THE QUANTUM UNIQUE ERGODICITY CONJECTURE FOR THIN SETS

MATTHEW P. YOUNG

Abstract. We consider some analogs of the quantum unique ergodicity conjecture for geodesics and “shrinking” families of sets. In particular, we prove the analog of the QUE conjecture for Eisenstein series restricted to the infinite geodesic connecting 0 and $\infty$ inside the modular surface.

1. Introduction

The quantum unique ergodicity (QUE) conjecture of Rudnick and Sarnak [RS] is an equidistribution statement for Hecke-Maass forms of large Laplace eigenvalue on the modular surface $\Gamma \setminus \mathbb{H}$, $\Gamma = \text{PSL}_2(\mathbb{Z})$. It says that if $U(z)$ is a Hecke-Maass form and $\phi$ is a fixed, smooth and compactly-supported function on $\Gamma \setminus \mathbb{H}$, then

$$\int_{\Gamma \setminus \mathbb{H}} |U(z)|^2 \phi(z) \frac{3}{\pi} \frac{dx dy}{y^2} \to \int_{\Gamma \setminus \mathbb{H}} \phi(z) \frac{3}{\pi} \frac{dx dy}{y^2},$$

as the Laplace eigenvalue of $U$ tends to infinity, provided $U$ is normalized with probability measure

$$\int_{\Gamma \setminus \mathbb{H}} |U(z)|^2 \frac{3}{\pi} \frac{dx dy}{y^2} = 1.$$

Lindenstrauss [L] recently proved the QUE conjecture in the compact setting (and proved it in the non-compact case except for the possibility of “escape of mass” at the cusp), and Soundararajan [So2] completed the proof in this non-compact case. The mass equidistribution conjecture [RS] is the analog of the QUE conjecture where $U(z) = y^{k/2} f(z)$ for $f$ a weight $k$ holomorphic Hecke cusp form, with $k \to \infty$. It was proven by Holowinsky and Soundararajan [HS].

In this paper, we investigate the possibility of equidistribution of $U(z)$ restricted to some “thin” sets, such as a geodesic or a shrinking family of discs. Since the rate of convergence in (1.1) is either not known (in the Maass case) or rather slow (in the holomorphic case, where the error term is a small power of $\log k$), it seems unlikely using current technology to prove QUE for the restrictions of Hecke cusp forms. Instead, one of our goals has been to find natural and (presumably) reliable conditions that imply QUE in these extreme cases. For Maass forms, questions of this type were raised and studied numerically by Hejhal and Rackner [HR]. Very recently, Ghosh, Reznikov, and Sarnak [GRS] have proven strong upper and lower bounds for the $L^2$ norm of Hecke-Maass cusp forms restricted to geodesics or horocycles, with the application of proving (on the Lindelöf hypothesis) that the number of nodal domains goes to infinity with the Laplace eigenvalue.

This material is based upon work supported by the National Science Foundation under agreement No. DMS-1101261. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

© 2015. This manuscript version is made available under the Elsevier user license
http://www.elsevier.com/open-access/userlicense/1.0/
Quantum ergodicity states that (1.1) holds for a density one subsequence of $U$'s $[Z1]$ $[Z2]$, and has been extended in many different directions. The analog of quantum ergodicity for restricted eigenfunctions has also recently been studied, for instance, see $[DZ]$ $[TZ1]$ $[TZ2]$.

One particularly interesting example, having extra structure, is the vertical geodesic connecting 0 and $i\infty$.

**Conjecture 1.1.** Suppose that $\psi : \mathbb{R}^+ \to \mathbb{R}$ is a smooth, compactly-supported function. Then

$$\lim_{j \to \infty} \int_{0}^{\infty} |u_j(iy)|^2 \psi(y) \frac{dy}{y} = \int_{0}^{\infty} 2\psi(y) \frac{dy}{y},$$

where $u_j$ runs over the even Hecke-Maass forms, normalized by (1.2). Furthermore,

$$\lim_{k \to \infty} \int_{0}^{\infty} y^k |f(iy)|^2 \psi(y) \frac{dy}{y} = \int_{0}^{\infty} \psi(y) \frac{dy}{y},$$

where $f(z)$ runs over weight $k$ holomorphic Hecke cusp forms, $L^2$-normalized with probability measure.

Remarks. Conjecture 1.1 says that QUE should hold for functions restricted to the geodesic joining 0 and $i\infty$. Notice that even Maass forms are predicted to be twice as big on this geodesic as on the fundamental domain. One might naturally speculate that the factor of 2 in (1.3) arises from the fact that the odd Maass forms all vanish along the geodesic and so the even Maass forms have to be twice as large (on average) to account for this disparity. One can see this type of behavior from the Selberg pre-trace formula where $\sum |u_j(z)|^2 h(t_j) \sim \sum h(t_j)$ for $z \in \mathbb{H}$ a fixed non-elliptic point, and for certain classes of weight functions $h$; cf. p.179 of [Iw4]. The point is that this spectral sum includes both the even and odd Maass forms, but if $u_j(z) = 0$ for all the odd forms, say, then $|u_j(z)|^2$ has to be 2 on average over the even forms. In the course of the derivation, the only difference between (1.3) and (1.4) is that the Fourier expansion of a Maass form has Fourier coefficients at both negative and positive integers. We do not have a conjecture for the precise size of the error terms in these two asymptotics, but we expect a power saving based on the integral moment conjectures of Conrey, Farmer, Keating, Rubinstein, and Snaith [CFKRS]. It would be interesting to prove an Omega-type result for the error terms.

Conjecture 1.1 is apparently more difficult than the usual QUE conjecture. Geometrically, since the domain is 1-dimensional (hence, of measure 0), it is not obviously implied by (1.1).

See [GRS] for upper and lower bounds on integrals of the form (1.3); these bounds are close to the right order of magnitude. For the holomorphic case, see [BKY] for an upper bound on (1.4) of the size $k^{1/4+\varepsilon}$ while a bound of the form $k^{\varepsilon}$ would prove that $L(f, 1/2) \ll k^{1/4+\varepsilon}$ which would be stronger than any known subconvexity bound for any $L$-function (here the conductor is $k^2$). J. Jung [Ju] has proven strong lower bounds on the $L^2$ restriction of a Hecke-Maass form restricted to a geodesic segment, for almost all forms $u_j$.

It is interesting to compare the tools needed to derive the usual QUE conjecture with the geodesic version. We briefly sketch a standard way to approach the proof of (1.1); see Section 4.1 for a more in-depth discussion. Using the spectral decomposition, write $\langle |U|^2, \phi \rangle = \langle |U|^2, \frac{3}{\pi} \langle 1, \phi \rangle + \sum_j \langle U, u_j \rangle \langle u_j, \phi \rangle + \ldots$ with the dots indicating the continuous spectrum. The constant eigenfunction provides a main term. An integration by parts argument shows that $|\langle u_j, \phi \rangle| \leq C(A)(1/4 + t_j^2)^{-A}$ where $t_j$ is the spectral parameter of $u_j$, and $A > 0$ arbitrary, so that the spectral sum can be effectively truncated almost instantly. Watson's
THE QUANTUM UNIQUE ERGODICITY CONJECTURE FOR THIN SETS

The quantum unique ergodicity conjecture concerns the distribution of eigenfunctions of the Laplacian on a compact Riemannian manifold. It posits that as the energy of the eigenfunctions increases, they become uniformly distributed across the manifold. This is a non-trivial statement, as the eigenfunctions can be highly concentrated near specific points or regions on the manifold.

The conjecture is supported by numerous results and partial results, as well as numerical simulations. It is a central problem in spectral geometry and has implications for number theory, quantum chaos, and other areas of mathematics.

The quantum unique ergodicity conjecture is closely related to the distribution of zeros of L-functions, which are central objects in analytic number theory. The conjecture predicts that the eigenfunctions are equidistributed with respect to the measure defined by the L-functions.

As an example, consider the equation
\[ \sum_j \langle |U|^2, u_j \rangle \int_0^\infty u_j(iy) \psi(y) \frac{dy}{y} + \ldots, \]

where \([W] \) relates \(|\langle |U|^2, u_j \rangle|^2\) to the central value of a triple product L-function. The convexity bound for the L-function then barely fails to prove (1.1) and a subconvexity bound would succeed. The Lindelöf hypothesis is known to give an optimal error term by work of Luo and Sarnak [LS1] [LS3]. It is a common theme, nicely illustrated in this example, that subconvexity and equidistribution are often closely related (almost equivalent).

As in the sketch of the proof of QUE outlined above, one is naturally inclined to apply the spectral decomposition to approach Conjecture 1.1, writing
\[ \int_0^\infty |U(iy)|^2 \psi(y) \frac{dy}{y} = \frac{3}{\pi} \int_0^\infty \psi(y) \frac{dy}{y} + \sum_j \langle |U|^2, u_j \rangle \int_0^\infty u_j(iy) \psi(y) \frac{dy}{y} + \ldots, \]

but this is not effective. The problem can already be anticipated in (1.3) where the constant eigenfunction only gives half the main term so one must expect to extract a main term from the spectral sum. The other issue is that the spectral sum is very “long” since the integrals \(\int_0^\infty u_j(iy) \psi(y) \frac{dy}{y}\) do not immediately rapidly decay, and there must be cancellation in the spectral sum. The square of \(\langle |U|^2, u_j \rangle\) is given as a triple product L-function, and the change in sign of these coefficients is difficult to detect. Rather than using the spectral decomposition, one may use Parseval (for \(\mathbb{R}^+\)) to relate the geodesic integral to a second moment of L-functions. Lindelöf here gives an upper bound of the right order of magnitude but the asymptotic is a more subtle issue and is not known to follow from GRH. See [H] following [So1] for the best results on upper bounds on moments assuming GRH. The five authors’ conjecture [CFKRS] predicts an asymptotic for this second moment of L-functions.

We noticed that the main term in this asymptotic could then be expressed as \(\langle |U|^2, H(z, \psi) \rangle\) for some nice \(\Gamma\)-invariant function \(H\) (almost an incomplete Eisenstein series– see (3.10) for the exact formula). Finally, we use the QUE/mass equidistribution theorems to relate this inner product to \(\langle 1, H(z, \psi) \rangle\) which then leads immediately to (1.3) and (1.4).

Our main (unconditional) result in this paper is the following

**Theorem 1.2.** Suppose \(\psi\) is a smooth, compactly-supported function on \(\mathbb{R}^+\). We have
\[ \pi \frac{1}{3 \log(1/4 + T^2)} \int_0^\infty |E(iy, 1/2 + iT)|^2 \psi(y) \frac{dy}{y} \sim \int_0^\infty 2\psi(y) \frac{dy}{y}, \]
as \(T \to \infty\).

Theorem 1.2 is inspired by an analogous result of Luo-Sarnak [LS1] who showed that \(\frac{\pi}{12} \log(1/4 + T^2)\) is the average value of \(|E(z, 1/2 + iT)|^2\) restricted to any fixed compact Jordan measurable subset of \(\Gamma \setminus \mathbb{H}\) having positive measure (here we have corrected the constant in place of \(\frac{48}{5} \log T\) as stated by [LS1]; for a correct statement, see [HR], (7.9) or [Sp], (1.1)). Then we interpret Theorem 1.2 as saying the Eisenstein series (absolute value squared) is twice as big (on average) on any segment of the geodesic as on the rest of the fundamental domain, just as is predicted by Conjecture 1.1 for even Maass forms. In fact, we appeal to the Luo-Sarnak result just as in the derivation of Conjecture 1.1 we appeal to the QUE theorem. Moreover, we prove a more precise version of the asymptotic with a power saving in the error term; see Theorem 5.1 for this result. Since we have a power saving, this indicates that we could probably estimate some variations such as by having \(\psi\) vary with \(T\) in some way–see Theorem 1.4 for an example of what is meant here. See also [Ja] for the equidistribution on \(PSL_2(\mathbb{Z}) \setminus PSL_2(\mathbb{R})\), and [K] for the level aspect of QUE for Eisenstein series.

---

1For odd Maass forms, the constant term gives a main term which must be cancelled by the spectral sum.
The Eisenstein series is often an interesting test case for more advanced problems with Maass forms. For instance, Spinu [Sp] bounded the $L^4$ norm of the Eisenstein series when restricted to a compact set.

Our proof of Theorem 1.2 requires the full spectral theory of automorphic forms. As mentioned earlier in the introduction, it is not effective to directly apply the spectral decomposition to $|E(iy, 1/2 + iT)|^2$ and then integrate over $y$, because the spectral sum is very “long.” Instead, we use Parseval to relate it to the fourth moment of the Riemann zeta function which of course has been extensively studied over many years, e.g. see [In] [R] [H-B] [Iw1] [Mo2] for some notable results. However, the present case is somewhat unusual in that it samples the zeta function at widely-separated points. Here it is roughly of the form

$$\int_{|t| \leq T} (1 + |t - T|)^{-1/2} (1 + |t + T|)^{-1/2} |\zeta(1/2 + it + iT)\zeta(1/2 + it - iT)|^2 dt.$$  

Furthermore, there are ranges in the integral where the conductor drops, e.g. for $T - 2\Delta \leq t \leq T - \Delta$ with $\Delta = o(T)$, in which case the weight becomes larger. Luckily, the measure of the set where the conductor drops is also relatively small, so these effects partially negate each other. In fact, one may use Cauchy’s inequality and short-interval bounds on the fourth moment of zeta (e.g., see (5.53) below) to dispense with the conductor-dropping ranges, so the main issue is to understand the range $|t| \leq (1 - o(1))T$. In a related direction, Bettin [Be] considered large shifts in the second moment of zeta, and Chandee [C] obtained strong bounds on shifted moments assuming GRH.

Our basic approach is to use an approximate functional equation for the product of zeta functions, leading to the problem of asymptotically evaluating the shifted divisor sum $\sum_{n \approx T} \tau_T(n)\tau_{iT}(n + m)$, with a smooth weight. In turn, for this we use spectral theory in the guise of the Kuznetsov formula. It is an important point that we only use the spectral tools on the favorable part of the integral with $|t| \leq (1 - o(1))T$. It might be illuminating to relate this portion of the $t$-integral directly to an inner product of $|E(z, 1/2 + iT)|^2$ with Poincare series. See [G1] [G2] [EHS] for some clues in this direction. One curious feature of the proof is that after applying the spectral theory to the shifted divisor sum, its estimation comes down to subconvexity for the Riemann zeta function and Hecke-Maass $L$-functions.

We are nevertheless reluctant to say that the geodesic QUE theorem for Eisenstein series is equivalent to subconvexity, because in fact we require the full spectral machinery to even get to this point, and because these tools themselves lead to strong subconvexity bounds. Furthermore, to treat the ranges with $\Delta = o(T)$ requires a short interval mean value bound for the zeta function which itself implies subconvexity (but not vice-versa).

