convergent integral representation

$$W\big(a(y)wn(z)\big) = \int_{\Re(\chi)=\sigma} \chi(y)\gamma(1/2, \pi \otimes \chi) \times \Big(\int_{t\in\mathbb{R}^\times} e(tz)W\big(a(t)\big)\chi(t)\, d^\times t\Big)\, d\chi. \tag{4.5}$$

Proof

By Mellin inversion (2.1)—using the estimate (2.25) to verify its hypotheses—we have

$$W\big(a(y)wn(z)\big) = \int_{\Re(\chi)=\sigma} \chi(y) \int_{t\in\mathbb{R}^\times} W\big(a(t)wn(z)\big)\chi^{-1}(t)\, d^\times t\, d\chi.$$

By the $\mathrm{PGL}(2) \times \mathrm{GL}(1)$ local functional equation (see [12, Theorem 3.1]), the inner integral evaluates to

$$\gamma(1/2, \pi \otimes \chi) \int_{t\in\mathbb{R}^\times} W\big(a(t)n(z)\big)\chi(t)\, d^\times t.$$

We conclude by calculating that $W(a(t)n(z)) = W(n(tz)a(t)) = e(tz)W(a(t))$. □

LEMMA 9

Let π be tempered of conductor Q, and let $W \in \pi$ be an analytic newvector. For fixed $N > 0$ and $j \in \mathbb{Z}_{\geq 0}$, we have

$$(x\partial_x)^j W\big(a(Qx)wn(z)\big) \preccurlyeq \min\big(|x|^{-N}, |x|^N + |x|^{1/2}Q^{-N}\big) \tag{4.6}$$

uniformly in $z \preccurlyeq 1$.

Proof

We write the proof in the case $j = 0$. The general case is treated similarly, using that $(y\partial_y)^j \chi(y) = s^j \chi(y)$ for $\chi = |\cdot|^s \operatorname{sgn}^a$.

By (4.5), we have the Mellin expansion

$$W\big(a(Qx)wn(z)\big) = \int_\chi \chi(Qx)\gamma(1/2, \pi \otimes \chi)V_z(\chi)\, d\chi,$$

where

$$V_z(\chi) := \int_{t\in\mathbb{R}^\times} e(tz)W(t)\chi(t)\, d^\times t.$$

Since $z \preccurlyeq 1$ and W is supported in a fixed compact set, we see that V_z is entire and $V_z(\chi) \preccurlyeq C(\chi)^{-N}$ in vertical strips. By Lemma 1 and [21, Lemma A.2] we have

$$\gamma(1/2, \pi \otimes \chi) \ll C(\pi \otimes \chi)^{-\Re(\chi)} \ll C(\pi)^{-\Re(\chi)} C(\chi)^{2|\Re(\chi)|},$$

which holds uniformly in any fixed vertical strip and for any χ separated by $\gg 1$ from any pole of $\gamma(1/2, \pi \otimes \chi)$. (In more detail, we apply Lemma 1 by writing $\chi = \chi_0 |\cdot|^{\Re(\chi)}$, with χ_0 unitary, and using that $\gamma(1/2, \pi \otimes \chi) = \gamma(1/2 + \Re(\chi), \pi \otimes \chi_0)$ and $C(\pi \otimes \chi_0) \asymp C(\pi \otimes \chi)$.) We obtain an adequate estimate in the case $|x| \geq 1$ by shifting the contour to $\Re(\chi) = -N$, passing no poles.

It remains to consider the case $|x| \leq 1$. We shift the contour to $\Re(\chi) = N$. Recall that Q is assumed sufficiently large in terms of "fixed" quantities; in particular, Q is much larger than any fixed power of N. It follows that if π belongs to the discrete series, then this contour shift passes no poles. The required estimate follows (in the stronger form obtained by omitting the term $Q^{-N}|x|^{1/2}$). Suppose now that π belongs to the principal series. By hypothesis, π is tempered. Thus each pole that we cross is of the form $\chi = \operatorname{sgn}^a |\cdot|^{\sigma+it}$ with $\sigma \geq 1/2 - \vartheta$ and $t \asymp Q^{1/2}$. The required estimate follows from the rapid decay of V_z. □

4.4. The easy cases: Estimates for F_w, $F_1^\flat$

PROPOSITION 10

The function F_w satisfies the conclusions of Theorem 4.

Proof

By the matrix identity

$$a(y_2')wn(z)a(Q) = a(y_2'/Q)wn(z/Q) \in G$$

and the relations $W_1 = W_1^0$, $W_2 = a(Q)W_2^0$, we write $F_w(y_1, y_2)$ as

$$\int_{z \in \mathbb{R}} W_1^0\big(a(y_1')wn(z)\big) W_2^0\Big(a\Big(\frac{y_2'}{Q}\Big)wn\Big(\frac{z}{Q}\Big)\Big) \times W_3\big(a(y_3')wn(z)\big)\phi_0\big(\arctan(z)\big)\,\frac{dz}{z^2+1}.$$

We use (2.25) to bound the first two factors of the integrand as

$$(y_1\partial_{y_1})^{j_1}(y_2\partial_{y_2})^{j_2} W_1^0\big(a(y_1')wn(z)\big) W_2^0\Big(a\Big(\frac{y_2'}{Q}\Big)wn\Big(\frac{z}{Q}\Big)\Big) \ll \big(1+|y_1|\big)^{-N}\Big(1+\frac{|y_2|}{Q}\Big)^{-N} \tag{4.7}$$

uniformly for $z \ll 1$. We bound the third factor by (4.6) with $x = y_3'/Q$ getting

$$(y_3\partial_{y_3})^j W_3\big(a(y_3')wn(z)\big) \ll \min\Big(\Big|\frac{y_3}{Q}\Big|^{-N}, \Big|\frac{y_3}{Q}\Big|^{N} + Q^{-N}\Big).$$

Recalling (4.1), it is easy to see that

$$(y_1\partial_{y_1})^{j_1}(y_2\partial_{y_2})^{j_2} F_w(y_1,y_2) \ll \big(1+|y_1|\big)^{-N} \min\Big(\Big|\frac{y_2}{Q}\Big|^{-N}, \frac{1+|y_2|^N}{Q^N}\Big).$$

If Ψ satisfies (4.3), then $e(\pm 2\sqrt{Q\Psi(y_1/y_2)})$ is flat for $y_1 \preccurlyeq 1$ and $Q^{1-\varepsilon} \ll |y_2| \ll Q^{1+\varepsilon}$, so up to a contribution that can go into $\mathcal{E}_3$ we have

$$\begin{aligned}(y_1\partial_{y_1})^{j_1}(y_2\partial_{y_2})^{j_2} e\big(\pm 2\sqrt{Q\Psi(y_1/y_2)}\big) F_w(y_1,y_2)\\ \preccurlyeq \big(1+|y_1|\big)^{-N} \min\Big(\Big|\frac{y_2}{Q}\Big|^{-N}, \frac{1+|y_2|^N}{Q^N}\Big),\end{aligned}$$

which is admissible for Theorem 4. □

PROPOSITION 11

The function $F_1^\flat$ satisfies the conclusions of Theorem 4.

Proof

By definition,

$$F_1^\flat(y_1,y_2) = \int_{z\in\mathbb{R}} \Big(\prod_{j=1,2,3} W_j\big(a(y_j')n'(z)\big)\Big)\phi^\flat(z)\phi_0\big(\arctan(z)\big)\,\frac{dz}{z^2+1}.$$

For $|z| \ll 1/Q$, we have

$$\begin{aligned}W_1\big(a(y_1')n'(z)\big) &\ll \big(1+|y_1|\big)^{-N},\\ W_2\big(a(y_2')n'(z)\big) &= W_2^0\big(a(Qy_2')n'(Qz)\big) \ll \big(1+|y_2|Q\big)^{-N},\end{aligned}$$

where we used (2.26). The required estimate $F_1^\flat \ll \mathcal{E}_2 \le \mathcal{E}$ follows now from the weak a priori bound $\|W_3\|_\infty \ll Q^{O(1)}$.

To deduce the latter, it suffices by the Iwasawa decomposition to estimate $W_3(a(y)k(\theta))$. We appeal first to Lemma 2, which gives for any $W \in \pi_3$ the estimate $W(a(y)) \ll \mathcal{S}_d(W)$ for some fixed d. Taking $W = k(\theta)W_3$, we obtain $W(a(y)) = W_3(a(y)k(\theta)) \ll \mathcal{S}_d(W) = \mathcal{S}_d(W_3)$. On the other hand, by [36, Section 3.2.5] and the fact that $y \mapsto W_3(a(y))$ is a fixed bump function, we have $\mathcal{S}_d(W_3) \ll Q^{d'}$ for some fixed d'. The required a priori bound follows. □

4.5. *The critical case: Estimates for* $F_1^\sharp$

PROPOSITION 12
The function $F_1^\sharp$ *satisfies the conclusions of Theorem 4.*

Proof
Let $Q^{-1} \ll z$. We use the identity

$$a(y)n'(z)a(Q) = n(y/z)a(y/Qz^2)wn(1/Qz) \in G$$

to write

$$W_2\big(a(y_2')n'(z)\big) = e(y_2'/z)W_2^0\big(a(y_2'/Qz^2)wn(1/Qz)\big).$$

Similarly, we have

$$\begin{aligned} W_3\big(a(y_3')n'(z)\big) &= W_3\big(n(y_3'/z)a(y_3'/z^2)wn(1/z)\big) \\ &= e(y_3'/z)W_3\big(a(y_3'/z^2)wn(1/z)\big). \end{aligned}$$

Using the consequence $e(y_2'/z)e(y_3'/z) = e(-y_1'/z)$ of the hypothesis (4.1), we may write $F_1^\sharp(y_1, y_2)$ in the form

$$\begin{aligned} &\int_{z\in\mathbb{R}} W\Big(y_1, \frac{y_2}{Qz^2}, z\Big)e\Big(-\frac{y_1(z^2+1)}{z}\Big) \\ &\qquad \times W_3\Big(a\Big(\frac{y_3'}{z^2}\Big)wn\Big(\frac{1}{z}\Big)\Big)\phi^\sharp(z)\phi_0\big(\arctan(z)\big)\,\frac{dz}{z^2+1}, \end{aligned}$$

where

$$W(x_1, x_2, z) := W_1^0\big(a\big(x_1(1+z^2)\big)n'(z)\big)W_2^0\big(a\big(x_2(1+z^2)\big)wn(1/Qz)\big).$$

From now on we will use extensively the conventions on smooth weight functions stated in Section 2.4. In particular, V denotes generally a flat function, *not necessarily the same at every occurrence*.

We apply the substitution $z \mapsto 1/z$ and a smooth dyadic partition of unity localizing $\pm z \asymp Z$, where Z runs over $\preccurlyeq 1$ values (e.g., powers of 2) satisfying $1 \ll Z \ll Q$. For notational simplicity we restrict to $z > 0$, the case $z < 0$ being essentially identical. Since W is flat in z, it suffices to estimate

$$\begin{aligned} I_0(y_1, y_2, y_3) = W\Big(y_1, \frac{y_2Z^2}{Q}\Big) &\\ \times \int_{z\in\mathbb{R}} e\big(-y_1(z+z^{-1})\big)W_3\big(a\big(y_3(1+z^2)\big)wn(z)\big) &\\ \times V\Big(\frac{z}{Z}\Big)\frac{dz}{|z|^2}, & \end{aligned} \tag{4.8}$$

where W satisfies the estimates

$$(x_1\partial_{x_1})^{j_1}(x_2\partial_{x_2})^{j_2}W(x_1,x_2)\ll\prod_{j=1}^{2}\big(1+|x_j|\big)^{-N}. \tag{4.9}$$

At this point, we can easily deal with the contribution $Z\gg Q^{1/2-\varepsilon}$. A trivial estimate returns the bound

$$\preccurlyeq Q^{-1/2}\big(1+|y_1|\big)^{-N}\big(1+|y_2|\big)^{-N}=\mathcal{E}_1(y_1,y_2)$$

which is acceptable.