Another interesting quantity along the lines of Conjecture 1.1 is the $L^2$-norm of $u_j(iy)$ along the geodesic, that is, the case $\psi(y) = 1$ for all $y$ (a problem considered by [GRS]). Although this problem should be easier than the geodesic integral weighted by $\psi$ (because it is a longer average), Conjecture 1.1 breaks down in this case. In this direction, we have

**Conjecture 1.3.** For $\alpha$ fixed, $|\text{Re}(\alpha)| < 1/2$, and $u_j$ an even Hecke-Maass cusp form, we have

$$\int_0^\infty |u_j(iy)|^2 y^\alpha dy \sim 2 \int_{\Gamma \setminus H} |u_j(z)|^2 (E(z, 1 + \alpha) + E(z, 1 - \alpha)) \frac{dx dy}{y^2},$$

for $0 < \alpha < 1/2$. 


as \( j \to \infty \). Similarly,

\[
(1.9) \quad \int_0^\infty y^k |f(iy)|^2 y e^y \frac{dy}{y} \sim \int_{\Gamma \setminus \mathbb{H}} y^k |f(z)|^2 (E(z, 1 + \alpha) + E(z, 1 - \alpha)) \frac{dxdy}{y^2},
\]
as \( k \to \infty \).

The right hand side of (1.8) has a removable singularity at \( \alpha = 0 \). When \( \alpha = 0 \), \( E(z, 1 + \alpha) + E(z, 1 - \alpha) \) becomes the constant term of \( 2E(z, s) \) in the Laurent expansion around \( s = 1 \) which can then be expressed in terms of \( \log(\sqrt{\eta}(z)) \); see (22.69) of [IK]. See also Remark 6.2 of [GRS] for an alternative formulation of the \( \alpha = 0 \) case. However, they did not provide a derivation of their argument so we briefly include one below with Conjecture 3.1.

A different way to study the QUE conjecture for thin sets is to restrict \( |U|^2 \) to (say) a small disc with fixed center but with radius that shrinks at some rate with \( U \). These questions are examined in Section 4. We were partially motivated to study this question based on an analogous problem for Heegner points of discriminant \( D \) and level \( q \) (prime) which was considered in [LMY]. A simple question asked in [LMY] is, given \( q \), how large does \( D \) have to be to guarantee that a Heegner point of discriminant \( D \) and level \( q \) lies in \( \omega SL_2(\mathbb{Z}) \setminus \mathbb{H} \) for any coset \( \omega SL_2(\mathbb{Z}) \in \Gamma_0(q) \setminus SL_2(\mathbb{Z}) \)? Obviously the number of Heegner points (\( \approx |D|^{1/2} \)) must exceed the index of \( \Gamma_0(q) \) in \( SL_2(\mathbb{Z}) \) (\( \approx q \)), so that \( q \ll D^{1/2-\delta} \) is a necessary condition. On the other hand, the Lindelöf hypothesis for certain Rankin-Selberg \( L \)-functions implies that \( q \ll D^{1/2-\delta} \) is sufficient, which is then best-possible (up to \( |D|^{\varepsilon} \)).

To this end, let \( \phi : SL_2(\mathbb{Z}) \setminus \mathbb{H} \to \mathbb{R} \) be smooth and compactly supported, and consider the inner product \( \langle |U|^2, \phi \rangle \) where we shall keep track of the dependence on \( \phi \), as it may vary with \( U \) in some way. One natural question is to understand how fast the support of \( \phi \) can shrink (which increases the size of the derivatives of \( \phi \)), and still expect QUE to hold. An obvious limitation to this is that a Maass form \( U \) typically oscillates at frequency \( T \), and is hence roughly constant at distances less than the de Broglie wavelength (or Planck scale) of size \( \approx 1/T \). See [HR, Sections 3 and 5.1] for elaboration here. So if \( \phi \) has support on a disc of radius \( T^{-1-\delta} \) and center \( z_0 \), then we expect that \( \langle U^2, \phi \rangle \approx U(z_0)^2 \langle 1, \phi \rangle \), and clearly QUE would not hold on this scale.

There are two standard approaches to QUE, one being Watson’s formula and bounds for triple product \( L \)-functions, and the other being Poincare series and bounds for shifted convolution sums. Unlike in the Heegner point case described above, assuming the Lindelöf hypothesis (for triple product \( L \)-functions) does not give a result valid on discs of radius \( T^{-1+\delta} \) (the smallest size discs upon which one might expect QUE to hold). See Proposition 1.5 for this result. Ghosh, Reznikov, and Sarnak [GRS, Appendix A], have formulated a version of “quantitative” QUE in terms of a Sobolev norm of \( \phi \). Our work here is complementary in that our goal is to find precise rates of convergence. J. Jung has shown a version of quantitative QUE for almost all forms \( U \); his setting is in terms of shifted convolution sums with explicit dependence on the weight function.

For the Eisenstein series case, Luo and Sarnak [LS1] showed that

\[
(1.10) \quad \langle |E(z, 1/2 + iT)|^2, \phi \rangle \sim \frac{3}{8} \log(1/4 + T^2) \langle 1, \phi \rangle
\]

for fixed \( \phi \). To allow varying test functions, we impose some conditions on \( \phi \) (or more accurately, the sequence of \( \phi \)'s depending on \( U \)). One natural choice is to pick a sequence of
numbers $C(k)$ and a constant $A = A(T)$, and consider $\phi$ satisfying

$$∥\Delta^k \phi∥_1 ≤ C(k)A^{2k},$$

for all $k = 0, 1, 2, \ldots$. As $A$ gets larger, this allows for more functions $\phi$, which can then be chosen to approximate the characteristic function of a disc of radius $A^{-1}$, for example. We shall show in Section 4.2 the following (unconditional)

**Theorem 1.4.** Suppose that a family of functions $\phi$ satisfies (1.11) with $A \ll T^{1-\delta}$ and that each $\phi$ in the family has support contained in a fixed compact set. Then

$$\langle |E(z, 1/2 + iT)|^2, \phi \rangle = \log(1/4 + T^2)\langle \phi, 3\pi \rangle + O(A^{1/2}T^{-1/6+\varepsilon}\|\phi\|_2)$$

$$+ O\left(\frac{\log T}{\log \log T}\|\phi\|_1\right) + O(T^{-100}).$$

As a corollary, we deduce that the Eisenstein series $E(z, 1/2 + iT)$ cannot be exceptionally small (nor large) on a disc of radius $\gg T^{-1/9+\delta}$ (the calculation for this exponent $1/9$ is that for $\phi$ approximating such a disc, $\langle 1, \phi \rangle \asymp A^{-2}$, while $\|\phi\|_2 \asymp A^{-1}$, and so one requires $A^{-2} \gg A^{-1/2}T^{-1/6+\varepsilon}$). We shall also give a more precise version of the main term with a power saving; see (4.13) below.

**Proposition 1.5.** Let $U$ be a Hecke-Maass cusp form, and suppose $\phi$ (possibly depending on $U$) satisfies (1.11) for some $A \leq T^{1-\delta}$. Assuming the Lindelöf hypothesis for triple product $L$-functions, we have

$$\langle U^2, \phi \rangle = \langle 1, \phi \rangle + O(A^{1/2}T^{-1/2+\varepsilon}\|\phi\|_2) + O(T^{-100}).$$

The implied constants depend on $\varepsilon > 0$ and the choice of constants $C(k)$ in (1.11).

We conclude from (1.13) that QUE should hold on a disc of radius $A^{-1}$ provided $A \ll T^{1/3-\delta}$ for some fixed $\delta > 0$.

It is implicit that even when QUE does not hold, the above work gives upper bounds on the $L^2$ norm of $U$ restricted to a shrinking family of sets, conditionally on the Lindelöf hypothesis. See Proposition 4.2 below for an unconditional upper bound on $\langle U^2, \phi \rangle$.

We also emphasize that we do not expect Proposition 1.5 to be optimal, even though it relies on the Lindelöf hypothesis. The reason is that in the derivation, there appears a sum over the spectrum, and it is possible that there is cancellation when combining the spectral coefficients. This situation is similar to the prime geodesic theorem where the analog of the Riemann Hypothesis holds (in this context this means there are no exceptional eigenvalues for $PSL_2(\mathbb{Z})$), yet from this one does not immediately deduce the presumably optimal “square-root” error term; see [Iw2]. Instead see for example p.139 of [Iw2] where a natural conjecture on sums of Kloosterman sums is stated, which would then give the optimal error term.

Since our paper is partially conjectural, we have organized it so that the conjectural results appear only in Sections 3 and 4.1. Within those sections, it should be clear from context whether a given formula is heuristically valid, conditional on some specific unproved hypothesis (e.g., GRH), or is valid unconditionally.
2. Notation and standard results

2.1. Fourier expansions and the standard $L$-functions. Let $U(z)$ be one of the three functions $u_j(z)$, $y^{k/2}f(z)$, $E(z,1/2 + iT)$. We recall the definition

$$E(z,s) = \sum_{\gamma \in \Gamma \setminus \Gamma} \text{Im}(\gamma z)^s,$$

where $\Gamma = PSL_2(\mathbb{Z})$ and $\Gamma_{\infty}$ is the stabilizer of $\infty$. Then $\text{Res}_{s=1}E(z,s) = \frac{3}{\pi}$. Each $U(z)$ has the Fourier expansion

$$U(z) = c_0(y) + \rho(1) \sum_{n \neq 0} \frac{\lambda(n)}{\sqrt{|n|}} e(nx)V(2\pi|n|y),$$

with additional notation as follows. Here $c_0(y) = c_0(y,s)$ is the constant term in the Fourier expansion which is nonzero only in the Eisenstein case for which $c_0(y,s) = y^s \varphi(s)y^{1-s}$ with $\varphi(s) = \frac{\theta(1-s)}{\theta(s)}$ and $\theta(s) = \pi^{-s}\Gamma(s)\zeta(2s)$. Note $|\varphi(1/2 + iT)| = 1$. Here $\lambda(n)$ are Hecke eigenvalues which on the Ramanujan conjecture are bounded in absolute value by the divisor function $d(n)$, and

$$V(y) = \begin{cases} V_T(y) := \sqrt{y} K_{iT}(y), & \text{Maass and Eisenstein cases} \\ V_k(y) := y^{k/2} \exp(-y), & \text{Holomorphic case}, \end{cases}$$

where $T$ is the spectral parameter of $u_j$. In the Maass and Eisenstein cases, $\lambda(-n) = \lambda(n)$ since $u_j$ is assumed to be even, while in the holomorphic case $\lambda(n) = 0$ for $n < 0$. In case $U$ is the Eisenstein series, $\lambda(n) = \tau_{iT}(n) = \sum_{ab=|n|} (a/b)^iT$, and with the above normalization of the constant term, we have

$$\theta(1/2 + iT)\rho(1) = (2/\pi)^{1/2}, \quad \theta(s) = \pi^{-s}\Gamma(s)\zeta(2s).$$

Then by Stirling’s formula and standard bounds for the Riemann zeta function, we deduce

$$|\rho(1)|^2 = \frac{2}{\pi} \frac{\cosh(\pi T)}{\zeta(1+2iT)^2} = T^{o(1)} \exp(\pi T).$$

Suppose initially that $\text{Re}(s) > \frac{1}{2}$. Then we define

$$L(1/2 + s) = \int_0^\infty (U(iy) - c_0(y)) y^s \frac{dy}{y},$$

This integral converges absolutely since $U(iy) - c_0(y,1/2+iT) \ll_T y^{-1/2}$ for $y \ll 1$ (it is $O(\exp(-y))$ for $y \gg 1$). If $U$ is cuspidal then the integral converges absolutely for all $s \in \mathbb{C}$. For $\text{Re}(s) > \frac{1}{2}$, we can reverse the order of integration and summation to show

$$L(1/2 + s) = (1 + \lambda(-1))\rho(1)L(1/2 + s, U)\gamma_V(1/2 + s),$$

where we write

$$L(s,U) = \sum_{n \geq 1} \frac{\lambda(n)}{n^s}, \quad \gamma_V(1/2 + s) = \int_0^\infty V(2\pi y) y^s \frac{dy}{y}.$$

It is sometimes useful to express $\gamma_V$ in terms of gamma functions. From [GR] (6.56.16),

$$\gamma_{V^2}(1/2 + s) = 2^{-3/2} \pi^{-s} \Gamma\left(\frac{1}{2} + s + iT\right) \Gamma\left(\frac{1}{2} + s - iT\right).$$
For the holomorphic case, we have by a direct calculation
\[(2.10)\] \[\gamma_{V_k}(1/2 + s) = (2\pi)^{-s}\Gamma(\frac{k}{2} + s).\]

When \(U(z) = E(z, 1/2 + iT) = E_T\), a short calculation shows
\[(2.11)\] \[L(s, E_T) = \zeta(s + iT)\zeta(s - iT).\]

### 2.2. Rankin-Selberg integrals.

**Lemma 2.1.** Suppose \(F\) and \(G\) are smooth functions on \(\mathbb{R}^+\) with rapid decay at 0 and \(\infty\). Then for any \(\alpha, \beta \in \mathbb{C}\), we have
\[(2.12)\] \[
\int_0^\infty F(y) y^\alpha G(y) y^\beta dy = \frac{1}{2\pi i} \int_{(c)} \tilde{F}(s + \alpha) \tilde{G}(\beta - s) ds,
\]
where \(\tilde{W}(s)\) denotes the Mellin transform of a function \(W:\)
\[(2.13)\] \[
\tilde{W}(s) = \int_0^\infty W(y) y^s dy, \quad W(y) = \frac{1}{2\pi i} \int_{(c)} \tilde{W}(s) y^{-s} ds.
\]

Lemma 2.1 is a special case of Mellin convolution (easily proved in the smooth setting by Mellin inversion) so we omit the proof.

Since we have the Mellin pair
\[(2.14)\] \[
\gamma_V(1/2 + s) = \int_0^\infty V(2\pi y) y^s dy, \quad V(2\pi y) = \frac{1}{2\pi i} \int_{(c)} \gamma_V(1/2 + s) y^{-s} ds,
\]
by Lemma 2.1, we have for \(\text{Re}(s) > -1\),
\[(2.15)\] \[
\gamma_{V^2}(1 + s) := \int_0^\infty V(2\pi y)^2 y^s dy = \frac{1}{2\pi i} \int_{(0)} \gamma_V(1/2 + s + v)\gamma_V(1/2 - v) dv.
\]

Technically, the conditions of Lemma 2.1 are not met, but the formula still holds in this range. It is also possible to find a closed formula for \(\gamma_{V^2}\). By (6.576.4) of [GR] we have for \(V_T(y) = \sqrt{y}K_iT(y)\) that
\[(2.16)\] \[
\gamma_{V_T^2}(1 + s) = 2^{-2} \pi^{-s} \frac{\Gamma(\frac{1+s+2iT}{2})\Gamma(\frac{1+s}{2})^2\Gamma(\frac{1+s-2iT}{2})}{\Gamma(1 + s)}.
\]

If \(V_k(y) = y^{k/2} \exp(-y)\), then by direct calculation
\[(2.17)\] \[
\gamma_{V_k^2}(1 + s) = 2^{-k} (4\pi)^{-s} \Gamma(k + s).
\]

In many cases throughout this paper we have preferred to use the definition (2.15) over the evaluations (2.16)–(2.17), in part because it allows us to treat the holomorphic and Maass cases simultaneously, but also because the exact evaluations are not necessary and it is easier to work directly with the definition.

The Dirichlet series analog of the above definitions is
\[(2.18)\] \[
Z(s, U) = \sum_{n=1}^\infty \frac{\lambda(n)^2}{n^s},
\]
which converges absolutely for \(\text{Re}(s) > 1\) by Rankin-Selberg theory (see below). For \(U = u_j\) or \(y^{k/2} f(z)\), we have
\[(2.19)\] \[
Z(s, U) = \frac{\zeta(s)}{\zeta(2s)} L(\text{sym}^2 U, s),
\]
while if $U = E(z, 1/2 + iT) = E_T$, then
\begin{equation}
Z(s, E_T) = \frac{\zeta^2(s)\zeta(s-2iT)\zeta(s+2iT)}{\zeta(2s)}.
\end{equation}

**Lemma 2.2.** Suppose that $U(z)$ is either $u_j(z)$ or $y^{k/2}f(z)$. Then for $\Re(s) > 0$, we have
\begin{equation}
\int_{\Gamma \backslash \mathbb{H}} |U(z)|^2 E(z, 1 + s) \frac{dxdy}{y^2} = (1 + \lambda(-1)^2)|\rho(1)|^2 Z(1 + s, U)\gamma_V^2(1 + s).
\end{equation}
The formula extends to hold for all $s \in \mathbb{C}$ by meromorphic continuation.

This is the standard unfolding argument but since we use it repeatedly we give the full proof.

**Proof of Lemma 2.2.** For $\Re(s) > 1$, we have
\begin{equation}
\langle |U|^2, E(\cdot, \pi) \rangle := \int_{\Gamma \backslash \mathbb{H}} |U(z)|^2 E(z, s) \frac{dxdy}{y^2} = \sum_{\gamma \in \Gamma \backslash \Gamma} \int_{\Gamma \backslash \mathbb{H}} |U(z)|^2 (\Im(\gamma z))^s \frac{dxdy}{y^2},
\end{equation}
which by unfolding gives
\begin{equation}
\langle |U|^2, E(\cdot, \pi) \rangle = \int_{0}^{1} \int_{0}^{\infty} |U(z)|^2 y^s \frac{dxdy}{y^2}.
\end{equation}
Inserting the Fourier expansion of $U$ and performing the $x$-integral, we obtain
\begin{equation}
\langle |U|^2, E(\cdot, \pi) \rangle = |\rho(1)|^2 \sum_{n \neq 0} \frac{\lambda(n)^2}{|n|} \int_{0}^{\infty} |V(2\pi|n|y)|^2 y^{s-1} \frac{dy}{y} = (1 + \lambda(-1)^2)|\rho(1)|^2 \left( \sum_{n \geq 1} \frac{\lambda(n)^2}{n^s} \right) \int_{0}^{\infty} |V(2\pi y)|^2 y^{s-1} \frac{dy}{y},
\end{equation}
which by (2.15) and (2.18), finishes the proof. \qed

When $U$ is the Eisenstein series then (2.21) does not converge. Instead, we consider the inner product against the incomplete Eisenstein series as do Luo-Sarnak in [LS1], but we briefly re-derive the formula to have a self-contained exposition. Suppose that $h : \mathbb{R}^+ \to \mathbb{R}$ is smooth and compactly-supported. By definition,
\begin{equation}
E(z, h) = \sum_{\gamma \in \Gamma \backslash \Gamma} h(\Im(\gamma z)),
\end{equation}
and Mellin inversion gives for $c > 1$,
\begin{equation}
E(z, h) = \frac{1}{2\pi i} \int_{(c)} E(z, s)\overline{h}(-s)ds.
\end{equation}
The formula (valid for $U$ Eisenstein series or a cusp form) states
\begin{equation}
\langle |U|^2, E(z, h) \rangle = \int_{0}^{\infty} |c_0(y)|^2 h(y) \frac{dy}{y^2} + (1 + \lambda(-1)^2)\frac{|\rho(1)|^2}{2\pi i} \int_{(c)} \overline{h}(-1 - s) Z(1 + s, U)\gamma_V^2(1 + s)ds.
\end{equation}
Proof. Unfolding as in the proof of Lemma 2.2, we obtain

\begin{equation}
\langle |U|^2, E(z, h) \rangle = \int_0^\infty \int_0^1 |U(z)|^2 h(y) \frac{dxdy}{y^2}.
\end{equation}

Inserting the Fourier expansion of $U$, we obtain

\begin{equation}
\langle |U|^2, E(z, h) \rangle = \int_0^\infty |c_0(y)|^2 h(y) \frac{dy}{y^2} + (1 + \lambda(-1)^2)|\rho(1)|^2 \sum_{n \geq 1} \frac{\lambda(n)^2}{n} \int_0^\infty |V(2\pi ny)|^2 h(y) \frac{dy}{y^2}.
\end{equation}

Using Mellin inversion on $h$, we obtain that the non-constant terms in (2.29) equal

\begin{equation}
(1 + \lambda(-1)^2)|\rho(1)|^2 \frac{1}{2\pi i} \int_{(c)} \bar{h}(-1 - s) \left( \sum_{n \geq 1} \frac{\lambda(n)^2}{n^{1+s}} \right) \int_0^\infty |V(2\pi y)|^2 y^s \frac{dy}{y} ds.
\end{equation}

Again using (2.15) and (2.18), we obtain (2.27). \qed

3. Derivation of Conjectures 1.1 and 1.3

In this section, we suppose $U$ is either $u_j(z)$ or $y^{k/2}f(z)$. By Lemma 2.1 and (2.6),

\begin{equation}
I(U, \alpha) := \int_0^\infty |U(iy)|^2 y^\alpha \frac{dy}{y} = \frac{1}{2\pi i} \int_{(0)} \mathcal{L}(1/2 + s + \alpha)\mathcal{L}(1/2 - s) ds.
\end{equation}

By (2.7), we have

\begin{equation}
I(U, \alpha) = (1 + \lambda(-1)^2)|\rho(1)|^2 \frac{1}{2\pi i} \int_{(c)} L(1/2 + \alpha + s, U) L(1/2 - s, U) \gamma_V(1/2 + \alpha + s) \gamma_V(1/2 - s) ds.
\end{equation}

This is effectively a shifted second moment of the $L$-function with a weight depending on $U$. The weight is very natural, being the Archimedean part of the completed $L$-function. Although such moments for $U$ fixed have been well-studied, the case of $U$ varying is more difficult, as mentioned in the introduction. Therefore we appeal to the recipe for computing the main term of a moment of $L$-functions due to [CFKRS]. The following formula should be understood in a formal sense:

\begin{equation}
I(U, \alpha) \sim (1 + \lambda(-1)^2)|\rho(1)|^2 \left( \sum_{m,n \geq 1} \frac{\lambda(m)\lambda(n)}{m^{1/2+\alpha}n^{1/2}} \frac{1}{2\pi i} \int_{-\infty}^\infty \frac{1}{\Gamma(n/m)} \frac{1}{\Gamma(1/2 + \alpha + it)} \gamma_V(1/2 + \alpha + it) \gamma_V(1/2 - it) dt + \ldots \right),
\end{equation}

where the dots indicate an identical term with $\alpha$ replaced by $-\alpha$, consistent with $I(U, \alpha) = I(U, -\alpha)$. This is the first step of the [CFKRS] recipe which calls for us to write a formal approximate functional equation for each $L$-function, and to discard terms where the “root number” is oscillating with the family. We view the $t$-integral as selecting only the diagonal term $m = n$, leading to

\begin{equation}
I(U, \alpha) \sim (1 + \lambda(-1)^2)|\rho(1)|^2 \left( \sum_{n \geq 1} \frac{\lambda(n)^2}{n^{1+\alpha}} \frac{1}{2\pi i} \int_{(0)} \gamma_V(1/2 + \alpha + s) \gamma_V(1/2 - s) ds + \ldots \right)
\end{equation}

\begin{equation}
= (1 + \lambda(-1)^2)|\rho(1)|^2 \left( Z(1 + \alpha, U) \gamma_V(1 + \alpha) + Z(1 - \alpha, U) \gamma_V(1 - \alpha) \right).
\end{equation}
Conveniently there was no need to invoke the evaluations (2.16)–(2.17). Using (2.21), we obtain the conjectured asymptotic

\[(3.5) \quad I(U, \alpha) \sim (1 + \lambda(-1)) \int_{\Gamma \setminus \mathbb{H}} |U(z)|^2 \left( E(z, 1 + \alpha) + E(z, 1 - \alpha) \right) \frac{dx dy}{y^2} \cdot\]

This is a compact way to express the two formulas in Conjecture 1.3, since \(\lambda(-1) = 1\) for \(U = u_j\) an even Maass form, and \(\lambda(-1) = 0\) for \(U = y^{k/2}f(z)\) coming from a holomorphic form.

Next we derive Conjecture 1.1 from Conjecture 1.3, supposing that the implied error term in Conjecture 1.3 has some uniformity in \(\alpha\). Let \(\psi(y)\) be a smooth function with compact support on \(\mathbb{R}^+\), and define

\[(3.6) \quad I(U, \psi) = \int_{0}^{\infty} |U(iy)|^2 \psi(y) \frac{dy}{y}.\]

We may suppose that \(\psi\) is even, i.e., \(\psi(y^{-1}) = \psi(y)\), because in general one can decompose \(\psi\) as the sum of an even and an odd function, and \(I(U, \psi)\) vanishes if \(\psi\) is odd. Thus \(\tilde{\psi}(s) = \tilde{\psi}(-s)\).

By Mellin inversion, and using (3.5), we have

\[(3.7) \quad I(U, \psi) = \frac{1}{2\pi i} \int_{(0)} \tilde{\psi}(-s) I(U, s) ds \sim (1 + \lambda(-1)) \frac{1}{2\pi i} \int_{(0)} \tilde{\psi}(-s) \int_{\Gamma \setminus \mathbb{H}} |U(z)|^2 \left( E(z, 1 + s) + E(z, 1 - s) \right) \frac{dx dy}{y^2} ds.\]

Here is where we assumed the error term in (3.5) holds with some uniformity.

Reversing the order of integration, and with the definition

\[(3.8) \quad H(z, \psi) = \frac{1}{2\pi i} \int_{(0)} \tilde{\psi}(-s) \left( E(z, 1 + s) + E(z, 1 - s) \right) ds,\]

we have the conjecture

\[(3.9) \quad I(U, \psi) \sim (1 + \lambda(-1)) \int_{\Gamma \setminus \mathbb{H}} |U(z)|^2 H(z, \psi) \frac{dx dy}{y^2}.\]

Finally we show unconditionally, using the QUE theorem, that the right hand side of (3.9) agrees with the statement of Conjecture 1.1.

Observe that \(H(z, \psi)\) is \(SL_2(\mathbb{Z})\)-invariant. In fact \(H\) is related to an incomplete Eisenstein series, namely

\[(3.10) \quad H(z, \psi) = -\frac{3}{\pi} \tilde{\psi}(0) + 2E(z, y\psi(y)).\]

We prove this now. First, shift the contour in (3.8) to \(\text{Re}(s) = \varepsilon\) and expand the integral as the sum of two integrals corresponding to the two terms \(E(z, 1 + s)\) and \(E(z, 1 - s)\). Then by shifting the latter of these contours to the left, we have

\[(3.11) \quad \frac{1}{2\pi i} \int_{(\varepsilon)} \tilde{\psi}(-s) E(z, 1 - s) ds = -\frac{3}{\pi} \tilde{\psi}(0) + \frac{1}{2\pi i} \int_{(-\varepsilon)} \tilde{\psi}(-s) E(z, 1 - s) ds,\]
which by the change of variables $s \to -s$ and the evenness of $\tilde{\psi}$ gives

$$H(z, \psi) = -\frac{3}{\pi} \tilde{\psi}(0) + \frac{2}{2\pi i} \int_{(\varepsilon)} \tilde{\psi}(-s) E(z, 1 + s) ds. \quad (3.12)$$

Finally we observe from (2.26) that

$$\frac{1}{2\pi i} \int_{(\varepsilon)} \tilde{\psi}(-s) E(z, 1 + s) ds = E(z, y\psi(y)). \quad (3.13)$$

Inserting (3.10) into (3.9), applying (1.1) to $I(U, \psi)$ (with $\psi$ fixed, the eigenvalue/weight of $U$ large), and recalling that $U$ is normalized by (1.2), we obtain

$$I(U, \psi) \sim (1 + \lambda(-1)) \int_{\Gamma \setminus \mathbb{H}} \left( -\frac{3}{\pi} \tilde{\psi}(0) + 2E(z, y\psi(y)) \right) \frac{dx dy}{y^2} = (1 + \lambda(-1)) \left( -\tilde{\psi}(0) + 2 \int_{0}^{1} \int_{0}^{\infty} y\psi(y) \frac{dx dy}{y^2} \right). \quad (3.14)$$

This simplifies as

$$I(U, \psi) \sim (1 + \lambda(-1)) \int_{0}^{\infty} \psi(y) \frac{dy}{y}, \quad (3.15)$$

which again is a compact way to write both asymptotics in Conjecture 1.1.

The case $\alpha = 0$ is particularly interesting, and Conjecture 1.3 is perhaps not the best form in this case.

**Conjecture 3.1.** We have

$$\int_{0}^{\infty} |u_j(iy)|^2 \frac{dy}{y} \sim 2 \log(1/4 + t_j^2), \quad (3.16)$$

as $t_j \to \infty$.

One way to guess the form of this answer is to imagine that $|u_j(iy)|^2 \sim 2$ (on average) until it begins to decay quickly, which is for $y + y^{-1} \gg t_j$.