We can also easily deal with the contribution $Z\preccurlyeq 1$. In this case, (4.6) yields

$$(y_3\partial_{y_3})^j W_3\big(a\big(y_3(1+z^2)\big)wn(z)\big)\ll\min\Big(\Big(\frac{|y_3|}{Q}\Big)^N+Q^{-N},\Big(\frac{|y_3|}{Q}\Big)^{-N}\Big).$$

Recalling (4.1), we argue as in the proof of Proposition 10 that under the present assumption $Z\preccurlyeq 1$, up to an error of size $\mathcal{E}_3$, we have

$$(y_1\partial_{y_1})^{j_1}(y_2\partial_{y_2})^{j_2}I_0(y_1,y_2,y_3)\preccurlyeq\big(1+|y_1|\big)^{-N}\min\Big(\Big|\frac{y_2}{Q}\Big|^{-N},\frac{1+|y_2|^N}{Q^N}\Big)$$

but then also

$$\begin{aligned}&(y_1\partial_{y_1})^{j_1}(y_2\partial_{y_2})^{j_2}e\big(\pm2\sqrt{Q\Psi(y_1/y_2)}\big)I_0(y_1,y_2,y_3)\\&\preccurlyeq\big(1+|y_1|\big)^{-N}\min\Big(\Big|\frac{y_2}{Q}\Big|^{-N},\frac{1+|y_2|^N}{Q^N}\Big)\end{aligned}$$

as required for the bound for $\mathcal{N}_\pm$. So from now on we assume that

$$Q^\varepsilon\le Z\le Q^{1/2-\varepsilon}.$$

We appeal to the (rapidly convergent) integral formula (4.5) to rewrite the integral over z in (4.8) as

$$\begin{aligned}I:=&\int_{\Re(\chi)=0}\gamma(1/2,\pi_3\otimes\chi)\\&\times\int_{z\in\mathbb{R}}\chi\big(y_3(z^2+1)\big)e\big(-y_1(z+z^{-1})\big)V\Big(\frac{z}{Z}\Big)I_1(z,\chi)\,\frac{dz}{|z|^2}\,d\chi,\end{aligned}\tag{4.10}$$

where

$$I_1(z,\chi):=\int_{t\in\mathbb{R}}e(tz)\chi(t)W_3\big(a(t)\big)\,\frac{dt}{|t|}.$$

We keep in mind that the bound (4.9) allows us to assume $y_1 \preccurlyeq 1$, otherwise we can bound I trivially by $\mathcal{E}_3$.

Let us write $\chi = |\cdot|^{it_\chi} \operatorname{sgn}^{a_\chi}$ with $t_\chi \in \mathbb{R}$ and $a_\chi \in \{0, 1\}$. We split the χ-integral in (4.10) according to the value of a_χ and regard that value as fixed from now on. We have expressed I as a triple integral in t, z, t_χ, and we will apply stationary phase in each of the variables, one at a time. Here $z \asymp Z$, $t \asymp 1$, and we will see in a moment that the t-integral is negligible unless $t_\chi \asymp Z$. Stationary phase saves a factor $Z^{1/2}$ in each of these integrals, so we expect that I is of size $Z^{-3/2}$, and we can explicitly compute the oscillatory behavior.

Step 1: The t-integral. We apply stationary phase analysis to find, for each fixed $N \in \mathbb{Z}_{\geq 0}$, a nice function V so that

$$I_1(z, \chi) = z^{-1/2} V\Big(\frac{-t_\chi}{z}, \frac{z}{Z}\Big)\chi\Big(\frac{t_\chi}{2\pi e z}\Big) + O(Q^{-N}). \tag{4.11}$$

To see this, we apply Lemma 4 if $-t_\chi/z \asymp 1$ by taking

$$\phi(t) = \phi(t; z, t_\chi) = 2\pi z t + t_\chi \log|t|, \quad t^* = -\frac{t_\chi}{2\pi z},$$
$$(X_1, X_2, X_3) = (1, Z, Z), \quad Y = Z$$

and otherwise the integral is negligible by Lemma 3 with $U = P = 1$, $S = Y = Z$. Here and henceforth the contribution of all negligible error terms is covered by $\mathcal{E}_3$ in the statement of Theorem 4. We deduce that I is given up to acceptable error (that can go into $\mathcal{E}_3$) by

$$\int_{\Re(\chi)=0} \gamma(1/2, \pi_3 \otimes \chi)\chi\Big(\frac{t_\chi y_3}{2\pi e}\Big) \times \int_{z\in\mathbb{R}} \chi(z) e(-y_1 z) V\Big(\frac{z}{Z}, \frac{-t_\chi}{Z}\Big)\chi(1 + z^{-2}) e(-y_1 z^{-1}) \frac{dz}{|z|^{5/2}}\, d\chi.$$

Note that for $|t_\chi| \asymp z \geq Q^\varepsilon$ and $y_1 \leq Q^\varepsilon$ the function $\chi(1 + z^{-2})e(-y_1 z^{-1})$ is flat in both z and t_χ, so we can incorporate this factor into V (and continue to call this new function V). Thus we may reduce further to estimating

$$\int_{\Re(\chi)=0} V\Big(\frac{-t_\chi}{Z}\Big)\gamma(1/2, \pi_3 \otimes \chi)\chi\Big(\frac{t_\chi y_3}{2\pi e}\Big) I_2(\chi, y_1)\, d\chi \tag{4.12}$$

with

$$I_2(\chi, y_1) := \int_{z\in\mathbb{R}} V\Big(\frac{z}{Z}\Big)\chi(z) e(-y_1 z) \frac{dz}{|z|^{5/2}}.$$

Step 2: The z-integral. By stationary phase analysis, we may find for each fixed N a nice function V (again potentially different from previous versions of V) so that

$$I_2(\chi, y_1) = Z^{-2} V\Big(\frac{t_\chi}{Z y_1}\Big)\chi\Big(\frac{t_\chi}{2\pi e y_1}\Big) + O(Q^{-N}), \tag{4.13}$$

whenever $-t_\chi \asymp Z$. We apply Lemma 4, taking

$$\phi(t) = \phi(t; y_1, t_\chi) = -2\pi y_1 t + t_\chi \log|t|, \quad t^* = \frac{t_\chi}{2\pi y_1},$$
$$(X_1, X_2, X_3) = (Z, 1, Z), \quad Y = Z$$

(if $|y_1| \asymp 1$, and otherwise the integral is negligible by Lemma 3 with $U = Y = P = Z$, $S = 1$). Thus we reduce to studying

$$Z^{-2} V(-y_1) \int_{\Re(\chi)=0} V\Big(\frac{-t_\chi}{Z}\Big)\gamma(1/2, \pi_3 \otimes \chi)\chi\Big(\frac{t_\chi^2 y_3}{(2\pi e)^2 y_1}\Big)\, d\chi, \tag{4.14}$$

for certain nice functions V.

Step 3: The t_χ-integral. We now evaluate $\gamma(1/2, \pi_3 \otimes \chi)$ asymptotically using Stirling's formula, and appeal in particular to (2.19) and the subsequent explicit formulas. Since $|t_\chi| \asymp Z \ll Q^{1/2-\varepsilon}$, both are applicable. We start with the principal series case. In view of (2.20), the relevant phase function may be written

$$\begin{aligned}\phi(t_\chi) &= \phi(t_\chi; r, y_1, y_3)\\ &= t_\chi \log\Big|\frac{t_\chi^2 y_3}{(2\pi e)^2 y_1}\Big|\\ &\quad + (r - t_\chi)\log(r - t_\chi) - (r + t_\chi)\log(r + t_\chi) + 2t_\chi \log(2\pi e).\end{aligned}$$

For convenience, we record some relevant derivatives:

$$\begin{aligned}&\frac{\partial}{\partial t_\chi}\phi(\cdots) = \log\Big|\frac{t_\chi^2 y_3}{y_1(r^2 - t_\chi^2)}\Big|, \qquad \frac{\partial}{\partial r}\phi(\cdots) = \log\frac{r - t_\chi}{r + t_\chi},\\ &\frac{\partial}{\partial y_1}\phi(\cdots) = -\frac{t_\chi}{y_1}, \qquad \frac{\partial}{\partial y_3}\phi(\cdots) = \frac{t_\chi}{y_3},\\ &\frac{\partial^2}{\partial t_\chi^2}\phi(\cdots) = \frac{2r^2}{t_\chi(r^2 - t_\chi^2)}, \qquad \frac{\partial^2}{\partial r^2}\phi(\cdots) = \frac{2t_\chi}{r^2 - t_\chi^2},\\ &\frac{\partial^2}{\partial r \partial t_\chi}\phi(\cdots) = -\frac{2r}{r^2 - t_\chi^2}.\end{aligned}$$

Using (4.1), we deduce that

$$|t_\chi^*| = r|y_1/y_2|^{1/2}.$$

The integral (4.14) is negligible by an application of Lemma 3 with $U = Y = P = Z$, $S = 1$ unless $\operatorname{sgn}(y_1) = -\operatorname{sgn}(y_2)$ and $|y_2| \asymp Q/Z^2$. In this case, we apply Lemma 4 with

$$(X_1, X_2, X_3, X_4) = (Z, Q^{1/2}, 1, Q/Z^2), \quad Y = Z.$$

We compute

$$\phi\Big(r\Big|\frac{y_1}{y_2}\Big|^{1/2}; r, y_1, y_3\Big) = 2r\Psi\Big(\frac{y_1}{y_2}\Big),$$
$$\Psi(y) = |y|^{1/2}\operatorname{arctanh}\big(|y|\big) - \operatorname{arctanh}\big(|y|^{1/2}\big),$$

where

$$\Psi(y) = -|y|^{1/2} + O\big(|y|^{3/2}\big), \qquad \Psi^j(y) \ll |y|^{1/2-j}. \tag{4.15}$$

The analysis in the discrete series case is similar. We write $\kappa = (k-1)/2$ and apply Lemma 4 and (2.21) with

$$\phi(t_\chi) = \phi(t_\chi; \kappa, y_1, y_3) = t_\chi \log\Big|\frac{t_\chi^2 y_3}{e^2 y_1 \kappa^2}\Big| - 2\kappa \arctan\frac{t_\chi}{\kappa} - t_\chi \log\frac{\kappa^2 + t_\chi^2}{e^2\kappa^2},$$
$$\frac{\partial}{\partial t_\chi}\phi(\cdots) = \log\Big|\frac{t_\chi^2 y_3}{y_1(\kappa^2 + t_\chi^2)}\Big|, \qquad \frac{\partial}{\partial \kappa}\phi(\cdots) = -2\arctan\frac{t_\chi}{\kappa},$$
$$\frac{\partial}{\partial y_1}\phi(\cdots) = -\frac{t_\chi}{y_1},$$
$$\frac{\partial}{\partial y_3}\phi(\cdots) = \frac{t_\chi}{y_3}, \qquad \frac{\partial^2}{\partial t_\chi^2}\phi(t_\chi) = \frac{2\kappa^2}{t_\chi(\kappa^2 + t_\chi^2)}.$$

We have $\phi^{(1,0,0,0)}(t_\chi^*; \kappa, y_1, y_3) = 0$ if and only if $\operatorname{sgn}(y_1) = \operatorname{sgn}(y_3)$ and

$$|t_\chi^*| = \kappa\sqrt{\frac{y_1}{y_3 - y_1}}.$$

Again this is in the support of the integrand if and only if $|y_2| \asymp Q/Z^2$, otherwise the integral is negligible. We compute

$$\phi\Big(\kappa\sqrt{\frac{y_1}{y_3 - y_1}}; \kappa, y_1, y_3\Big) = -2\kappa\arctan\sqrt{\frac{y_1}{y_3 - y_1}} = 2\kappa\tilde{\Psi}\Big(\frac{y_1}{y_2}\Big),$$
$$\tilde{\Psi}(y) = -\arctan\Big(\Big(\frac{1}{|x|} + 2\Big)^{-1/2}\Big),$$

where $\tilde{\Psi}$ satisfies the same formulas as in (4.15).