We set $t_j = T$ and continue with (3.4). With this notation, Lemma 2.2 shows

$$\text{Res}_{\alpha=0}(1 + \lambda(-1))|\rho(1)|^2 Z(1 + \alpha)\gamma_{V^2}(1 + \alpha) = 1. \quad (3.17)$$

Let $Z(1 + \alpha) = \frac{r_0}{\alpha}(1 + r_0\alpha + \ldots)$. By Theorem 5.17 of [IK] (conditional on GRH and Ramanujan), we have $r_0 \ll \log \log T$. By taking a Taylor expansion, we have the prediction

$$\int_{0}^{\infty} |u_j(iy)|^2 y^\alpha \frac{dy}{y} \sim (1 + \lambda(-1)) \left( 2 \frac{\gamma'_{V^2}}{\gamma_{V^2}}(1) + 2r_0 + O(\alpha) \right). \quad (3.18)$$

Using the explicit evaluation (2.16) and Stirling’s formula, we have

$$\frac{\gamma'_{V^2}}{\gamma_{V^2}}(1) = \frac{1}{2} \log(1/4 + T^2) + O(1). \quad (3.19)$$

This quickly leads to Conjecture 3.1.
4. The QUE conjecture for shrinking sets

This section contains proofs of Theorem 1.4 and Proposition 1.5. Recall that our notation is such that QUE is an asymptotic for \( \langle U^2, \phi \rangle \) where \( \phi \) is a family of functions that may vary with \( U \) (but the \( \phi \)'s are generally not as oscillatory as \( U \) is). In this section we take \( U = u_j \) or \( U = E_T \), but it would be interesting to study the holomorphic case too.

It is well-known that the QUE conjecture can be approached either via bounds for triple product \( L \)-functions via Watson’s formula [W], or alternatively by shifted convolution sums via Poincare series. These two methods are apparently not equivalent and indeed Holowinsky and Soundararajan [HS] exploit both approaches in their proof of the mass equidistribution conjecture. We shall work with the triple product approach below.

4.1. Maass forms. Here we prove Proposition 1.5.

Proof. The Plancherel formula gives

\[
\langle U^2, \phi \rangle = \langle U^2, \frac{3}{\pi} \rangle \langle 1, \phi \rangle + \sum_{j \geq 1} \langle U^2, u_j \rangle \langle u_j, \phi \rangle + \frac{1}{4\pi} \int_{-\infty}^{\infty} \langle U^2, E(\cdot, 1/2 + it) \rangle \langle E(\cdot, 1/2 + it), \phi \rangle dt,
\]

where recall the inner product is with respect to \( \frac{dx dy}{\pi} \) (not probability measure). Note that if \( U \) is normalized with (1.2), then \( \langle U^2, \frac{3}{\pi} \rangle = 1 \) and the constant eigenfunction gives the expected main term in the QUE conjecture. If \( u_j \) is even and \( U \) is a Maass form then Watson’s formula reads

\[
|\langle U^2, u_j \rangle|^2 = \frac{\pi}{8} \frac{|\Gamma(\frac{1+2iT+it_j}{2})|^2 |\Gamma(\frac{1+2iT-it_j}{2})|^2 |\Gamma(\frac{1+it_j}{2})|^4 |\Gamma(\frac{1+2it_j}{2})|^2 L(U \times U \times u_j, 1/2)}{|\Gamma(\frac{1+2iT}{2})|^4 |\Gamma(\frac{1+2it_j}{2})|^2 L(\text{sym}^2 U, 1)^2 L(\text{sym}^2 u_j, 1)}.
\]

A similar formula holds for the Eisenstein series by an unfolding argument along the lines of Lemma 2.2. We shall simply quote the work of Section 2 of [LS1] for the following:

\[
|\langle E(z, 1/2 + iT), u_j \rangle|^2 = c(T) \rho_j(1) \frac{|\Gamma(\frac{1+it_j}{2})|^2 |\Gamma(\frac{1-it_j-2iT}{2})|^2 |\Gamma(\frac{1+2iT}{2})|^2 |\Gamma(\frac{1}{2} + iT)^2 |\Gamma(\frac{1}{2} + it_j)^2 |\Gamma(1+2iT)|^2 L(u_j, 1/2)L(u_j, 1/2 - 2iT),
\]

where \( c(T) \) is such that \( |c(T)| \) is an absolute constant independent of \( T \). Note that \( |\langle E(z, 1/2 +iT), u_j \rangle|^2 \) is a close cousin to (4.2), recalling that \( |\rho_j(1)|^2 = c|\Gamma(\frac{1+it_j}{2})|^{-2}(L(1, \text{sym}^2 u_j))^{-1} \). In particular, the gamma factors are the same as in (4.2), and the triple product \( L \)-function is replaced by \( |L(u_j, 1/2)L(u_j, 1/2 - 2iT)|^2 \).

By Stirling’s formula, the ratio of gamma factors in (4.2) is

\[
\ll P(t_j, T) \exp(\frac{\pi}{2} Q(t_j, T)),
\]

where

\[
Q(t, T) = 4T - |t + 2T| - |t - 2T|,
\]

and

\[
P(t_j, T) = t_j^{-1}(1 + |2T + t_j|)^{-1/2}(1 + |2T - t_j|)^{-1/2}.
\]
By a simple calculation, $Q(t, T) = 0$ for $|t| \leq 2T$, and $Q(t, T) = 2(2T - |t|)$ for $|t| > 2T$. Therefore we can bound the terms in (4.1) with $t_j \geq 2T + C \log T$ with

$$ \|\phi\|_1 \sum_{t_j \geq 2T + C \log T} \text{Polynomial}(t_j, T) \exp(-\pi(t_j - 2T)), $$

(4.7)

using the trivial bound $|\langle u_j, \phi \rangle| \leq \|u_j\|_\infty \|\phi\|_1 \ll (1/4 + t_j^2)^{1/4} \|\phi\|_1$. Taking $C$ large enough compared to the degree of the unspecified polynomial, we can bound this error term by $\ll \|\phi\|_1 T^{-100}$, with an absolute implied constant. In summary, we have shown

$$ \langle U^2, \phi \rangle - \langle 1, \phi \rangle = \sum_{t_j \leq 2T + C \log T} \langle U^2, u_j \rangle \langle u_j, \phi \rangle + (\text{Eisenstein}) + O(\|\phi\|_1 T^{-100}). $$

Without some additional assumptions on $\phi$, it is not reasonable to expect that the terms with $t_j \leq 2T + C \log T$ also constitute an error term. For instance, if $\phi = U^2$ then in [BKY] it is conjectured that $3/16(U^2, U^2) \sim 3$. In an even more extreme direction, we could take a sequence of $\phi$'s tending to a delta function in which case one would not expect an asymptotic law for $\langle U^2, \phi \rangle$.

Assuming (1.11) holds, then by the self-adjointness of the Laplacian, we can bound the spectral coefficients by

$$ (1/4 + t_j^2)^k \langle u_j, \phi \rangle = \langle \Delta^k u_j, \phi \rangle = \langle u_j, \Delta^k \phi \rangle \ll A^{2k} \|u_j\|_\infty, $$

(4.9)

whence

$$ \langle u_j, \phi \rangle \ll (1/4 + t_j^2)^{1/4} \left( \frac{A^2}{1/4 + t_j^2} \right)^k. $$

(4.10)

Thus if $A \leq T^{1-\delta}$ for some fixed $\delta > 0$ (meaning in some sense that $u_j$ is more oscillatory than $\phi$) then in (4.8) we can truncate the sum at $t_j \leq AT^\epsilon$. By Cauchy’s inequality, we derive

$$ |\langle U^2, \phi \rangle - \langle 1, \phi \rangle| \ll \left( \sum_{t_j \leq AT^\epsilon} \langle U^2, u_j \rangle^2 \right)^{1/2} \left( \sum_{t_j} |\langle u_j, \phi \rangle|^2 \right)^{1/2} + (\text{Eisenstein}) + O(T^{-100}). $$

(4.11)

Bessel’s inequality implies $\sum_{t_j} |\langle u_j, \phi \rangle|^2 \leq \langle \phi, \phi \rangle$. Assuming the Lindelöf Hypothesis, we have by Watson’s formula that

$$ \sum_{t_j \leq AT^\epsilon} |\langle U^2, u_j \rangle|^2 \ll \sum_{t_j \leq AT^\epsilon} T^{-1+\epsilon} t_j^{-1+\epsilon} \ll AT^{-1+\epsilon}. $$

(4.12)

Similar estimates hold for the Eisenstein series so we suppress those arguments. Thus we derive (1.13).

4.2. QUE with shrinking sets for Eisenstein series. In this section we prove Theorem 1.4. Our approach most naturally shows a more precise form

$$ \langle |E(z, 1/2 + iT)|^2, \phi \rangle = \lim_{\alpha \to 0} \langle D_\alpha, \phi \rangle + O(A^{1/2} T^{-1/6+\epsilon} \|\phi\|_2), $$

(4.13)

where

$$ D_\alpha(z) = E(z, 1 + \alpha) + \Phi_T(\alpha) E(z, 1 - \alpha), \quad \Phi_T(\alpha) = \frac{\theta(1/2 - iT - \alpha) \theta(1/2 + iT)}{\theta(1/2 + iT + \alpha) \theta(1/2 - iT)}. $$

(4.14)
Write \( E(z, 1 + \alpha) = \frac{3\pi}{\alpha} + a(z) + O(\alpha) \). The constant term in the Taylor expansion for \( D_\alpha(z) \) around \( \alpha = 0 \) is
\[
(4.15) \quad 2a(z) - \frac{3}{\pi} \Phi'(0).
\]
One can find expressions for \( a(z) \) via (22.69) of [IK], and by a calculation,
\[
(4.16) \quad \Phi'(0) = \Phi'(T) = -\sum \pm \left( \frac{\Gamma'}{\Gamma}(1/2 \pm iT) + 2\frac{\zeta'}{\zeta}(1 \pm 2iT) - \log \pi \right).
\]
By Stirling’s formula, \( \frac{\Gamma'}{\Gamma}(1/2 + iT) + \frac{\Gamma'}{\Gamma}(1/2 - iT) = \log(1/4 + T^2) + O(T^{-2}) \), while \( \frac{\zeta'}{\zeta}(1 + \pm 2iT) = O(T^{\log T}) \), which explains how (1.12) follows from (4.13).

One can also compare Theorem 1.4 with the Maass-Selberg relation (cf. Proposition 6.8 of [Iw4]).

The Luo-Sarnak [LS1] approach proceeds by showing (1.10) for \( \phi \) a Maass form, or \( \phi \) an incomplete Eisenstein series. The incomplete Eisenstein series span the space \( \mathcal{E}(\Gamma \backslash \mathbb{H}) \) which is the orthogonal complement of the span of the Maass forms \( \mathcal{C}(\Gamma \backslash \mathbb{H}) \) (here we use notation as in Iwaniec’s book [Iw4]), so this suffices to show (1.10), though with an inexplicit error term. On the other hand, the error term in (1.10) when \( \phi \) is a fixed Maass form gives a power saving in \( T \) (and similarly for a fixed incomplete Eisenstein series, but in this case there is a lower-order term that must be included as in (4.13)). However, it seems difficult to constructively approximate the projection of \( \phi \) onto \( \mathcal{E}(\Gamma \backslash \mathbb{H}) \) in terms of incomplete Eisenstein series. Of course, the spectral expansion does give a decomposition of the projection of \( \phi \), but in terms of Eisenstein series themselves which have moderate growth at the cusp, even though the projection of \( \phi \) has rapid decay at the cusp. This is the main technical difficulty in analyzing (1.10) using the spectral decomposition and Parseval’s formula, because \( |E(z, 1/2 + iT)|^2 \) grows too fast at the cusp. However, there is a way around this problem of convergence that was discovered by Zagier [Za], namely, to work with renormalized integrals. Michel and Venkatesh [MV] have recently given an interpretation of this renormalization in the language of representation theory. In addition they give a regularized Plancherel formula that we shall use here; see the Proposition on p.243 of [MV]. See also [Ze2] for an application of regularization to quantum ergodicity with Eisenstein series.

First we define the regularized inner product. Suppose that \( F \) is a function of moderate growth which by definition means that
\[
(4.17) \quad F(z) = \phi(y) + O(y^{-N})
\]
as \( y \to \infty \), for any \( N > 0 \), where
\[
(4.18) \quad \phi(y) = \sum_{i=1}^{l} \frac{c_i}{n_i!} y^{\alpha_i} \log^{n_i} y,
\]
for \( c_i, \alpha_i \in \mathbb{C} \), and \( n_i \geq 0 \) an integer. Suppose that no \( \alpha_i = 1 \). Let \( \mathcal{E}(z) \) denote a linear combination of Eisenstein series \( E(z, \alpha) \) (possibly including derivatives with respect to \( \alpha \)) with \( \text{Re}(\alpha) > 1/2 \) such that \( F(z) - \mathcal{E}(z) = O(y^{1/2}) \). The regularized integral of \( F \) is then defined to be
\[
(4.19) \quad \int_{\Gamma \backslash \mathbb{H}}^{\text{reg}} F(z) \frac{dx dy}{y^2} := \int_{\Gamma \backslash \mathbb{H}} (F(z) - \mathcal{E}(z)) \frac{dx dy}{y^2}.
\]
Zagier computed some special cases of these regularized integrals that we require. It is obvious from the definition that \( \langle E(z, s), 1 \rangle_{\text{reg}} = 0 \). For \( 0 < \text{Re}(s) < 1 \), \( E(z, s) \in L^1(\Gamma \backslash \mathbb{H}) \), and it is easy to see that the integral (without regularization) vanishes using the fact that the Laplacian is self-adjoint, yet \( E(z, s) \) and \( 1 \) have different Laplace eigenvalues. For the case of two Eisenstein series, we have \( \langle E(z, s_1), E(z, s_2) \rangle_{\text{reg}} = 0 \); see p.428 of [Za]. This formula should be heuristically natural because again \( E(z, s_1) \) and \( E(z, s_2) \) have distinct Laplace eigenvalues (for \( s_1(1-s_1) \neq s_2(1-s_2) \)). Technically, to even define the regularized integral we require that \( s_1 \neq s_2 \) and \( s_1 \neq 1-s_2 \). Finally, for the case of the product of three Eisenstein series, Zagier computes (see p.430 of [Za]) that

\[
\langle E(z, \frac{1}{2} + s_1), E(z, \frac{1}{2} + s_2), E(z, \frac{1}{2} + s) \rangle_{\text{reg}} = \frac{c}{\theta(\frac{1}{2} + s) \theta(\frac{1}{2} + s_1) \theta(\frac{1}{2} + s_2)} \prod_{\delta_1, \delta_2 \in \{\pm 1\}} \zeta(\frac{1}{2} + s + \delta_1 s_1 + \delta_2 s_2) \Gamma(\frac{1}{2} + s + \delta_1 s_1 + \delta_2 s_2). 
\]

**Lemma 4.1 ([MV]).** Suppose that \( F \) and \( G \) are smooth functions on \( \Gamma \backslash \mathbb{H} \), \( G \) with compact support, \( F \) of moderate growth with \( \text{Re}(\alpha_i) \neq 1/2 \) for all \( i \) and no \( \alpha_i = 1 \). Then the following regularized version of Parseval’s formula holds:

\[
\langle F, G \rangle = \sum_j \langle F, u_j \rangle \langle u_j, G \rangle + \langle F, \frac{3}{\pi} \rangle_{\text{reg}} \langle 1, G \rangle + \frac{1}{4\pi} \int_{-\infty}^\infty \langle F, E_t \rangle_{\text{reg}} \langle E_t, G \rangle dt + \langle \mathcal{E}, G \rangle,
\]

where \( E_t \) denotes \( E(z, 1/2 + it) \).

**Proof.** We give a minor variation of the proof of Michel and Venkatesh [MV]. Suppose that \( F_1(z) := F(z) - \mathcal{E}(z) = O(y^{1/2-\delta}) \) for some \( \delta > 0 \), whence \( F_1 \in L^2(\Gamma \backslash \mathbb{H}) \). Since \( G \) has rapid decay, we have \( \langle F, G \rangle = \langle F_1, G \rangle + \langle \mathcal{E}, G \rangle \) as absolutely convergent integrals. Then by the usual Plancherel formula, we have

\[
\langle F_1, G \rangle = \sum_j \langle F_1, u_j \rangle \langle u_j, G \rangle + \langle F_1, \frac{3}{\pi} \rangle_{\text{reg}} \langle 1, G \rangle + \frac{1}{4\pi} \int_{-\infty}^\infty \langle F_1, E_t \rangle \langle E_t, G \rangle dt.
\]

Note that \( \langle F_1, u_j \rangle = \langle F, u_j \rangle \) since \( \langle \mathcal{E}, u_j \rangle = 0 \), and \( \langle F_1, \frac{3}{\pi} \rangle_{\text{reg}} = \langle F, \frac{3}{\pi} \rangle_{\text{reg}} \) by definition. As mentioned above, \( \langle \mathcal{E}, E_t \rangle_{\text{reg}} = 0 \), so \( \langle F_1, E_t \rangle = \langle F, E_t \rangle_{\text{reg}} \). Gathering the terms finishes the proof. \( \square \)

**Proof of Theorem 1.4.** We shall apply Lemma 4.1 with \( G = \phi \) and \( F(z) = E(z, s_1)E(z, s_2) \). We shall eventually let \( s_1 = 1/2 + iT \) and \( s_2 = 1/2 - iT \) by analytic continuation. In fact, we shall be able to do this for each of the four terms appearing in (4.21).

By following the arguments in Section 4.1, we have that

\[
\sum_j |\langle E(z, 1/2 + iT), u_j \rangle| \langle u_j, \phi \rangle \leq \|\phi\|_2 \left( \sum_{t_j \leq AT^\varepsilon} |\langle E(z, 1/2 + iT), u_j \rangle|^2 \right)^{1/2} + O(T^{-100}).
\]

Consulting (4.3), (4.4) and (4.6), we then have

\[
\sum_j |\langle E(z, 1/2 + iT), u_j \rangle|^2 \langle u_j, \phi \rangle \ll \|\phi\|_2 T^{-1/2+\varepsilon} \left( \sum_{t_j \leq AT^\varepsilon} t_j^{-1} L(u_j, 1/2)^2 L(u_j, 1/2 - 2iT)^2 \right)^{1/2} + O(T^{-100}).
\]
Next we apply the uniform subconvexity bound \( L(u_j, 1/2-2iT) \ll T^{1/3+\varepsilon} \) of Jutila-Motohashi [JM], and the following bound which follows from the spectral large sieve inequality:

\[
(4.25) \sum_{t_j \leq AT^c} L(u_j, 1/2)^2 \ll A^2 T^c.
\]

In this way, we obtain

\[
(4.26) \sum_j \langle |E(z, 1/2 + iT)|^2, u_j \rangle \langle u_j, \phi \rangle \ll T^{-1/6+\varepsilon} A^{1/2} \|\phi\|_2,
\]

which is the error term stated in (4.13).

Next we examine the regularized projections of \( |E|^2 \) onto the constant eigenfunction and the Eisenstein series. The inner product with the constant eigenfunction vanishes as remarked following (4.19) (after taking a limit to treat the case \( s_1 = -s_2 = iT \)). For the Eisenstein contribution, we use (4.20). Note that \( \|\langle |E(z, 1/2 + iT)|^2, E_{t, \text{reg}} \rangle \|^2 \) takes the form

\[
(4.27) \left| \frac{\Gamma(\frac{1+2iT+iT}{2})^2 |\Gamma(\frac{1+2iT-iT}{2})|^2 |\zeta(\frac{1}{2} + it + 2iT)|^2 |\zeta(\frac{1}{2} + it)|^4 |\zeta(\frac{1}{2} + it - 2iT)|^2}{|\Gamma(1+2iT)|^4 |\Gamma(\frac{1+2iT}{2})|^2} \right| \ll |\zeta(1+2iT)|^2 |\zeta(1+2iT)|^4,
\]

for some absolute constant \( c \). This, as expected, has the exact shape as (4.2). By a similar argument as (4.9), we can truncate the integral at \( |t| \leq AT^c \) with a very small error. In this way, we obtain the bound

\[
(4.28) \frac{1}{4\pi} \int_{-\infty}^{\infty} \langle |E(z, 1/2 + iT)|^2, E_{t, \text{reg}} \rangle dt \ll \|\phi\|_2 T^{-100}
\]

\[+ \|\phi\|_2 T^{-1/2+\varepsilon} \left( \int_{|t| \leq AT^c} (1+|t|)^{-1} |\zeta(\frac{1}{2} + it)|^4 |\zeta(\frac{1}{2} + it + 2iT)|^2 |\zeta(\frac{1}{2} + it - 2iT)|^2 \right)^{1/2}.
\]

Using Weyl's bound \( \zeta(1/2 + it \pm 2iT) \ll T^{1/6+\varepsilon} \), and a bound for the fourth moment of zeta (on the level of the mean value theorem for Dirichlet polynomials), we obtain

\[
(4.29) \frac{1}{4\pi} \int_{-\infty}^{\infty} \langle |E(z, 1/2 + iT)|^2, E_{t, \text{reg}} \rangle dt \ll T^{-1/6+\varepsilon} \|\phi\|_2.
\]

This bound is better than (4.26) due to the smaller spectral measure of the Eisenstein series.

Finally, we evaluate \( \langle \mathcal{E}, \phi \rangle \). First we need to identify \( \mathcal{E}(z) \). The constant term of \( E(z, 1/2+s_1)E(z, 1/2+s_2) \) is

\[
(4.30) y^{1+s_1+s_2} + c_1 y^{1-s_1+s_2} + c_2 y^{1+s_1-s_2} + c_1 c_2 y^{1-s_1-s_2}, \quad c_1 = \frac{\theta(1/2-s_1)}{\theta(1/2+s_1)}, \quad c_2 = \frac{\theta(1/2-s_2)}{\theta(1/2+s_2)},
\]

so

\[
(4.31) \mathcal{E}(z) = E(z, 1+s_1+s_2) + c_1 E(z, 1-s_1+s_2) + c_2 E(z, 1+s_1-s_2) + c_1 c_2 E(z, 1-s_1-s_2).
\]

Note that by unfolding (and analytic continuation), \( \langle E(z, 1-s_1+s_2), \phi \rangle = \widetilde{\phi}_0(-s_1+s_2) \), where \( \phi_0(y) = \int_0^1 \phi(x + iy) dx \). In our application, we have \( s_2 = -iT \) and \( s_1 = iT + \alpha \) with \( \alpha \to 0 \), and we may use the rapid decay of \( \widetilde{\phi}_0 \), which implies \( \widetilde{\phi}_0(-2iT) \ll T^{-1000} \), to absorb this term into the error term (and similarly for \( E(z, 1+s_1-s_2) \)). Thus, we have

\[
(4.32) \langle \mathcal{E}, \phi \rangle = \langle D_\alpha, \phi \rangle + O(T^{-100}),
\]
where $D_u$ is defined by (4.14). This completes the proof of Theorem 1.4 in the form of (4.13). \qed

4.3. Unconditional upper bounds.

**Proposition 4.2.** Let $U$ be a Hecke-Maass cusp form, and suppose the family of $\phi$’s have support in a fixed compact set, and satisfy (1.11) for some $A \leq T^{1-\delta}$. Then

\begin{equation}
(4.33)
\langle U^2, \phi \rangle \ll T^\varepsilon A^{1/4} \|\phi\|_2 + T^{-100}.
\end{equation}

Here we can take $\phi$ nonnegative such that $\chi_A(z) \leq \phi(z) \leq \chi_{2A}(z)$ where $\chi_r(z)$ is the characteristic function of the disc $D_{z_0}(r)$ centered at a fixed point $z_0$ having radius $r$, and (4.33) gives an upper bound on the $L^2$ norm of $U$ restricted to such a disc, namely

\begin{equation}
(4.34)
\int_{D_{z_0}(A^{-1})} |U(z)|^2 \frac{dxdy}{y^2} \ll z_0 A^{-3/4+\varepsilon} T^\varepsilon.
\end{equation}

One may also wonder about bounds on average. It is easy to produce the following strong average bound:

\begin{equation}
(4.35)
\sum_{T \leq t_j \leq T+1} \langle u_j^2, \phi \rangle \ll T \|\phi\|_1.
\end{equation}

The proof follows immediately upon using (13.8) of [Iw4] which says $\sum_{T \leq t_j \leq T+1} |u_j(z)|^2 \ll T$, uniformly for $z$ in some fixed compact set.

**Proof.** As in the proof of Proposition 1.5, we have

\begin{equation}
(4.36)
\langle U^2, \phi \rangle = \sum_j \langle U^2, u_j \rangle \langle u_j, \phi \rangle + \ldots,
\end{equation}

with the dots indicating the constant eigenfunction and the Eisenstein series contributions.

We can truncate the spectral sum at $t_j \leq AT^\varepsilon$ with an error of size $O(T^{-100}\|\phi\|_1)$. By Watson’s formula (4.2), for $t_j = o(T)$,

\begin{equation}
(4.37)
\langle U^2, u_j \rangle = \theta_{U,j} T^{-1/2} t_j^{-1/2} \sqrt{L(1/2, \text{sym}^2U \times u_j)} \sqrt{L(1/2, u_j)},
\end{equation}

where $\theta_{U,j}$ is a real number satisfying $|\theta_{U,j}| \ll T^\varepsilon$. The conductor of $L(1/2, U \times U \times u_j) = L(1/2, \text{sym}^2U \times u_j)L(1/2, u_j)$ is, for $t_j = o(T)$, $T^{4t_j^2}$. The conductor of $L(1/2, u_j)$ is $t_j^2$, so the conductor of the degree 6 factor is $T^{4t_j^2}$.

By Hölder’s inequality (with exponents 4, 4, 2), we obtain

\begin{equation}
(4.38)
\langle U^2, \phi \rangle \ll T^{-1/2+\varepsilon} \left( \sum_{t_j \leq AT^\varepsilon} t_j^2 L(1/2, u_j)^2 \right)^{1/4} \left( \sum_{t_j \leq AT^\varepsilon} L(1/2, \text{sym}^2U \times u_j)^2 \right)^{1/4} \times \left( \sum_{t_j \leq AT^\varepsilon} |\langle u_j, \phi \rangle|^2 \right)^{1/2} + \ldots.
\end{equation}

By Bessel’s inequality, $\sum_{t_j} |\langle u_j, \phi \rangle|^2 \leq \langle \phi, \phi \rangle$. As noted earlier, (4.25) holds, and in addition we claim that the spectral large sieve inequality proves

\begin{equation}
(4.39)
\sum_{t_j \leq AT^\varepsilon} L(1/2, \text{sym}^2U \times u_j)^2 \ll AT^{2+\varepsilon},
\end{equation}

which completes the proof of Proposition 4.2.
which taken together leads to (4.33). A case similar to (4.39) (varying the level) appeared in [LMY], Proposition 6.3, so we omit a detailed explanation and instead give a sketch that displays the main ideas. The rough idea of the proof is to apply the approximate functional equation to see

\[
\sum_{t_j \leq AT^\varepsilon} L(1/2, \text{sym}^2 U \times u_j)^2 \approx \sum_{t_j \leq AT^\varepsilon} \left| \sum_{n \leq AT^{2+\varepsilon}} \frac{a_n \lambda_j(n)}{\sqrt{n}} \right|^2,
\]

where \(a_n = A(1, n)\) occur as Fourier coefficients of \(\text{sym}^2 U\). Then the spectral large sieve inequality shows

\[
\sum_{t_j \leq V} \left| \sum_{n \leq N} a_n \lambda_j(n) \right|^2 \ll (VN)^\varepsilon (V^2 + N) \sum_{n \leq N} |a_n|^2.
\]

In our application, we have

\[
\sum_{n \leq N} |A(1, n)|^2 \ll (NT)^\varepsilon \text{ by the convexity bound for Rankin-Selberg } L\text{-functions [Li].}
\]

\[\square\]

5. The geodesic QUE theorem for Eisenstein series

Here we present the proof of our main result, Theorem 1.2. We shall prove the following more precise version of Theorem 1.2 which has a power saving in the error term.

**Theorem 5.1.** Suppose \(\psi\) is a fixed smooth, compactly-supported function on \(\mathbb{R}^+\). Then

\[
\int_0^\infty \psi(y) |E(iy, 1/2 + iT)|^2 \frac{dy}{y} = 2 \langle |E(z, 1/2 + iT)|^2, E(z, y\psi(y)) \rangle + a + b(T) + c + O_\delta(T^{-\delta}),
\]

where \(a\) is defined by (5.22), \(b(T) = O(1)\) is defined by (5.8), \(c\) is defined by (5.23) and \(\delta < 1/33\).

Indeed, they show (see their Proposition 2.2) that

\[
\frac{3}{\pi} \log(1/4 + T^2) \int_0^\infty \psi(y) \frac{dy}{y} + O\left(\frac{\log T}{\log \log T}\right).
\]

Indeed, they show (see their Proposition 2.3) that \(\frac{3}{\pi} \log(1/4 + T^2)\) is the average size of \(|E(z, 1/2 + iT)|^2\) restricted to any fixed compact Jordan measurable subset of \(\Gamma \setminus \mathbb{H}\) having positive measure. Thus Theorem 1.2 follows from Theorem 5.1. See also Theorem 1.4 or the more precise version appearing in Section 4.2.