By the definition of the conductor in (2.20) and (2.21), we conclude that (4.14) equals (up to a negligible error)

$$Z^{-3/2}V\Big(-y_1,\frac{y_2Z^2}{Q}\Big)e\Big(-2\sqrt{Q}\Psi\Big(\frac{y_1}{y_2}\Big)\Big)$$

in the principal series case and

$$Z^{-3/2}V\Big(-y_1,\frac{y_2Z^2}{Q}\Big)e\Big(-2\sqrt{Q}\tilde{\Psi}\Big(\frac{y_1}{y_2}\Big)\Big)$$

in the discrete series case.

We made in the beginning the assumption $z>0$. The case $z<0$ leads to an analogous expression, with the minus sign in the exponential removed. This completes the proof. □

5. A shifted convolution problem

5.1. Some preparation

5.1.1. Bessel functions

We need the following uniform asymptotic formulas for Bessel functions. For $t\in\mathbb{R}$, $|t|>1$ and $x>0$, we have (see [14, 7.13.2(17)])

$$\frac{J_{2it}(2x)}{\cosh(\pi t)}=\sum_{\pm}e^{\pm 2i\omega(x,t)}\frac{f_N^{\pm}(x,t)}{x^{1/2}+|t|^{1/2}}+O_N\big((x+|t|)^{-N}\big),$$

$$\omega(x,t)=|t|\cdot\operatorname{arcsinh}\frac{|t|}{x}-\sqrt{t^2+x^2}, \tag{5.1}$$

where for any fixed $N>0$ the function $f_N^{\pm}$ is flat. The error term estimate stated in [14] is $O(x^{-N})$, but for $x\le t^{1/3}$, say, the estimate $O(|t|^{-N})$ follows from the power series expansion [18, 8.402] of $J_{2it}(x)$. When we apply this formula in practice, we first extract the negligible error $O_N((x+|t|)^{-N})$ in the series expansion given by [14, 7.13.2(17)] and [18, 8.402], without pausing to estimate any derivatives of that error. We then differentiate the remaining series expansion to verify the flatness condition.

For future reference, we note the identities

$$\begin{aligned}\frac{\partial}{\partial t}\omega(x,t)&=\operatorname{arcsinh}\frac{t}{x},&\qquad\frac{\partial^2}{\partial t^2}\omega(x,t)&=\frac{1}{x^2+t^2},\\ \frac{\partial}{\partial x}\omega(x,t)&=-\frac{\sqrt{x^2+t^2}}{x},&\qquad\omega(x,t)&=-x+\frac{t^2}{2x}+O\Big(\frac{t^4}{x^3}\Big).\end{aligned} \tag{5.2}$$

For $|t|\ge 2x>0$, we have by [14, 7.13.2(19)] (again coupled with the power series expansion [18, 8.445, 8.485] for $x\le t^{1/3}$) or [2, (20) with $z\le 1/2$]

$$K_{2it}(2x)\cosh(\pi t) = |t|^{-1/2}\sum_{\pm} e^{\pm 2i\omega^*(x,t)} g_N^{\pm}(x,t) + O_N\big((x+|t|)^{-N}\big),$$
$$\omega^*(x,t) = |t|\cdot \operatorname{arccosh}\frac{|t|}{x} - \sqrt{t^2-x^2}, \tag{5.3}$$

where for any fixed $N>0$ the function $g_N^{\pm}$ is flat.

Similarly, for $x \geq 2k$ we have (see [41, (4.24)])

$$J_{2k-1}(2x) = \sum_{\pm} e^{\pm 2ik\tilde{\omega}(x,k)} \frac{h_N^{\pm}(x,k)}{x^{1/2}} + O_N\big((x+k)^{-N}\big),$$
$$\tilde{\omega}(x,k) = -k\arctan\sqrt{\frac{x^2}{k^2}-1} + \sqrt{x^2-k^2}, \tag{5.4}$$

where $h_N^{\pm}$ is flat. For fixed index, these formulas simplify greatly, and we have

$$\frac{J_{2it}(2x)}{\cosh(\pi t)} = \frac{1}{x^{1/2}+x^{2|\Im t|}}\sum_{\pm} e^{\pm 2ix} f^{\pm}(x),$$
$$J_{2k-1}(2x) = \frac{1}{x^{1/2}+1}\sum_{\pm} e^{\pm 2ix} h^{\pm}(x), \tag{5.5}$$
$$K_{2it}(2x)\cosh(\pi t) = \frac{e^{-2x}}{x^{1/2}+x^{2|\Im t|}} g(x)$$

for $t\in\mathbb{C}, k\in\mathbb{N}, x>0$ with $|\Im t|\leq 1/4$ (for simplicity) and t,k in a fixed compact set, where $f^{\pm}$, g, $h^{\pm}$ can be chosen to be flat (depending on t or k). (See [8, Lemma 15] for details on how to glue together the asymptotic formulas for $x>1$ and $x<1$.)

Remark

As the referee remarked, [14] contains no proofs. The expansions (5.1), (5.3), and (5.4) are all relatively simple to obtain, since we are in the so-called *oscillatory range* away from possible degenerate points ($t=x$ for the K-function and $k=x$ for the J-function with real order). All three uniform asymptotic expansions can be obtained from the integral representations (see [18, 8.421.1/2, 8.405], [18, 8.432.4], [18, 8.411])

$$K_{it}(x) = \frac{1}{2\cosh(\pi t/2)} \int_{-\infty}^{\infty} \cos(x \sinh v) \exp(itv)\, dv,$$

$$J_{it}(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \big(\cosh(\pi t/2) \sin(x \cosh v) - i \sinh(\pi t/2) \cos(x \cosh v)\big) \times \exp(itv)\, dv,$$

$$J_k(x) = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \exp(-ki\theta + ix \sin\theta)\, d\theta$$

by an application of Lemma 4. For the improper integrals, note that the tail can be estimated by partial integration using Lemma 3 (cf. e.g., [5, Section 4.4]).

5.1.2. Jutila's circle method

We quote Jutila's circle method (see [28]).

LEMMA 13

Let $Q \geq 1$, and let V be a smooth, nonnegative, nonzero function with support in $[1,2]$. For $r \in \mathbb{Q}$, write $I_r(\alpha)$ for the characteristic function of the interval $[r - 1/Q, r + 1/Q]$, and define

$$\Lambda := \sum_q V\Big(\frac{q}{Q}\Big)\phi(q), \qquad I(\alpha) = \frac{Q}{2\Lambda} \sum_q V\Big(\frac{q}{Q}\Big) \sum_{\substack{d \pmod q \\ (d,q)=1}} I_{d/q}(\alpha).$$

Then $I(\alpha)$ is a good approximation to the characteristic function on $[0,1]$ in the sense that

$$\int_0^1 \big(1 - I(\alpha)\big)^2\, d\alpha \ll_\varepsilon Q^{\varepsilon-1}$$

for any $\varepsilon > 0$.

5.2. Notation and setup

Let T be a large positive real number. We recall once again the notation and conventions from Section 2.4, in particular with respect to weight functions V. Moreover, $A \preccurlyeq B$ denotes $A \ll T^\varepsilon B$. We consider two more parameters M and H satisfying

$$M \preccurlyeq T^2, \qquad H = T^{1/3+\varepsilon}, \tag{5.6}$$

and $\nu \in \mathbb{Z} \setminus \{0\}$ with $\nu \preccurlyeq 1$. The choice of H will eventually turn out to be the optimal choice, and it simplifies the argument if we make it right away at this point. With slightly more extra work, we could run the same argument for $T^{1/3} \ll H \ll T^{1-\varepsilon}$.

Let g be a fixed holomorphic or Maass Hecke eigenform for $\mathrm{SL}_2(\mathbb{Z})$ with Hecke eigenvalues $\lambda_g(n)$ of weight k_g or spectral parameter t_g. Let $\sigma_t(m) = \sum_{ab=n} a^{it} b^{-it}$.

We fix a choice of sign $\pm$. With these notational conventions, we consider

$$\begin{aligned}\mathcal{L} := \frac{1}{TM^{1/2}} \sum_{2\pi T \le t_j \le 2\pi(T+H)} \frac{1}{L(\mathrm{Ad}^2 u_j, 1)} \\ \times \Big| \sum_m V\Big(\frac{m}{M}\Big) \lambda_j(m) \lambda_g(m+\nu) \exp\Big(\pm 2i \frac{t_j \sqrt{|\nu|}}{\sqrt{m}}\Big)\Big|^2 \\ + \frac{1}{TM^{1/2}} \int_{2\pi T}^{2\pi(T+H)} \frac{1}{|\zeta(1+2it)|^2} \\ \times \Big| \sum_m V\Big(\frac{m}{M}\Big) \sigma_t(m) \lambda_g(m+\nu) \exp\Big(\pm 2i \frac{t \sqrt{|\nu|}}{\sqrt{m}}\Big)\Big|^2 \frac{dt}{2\pi}, \end{aligned} \tag{5.7}$$

where u_j runs through Hecke Maass cusp forms for $\mathrm{SL}_2(\mathbb{Z})$ with Hecke eigenvalues $\lambda_j(n)$ and spectral parameter $t_j \in [2\pi T, 2\pi(T+H)]$. The right-hand side of (5.7) is essentially a combination of (1.4) and (1.9); this explains its relevance.

Analogously, we also consider the holomorphic analogue

$$\begin{aligned}\tilde{\mathcal{L}} := \frac{1}{TM^{1/2}} \sum_{\substack{4\pi T \le k \le 4\pi(T+H) \\ k \text{ even}}} \sum_{u_j \in B_k} \frac{1}{L(\mathrm{Ad}^2 u_j, 1)} \\ \times \Big| \sum_m V\Big(\frac{m}{M}\Big) \lambda_j(m) \lambda_g(m+\nu) \exp\Big(\pm 2i \frac{k \sqrt{|\nu|}}{\sqrt{m}}\Big)\Big|^2, \end{aligned} \tag{5.8}$$

where u_j runs over a Hecke eigenbasis B_k of cusp forms of weight k. For simplicity let us assume $\nu > 0$, the other case being essentially identical. The aim of this section is to prove the following theorem.

THEOREM 5
Let T be a large parameter, and let $\mathcal{L}$, $\tilde{\mathcal{L}}$ be defined as in (5.7) and (5.8) with M, H as in (5.6). Then $\mathcal{L}, \tilde{\mathcal{L}} \preccurlyeq TH$.

The proof of the theorem follows the steps outlined in Section 1.4. In particular, we will eventually transform the spectral sum (5.7) into a "reciprocal" spectral sum of length T/H in Section 5.9 to which we apply the large sieve.