The proof is long and we have presented some auxiliary results in Sections 6 and 7, so here we map out the strategy. The first step, appearing in Section 5.1, is to use harmonic analysis on the positive reals to relate the integral of the Eisenstein series to a shifted fourth moment of the Riemann zeta function, similarly to the method used in Section 3. The main difference here is that this requires some regularization. In Section 5.2, working on the assumption that the asymptotic formula for the shifted fourth moment with large shifts takes the same form as for small shifts, we relate the main term in the fourth moment to an inner product of the Eisenstein series with an incomplete Eisenstein series, again similarly to the approach in Section 3; this appears as Theorem 5.3. This asymptotic for the fourth moment of zeta requires, as one ingredient, the asymptotic behavior of a shifted divisor sum. We present this part separately as Theorem 7.1; in fact all of Section 7 is devoted to the proof of this estimate. In Section 6, which is also self-contained, we show (loosely speaking) that if one has
an asymptotic for a shifted divisor sum with a power saving, then one can evaluate the fourth moment of zeta including all the lower-order main terms. This amounts to a calculation of various main terms and is a generalization of work of Hughes-Young [HY] which assumed that the shifts are all small. Finally, in Section 5.3 we collect these auxiliary results and show that indeed the shifted fourth moment with large shifts does take the expected form, thus completing the proof of Theorem 5.1.

5.1. Reduction to a fourth moment. The Fourier expansion for Eisenstein series, i.e., (2.2), may be alternatively expressed in the symmetric form

\[ \theta(s)E(z, s) = \theta(s)y^s + \theta(1 - s)y^{1-s} + \ldots \] analogously to the definition of the Hardy Z-function, set \( E(s, z, s) = \frac{\theta(s)}{\theta(1-s)} \). The Fourier expansion of \( E^* \) takes the form

\[ E^*(z, 1/2 + iT) = c_0^*(y) + \rho^*(1) \sum_{n \neq 0} \frac{\tau_T(n)}{\sqrt{|n|}} e(nx)V_T(2\pi|n|y), \]

where now \( c_0^*(y) = \mu y^{1/2+iT} + \pi y^{1/2-iT} \), with

\[ \mu = \frac{\theta(1/2 + iT)}{|\theta(1/2 + iT)|}. \]

Furthermore, by comparison with (2.4), we have

\[ \rho^*(1) = |\rho(1)| = (2/\pi)^{1/2}|\theta(1/2 + iT)|^{-1}, \]

and we see that \( E^*(z, 1/2 + iT) \) is real-valued.

Our main goal in this subsection is to prove the following

**Proposition 5.2.** Let \( I(T; \psi) \) denote the left hand side of (5.1). Then

\[ I(T; \psi) = 4 \frac{\rho^*(1)^2}{\cosh(\pi T)} \frac{1}{2\pi i} \int_{(0)} \tilde{\psi}(-v)M(v)dv + b(T) + O(T^{-1/12+\varepsilon}), \]

where

\[ M(v) = \frac{\cosh(\pi T)}{2\pi i} \int_{(0)} L(1/2 + s + v, E_T)L(1/2 - s, E_T)\gamma_{V_T}(1/2 + s + v)\gamma_{V_T}(1/2 - s)ds. \]

and with \( \mu \) defined above by (5.4), we set

\[ b(T) = 2 \int_0^\infty \psi(y)(y + y^{-1} + \mu^2 + \mu^{-2}) \frac{dy}{y}. \]

Remark. We can interpret \( b(T) \) alternatively via

\[ b(T) = \int_0^\infty (c_0^*(y) + c_0^*(1/y))^2 \psi(y) \frac{dy}{y} + O(T^{-100}). \]

Note also that \( M(v) = M(-v) \), which is a symmetry corresponding to the fact that the integrands in \( I(T, \psi) \) and \( b(T) \) are invariant under \( y \rightarrow 1/y \).

**Proof.** As a simple method of regularization, write

\[ I(T; \psi) = \int_0^\infty E_T(iy)^2 \psi(y) \frac{dy}{y} = \int (E_T^* - c_0^*)^2 + 2 \int c_0^*(E_T^* - c_0^*) + \int (c_0^*)^2, \]

which we denote as \( I_1 + I_2 + I_3 \), respectively. We compute the three terms in turn.
A very short calculation shows

\[ I_3 = 2 \int_0^\infty \psi(y) y \frac{dy}{y} + O(T^{-100}), \]

which gives one of the four terms making up \( b(T) \).

Using the Fourier expansion of \( E_T^* \), we have

\[ I_2 = 4 \rho^*(1) \sum_{n \geq 1} \frac{\tau_T(n)}{\sqrt{n}} \int_0^\infty V_T(2\pi ny)(\mu y^{1/2+it} + \bar{\mu} y^{1/2-it}) \psi(y) \frac{dy}{y}. \]

Applying the Mellin inversion formula to \( \psi \), changing variables \( y \rightarrow y/n \), and reversing orders of integration and summation, we obtain (the reader may wish to recall the definitions (2.9) and (2.11))

\[ I_2 = 4 \rho^*(1) \int (1) \tilde{\psi}(-s) \left( \mu L(1 + s + iT, E_T) \gamma_{V_T}(1 + s + iT) \right. \\
+ \bar{\mu} L(1 + s - iT, E_T) \gamma_{V_T}(1 + s - iT) \right) ds. \]

Next we move the contour of integration to \( \text{Re}(s) = -1/2 \), crossing poles at \( s = 0 \), and \( s = \pm 2iT \). The residues at \( s = \pm 2iT \) contribute \( O(T^{-100}) \) to \( I_2 \) since \( \tilde{\psi}(-s) \ll (1 + |s|)^{-200} \).

It will be useful to record that

\[ \text{Res}_{s=0} L(1 + s + iT, E_T) \gamma_{V_T}(1 + s + iT) = \frac{1}{2} \sqrt{\frac{\pi}{2}} \theta(1/2 + iT). \]

We also have that the size of the new contour integral is

\[ \ll T^\varepsilon \int_{|t| \ll T^\varepsilon} |\zeta(1/2 + it)\zeta(1/2 + it + 2iT)|T^{-1/4} dt + O(T^{-100}), \]

and using Weyl’s bound shows this term is \( O(T^{-1/12+\varepsilon}) \). Gathering these estimates, and using (5.5), we have

\[ I_2 = 2 \int_0^\infty \psi(y)(\mu^2 + \bar{\mu}^2) \frac{dy}{y} + O(T^{-1/12+\varepsilon}), \]

which contributes two of the four terms making up \( b(T) \).

Finally we examine \( I_1 \). By Lemma 2.1 (the Mellin convolution theorem) and (2.7),

\[ I_1 = \frac{4\rho^*(1)^2}{(2\pi i)^2} \int (2) \tilde{\psi}(-v) \\
\int (1) L(1/2 + s, E_T)L(1/2 + v - s, E_T) \gamma_{V_T}(1/2 + s) \gamma_{V_T}(1/2 + v - s) ds dv. \]

As a first step, we move the \( s \)-integral to the line \( \text{Re}(s) = 0 \), crossing poles at \( s = 1/2 \pm iT \) (and only those, since \( \text{Re}(v) = 2 \) at this time). We defer treatment of these residues for a moment. For the new integral, we reverse the orders of integration and move the \( v \)-integral to the 0-line, crossing poles at \( v = 1/2 + s \pm iT \). The new double integral equals the integral main term in Proposition 5.2 after changing variables \( s \rightarrow -s \). The residues of these poles at \( v = 1/2 + s \pm iT \) contribute \( O(T^{-1/12+\varepsilon}) \) to \( I_1 \), by a calculation very similar to (5.15).
By (5.14), the residues at $s = 1/2 + iT$ give

$$
\frac{2\rho^*(1)^2}{2\pi i} \sqrt{\frac{\pi}{2}} \sum_{\pm} \theta(1/2 \pm iT) \int_{(2)} \tilde{\psi}(-v)L(v \mp iT, E_T)\gamma_{V_T}(v \mp iT)dv.
$$

We shift this contour to Re($v$) = 1/2, crossing poles at $v = 1$ and $v = 1 \pm 2iT$. The residues with $v = 1 \pm 2iT$ are very small by the rapid decay of $\tilde{\psi}$, and using (5.14) again as well as (5.5), the residues at $v = 1$ give

$$
\rho^*(1)^2 \frac{\pi}{2} \sum_{\pm} \theta(1/2 \pm iT)\theta(1/2 \mp iT)\tilde{\psi}(-1) = 2 \int_{0}^{\infty} \psi(y)y^{-1}dy,
$$

which is the remaining term of $b(T)$. The new integral along Re($v$) = 1/2 is again $O(T^{-1/12+\varepsilon})$ by Weyl’s bound.

We have accounted for all the terms of (5.6), so the proof is complete. □

5.2. The main term. Recall (2.5), that is, $\rho^*(1)^2 = \cosh(\pi T)\Theta^{(1)}$, and let

$$
I_J(T; \psi) = 4\frac{\rho^*(1)^2}{\cosh(\pi T)} \frac{1}{2\pi i} \int_{(0)} \tilde{\psi}(-v)M(v)dv.
$$

Since $\tilde{\psi}(-v)$ has rapid decay, understanding $I_J(T; \psi)$ (and hence $I(T; \psi)$) is, to first approximation, the same as understanding $M(v)$ for a fixed $v$.

**Theorem 5.3.** We have the asymptotic

$$
I_J(T; \psi) = 2\langle |E(z, 1/2 + iT)|^2, E(z, y\psi(y)) \rangle + a + c + O(T^{-1/33+\varepsilon}),
$$

where

$$
a = -2 \int_{0}^{\infty} (y + y^{-1})\psi(y)\frac{dy}{y},
$$

and where with $f(z) = y^{1/2}|\eta(z)|^2$ (so in particular this $f$ is $\Gamma$-invariant), we define

$$
c = -4 \int_{0}^{\infty} \psi(y)(y + y^{-1} + \frac{3}{\pi}\log f(iy))\frac{dy}{y}.
$$

We deduce Theorem 5.1 from Proposition 5.2 and Theorem 5.3, and Theorem 1.2 follows from Theorem 5.1 via (5.2). See also Section 4.2 for a more precise form of the main term (compared to (5.2)).

We begin by indicating how this main term emerges by a rigorous analog of the work in Section 3. First observe that $M(v)$ is a weighted shifted fourth moment of zeta. That is, it takes the form

$$
M(v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \zeta(1/2 + \alpha + iT)\zeta(1/2 + \beta + iT)\zeta(1/2 + \gamma - iT)\zeta(1/2 + \delta - iT)w(t)dt,
$$

where

$$
\alpha = v + iT, \quad \beta = v - iT, \quad \gamma = iT, \quad \delta = -iT,
$$

Re($v$) = 0, and we may assume Im($v$) $\ll T^\varepsilon$. Furthermore, via (2.9),

$$
w(t) = \frac{\cosh(\pi T)}{2^3\pi^{\frac{\alpha + \beta + \gamma + \delta}{2}}} \Gamma\left(\frac{\frac{1}{2} + \alpha + iT}{2}\right)\Gamma\left(\frac{\frac{1}{2} + \beta + iT}{2}\right)\Gamma\left(\frac{\frac{1}{2} + \gamma - iT}{2}\right)\Gamma\left(\frac{\frac{1}{2} + \delta - iT}{2}\right).
$$
As stated in Theorem 6.1 below, (which is the large shift analog of Theorem 1.1 of [HY]),
the main term of (5.24) is the sum of six terms. In this paper we shall extend this result
to hold for shifts of the form (5.25) and $T$ very large; see Theorem 5.4 below. The full main
term in (6.8) is holomorphic in terms of the shift parameters as long as they are all in the
strip $-1/2 < \text{Re}(z) < 1/2$, even though each of the terms in the sum has poles. One of the
six terms is

$$
\mathcal{M}_0(v) := \frac{\zeta(1 + \alpha + \gamma)\zeta(1 + \alpha + \delta)\zeta(1 + \beta + \gamma)\zeta(1 + \beta + \delta)}{\zeta(2 + \alpha + \beta + \gamma + \delta)} \frac{1}{2\pi} \int_{-\infty}^{\infty} w(t)dt.
$$

(5.27)

To match more closely with the arguments in Section 3, it is helpful to know that the term
(5.27) arises from the “first part” of the approximate functional equation and taking the
diagonal analogously to (3.4). One can see this by following the proof of Theorem 6.1,
specifically by looking at the residue at $s = 0$ of (6.10). In this way, we have for $\text{Re}(v) > 0$,

$$
\mathcal{M}_0(v) = \sum_{n=1}^{\infty} \tau_T(n^2) n^{1+v} \frac{1}{2\pi i} \int_{(\varepsilon)} \gamma_n \gamma_T(1/2 + s + v) \gamma_T(1/2 - s) ds,
$$

and the formula extends by meromorphic continuation. Note the similarity to (3.4). To work
with only $\mathcal{M}_0(v)$ we cannot integrate along the line $\text{Re}(v) = 0$ because it passes through
poles. Therefore in (5.20) we first shift the contour slightly to the right to $\text{Re}(v) = \varepsilon$ with $\varepsilon < 1/2$, and then insert (5.28) into (5.20). Using (2.18) and (2.15), we then obtain

$$
4 \frac{\rho^*(1)^2}{\cosh(\pi T)} \frac{1}{2\pi i} \int_{(\varepsilon)} \tilde{\psi}(-v) \mathcal{M}_0(v) dv = 4|\rho(1)|^2 \frac{1}{2\pi i} \int_{(\varepsilon)} \tilde{\psi}(-v) Z(1 + v) \gamma_T(1 + v) dv.
$$

(5.29)

By (2.27), with $h(y) = y\psi(y)$ so that $\tilde{h}(-1 - s) = \tilde{\psi}(-s)$, we can recognize this as the pretty
formula

$$
\frac{4\rho^*(1)^2}{\cosh(\pi T)} \frac{1}{2\pi i} \int_{(1)} \tilde{\psi}(-v) \mathcal{M}_0(v) dv = 2|E(z, 1/2 + iT)|^2, E(z, y\psi(y))) - 2 \int_0^{\infty} c_0^*(y) \psi(y) \frac{dy}{y}.
$$

(5.30)

This second integral we can write more symmetrically, using the assumption $\psi(y) = \psi(y^{-1}), as

$$
-\int_0^{\infty} (c_0^*(y)^2 + c_0^*(y^{-1})^2) \psi(y) \frac{dy}{y} = -2 \int_0^{\infty} (y + y^{-1}) \psi(y) \frac{dy}{y} + O(T^{-100}).
$$

(5.31)

One immediately reads off the inner product appearing in (5.21) as well as the constant $a$
defined by (5.22).