For $M \le T^{2/3+\varepsilon}$, we can estimate trivially (using a standard Rankin–Selberg bound)

$$\mathcal{L} \preccurlyeq \frac{1}{TM^{1/2}} THM^2 = HM^{3/2} \preccurlyeq TH$$

in agreement with Theorem 5. From now on, we assume that $M \ge T^{2/3+\varepsilon}$. Then

$$\exp\Big(\pm 2i\frac{(t-2\pi T)\sqrt{\nu}}{\sqrt{m}}\Big)$$

is flat for $t \in [2\pi T, 2\pi(T+H)]$ by our choice of H in (5.6), so we can replace t_j and t with $2\pi T$ in the exponential (using the by now familiar device to separate variables in nice functions after having multiplied by a suitable function with compact support in $2\pi T + O(H)$). We restrict to the positive sign in the exponential, the negative sign being essentially identical.

We can majorize the characteristic function on $[2\pi T, 2\pi(T+H)]$ by $h(t) = \exp(-(\frac{t-2\pi T}{H})^2)$ and then symmetrize with respect to $t \mapsto -t$ in order to make the expression amenable for the Kuznetsov formula. The exact shape of the function plays no role.

5.3. Application of the spectral summation formula

We open the square in (5.7) and (5.8) and apply the Kuznetsov formula from [23, Theorem 16.3] (along with a conversion from Fourier coefficients to Hecke eigenvalues). Since $\int h(t)t\tanh(\pi t)\,dt \ll TH$ by our choice of h, the diagonal term is bounded by

$$\preccurlyeq \frac{1}{TM^{1/2}}TH\cdot M = M^{1/2}H \preccurlyeq TH$$

by (5.6) in agreement with Theorem 5. For the off-diagonal term, we must understand the Bessel transform

$$\int_{t\in\mathbb{R}} \exp\Big(-\Big(\frac{t-2\pi T}{H}\Big)^2\Big)\frac{J_{2it}(x)}{\cosh(\pi t)}t\,dt. \tag{5.9}$$

As usual, we use holomorphicity to shift the contour a bit; in this way, we can truncate the c-sum in (5.12) by some large power of T (cf. [29, p. 75]). Having this done, we may smoothly truncate the integral to the interval $[2\pi T - HT^{\varepsilon}, 2\pi T + HT^{\varepsilon}]$. For $x \le T^{1-\varepsilon}H$, we apply Lemma 3 and the uniform asymptotic formula (5.1) (along with (5.2)) with

$$S = \min(1, T/x), \qquad Q = Y = T + x, \qquad U = H$$

to see that the integral is negligible. Having recorded the condition

$$x \ge T^{1-\varepsilon}H, \tag{5.10}$$

we do not exploit any further cancellation in the integral. Using (5.1) along with a Taylor expansion, we have, for $t \in [2\pi T - HT^{\varepsilon}, 2\pi T + HT^{\varepsilon}]$ and $x \ge T^{1-\varepsilon}H$, the approximation

$$\frac{J_{2it}(2x)}{\cosh(\pi t)} = x^{-1/2}\sum_{\pm} e^{\pm 2i(-x+\frac{1}{2}(2\pi T)^2/x)}F^{\pm}(x,t) + O\big(|t|^{-N}\big) \tag{5.11}$$

for a flat function $F^{\pm}$. We substitute this into (5.9) and integrate trivially over t. Thus it suffices to estimate the off-diagonal term

$$\frac{TH}{TM^{1/2}}\sum_{m_1,m_2} V\Big(\frac{m_1}{M},\frac{m_2}{M}\Big)\sum_c V\Big(\frac{c}{C}\Big)\frac{S(m_1,m_2,c)}{C}\lambda_g(m_1+\nu)\lambda_g(m_2+\nu)$$
$$\times\frac{C^{1/2}}{M^{1/2}}e\Big(\pm\Big(\frac{2\sqrt{m_1m_2}}{c}-\frac{T^2c}{\sqrt{m_1m_2}}\Big)\pm\frac{2T\nu^{1/2}(\sqrt{m_2}-\sqrt{m_1})}{\sqrt{m_1m_2}}\Big) \quad (5.12)$$

by $\preccurlyeq TH$ (with all sign combinations), as required in Theorem 5, where C runs through $\preccurlyeq 1$ numbers (e.g., powers of 2) satisfying

$$1\le C\preccurlyeq\frac{M}{TH} \quad (5.13)$$

and each weight function V is nice.

We pause for a moment and consider the average $\tilde{\mathcal{L}}$ over holomorphic forms, in which we replace the Kuznetsov formula with the Petersson formula (see [23, Theorem 14.5]). Here the analogue of (5.9) is

$$\sum_{k\in 2\mathbb{Z}} i^k V\Big(\frac{k-4\pi T}{H}\Big)kJ_{k-1}(2\pi x).$$

Using the Fourier representation (see [18, 8.411.1]) of the Bessel function coupled with Poisson summation, this equals

$$\sum_{k\in 2\mathbb{Z}} i^k V\Big(\frac{k-4\pi T}{H}\Big)k\int_{-1/2}^{1/2} e\big((1-k)\theta+x\sin 2\pi\theta\big)\,d\theta$$
$$=\frac{1}{2}\sum_{\substack{h\in\mathbb{Z}\\ h\text{ odd}}}\int_{-1/2}^{1/2}\int_{-\infty}^{\infty} V\Big(\frac{y-4\pi T}{H}\Big)ye\Big(\frac{yh}{4}+(1-y)\theta+x\sin 2\pi\theta\Big)\,dy\,d\theta.$$

The y-integral is negligible unless $h=\pm 1$ and $\theta=\pm 1/4+O(T^{\varepsilon}H^{-1})$, but then the remaining θ-integral, smoothly truncated to the latter range, is negligible unless $x\ge T^{1-\varepsilon}H$. This last condition is the analogue of (5.10). The analogue of (5.11), derived from (5.4), is

$$i^kJ_{2k}(x)=x^{-1/2}\sum_{\pm} e^{\pm 2i(x+\frac{1}{2}(2\pi T)^2/x)}\tilde{F}^{\pm}(x,k)+O(k^{-N})$$

for the present range of variables which leads to the same expression as (5.12) except for a sign in the exponential which will not play a role later.

We now continue with the analysis of (5.12). In preparation for an application of Voronoi summation, we shift the variables m_1, m_2 by ν. By a Taylor expansion, this makes no difference in the exponential, the resulting correction term being flat. It

therefore suffices to bound

$$\frac{H}{M}\sum_{m_1,m_2} V\Big(\frac{m_1}{M},\frac{m_2}{M}\Big)\sum_c V\Big(\frac{c}{C}\Big)\frac{S(m_1-\nu,m_2-\nu,c)}{\sqrt{C}}\lambda_g(m_1)\lambda_g(m_2)$$
$$\times e\Big(\pm\Big(\frac{2\sqrt{m_1m_2}}{c}-\frac{T^2c}{\sqrt{m_1m_2}}\Big)\pm\frac{2T\nu^{1/2}(\sqrt{m_2}-\sqrt{m_1})}{\sqrt{m_1m_2}}\Big). \tag{5.14}$$

5.4. Preparatory interlude

We pause to recall the Voronoi formula (see, e.g., [8, Lemmas 25 and 6]): for $c\in\mathbb{N}$ and $(b,c)=1$, we have

$$\sum_n V(n)\lambda_g(n)e\Big(\frac{bn}{c}\Big)=\frac{1}{c}\sum_{\pm}\sum_n V_\pm^\wedge\Big(\frac{n}{c^2}\Big)\lambda_g(n)e\Big(\mp\frac{\bar b n}{c}\Big),$$
$$V_\pm^\wedge(y)=\int_0^\infty V(x)\mathcal{J}^\pm(4\pi\sqrt{xy})\,dx,$$

where $\mathcal{J}^\pm=\mathcal{J}_g^\pm$ is given for g a Maass form of eigenvalue $1/4+t_g^2$ by

$$\mathcal{J}^+(x)=\pi i\,\frac{J_{2it_g}(x)-J_{-2it_g}(x)}{\sinh(\pi t_g)},\qquad \mathcal{J}^-(x)=4\cosh(\pi t_g)K_{2it_g}(x)$$

and for g a holomorphic form of weight k_g by

$$\mathcal{J}^+(x)=2\pi i^{k_g}J_{k_g-1}(x),\qquad \mathcal{J}^-(x)=0.$$

In the following sections, we will twice have to compute integrals of the form

$$\int_{x\in\mathbb{R}} V\Big(\frac{x}{M}\Big)e(\alpha x^{1/2}+\beta x^{-1/2})\,dx \tag{5.15}$$

for certain $\alpha,\beta\in\mathbb{R}$ satisfying

$$|\alpha|M^{1/2}+|\beta|M^{-1/2}\gg M^\varepsilon. \tag{5.16}$$

In this case, it follows from Lemma 3 with

$$U=M^{1-\varepsilon/2},\qquad P=M,\qquad S=|\alpha|M^{-1/2}+|\beta|M^{-3/2},\qquad Y=MS$$

that (5.15) is negligible unless $\beta/\alpha\asymp M$ (which, by the conventions of Section 2.4, implies in particular that $\operatorname{sgn}(\alpha)=\operatorname{sgn}(\beta)$) in which case by Lemma 4 (after restricting to dyadic ranges $\alpha\asymp A$ and $\beta\asymp B$ and also possibly restricting the support of V to a neighborhood of t^*) with

$$\phi(t) = \phi(t;\alpha,\beta) = \alpha t^{1/2} + \beta t^{-1/2}, \qquad t^* = \frac{\beta}{\alpha},$$
$$\frac{\partial}{\partial t}\phi(t^*;\alpha,\beta) = \frac{\beta}{2t_0^{5/2}} \asymp \frac{\beta}{M^{5/2}},$$

it equals

$$\frac{M^{5/4}}{|\beta|^{1/2}} V\Big(\frac{\alpha}{A}\Big) V\Big(\frac{\beta}{B}\Big) V\Big(\frac{\beta/\alpha}{M}\Big) e\big(2\operatorname{sgn}(\alpha)\sqrt{\alpha\beta}\big), \tag{5.17}$$

as usual with different functions V, and up to a negligible error.

5.5. Application of Voronoi summation

We open the Kloosterman sum in (5.14) and apply the Voronoi formula to the following m_2-sum:

$$\sum_{m_2} V\Big(\frac{m_2}{M}\Big)\lambda_g(m_2) e\Big(\frac{\bar{d}m_2}{c}\Big) e\Big(\pm\Big(\frac{2\sqrt{m_1m_2}}{c} - \frac{T^2c}{\sqrt{m_1m_2}}\Big) \mp \frac{2Tv^{1/2}}{\sqrt{m_2}}\Big),$$

where $(d,c) = 1$. We analyze the integral transforms using (5.4) and (5.5). In the Maass case, we see that the $\mathcal{J}^-$-term is negligible thanks to the rapid decay of the Bessel K-function and the consequence $M/C^2 \gtrsim H^2 \geq T^{2/3+\varepsilon}$ of our hypotheses (5.6) and (5.13). In either case, the $\mathcal{J}^+$-term contributes

$$\sum_{m_2} \lambda_g(m_2) e\Big(\frac{-dm_2}{c}\Big) \frac{1}{c} \int_{x\in\mathbb{R}} V\Big(\frac{x}{M}\Big) \Big(\Big(\frac{\sqrt{m_2x}}{c}\Big)^{1/2} + \Big(\frac{\sqrt{m_2x}}{c}\Big)^{2|\Im t_g|}\Big)^{-1}$$
$$\times e\Big(\sigma_1\Big(\frac{2\sqrt{m_1x}}{c} - \frac{T^2c}{\sqrt{m_1x}}\Big) + \sigma_2\frac{2Tv^{1/2}}{\sqrt{x}}\Big) e\Big(\sigma_3\frac{2\sqrt{xm_2}}{c}\Big)\, dx \tag{5.18}$$

with $\sigma_1,\sigma_2,\sigma_3 \in \{\pm 1\}$, and where as usual the meaning of V may have changed. The x-integral is of the shape (5.15) with

$$\alpha = \frac{2(\sigma_1\sqrt{m_1} + \sigma_3\sqrt{m_2})}{c}, \qquad \beta = -\sigma_1\frac{T^2c}{\sqrt{m_1}} + \sigma_2 2Tv^{1/2}.$$

For the analysis of this integral, it is important to note that

$$\frac{\sqrt{m_1x}}{c} \asymp \frac{M}{C}$$

is by at least a factor $H^2T^{-\varepsilon}$ larger than

$$\frac{T^2c}{\sqrt{m_1x}} \asymp \frac{T^2C}{M}$$

by (5.6) and (5.13), and the latter is larger than $Tv^{1/2}x^{-1/2} \preccurlyeq TM^{-1/2}$ unless $C \preccurlyeq 1$ and $T^2 \preccurlyeq M$. Let us define

$$R := T^2C^2/M. \tag{5.19}$$

If $R \geq T^{\varepsilon}$, then (5.16) is satisfied, and we conclude from the discussion in the previous subsection that the x-integral is negligible unless $\sigma_1 = -\sigma_3$ and[4]

$$m_2 - m_1 \asymp R.$$

If $R \preccurlyeq 1$, then the same argument still shows that $m_2 - m_1 \preccurlyeq R \preccurlyeq 1$ (which allows in particular $m_1 = m_2$). In particular, $\sqrt{m_2x}/c \asymp M/C \gg T^{1-\varepsilon}H \gg 1$ is large and therefore we can drop the term containing $|\Im t_g|$ in (5.18).

Before we proceed, we estimate the total contribution of $R \preccurlyeq 1$ in (5.18) when substituted into (5.14) trivially by

$$\preccurlyeq \frac{H}{M} \cdot M \cdot M^{1/2} = HM^{1/2} \preccurlyeq HT,$$

which is acceptable for Theorem 5. So from now on we assume that

$$R \geq T^{\varepsilon}. \tag{5.20}$$

In view of (5.6), we can then also assume that

$$T^2C/M \geq T^{\varepsilon}. \tag{5.21}$$

For such R and $m_2 - m_1 \asymp R$ and $\sigma_1 = -\sigma_3$, we see from (5.17) that (5.18) can be recast as

$$\frac{M/T}{c} \sum_{m_2} V\Big(\frac{m_2 - m_1}{R}\Big)\lambda_g(m_2)e\Big(-\frac{dm_2}{c}\Big)$$
$$\times e\Big(\pm\sqrt{8(\sqrt{m_2} - \sqrt{m_1})\Big|\frac{T^2}{\sqrt{m_1}} \pm \frac{2Tv^{1/2}}{c}\Big|}\Big),$$

up to a negligible error and for suitable sign combinations. Plugging back into (5.14) and calling $r = m_2 - m_1$, $m = m_1$, we obtain a total contribution

[4]This implies in particular that $m_2 - m_1 > 0$. In the holomorphic average $\tilde{\mathcal{L}}$, the term $-T^2c/\sqrt{m_1x}$ would not have a minus sign, so that here $m_2 - m_1 < 0$. This sign is responsible for the choice of the integral transform in the final application of the Kuznetsov formula in Section 5.9. We mention this only for the sake of clarity. In the following, all sign combinations are treated uniformly.

$$\frac{H}{TC^{1/2}}\sum_r\sum_m\sum_c V\Big(\frac{r}{R},\frac{m}{M},\frac{c}{C}\Big)\frac{S(-r-\nu,-\nu,c)}{c}\lambda_g(m)\lambda_g(m+r)$$
$$\times e\Big(\pm\sqrt{8(\sqrt{m+r}-\sqrt{m})\Big|\frac{T^2}{\sqrt{m}}\pm\frac{2T\nu^{1/2}}{c}\Big|}\pm\frac{2T\nu^{1/2}}{\sqrt{m}}\Big).$$

For notational simplicity, the previous display deals only with the case $r>0$. The case $r<0$, coming from the holomorphic average $\tilde{\mathcal{L}}$ can be treated in the same way. We simplify the exponential a bit using suitable Taylor expansions. First, using the expansion $\sqrt{\sqrt{1+2x}-1}=x^{1/2}-x^{3/2}/4+\cdots$, we replace $\sqrt{|\sqrt{m+r}-\sqrt{m}|}$ with $(r/2)^{1/2}m^{-1/4}$, the error being flat since, by (5.6) and (5.13),

$$\frac{1}{M^{1/4}}\cdot\frac{R^{3/2}}{M^{3/2}}\cdot\frac{T}{M^{1/4}}=\frac{T^4C^3}{M^3}\preccurlyeq\frac{T}{H^3}\preccurlyeq 1.$$

Similarly, we can replace

$$\Big|\frac{T^2}{\sqrt{m}}\pm\frac{2T\nu^{1/2}}{c}\Big|^{1/2}\quad\text{with}\quad\Big|\frac{T}{m^{1/4}}\pm\frac{m^{1/4}\nu^{1/2}}{c}\Big|$$

up to a flat function. Thus it suffices to bound (with various sign combinations)

$$\frac{H}{TC^{1/2}}\sum_r\sum_m\sum_c V\Big(\frac{r}{R},\frac{m}{M},\frac{c}{C}\Big)\frac{S(-r-\nu,-\nu,c)}{c}$$
$$\times\lambda_g(m)\lambda_g(m+r)e\Big(\pm\frac{2Tr^{1/2}}{m^{1/2}}\pm\frac{2(r\nu)^{1/2}}{c}\pm\frac{2T\nu^{1/2}}{m^{1/2}}\Big)$$

by $\preccurlyeq TH$. We write the previous display as

$$\frac{H}{TC^{1/2}}\sum_r F(r)\sum_m V\Big(\frac{r}{R},\frac{m}{M}\Big)\lambda_g(m)\lambda_g(m+r)e\Big(\pm\frac{2T(r^{1/2}\pm\nu^{1/2})}{m^{1/2}}\Big),\quad(5.22)$$

where

$$F(r)=\sum_c V\Big(\frac{c}{C}\Big)\frac{S(-r-\nu,-\nu,c)}{c}e\Big(\pm\frac{2(r\nu)^{1/2}}{c}\Big).$$

5.6. An average of Kloosterman sums

We pause for a moment and prove that

$$\sum_{r\asymp R}\big|F(r)\big|^2\preccurlyeq R.\quad(5.23)$$

It is tempting to use the Kuznetsov formula, but we can argue in an elementary way. We insert a smooth weight and open the square getting

$$\sum_r V\Big(\frac{r}{R}\Big)\big|F(r)\big|^2 = \sum_{c_1,c_2\asymp C}\frac{1}{c_1c_2}V\Big(\frac{c_1}{C},\frac{c_2}{C}\Big)$$
$$\times \sum_{d_1 \,(\mathrm{mod}\, c_1)}^{*}\sum_{d_2 \,(\mathrm{mod}\, c_2)}^{*} e\Big(-\nu\Big(\frac{d_1+\bar{d}_1}{c_1}+\frac{d_2+\bar{d}_2}{c_2}\Big)\Big)$$
$$\times \sum_r V\Big(\frac{r}{R}\Big)e\Big(-\frac{(d_1c_2+d_2c_1)r}{c_1c_2}\Big)e\Big(\pm 2(r\nu)^{1/2}\frac{c_2-c_1}{c_1c_2}\Big).$$

We split the r-sum into residue classes modulo c_1c_2 and apply Poisson summation. The combined conductor of the exponentials is $\preccurlyeq C^2R^{1/2}/C$, and since $R \gg T^{-\varepsilon}C^2R^{1/2}/C$, it is easy to see that the dual sum picks up at most $\preccurlyeq 1$ terms. Hence we obtain the upper bound

$$\sum_{r\asymp R}\big|F(r)\big|^2 \ll R\sum_{h\preccurlyeq 1}\sum_{c_1,c_2\asymp C}\frac{1}{c_1c_2}\sum_{\substack{d_1 \,(\mathrm{mod}\, c_1)}}^{*}\sum_{\substack{d_2 \,(\mathrm{mod}\, c_2)\\ d_1c_2+d_2c_1\equiv h \,(\mathrm{mod}\, c_1c_2)}}^{*} 1.$$

If $h=0$, then the inner double sum vanishes unless $c_1=c_2$. If $h\neq 0$, then the congruence fixes d_j modulo $c_j/(c_j,h)$. In either case, we confirm (5.23).

Remark
It is clear from the proof that the smooth weight function $V(c/C)$ in the definition of $F(r)$ plays no role here and could be replaced with arbitrary bounded weights $\alpha_c \ll 1$ for $c\asymp C$. The only assumption on R and C used in the proof is $C\preccurlyeq R^{1/2}$.

5.7. Application of the circle method
We now return to (5.22) and treat the m-sum as a shifted convolution problem:

$$\sum_{n,m} V\Big(\frac{n}{M},\frac{m}{M}\Big)\lambda_g(n)\lambda_g(m)e\Big(\pm\frac{2T(r^{1/2}\pm\nu^{1/2})}{n^{1/2}}\Big)\int_0^1 e\big((m-n-r)\alpha\big)\,d\alpha.$$

We choose a gigantic parameter $Q=T^{1000}$ and replace the characteristic function on $[0,1]$ with $I(\alpha)$, defined in Lemma 13. By Cauchy–Schwarz and trivial bounds, this introduces an error at most $M^2Q^{\varepsilon-1/2}$ which can be neglected. In this way, the α-integral becomes

$$\frac{Q}{2\Lambda}\sum_q V\Big(\frac{q}{Q}\Big)\sum_{\substack{d \,(\mathrm{mod}\, q)\\ (d,q)=1}}\int_{-1/Q}^{1/Q} e\Big(\Big(\frac{d}{q}+\alpha\Big)(m-n-r)\Big)\,d\alpha.$$

The portion $e(\alpha(m-n-r))$ is obviously flat, so we end up with bounding

$$\frac{1}{\Lambda}\sum_q V\Big(\frac{q}{Q}\Big)\sum_{\substack{d \pmod q\\(d,q)=1}}\sum_{n,m} V\Big(\frac{n}{M},\frac{m}{M}\Big)$$

$$\times \lambda_g(n)\lambda_g(m)e\Big(\pm\frac{2T(r^{1/2}\pm\nu^{1/2})}{n^{1/2}}\Big)e\Big(\frac{d}{q}(m-n-r)\Big), \tag{5.24}$$

where $\Lambda = \sum_q V(q/Q)\phi(q) = Q^{2+o(1)}$. Recall that this represents the m-sum in (5.22).

5.8. Voronoi again

Having separated the variables m, n by the circle method, we apply the Voronoi formula (Section 5.5) to both sums in (5.24). This is simple for the m-sum

$$\sum_m V\Big(\frac{m}{M}\Big)\lambda_g(m)e\Big(\frac{dm}{q}\Big)$$

because it contains no Archimedean oscillation. If g is a Maass form, then we get two terms, one with a Bessel J-function and one with a Bessel K-function, and as before we use (5.5) for the analysis. The dual variable can be truncated at $\preccurlyeq Q^2/M$ at the cost of a negligible error, by the oscillatory behavior of the Bessel J-function (5.1) with $t \ll 1$ and the rapid decay of the Bessel K-function. Using a smooth partition of unity, we obtain $\preccurlyeq 1$ partial sums of the shape

$$\frac{M}{Q}\sum_m V\Big(\frac{m}{M'}\Big)\Big(\frac{M'M}{Q^2}\Big)^{-|\Im t_g|}\lambda_g(m)e\Big(\pm\frac{\bar{d}m}{q}\Big), \quad M' \preccurlyeq Q^2/M. \tag{5.25}$$

The same analysis (slightly simpler) applies if g is holomorphic.