Next we shall analyze the other five terms making up the main term of (5.24), which are
given in the general form by Theorem 6.1. We need to show that these give the constant $c$
up to a satisfactory error term. The underlying principle is that the original fourth moment
has symmetries arising from applying the functional equation to one of $\zeta(1/2 + \alpha + it)$ or
$\zeta(1/2 + \beta + it)$, and one of $\zeta(1/2 + \gamma - it)$, $\zeta(1/2 + \delta - it)$ (or both). There are five such
symmetries. If we apply the functional equation to $\zeta(1/2 + \alpha + it)$ and to $\zeta(1/2 + \gamma - it)$,
for instance, then this amounts to switching $\alpha$ with $-\gamma$ and multiplying by $X_{\alpha,\gamma,t}$, in the
notation of Section 6. However, our calculations are simplified here by noting that the weight
function $w(t)$ also depends on the shifts and in fact $w(t)$ times the product of four zetas gives
the completed zeta functions which are then invariant under changes of variable $\alpha \leftrightarrow -\gamma$.

\[ \text{Of course we need to show that the error term is indeed smaller than this “main term.”} \]
etc. That is, the other five terms are obtained by changing variables (1) \( \alpha \leftrightarrow -\gamma \) (meaning \( \alpha \) is replaced by \(-\gamma \) and \( \gamma \) is replaced by \(-\alpha \)) or (2) \( \alpha \leftrightarrow -\delta \) or (3) \( \beta \leftrightarrow -\gamma \) or (4) \( \beta \leftrightarrow -\delta \) or (5) \( \alpha \leftrightarrow -\gamma \) and \( \beta \leftrightarrow -\delta \).

We need to explicitly evaluate the \( t \)-integral in terms of \( \alpha, \beta, \gamma, \delta \). By (6.412) of [GR],

\[
(5.32) \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} w(t) dt = \frac{\cosh(\pi T)}{2\pi} \frac{\Gamma\left(\frac{1+\alpha+\gamma}{2}\right)\Gamma\left(\frac{1+\alpha+\delta}{2}\right)\Gamma\left(\frac{1+\beta+\gamma}{2}\right)\Gamma\left(\frac{1+\beta+\delta}{2}\right)}{\Gamma\left(\frac{2}{2}\alpha+\beta+\gamma+\delta\right)}.
\]

Letting \( \Lambda(s) = \pi^{-s/2} \Gamma(s/2) \zeta(s) \), and inserting (5.32) into (5.27), we derive

\[
(5.33) \quad \frac{M_0(v)}{\cosh(\pi T)} = \frac{\pi \Lambda(1+\alpha+\gamma)\Lambda(1+\alpha+\delta)\Lambda(1+\beta+\gamma)\Lambda(1+\beta+\delta)}{4\Lambda(2+\alpha+\beta+\gamma+\delta)}.
\]

Thus we arrive at

\[
(5.34) \quad I_j(T; \psi) = 4 - \frac{1}{2\pi i} \int_{(\epsilon)} \tilde{\psi}(-v) \frac{\rho^*(1)^2}{\cosh(\pi T)} \sum_{k=0}^{5} M_k(v) dv + \mathcal{E}(T)
\]

where \( \mathcal{E}(T) \) is an error term that we shall estimate with Theorem 5.4 (showing \( \mathcal{E}(T) \ll T^{-1/33+\epsilon} \)), and \( \sum_{k=0}^{5} \frac{4}{\pi \cosh(\pi T)} M_k(v) \) is tediously calculated to be (applying the changes of variables (1)-(5) in terms of the shift parameters \( \alpha, \beta, \gamma, \delta \), and then substituting back into \( v \) and \( T \) with (5.25))

\[
(5.35) \quad \frac{\Lambda(1+v+2iT)\Lambda(1+v)^2\Lambda(1+v-2iT)}{\Lambda(2+2v)} + \frac{\Lambda(1-v-2iT)\Lambda(1-2iT)^2\Lambda(1+v-2iT)}{\Lambda(2-4iT)}
\]

\[
+ \frac{\Lambda(1+2iT)\Lambda(1+v)\Lambda(1-v)\Lambda(1-2iT)}{\Lambda(2)} + \frac{\Lambda(1+2iT)\Lambda(1+2iT)^2\Lambda(1-v+2iT)}{\Lambda(2+4iT)} + \frac{\Lambda(1-v-2iT)\Lambda(1-v)^2\Lambda(1-v+2iT)}{\Lambda(2-2v)}.
\]

Here the displayed terms are respective to the index \( k \) in \( M_k(v) \) which in turn is with respect to the labelling of the changes of variable in the paragraph immediately preceding (5.32). As a consistency check, we remark that (5.35) is symmetric under \( v \to -v \), as it should be, recalling that \( M(v) = M(-v) \). It is holomorphic for \(-1/2 < \text{Re}(v) < 1/2\), and Stirling’s formula shows that \( M_k(v) \) is bounded by a polynomial in \( v \) and \( T \).

Let \( \mathcal{R}(T) \) denote the integral on the right hand side of (5.34). We will presently show

\[
(5.36) \quad \mathcal{R}(T) = 2 |\langle E(z, 1/2 + iT) \rangle|^2 + |\langle E(z, y\psi(y)) \rangle| + a + c + O(T^{-1/6+\epsilon}).
\]

Write \( \mathcal{R}(T) = \sum_{k=0}^{5} \mathcal{R}_k(T) \) according to the sum in (5.34). We have already seen that \( \mathcal{R}_0(T) \) gives the inner product and the constant \( a \) on the right hand side of (5.36). We will presently show that \( \mathcal{R}_k(T) \) for \( k = 1, 4, 5 \) is bounded by \( O(T^{-1/6+\epsilon}) \), while \( \mathcal{R}_2(T) + \mathcal{R}_3(T) \) give the constant \( c \) defined by (5.23).

By a trivial estimation, for \( k = 1, 4 \) we have \( \mathcal{R}_k(T) \ll T^{-1/2+\epsilon} \). This bound arises from the fact that \( \Lambda(2 \pm 4iT) \gg T^{1/2} \exp(-\pi T) \). For \( \mathcal{R}_5(T) \), we move the contour to \( \text{Re}(v) = \frac{1}{2} \), and use \( \Lambda(\frac{1}{2} + it) \ll t^{-1/12+\epsilon} \) by Weyl’s bound \( \zeta(1/2 + it) \ll t^{1/6+\epsilon} \). In this way we have

\[
(5.37) \quad \int_{(\epsilon)} \tilde{\psi}(-v) \frac{\rho^*(1)^2}{\cosh(\pi T)} M_5(v) dv = \int_{(\frac{1}{2})} \tilde{\psi}(-v) \frac{\rho^*(1)^2}{\cosh(\pi T)} M_5(v) dv \ll T^{-1/6+\epsilon}.
\]
For $\mathcal{R}_2(T) = \mathcal{R}_3(T)$, by direct substitution we have
\begin{equation}
\mathcal{R}_2(T) + \mathcal{R}_3(T) = 2\pi |\rho^*(1)|^2 \frac{\Lambda(1+2iT)\Lambda(1-2iT)}{\Lambda(2)} \frac{1}{2\pi i} \int_{(c)} \tilde{\psi}(-v)\Lambda(1+v)\Lambda(1-v)dv.
\end{equation}
We work with this inner integral. By the functional equation,
\begin{equation}
\frac{1}{2\pi i} \int_{(c)} \tilde{\psi}(-v)\Lambda(1+v)\Lambda(1-v)dv = \frac{1}{2\pi i} \int_{(c)} \tilde{\psi}(-v)\Lambda(1+v)\Lambda(v)dv,
\end{equation}
which equals
\begin{equation}
\int_0^\infty \psi(y)\frac{1}{2\pi i} \int_{(c)} \Lambda(1+v)\Lambda(v)y^{-v}dv\frac{dy}{y} = \int_0^\infty \psi(y)\left(-\Lambda(2)y^{-1} + \frac{1}{2\pi i} \int_{(2)} \Lambda(1+v)\Lambda(v)y^{-v}dv\right)\frac{dy}{y}.
\end{equation}
Then we write $\Lambda(1+v)\Lambda(v) = \Gamma_R(1+v)\Gamma_R(v)\zeta(1+v)\zeta(v)$, use the identity $\Gamma_R(1+v)\Gamma_R(v) = \Gamma_C(v) = 2(2\pi)^{-\frac{1}{2}}\Gamma(v)$, and reverse the orders of summation and integration, so we see that (5.39) equals
\begin{equation}
\int_0^\infty \psi(y)\left(-\Lambda(2)y^{-1} + 2 \sum_{m,n} m^{-1} \exp(-2\pi mny)\right)\frac{dy}{y}.
\end{equation}
The inner sum equals $2 \sum_{n \geq 1} \sigma_1(n) \exp(-2\pi mny)$ which is $D(iy)$ defined by (22.64) of [IK]. This function $D(iy)$ makes up the tail of the Fourier expansion in the constant term of the Laurent expansion of the Eisenstein series at $s = 1$. Inserting this calculation into (5.38), we have
\begin{equation}
\mathcal{R}_2(T) + \mathcal{R}_3(T) = 2\pi \rho^*(1)^2 \frac{\Lambda(1+2iT)\Lambda(1-2iT)}{\Lambda(2)} \int_0^\infty \psi(y)(-\Lambda(2)y^{-1} + D(iy))\frac{dy}{y}.
\end{equation}
By (22.68) of [IK], $D(iy) = -\frac{\pi}{6} y + \frac{1}{2} \log y - \frac{1}{2} \log f(iy)$ where $f(z) = y^{1/2}|\eta(z)|^2$ is $\Gamma$-invariant. Also since log is odd under $y \to y^{-1}$, and we assumed $\psi$ is even, (5.42) then becomes
\begin{equation}
-2\pi \rho^*(1)^2 \frac{\Lambda(1+2iT)\Lambda(1-2iT)}{\Lambda(2)} \int_0^\infty \psi(y)(y^{-1} + y + \frac{1}{2\Lambda(2)} \log f(iy))\frac{dy}{y}.
\end{equation}
Using (5.5) and $\theta(1/2 + iT) = \Lambda(1+2iT)$, we have
\begin{equation}
\mathcal{R}_2(T) + \mathcal{R}_3(T) = -4 \int_0^\infty \psi(y)(y + y^{-1} + \frac{3}{\pi} \log f(iy))\frac{dy}{y},
\end{equation}
which is the constant $c$ defined by (5.23). This finishes the proof of (5.36).

5.3. The fourth moment of the zeta function with large shifts.

Theorem 5.4. Let $M(v)$ be defined by (5.24), with shifts as in (5.25). Then
\begin{equation}
M(v) = M.T. + O(T^{-\frac{1}{3} + \varepsilon}),
\end{equation}
where $M.T. = \sum_{k=0}^{5} M_k(v)$ denotes the main term defined by (5.33) and following discussion.
We begin the analysis of $M(v)$ by describing the properties of the weight function $w(t)$ defined by (5.26). By Stirling’s approximation,

$$w(t) \ll_v \exp\left(\frac{2}{7}Q'(t, v, T)(1 + |t - T|)^{-1/2}(1 + |t + T|)^{-1/2}\right),$$

where the implied constant depends polynomially on $v$, and

$$Q'(t, v, T) = 4T - |v + t + T| - |v + t - T| - |t - T| - |t + T|.$$

We claim that $Q'(t, v, T) \leq 0$ for all $t, v, T \in \mathbb{R}$. For a given $t$ and $T$, the maximum of $Q'$ in terms of $v$ must occur at $v = -t - T$ or $v = -t + T$. At $v = -t - T$ (which it suffices to check by symmetry), $Q'$ specializes as

$$2T - |t - T| - |t + T|,$$

which happens to equal $Q(t, T/2)$ where recall we originally encountered $Q(t, T)$ in (4.5). We know $Q(t, T/2) \leq 0$ for all $t, T$ and $Q(t, T/2) \leq 2(T - |t|)$ for $|t| \geq T$. Thus for any $v \in \mathbb{R}$ and $|t| \geq T$, we have

$$Q'(t, v, T) \leq 2(T - |t|).$$

By logarithmic differentiation and the asymptotic expansion of $\Gamma(z) = \log(z) + \frac{a_1}{z} + \ldots$, one can derive the bounds

$$w^{(j)}(t) \ll_v (1 + |t - T|)^{-1/2}(1 + |t + T|)^{-1/2}\left(\frac{1}{1 + |t - T|} + \frac{1}{1 + |t + T|}\right)^j.$$

Let $M(v) = M_+(v) + M_-(v)$, according to $t \geq 0$ or $t \leq 0$. By symmetry, we focus on $M_+(v)$. Next we apply a smooth partition of unity supported on sets of the form $T - 2\Delta \leq t \leq T - \Delta$ with $1 \leq \Delta \leq T/2$, as well as the region $t \geq T - 1$. Accordingly, write $M_+(v) = \sum\Delta M_\Delta(v) + (T^{-1/6+\varepsilon})$, where we now explain the origin of this error term. It accounts for $0 \leq t \leq 1$, say, and $t \geq T - 1$. For $t = O(1)$, the contribution to $M_+(v)$ is $\ll (T^{1/6+\varepsilon})^4T^{-1} \ll T^{-1/3+\varepsilon}$, by Weyl’s bound and the fact that $w(t) \ll T^{-1}$ for such $t$. For $T - 1 \leq t \leq T + T^\varepsilon$, the error is $\ll (T^{1/6+\varepsilon})^2T^{-1/2} \ll T^{-1/6+\varepsilon}$, by the same type of reasoning, except here $w(t) \ll T^{-1/2}$, and only two of the zeta functions are evaluated at height $T$, the other two having height $T^\varepsilon$. The exponential decay of $w$ overwhelms the polynomial growth of the zeta functions for $t \geq T + T^\varepsilon$, using (5.49).

We need to treat $M_\Delta(v)$ in two different ways depending on the size of $\Delta$.

**Lemma 5.5.** We have

$$M_\Delta(v) \ll T^{-1/6+\varepsilon} + \Delta^{1/2}T^{-1/2+\varepsilon}.$$

This bound is satisfactory for $\Delta \ll T^{1-\delta}$. For $\Delta$ large, we need to extract the main term, and to this end we have

**Lemma 5.6.** We have

$$M_\Delta(v) = M.T.(\Delta) + O(T^\varepsilon(T^{25/12}\Delta^{-9/4} + T^{13/24}\Delta^{-5/8}).$$

Here the main term is a sum of the form $\sum_{k=0}^5 M_k^{(\Delta)}(v)$ where for example $M_0^{(\Delta)}$ is defined by (5.27), but with $w$ multiplied by the appropriate constituent of the partition of unity.

Here the main term depends on $\Delta$, but is bounded from above by $\Delta^{1/2}T^{-1/2+\varepsilon}$. Thus in Lemma 5.5 we can freely claim that $M_\Delta(v) = M.T.(\Delta) + (\text{error})$. Also, $\sum_\Delta M.T.(\Delta)$ forms the main term in Theorem 5.4, as summing over $\Delta$ simply eliminates the partition of unity.
Taking the optimal choice of using Lemma 5.5 for $\Delta \leq T^{31/33}$, and Lemma 5.6 for $\Delta \geq T^{31/33}$, we obtain the error term stated in Theorem 5.4.