For the n-sum

$$\sum_n V\Big(\frac{n}{M}\Big)\lambda_g(n)e\Big(\pm\frac{2T(r^{1/2}\pm\nu^{1/2})}{n^{1/2}}\Big)e\Big(-\frac{dn}{q}\Big),$$

we argue as follows. We note that $Tr^{1/2}x^{-1/2} \asymp TR^{1/2}M^{-1/2} \geq T^{\varepsilon}$ by (5.6) and (5.20), so there is sizeable oscillation. In particular, we see from the last formula in (5.5) and Lemma 3 with $U = P = M$, $Y = TR^{1/2}M^{-1/2}$, $S = Y/M$ (so that $PS/\sqrt{Y}$ and SU are both $\gg T^{\varepsilon}$) that the Bessel K-term is negligible. For the Bessel J-term, again by (5.5) we need to understand the transform

$$\int_{x\in\mathbb{R}} V\Big(\frac{x}{M}\Big)e\Big(\pm\frac{2T(r^{1/2}\pm\nu^{1/2})}{x^{1/2}}\Big)e\Big(\frac{\pm 2\sqrt{nx}}{q}\Big)$$

$$\times\Big(\Big(\frac{\sqrt{nx}}{q}\Big)^{1/2}+\Big(\frac{\sqrt{nx}}{q}\Big)^{2|\Im t_g|}\Big)^{-1}dx,$$

which is of the shape (5.15) with

$$\alpha = \pm\frac{2\sqrt{n}}{q}, \qquad \beta = \pm 2T(r^{1/2} \pm \nu^{1/2}).$$

The condition (5.16) is satisfied in view of (5.6) and (5.20). From the discussion in Section 5.4 we conclude that the x-integral is negligible unless

$$n \asymp \frac{Q^2T^2R}{M^2},$$

in which case it equals, up to a negligible error,

$$V\Big(\frac{n}{Q^2T^2R/M^2}\Big)e\Big(\pm\frac{4T^{1/2}(r^{1/2}\pm\nu^{1/2})^{1/2}n^{1/4}}{q^{1/2}}\Big)$$
$$\times\Big(\Big(\frac{TR^{1/2}}{M^{1/2}}\Big)^{1/2}+\Big(\frac{TR^{1/2}}{M^{1/2}}\Big)^{2|\Im t_g|}\Big)^{-1}\Big(\frac{TR^{1/2}}{M^{5/2}}\Big)^{-1/2}.$$

Here we can afford to drop the term involving $|\Im t_g|$ and see that the n-sum is of the form

$$\frac{1}{q}\Big(\frac{TR^{1/2}}{M^{3/2}}\Big)^{-1}\sum_n V\Big(\frac{n}{Q^2T^2R/M^2}\Big)\lambda_g(n)e\Big(\frac{\bar{d}n}{q}\Big)$$
$$\times e\Big(\pm\frac{4T^{1/2}(r^{1/2}\pm\nu^{1/2})^{1/2}n^{1/4}}{q^{1/2}}\Big).$$

Substituting this and (5.25) back into (5.24), we can replace (5.24) with terms of the form

$$\frac{M}{Q}\Big(\frac{M'M}{Q^2}\Big)^{-|\Im t_g|}\frac{M^{3/2}}{TR^{1/2}}\frac{1}{Q^2}\sum_m\sum_n V\Big(\frac{m}{M'}\Big)V\Big(\frac{n}{Q^2T^2R/M^2}\Big)\lambda_g(n)\lambda_g(m)$$
$$\times\sum_q V\Big(\frac{q}{Q}\Big)\frac{S(\pm m+n,-r,q)}{q}e\Big(\pm\frac{4T^{1/2}(r^{1/2}\pm\nu)^{1/2}n^{1/4}}{q^{1/2}}\Big). \tag{5.26}$$

We note that n is always substantially bigger than m, since $R \geq T^{\varepsilon}$, so the arguments of the Kloosterman sum never vanish.

5.9. Kuznetsov again

Eventually (5.26) has to be inserted into (5.22), but before we do this we focus on the q-sum which calls for an application of the Kuznetsov formula (see [23, Theorems 16.5 and 16.6]). Before we carry this out, we simplify the exponential a bit. By a Taylor expansion we can replace $n^{1/4}$ with $(n \pm m)^{1/4}$, the error being flat since

$$\frac{T^{1/2}R^{1/2}m}{n^{3/4}Q^{1/2}} \preccurlyeq \frac{T^{1/2}R^{1/2}Q^2M^{3/2}}{MQ^{3/2}T^{3/2}R^{3/4}Q^{1/2}} = \frac{M^{1/2}}{TR^{1/4}} \preccurlyeq \frac{1}{R^{1/4}} \leq 1.$$

We can also replace $(r^{1/2} \pm \nu^{1/2})^{1/2}$ with $r^{1/4} \pm \frac{1}{2}\nu^{1/2}r^{-1/4}$, the total error being flat since

$$\frac{T^{1/2}n^{1/4}}{Q^{1/2}R^{3/4}} \preccurlyeq \frac{T^{1/2}Q^{1/2}T^{1/2}R^{1/4}}{M^{1/2}Q^{1/2}R^{3/4}} = \frac{1}{C} \leq 1.$$

Therefore, the q-sum in (5.26) becomes

$$\sum_q V\Big(\frac{q}{Q}\Big)\frac{S(\pm m+n,-r,q)}{q}e\Big(\pm\frac{4T^{1/2}r^{1/4}(n\pm m)^{1/4}}{q^{1/2}}\Big) \times e\Big(\pm\frac{2(\nu T)^{1/2}(n\pm m)^{1/4}}{r^{1/4}q^{1/2}}\Big).$$

The first exponential fits very well into the shape of the Kuznetsov formula, the second does not. Unfortunately it is not (always) flat, but we can afford to open it by Mellin inversion. We first add a redundant weight function $V(r/R)$ and then write

$$V\Big(\frac{r}{R}\Big)e\Big(\pm\frac{2(\nu T)^{1/2}(n\pm m)^{1/4}}{r^{1/4}q^{1/2}}\Big) = \int_{\Re(s)=0}\Big(\int_{x\in\mathbb{R}_{>0}} V\Big(\frac{x}{R}\Big)e\Big(\pm\frac{2(\nu T)^{1/2}r^{1/4}(n\pm m)^{1/4}}{x^{1/2}q^{1/2}}\Big)x^s\frac{dx}{x}\Big)r^{-s}\frac{ds}{2\pi i}.$$

The outer s-integral can be truncated at

$$\Im s \preccurlyeq \frac{T^{1/2}(Q^2T^2R/M^2)^{1/4}}{R^{1/4}Q^{1/2}} = \frac{T}{M^{1/2}}.$$

We sacrifice all cancellation in the x-, s-integrals and pull them outside of all sums, including the r-, m-, n-sums in (5.26) and (5.22) which are currently not displayed. (We remark that in the "generic" range $M \approx T^2$, we sacrifice nothing here.) The remaining q-sum is of the form

$$\frac{T}{M^{1/2}}\sum_q \frac{S(\pm m+n,-r,q)}{q}\Phi_x\Big(\frac{\sqrt{(n\pm m)r}}{q}\Big), \tag{5.27}$$

where $x \asymp R$ and

$$\Phi_x(z) = V\Big(\frac{z}{Z}\Big)e\Big(\pm 4(Tz)^{1/2}\Big(1\pm\frac{\nu^{1/2}}{2x^{1/2}}\Big)\Big)$$

with

$$Z = \frac{R^{1/2}(Q^2T^2R/M^2)^{1/2}}{Q} = \frac{RT}{M} = \frac{T^3C^2}{M^2} \tag{5.28}$$

by (5.19). It is important that Φ_x does not depend on any of the variables n, m, r, q. We can now apply the Kuznetsov formula to the q-sum. For a quantitative analysis, we need to understand the three integral transforms

$$\check{\Phi}_+(t) = \int_0^\infty \Phi_x(z)\frac{J_{2it}(z) - J_{-2it}(z)}{\sinh(\pi t)}\frac{dz}{z},$$
$$\check{\Phi}_-(t) = \int_0^\infty \Phi_x(z)K_{2it}(z)\cosh(\pi t)\frac{dz}{z},$$
$$\check{\Phi}_{\mathrm{hol}}(k) = \int_0^\infty \Phi_x(z)J_{k-1}(z)\frac{dz}{z}.$$

To this end, it is important to note that Φ_x has sizeable oscillation since

$$T^{1/2}z^{1/2} \asymp T^{1/2}Z^{1/2} = T^2C/M \gg T^\varepsilon$$

by (5.21). We note also, by (5.13) and (5.28), that

$$Z/T = (TC/M)^2 \preccurlyeq 1/H^2, \tag{5.29}$$

so that Z is much smaller than T.

We start with an analysis of $\check{\Phi}_+(t)$, recalling the formula (5.1). First, we observe that it is negligible unless $t \asymp T^{1/2}Z^{1/2}$, otherwise we may apply Lemma 3 with

$$U = ZT^{-\varepsilon}, \qquad P = Z, \qquad Y = \max\bigl(|t|, T^{1/2}Z^{1/2}\bigr), \qquad S = Y/Z.$$

If $t \asymp T^{1/2}Z^{1/2}$, then we can apply Lemma 4 (or in fact the 1-dimensional version of [6, Proposition 8.2]). We will not compute the stationary point x_0 and the shape of the resulting phase (although it can be done algebraically and leads to a quadratic equation with a unique solution if potentially the support of V is slightly restricted), but only bound the size of the integral to be

$$\ll \Bigl(\frac{T^{1/2}}{Z^{3/2}}\Bigr)^{-1/2}\frac{1}{(TZ)^{1/4}}\frac{1}{Z} = \frac{1}{(TZ)^{1/2}}.$$

Here the first factor comes from the stationary phase analysis, the second factor from (5.1) (noting that $Z^{1/2} + |t|^{1/2} \asymp |t|^{1/2} \asymp (TZ)^{1/4}$), and the last factor from the measure dz/z.

For $\check{\Phi}_-(t)$, we can argue in the same way, using (5.3), if $|t| \gg Z$ with a sufficiently large implied constant. For $|t| \ll Z$ with a sufficiently small implied constant,

the Bessel K-function is negligible (cf. e.g., [7, (A.3)]), and for $|t| \asymp Z$ we simply regard the Bessel kernel as part of the weight function and use Lemma 3 with

$$U = \min(ZT^{-\varepsilon}, 1), \qquad P = Z, \qquad Y = T^{1/2}Z^{1/2}, \qquad S = Y/Z$$

(cf. e.g., [7, (A.1), (A.3)] for the relevant bounds for Bessel functions) to show that this contribution is negligible, too.

Similarly, we see that $\check{\Phi}_{\mathrm{hol}}(k)$ is negligible in all cases.

As an aside, we note that this analysis is independent of potential exceptional eigenvalues, whose contribution would always be negligible (because of (5.29)).

Summarizing the previous discussion, we can rewrite (5.27), up to a negligible error, as

$$\frac{T}{M^{1/2}}\frac{1}{TZ}\sum_{t_j \asymp (TZ)^{1/2}} \frac{\rho_j(n \pm m)\rho_j(-r)}{\cosh(\pi t_j)}\Psi(t_j) + \text{continuous spectrum}, \tag{5.30}$$

where Ψ is some function of which we only need to know $\Psi \preccurlyeq 1$, and the sum runs over the Fourier coefficients (in usual normalization) of Hecke Maass cusp forms for $\mathrm{SL}_2(\mathbb{Z})$ with spectral parameter t_j. We do not need to be more precise since the spectral sum (including the continuous contribution) will disappear in a moment when we apply the Cauchy–Schwarz inequality and the large sieve.

5.10. Cauchy–Schwarz and the large sieve

We recall that the previous display represents the q-sum in (5.26) which itself is the m-sum in (5.22). Applying the Cauchy–Schwarz inequality, we deduce that the total contribution to $\mathscr{L}$ and $\tilde{\mathscr{L}}$ is $\preccurlyeq \Delta\sqrt{\Sigma_1\Sigma_2}$, where

$$\Delta := \frac{H}{TC^{1/2}}\frac{M}{Q}\Big(\frac{M'M}{Q^2}\Big)^{-|\Im t_g|}\frac{M^{3/2}}{TR^{1/2}}\frac{1}{Q^2}\frac{T}{M^{1/2}}\frac{1}{TZ},$$

$$\Sigma_1 := \sum_{t_j \asymp T^2C/M} \frac{1}{\cosh(\pi t_j)}\Big|\sum_r V\Big(\frac{r}{R}\Big)F(r)\rho_j(r)\Big|^2 + (\cdots),$$

$$\Sigma_2 := \sum_{t_j \asymp T^2C/M} \frac{1}{\cosh(\pi t_j)}\Big|\sum_s \rho_j(s)G(s)\Big|^2 + (\cdots),$$

where $(\cdots)$ denotes the continuous spectrum contribution and

$$G(s) := \sum_{n \pm m = s} V\Big(\frac{m}{M'}\Big)V\Big(\frac{M^2 n}{Q^2T^2R}\Big)\lambda_g(n)\lambda_g(m).$$

We prepare for the large sieve by estimating the 2-norm

$$\sum_s |G(s)|^2 = \sum_{n_1 \pm m_1 = n_2 \pm m_2} V\Big(\frac{m_1}{M'}, \frac{m_2}{M'}\Big) \times V\Big(\frac{M^2 n_1}{Q^2 T^2 R}, \frac{M^2 n_2}{Q^2 T^2 R}\Big) \lambda_g(n_1)\lambda_g(m_1)\lambda_g(n_2)\lambda_g(m_2).$$

We detect the condition $n_1 \pm m_1 = n_2 \pm m_2$ by a Fourier integral $\int_0^1 e((n_1 \pm m_1 - (n_2 \pm m_2))\alpha)\, d\alpha$ and use Wilton's bound to conclude that

$$\sum_s |G(s)|^2 \preccurlyeq M' \frac{Q^2 T^2 R}{M^2}.$$

Now the scene has been prepared for the endgame with the spectral large sieve of Deshouillers and Iwaniec [13]. Using this and recalling that Q is very large, we deduce that $\Delta\sqrt{\Sigma_1 \Sigma_2}$ is

$$\preccurlyeq \frac{H}{TC^{1/2}} \frac{M}{Q} \Big(\frac{M'M}{Q^2}\Big)^{-|\Im t_g|} \frac{M^{3/2}}{TR^{1/2}} \frac{1}{Q^2} \frac{T}{M^{1/2}} \frac{1}{TZ} \times \Big(\Big(\frac{T^4 C^2}{M^2} + R\Big) R\Big)^{1/2} \Big(\frac{Q^2 T^2 R}{M^2} \cdot \frac{Q^2 T^2 R}{M^2} \cdot M'\Big)^{1/2}.$$

Since $|\Im t_g| < 1/2$, this expression is increasing in M', so we can replace M' with its largest value Q^2/M (up to T^ε) (cf. (5.25)), so that we can drop the term $(M'M/Q^2)^{-|\Im t_g|}$. Simplifying and using (5.28), (5.19), (5.13), and (5.6), we obtain the final upper bound

$$\preccurlyeq \frac{HM^3}{Q^3 R^{1/2} C^{3/2} T^3} \cdot \frac{T^2 C}{M} R^{1/2} \cdot \frac{Q^3 T^2 R}{M^{5/2}} = \frac{HTR}{(MC)^{1/2}} = \frac{HT^3 C^{3/2}}{M^{3/2}} \preccurlyeq \frac{T^{3/2}}{H^{1/2}} \preccurlyeq TH.$$

This finishes the proof of Theorem 5.

6. Proof of the main results

We deduce Theorems 1 and 2 by applying the combination of Theorems 3, 4, and 5 to the triple product formula (1.5). Recall the setup and the choice of test vectors from the beginning of Section 4. Ichino's formula in [22] says that the identity (1.5) holds with $\mathcal{L}_\infty$ a constant multiple of the matrix coefficient integral as in Theorem 3. Therefore,

$$\frac{L(1/2, \pi_1 \otimes \pi_2 \otimes \pi_3)}{L(1, \mathrm{Ad}^2 \pi_1) L(1, \mathrm{Ad}^2 \pi_2) L(1, \mathrm{Ad}^2 \pi_3)} \ll Q \Big| \int_{g \in \Gamma \backslash G} v_1 v_2 v_3(g)\, dg \Big|^2$$

with $\Gamma = \mathrm{PGL}_2(\mathbb{Z})$. We now appeal to the following regularized version of (1.11) (cf. [27, (2.10)], [38, Theorem 5.6]). Recall from Section 2.1 our notation for Iwasawa coordinates on G. For a Γ-invariant function Φ on G of rapid decay near the cusps, we have

$$\int_{g\in\Gamma\backslash G} \Phi(g)\,dg = 2\int_{\Re(s)=a} \pi^{-s}\Gamma(s)\zeta(2s)(2s-1) \times \Big(\int_{\Gamma_N\backslash G^+} \Phi\big(n(x)a(y)k(\theta)\big)y^s\,dx\frac{dy}{y^2}\,d\theta\Big)\frac{ds}{2\pi i},$$

where $\Gamma_N := \Gamma\cap N$ denotes the upper triangular unipotent subgroup of Γ, G^+ denotes the positive-determinant subgroup of G, and the parameter $a>1$ is at our disposal. Note that the s-integral is rapidly convergent due to the decay of $\Gamma(s)$. We apply this formula with $\Phi = v_1v_2v_3$, insert the Fourier expansion and integrate over x. This gives

$$\int_{\Gamma_N\backslash G} v_1v_2v_3\big(n(x)a(y)k(\theta)\big)y^s\,dx\frac{dy}{y^2}\,d\theta = \sum_{\substack{n_1,n_2,n_3\neq 0\\ n_1+n_2+n_3=0}} \frac{\lambda_{\pi_1}(n_1)\lambda_{\pi_2}(n_2)\lambda_{\pi_3}(n_3)}{\sqrt{n_1n_2n_3}}\int_0^\infty F(n_1y,n_2y)y^s\frac{dy}{y^2}$$

with F as in (4.2). Shifting the s-contour to the far right, we can restrict the y-integral to $y\gg Q^{-\varepsilon}$, the remaining error being $O(Q^{-N})$. On the other hand, since $n_1\in\mathbb{Z}\setminus\{0\}$, the upper bound for F in Theorem 4 implies that we can also restrict to $y\ll Q^{\varepsilon}$ and $n_1\preccurlyeq 1$ at the cost of a negligible error. We insert the asymptotic formula from Theorem 4. The error terms $\mathcal{E}_2$, $\mathcal{E}_3$ contribute negligibly, while $\mathcal{E}_1$ contributes $\preccurlyeq Q^{-1/2}$. It remains to consider the contribution of $\mathcal{N}$. We smoothly decompose the sum over n_2 into dyadic ranges $n_2\asymp M\preccurlyeq Q$. We focus on the contribution from $M>0$; the case $M<0$ may be treated similarly. We estimate the contribution from $M\preccurlyeq Q^{1/3}$ trivially (using Cauchy–Schwarz and standard Rankin–Selberg bounds) by

$$\sum_{n_1\preccurlyeq 1}\big|\lambda_{\pi_1}(n_1)\big|\sum_{n_2\asymp M}\frac{|\lambda_{\pi_2}(n_2)\lambda_{\pi_3}(n_1+n_2)|}{|n_2|}\Big(\frac{|n_2|}{Q}\Big)^{3/4}\preccurlyeq\frac{M^{3/4}}{Q^{3/4}}\preccurlyeq Q^{-1/2}.$$

For $M\geq Q^{1/3+\varepsilon}$, we insert the full asymptotic formula for $\mathcal{N}$, giving

$$\frac{L(1/2,\pi_1\otimes\pi_2\otimes\pi_3)}{L(1,\mathrm{Ad}^2\pi_1)L(1,\mathrm{Ad}^2\pi_2)L(1,\mathrm{Ad}^2\pi_3)}$$
$$\ll 1+Q\sup_{Q^{1/3+\varepsilon}\le M\ll Q}\sum_{\nu\ll 1}\Big|\sum_m V\Big(\frac{m}{M}\Big)\frac{\lambda_{\pi_2}(m)\lambda_{\pi_3}(m+\nu)}{\sqrt{|m(m+\nu)|}}\Big(\frac{|m|}{Q}\Big)^{3/4}$$
$$\times e\Big(\pm 2\sqrt{Q}\Psi\Big(\frac{\nu}{m}\Big)\Big)\Big|^2$$

for some nice function V (cf. Section 2.4). We implicitly restrict the sum over m to the support of $V(m/M)$; in particular, $m\neq 0,-\nu$. In the given range of M, we can replace $\Psi(y)$ with $|y|^{1/2}$, the error being flat. By the usual procedure (see Section 2.4) of separating variables and changing the weight function, we arrive at the upper bound

$$1+\sup_{Q^{1/3+\varepsilon}\le M\le Q}\frac{1}{(MQ)^{1/2}}\sum_{\nu\ll 1}\Big|\sum_m V\Big(\frac{m}{M}\Big)\lambda_{\pi_2}(m)\lambda_{\pi_3}(m+\nu)e\big(\pm 2\sqrt{Q|\nu/m|}\big)\Big|^2.$$

For π_1, π_2 fixed, we average this over π_3 in a spectral window $T\le\sqrt{Q}\le T+H$ with $H=T^{1/3+\varepsilon}$. From Theorem 5 we obtain the first bounds in Theorems 1 and 2. The second bound in Theorem 1 follows directly by dropping all but one term (using positivity of central triple product values), while the second bound in Theorem 2 follows from a standard argument based on the functional equation (see, e.g., [17, p. 63]).

Acknowledgments. The second author thanks ETH Zürich and Max Planck Institute for Mathematics, where parts of this article were worked out, for providing a perfect research atmosphere. The third author thanks the Institute for Advanced Study for its hospitality. All authors would like to thank the referees for a careful reading of the manuscript.

The first author's work was partially supported by DFG-SNF lead agency program grant BL 915/2-2 as well as Germany's Excellence Strategy grant EXC-2047/1-390685813. The third author's work was partially supported by National Science Foundation (NSF) grant DMS-1926686.

References

[1] R. ACHARYA, P. SHARMA, and S. SINGH, *t-aspect subconvexity for* GL(2) × GL(2) *L-function*, J. Number Theory **240** (2022), no. 1, 296–324. MR 4458242. DOI 10.1016/j.jnt.2022.01.011. *(1176)*

[2] C. B. BALOGH, *Asymptotic expansions of the modified Bessel function of the third kind of imaginary order*, SIAM J. Appl. Math. **15** (1967), no. 5, 1315–1323. MR 0222354. DOI 10.1137/0115114. *(1212)*

[3] J. BERNSTEIN and A. REZNIKOV, *Periods, subconvexity of L-functions and representation theory*, J. Differential Geom. **70** (2005), no. 1, 129–142. MR 2192063. DOI 10.4310/jdg/1143572016. *(1174, 1177)*

[4] ———, *Subconvexity bounds for triple L-functions and representation theory*, Ann. of Math. (2) **172** (2010), no. 3, 1679–1718. MR 2726097. DOI 10.4007/annals.2010.172.1679. *(1174, 1175, 1176, 1177, 1178)*

[5] V. BLOMER and J. BUTTCANE, *On the subconvexity problem for L-functions on* GL(3), Ann. Sci. Éc. Norm. Supér. (4) **53** (2020), no. 6, 1441–1500. MR 4203038. DOI 10.24033/asens.2451. *(1214)*

[6] V. BLOMER, R. KHAN, and M. YOUNG, *Distribution of mass of holomorphic cusp forms*, Duke Math. J. **162** (2013), no. 14, 2609–2644. MR 3127809. DOI 10.1215/00127094-2380967. *(1192, 1226)*

[7] V. BLOMER, X. LI, and S. D. MILLER, *A spectral reciprocity formula and non-vanishing for L-functions on* GL(4) × GL(2), J. Number Theory **205** (2019), 1–43. MR 3996341. DOI 10.1016/j.jnt.2019.05.011. *(1177, 1227)*

[8] V. BLOMER and D. MILIĆEVIĆ, *The second moment of twisted modular L-functions*, Geom. Funct. Anal. **25** (2015), no. 2, 453–516. MR 3334233. DOI 10.1007/s00039-015-0318-7. *(1213, 1218)*

[9] J. BOURGAIN, *Decoupling, exponential sums and the Riemann zeta function*, J. Amer. Math. Soc. **30** (2017), no. 1, 205–224. MR 3556291. DOI 10.1090/jams/860. *(1174)*

[10] D. BUMP, *Automorphic Forms and Representations*, Cambridge Stud. Adv. Math. **55** (1997), Cambridge Univ. Press, Cambridge, 1997. MR 1431508. DOI 10.1017/CBO9780511609572. *(1194, 1195)*

[11] W. CASSELMANN, *On some results of Atkin and Lehner*, Math. Ann. **201** (1973), 301–314. MR 0337789. DOI 10.1007/BF01428197. *(1182, 1183)*

[12] J. W. COGDELL, "Analytic theory of L-functions for GL_n" in *An Introduction to the Langlands Program (Jerusalem, 2001)*, Birkhäuser Boston, Boston, 2003, 197–228. MR 1990380. *(1195, 1204)*

[13] J.-M. DESHOUILLERS and H. IWANIEC, *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. Math. **70** (1982), no. 2, 219–288. MR 0684172. DOI 10.1007/BF01390728. *(1228)*

[14] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER, and F. TRICOMI, *Higher Transcendental Functions, Vols. I, II*, McGraw-Hill, New York, 1953. MR 0058756. *(1212, 1213)*

[15] S. GELBART and H. JACQUET, *A relation between automorphic representations of* GL(2) *and* GL(3), Ann. Sci. Éc. Norm. Supér. (4) **11** (1978), no. 4, 471–542. MR 0533066. *(1196)*

[16] A. GHOSH and P. SARNAK, *Real zeros of holomorphic Hecke cusp forms*, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 2, 465–487. MR 2881302. DOI 10.4171/JEMS/308. *(1174)*

[17] A. GOOD, *The square mean of Dirichlet series associated with cusp forms*, Mathematika **29** (1982), no. 2, 278–295. MR 0696884. DOI 10.1112/S0025579300012377. *(1174, 1181, 1230)*

[18] I. S. GRADSHTEYN and I. M. RYZHIK, *Tables of Integrals, Series, and Products*, 7th ed., Academic Press, New York, 2007. MR 2360010. *(1212, 1213, 1217)*

[19] R. HOLOWINKSY, *A sieve method for shifted convolution sums*, Duke Math. J. **146** (2009), no. 3, 401–448. MR 2484279. DOI 10.1215/00127094-2009-002. *(1182)*

[20] Y. HU, *Triple product formula and the subconvexity bound of triple product L-function in level aspect*, Amer. J. Math. **139** (2017), no. 1, 215–259. MR 3619914. DOI 10.1353/ajm.2017.0004. *(1176)*

[21] P. HUMPHRIES and F. BRUMLEY, *Standard zero-free regions for Rankin-Selberg L-functions via sieve theory*, with appendix "Standard zero-free regions when at least one factor is self-dual" by F. Brumley, Math. Z. **292** (2019), no. 3–4, 1105–1122. MR 3980284. DOI 10.1007/s00209-018-2136-8. *(1196, 1205)*

[22] A. ICHINO, *Trilinear forms and the central values of triple product L-functions*, Duke Math. J. **145** (2008), no. 2, 281–307. MR 2449948. DOI 10.1215/00127094-2008-052. *(1174, 1177, 1228)*

[23] H. IWANIEC and E. KOWALSKI, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. **53** (2004), Amer. Math. Soc., Providence, 2004. MR 2061214. DOI 10.1090/coll/053. *(1185, 1216, 1217, 1224)*

[24] H. IWANIEC and P. SARNAK, *Perspectives on the analytic theory of L-functions*, Geom. Funct. Anal. **2000**, 705–741. MR 1826269. DOI 10.1007/978-3-0346-0425-3_6. *(1185)*

[25] H. JACQUET, I. PIATETSKI-SHAPIRO, and J. SHALIKA, *Conducteur des représentations du groupe linéaire*, Math. Ann. **256** (1981), no. 2, 199–214. MR 0620708. DOI 10.1007/BF01450798. *(1182)*

[26] S. JANA and P. NELSON, *Analytic newvectors for* $\mathrm{GL}_n(\mathbb{R})$, preprint, arXiv:1911.01880v2 [math.NT]. *(1182, 1183, 1199, 1200)*

[27] M. JUTILA, *The additive divisor problem and its analogs for Fourier coefficients of cusp forms, I*, Math. Z. **223** (1996), no. 3, 435–461. MR 1417854. DOI 10.1007/PL00004270. *(1229)*

[28] ———, "A variant of the circle method" in *Sieve Methods, Exponential Sums, and Their Application in Number Theory (Cardiff, 1995)*, London Math. Soc. Lecture Note Ser. **237**, Cambridge Univ. Press, Cambridge, 1997, 245–254. MR 1635766. DOI 10.1017/CBO9780511526091.016. *(1214)*

[29] M. JUTILA and Y. MOTOHASHI, *Uniform bound for Hecke L-functions*, Acta Math. **195** (2005), 61–115. MR 2233686. DOI 10.1007/BF02588051. *(1174, 1216)*

[30] ———, "Uniform bounds for Rankin-Selberg L-functions" in *Multiple Dirichlet Series, Automorphic Forms, and Analytic Number Theory*, Proc. Sympos. Pure Math. **75**, Amer. Math. Soc., Providence, 2006, 243–256. MR 2279941. DOI 10.1090/pspum/075/2279941. *(1174)*

[31] E. M. KIRAL, I. PETROW, and M. P. YOUNG, *Oscillatory integrals with uniformity in parameters*, J. Théor. Nombres Bordeaux **31** (2019), no. 1, 145–159. MR 3994723. *(1192, 1193)*

[32] E. LANDAU, *Über die ζ-Funktion und die L-Funktionen*, Math. Z. **20** (1924), no. 1, 105–125. MR 1544665. DOI 10.1007/BF01188074. *(1174)*

[33] Y.-K. LAU, J. LIU, and Y. YE, *A new bound $k^{2/3+\varepsilon}$ for Rankin-Selberg L-functions for Hecke congruence subgroups*, IMRP Int. Math. Res. Pap. **2006**, art. ID 35090. MR 2235495. *(1174)*

[34] J. E. LITTLEWOOD, "Researches in the theory of the Riemann ζ-function" in *Records of Proceedings at Meetings*, Proc. Lond. Math. Soc. (2) **20** (1922), xxiv. *(1174)*

[35] K. MATOMÄKI, *Real zeros of holomorphic Hecke cusp forms and sieving short intervals*, J. Eur. Math. Soc. (JEMS) **18** (2016), no. 1, 123–146. MR 3438381. DOI 10.4171/JEMS/585. *(1174)*

[36] P. MICHEL and A. VENKATESH, *The subconvexity problem for* GL_2, Publ. Math. Inst. Hautes Études. Sci. **111** (2010), 171–271. MR 2653249. DOI 10.1007/s10240-010-0025-8. *(1177, 1178, 1182, 1183, 1185, 1191, 1197, 1206)*

[37] D. MILIĆEVIĆ, *Sub-Weyl subconvexity for Dirichlet L-functions to prime power moduli*, Compos. Math. **152** (2016), no. 4, 825–875. MR 3484115. DOI 10.1112/S0010437X15007381. *(1174)*

[38] P. D. NELSON, *Evaluating modular forms on Shimura curves*, Math. Comp. **84** (2015), no. 295, 2471–2503. MR 3356036. DOI 10.1090/S0025-5718-2015-02943-3. *(1229)*

[39] ———, *Subconvex equidistribution of cusp forms: Reduction to Eisenstein observables*, Duke Math. J. **168** (2019), no. 9, 1665–1722. MR 3961213. DOI 10.1215/00127094-2019-0005. *(1197, 1198)*

[40] ———, *Eisenstein series and the cubic moment for* PGL_2, preprint, arXiv:1911.06310v3 [math.NT]. *(1174)*

[41] F. W. J. OLVER, *The asymptotic expansion of Bessel functions of large order*, Philos. Trans. Roy. Soc. Lond. Ser. A **247** (1954), no. 930, 328–368. MR 0067250. DOI 10.1098/rsta.1954.0021. *(1213)*

[42] I. PETROW and M. P. YOUNG, *The Weyl bound for Dirichlet L-functions of cube-free conductor*, Ann. of Math. (2) **192** (2020), no. 2, 437–486. MR 4151081. DOI 10.4007/annals.2020.192.2.3. *(1174)*

[43] ———, *The fourth moment of Dirichlet L-functions along a coset and the Weyl bound*, preprint, arXiv:1908.10346v3 [math.NT]. *(1174)*

[44] P. SARNAK, *Integrals of products of eigenfunctions*, Int. Math. Res. Not. IMRN **1994**, no. 6, art. ID 251. MR 1277052. DOI 10.1155/S1073792894000280. *(1175)*

[45] E. SUVITIE, *On inner products involving holomorphic cusp forms and Maass forms*, Šiauliai Math. Semin. **3** (2008), no. 11, 221–233. MR 2543461. *(1175, 1182)*

[46] ———, *On a short spectral sum involving inner products of a holomorphic cusp form and Maass forms*, Acta Arith. **144** (2010), no. 4, 395–418. MR 2684289. DOI 10.4064/aa144-4-5. *(1174, 1181)*

[47] A. VENKATESH, *Sparse equidistribution problems, period bounds and subconvexity*, Ann. of Math. (2) **172** (2010), no. 2, 989–1094. MR 2680486. DOI 10.4007/annals.2010.172.989. *(1176)*

[48] T. C. WATSON, *Rankin triple products and quantum chaos*, Ph.D. dissertation, Princeton University, Princeton, 2002. MR 2703041. *(1175, 1185)*

[49] H. WEYL, *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. **77** (1916), no. 3, 313–352. MR 1511862. DOI 10.1007/BF01475864. *(1174)*

[50] D. ZAGIER, *The Rankin-Selberg method for automorphic functions which are not of rapid decay*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **28** (1981), no. 3, 415–437. MR 0656029. *(1182)*

Blomer

Mathematisches Institut, Universität Bonn, Bonn, Germany; blomer@math.uni-bonn.de

Jana

Queen Mary University of London, London, United Kingdom; subhajit@mpim-bonn.mpg.de

Nelson

Aarhus University, Aarhus, Denmark; nelson.paul.david@gmail.com