**Proof of Lemma 5.5.** The key tool here is Iwaniec’s [Iw1] upper bound on the fourth moment:

\[(5.53) \int_X^{X+Y} |\zeta(1/2 + it)|^4 dt \ll (X^{2/3} + Y)(XY)^{\epsilon}.\]

By applying Hölder’s inequality to $M_\Delta(v)$ with exponents $4, 4, 4, 4$, and using $w(t) \ll (T\Delta)^{-1/2}$ on this range of $t$, we obtain

\[(5.54) M_\Delta(v) \ll (\Delta T)^{-1/2} \max_{|u| \leq T^\epsilon} |\zeta(1/2 + v + 2iT + iu)\zeta(1/2 + v + iu)^2\zeta(1/2 + v - 2iT + iu)|.\]

By Weyl’s bound, $E_1 \ll T^{-2/3+\epsilon}$.

The other error term $E_2$ arises as the error term in the shifted divisor problem. The off-diagonal terms are given by (6.14) below, which we write here as

\[(5.57) M_\Delta^{OD}(v) := \sum_{r \neq 0} \sum_{m-n=r} \frac{\sigma_{\alpha,\beta}(m)\sigma_{\gamma,\delta}(n)}{(mn)^{1/2}} f(m,n),\]

with

\[(5.58) f(x,y) = \int_{-\infty}^{\infty} \left( \frac{x}{y} \right)^{-it} V_{\alpha,\beta,\gamma,\delta,t}(\pi^2 xy) w_\Delta(t) dt,\]

and where the notation is given below by (6.4), (6.3). By a short calculation, using (5.25), we have that

\[(5.59) \sigma_{\alpha,\beta}(n) = n^{-v} \tau_T(n), \quad \sigma_{\gamma,\delta}(n) = \tau_T(n).\]

We wish to apply Theorem 7.1, which is an asymptotic formula for a shifted divisor sum. For this, we need to know the sizes of the derivatives of $f$. First we note that Stirling’s formula implies

\[(5.60) g_{\alpha,\beta,\gamma,\delta}(s,t) \ll (\Delta T)^{s},\]
for $\text{Im}(s) \ll T^\varepsilon$, for $t$ in the support of $w_\Delta$. The weight function $V$ decays quickly for $xy \geq (\Delta T)^{1+\varepsilon}$ (as this is the square-root of the conductor), by shifting the contour far to the right if necessary, using (5.60). In terms of $t$, we have

\begin{equation}
(5.61) \quad \frac{\partial}{\partial t} V_{\alpha,\beta,\gamma,\delta,t}(xy) \ll_{j,s} (xy)^{\varepsilon j} \Delta^{-j},
\end{equation}

where the implied constant depends polynomially on $s$; this is a slight generalization of (5.50). Hence $H(t)$ defined by $H(t) = V_{\alpha,\beta,\gamma,\delta,t}(xy)w_\Delta(t)$ satisfies a bound similar to (5.50), except bigger by a small factor $(xy)^{\varepsilon j}$ (which has no practical effect for our work). Furthermore, by integration by parts, $f(x,y)$ is small unless $\Delta|\log(x/y)| \leq T^\varepsilon$. We can conclude that $f$ is very small unless $|x-y| \ll \Delta^{-1/2}T^{1/2+\varepsilon}$. Furthermore, since $|x-y| \geq 1$, we have that $f$ is very small unless $x,y \gg \Delta T^{-\varepsilon}$.

By these observations, we have

\begin{equation}
(5.62) \quad M_\Delta^{OD}(v) = \sum_{r \neq 0} \sum_{n \geq 1} \tau_{IT}(n) \tau_{IT}(n+r) F(n),
\end{equation}

where $F(n) = F_{r,v}(n) = n^{-v}f(n,n+r)$ is a smooth function satisfying the following bound

\begin{equation}
(5.63) \quad \frac{d^j}{dx^j} F(x) \ll (1 + |v|)^j x^{-1-j} T^{-1/2} \Delta^{1/2}.
\end{equation}

In order to apply Theorem 7.1 to the shifted convolution sum appearing in (5.62), we need two minor modifications to meet the conditions of the theorem. Firstly, we need to apply a dyadic partition of unity to say $[N,2N]$, summing over $1 \ll N \ll (\Delta T)^{1/2+\varepsilon}$, and then we need to extract the constant factor $N^{-1}T^{-1/2} \Delta^{1/2}$ from the weight function (a simple re-scaling of the weight function). Having done this, we obtain

\begin{equation}
(5.64) \quad M_\Delta^{OD}(v) = M.T. + E.T.,
\end{equation}

where, in the notation of Theorem 7.1, we have $Y = N$, $P = T^\varepsilon$, $R \ll \frac{|r|T}{N} T^\varepsilon$. Thus

\begin{equation}
(5.65) \quad E.T. \ll \sum_{N \text{ dyadic}} \frac{\Delta^{1/2}T^\varepsilon}{T^{1/2}N} \sum_{1 \leq |r| \ll \frac{N}{\Delta T^\varepsilon}} \left( T^{1/3}N^{1/2} \left( \frac{|r|T}{N} \right)^2 + T^{1/6}N^{3/4} \left( \frac{|r|T}{N} \right)^{1/2} \right).
\end{equation}

We do not need to work with the main term at this point because that is the purview of Theorem 6.1. To simplify (5.65), it is useful to notice that the “worst” value of $N$ is at $N = \Delta^{1/2}T^{1/2+\varepsilon}$ (the maximal value we need to consider), since after summing over $r$, $N$ does not appear in a denominator. Thus, after simplifications, we obtain

\begin{equation}
(5.66) \quad E.T. \ll T^\varepsilon (T^{25/12} \Delta^{-9/4} + T^{13/24} \Delta^{-5/8}),
\end{equation}

which is the error claimed in Lemma 5.6. This term $E.T.$ is precisely one of the two terms making up $E_2$ in Theorem 6.1, while the other term in $E_2$ has a bound of the same size by symmetry.

6. The main terms in the shifted fourth moment of the zeta function

In [HY], we proved an asymptotic formula for the twisted fourth moment of the Riemann zeta function, with small shifts. One standard approach to this type of problem is to reduce it to an asymptotic evaluation of a shifted divisor problem. The main term in the fourth moment comes about through a complicated process of matching of terms. The method used in [HY] to simplify the main terms unfortunately does not carry over to large shift
parameters because we used Stirling’s approximation at different stages of the calculations. Here we modify the calculations of the main terms of [HY] to carry over to arbitrary shifts. This section is independent from the rest of this paper but heavily relies on the work of [HY]. For the purposes of this paper, we do not require the full “twisted” version, so the reader may wish to consider the simpler special case $h = k = 1$ below.

Begin with some smooth weight function $w$ having compact support, and let

$$I(h, k) = \int_{-\infty}^{\infty} \left( \frac{h}{k} \right)^{-it} \zeta \left( \frac{1}{2} + \alpha + it \right) \zeta \left( \frac{1}{2} + \beta + it \right) \zeta \left( \frac{1}{2} + \gamma - it \right) \zeta \left( \frac{1}{2} + \delta - it \right) w(t) dt,$$

where $(h, k) = 1$. We set some more notation. Let $\sigma_{\alpha, \beta}(n) = \sum_{ab=n} a^{-\alpha} b^{-\beta}$. Define

$$X_{\alpha, \beta, \gamma, \delta, t} = \pi^{\alpha+\beta+\gamma+\delta} \frac{\Gamma \left( \frac{1}{2} - \alpha - it \right) \Gamma \left( \frac{1}{2} - \beta - it \right) \Gamma \left( \frac{1}{2} - \gamma + it \right) \Gamma \left( \frac{1}{2} - \delta + it \right)}{\Gamma \left( \frac{1}{2} + \alpha + it \right) \Gamma \left( \frac{1}{2} + \beta + it \right) \Gamma \left( \frac{1}{2} + \gamma - it \right) \Gamma \left( \frac{1}{2} + \delta - it \right)},$$

and similarly write $X_{\alpha, \beta, \gamma, \delta, t} = X_{\alpha, \gamma, t} X_{\beta, \delta, t}$ (with hopefully obvious meaning). Let

$$g_{\alpha, \beta, \gamma, \delta}(s, t) = \frac{\Gamma \left( \frac{1}{2} + \alpha + s + it \right) \Gamma \left( \frac{1}{2} + \beta + s + it \right) \Gamma \left( \frac{1}{2} + \gamma - s + it \right) \Gamma \left( \frac{1}{2} + \delta - s + it \right)}{\Gamma \left( \frac{1}{2} + \alpha + it \right) \Gamma \left( \frac{1}{2} + \beta + it \right) \Gamma \left( \frac{1}{2} + \gamma - it \right) \Gamma \left( \frac{1}{2} + \delta - it \right)},$$

and with $G(s) = e^{s^2}$ (or any even holomorphic function with rapid decay in vertical strips, satisfying $G(0) = 1$), set

$$V_{\alpha, \beta, \gamma, \delta, t}(x) = \frac{1}{2\pi i} \int_{(1)} \frac{G(s)}{s} g_{\alpha, \beta, \gamma, \delta}(s, t) x^{-s} ds.$$

Let

$$A_{\alpha, \beta, \gamma, \delta}(s) = \frac{\zeta(1 + s + \alpha + \gamma) \zeta(1 + s + \alpha + \delta) \zeta(1 + s + \beta + \gamma) \zeta(1 + s + \beta + \delta)}{\zeta(2 + 2s + \alpha + \beta + \gamma + \delta)},$$

and supposing $p^{h_T} | h$ and $p^{k_T} | k$, we define

$$B_{\alpha, \beta, \gamma, \delta, h, k}(s) = \prod_{p | h} \left( \frac{\sum_{j=0}^{\infty} \sigma_{\alpha, \beta}(p^j) \sigma_{\gamma, \delta}(p^{j+h_T}) p^{-j(s+1)}}{\sum_{j=0}^{\infty} \sigma_{\alpha, \beta}(p^j) \sigma_{\gamma, \delta}(p^j) p^{-j(s+1)}} \right) \times \prod_{p | k} \left( \frac{\sum_{j=0}^{\infty} \sigma_{\alpha, \beta}(p^{j+k_T}) \sigma_{\gamma, \delta}(p^{j}) p^{-j(s+1)}}{\sum_{j=0}^{\infty} \sigma_{\alpha, \beta}(p^{j}) \sigma_{\gamma, \delta}(p^j) p^{-j(s+1)}} \right).$$

Let

$$Z_{\alpha, \beta, \gamma, \delta, h, k}(s) = A_{\alpha, \beta, \gamma, \delta}(s) B_{\alpha, \beta, \gamma, \delta, h, k}(s).$$

Our result is the following.

**Theorem 6.1.** Let $T > 0$, and suppose that $w(t)$ is supported on $|t| \leq T$. Assume that the shifts $\alpha, \beta, \gamma, \delta$ have real part $O(1/\log T)$, that $|\alpha + it|, |\beta + it|, |\gamma - it|, |\delta - it| \gg T^\varepsilon$ for all $t$ in the support of $w$, and that all the shifts have imaginary parts bounded by some fixed
polynomial in $T$. Then

$$I(h, k) = \frac{1}{\sqrt{hk}} \int_{-\infty}^{\infty} w(t) \left( Z_{\alpha, \beta, \gamma, \delta, h, k}(0) + X_{\alpha, \beta, \gamma, \delta, t}Z_{-\gamma, -\delta, -\alpha, -\beta, h, k}(0) 
+ X_{\alpha, \gamma, t}Z_{-\gamma, -\alpha, -\delta, h, k}(0) + X_{\alpha, \delta, t}Z_{-\beta, \gamma, -\alpha, h, k}(0) 
+ X_{\beta, \gamma, t}Z_{\alpha, -\gamma, -\delta, h, k}(0) + X_{\beta, \delta, t}Z_{\alpha, -\delta, -\beta, h, k}(0) \right) dt + E_1 + E_2 + O_A(T^{-A}),$$

where $E_1$ is bounded by (6.13) below, and $E_2$ is defined by (6.18) below.

We do not attempt to bound the error terms with large shifts since our purpose here is solely the calculation of the main terms.

**Proof.** As in (37) of [HY] (the approximate functional equation), we have

$$I(h, k) = \sum_{m,n} \frac{\sigma_{\alpha, \beta}(m)\sigma_{\gamma, \delta}(n)}{\sqrt{mn}} \int_{-\infty}^{\infty} \left( \frac{hm}{kn} \right)^{-it} V_{\alpha, \beta, \gamma, \delta, t}(\pi^2 mn) w(t) dt$$

$$+ \sum_{m,n} \sigma_{-\gamma, -\delta}(m)\sigma_{-\alpha, -\beta}(n) \int_{-\infty}^{\infty} \left( \frac{hm}{kn} \right)^{-it} X_{\alpha, \beta, \gamma, \delta, t}V_{-\gamma, -\delta, -\alpha, -\beta, t}(\pi^2 mn) w(t) dt,$$

and we write $I(h, k) = I^{(1)}(h, k) + I^{(2)}(h, k)$ correspondingly. We will work mainly with $I^{(1)}$ because we can derive analogous formulas for $I^{(2)}$ by first switching $\alpha$ with $-\gamma$ and $\beta$ with $-\delta$, and then replacing $w(t)$ by $X_{\alpha, \beta, \gamma, \delta, t}w(t)$.

Let $I^{(1)}_{D}(h, k)$ denote the contribution to $I^{(1)}(h, k)$ from $hm = kn$. By (46) of [HY], we have

$$I^{(1)}_{D}(h, k) = \frac{1}{\sqrt{hk}} \int_{-\infty}^{\infty} w(t) \frac{1}{2\pi i} \int_{\epsilon} G(s) s (\pi^2 hk)^{-s} g_{\alpha, \beta, \gamma, \delta}(s, t) Z_{\alpha, \beta, \gamma, \delta, h, k}(2s) ds dt.$$

We shift the contour to $\text{Re}(s) = -\frac{1}{4} + \epsilon$, crossing a pole at $s = 0$ as well as four poles at $2s = -\alpha - \gamma$, etc. The pole at $s = 0$ gives the first main term in (6.8). The other four poles give “junk terms” which surely cannot persist in the final answer since they depend on $G(s)$ which is chosen from a wide class of functions. In fact it is possible to choose $G$ to vanish at these four points. However, we do not need to impose this condition. The new contour, which we view as an “error term,” is

$$\ll (hk)^{-1/4+\epsilon} \int_{-\infty}^{\infty} |w(t)| \int_{(-1/4+\epsilon)} G(s) |g_{\alpha, \beta, \gamma, \delta}(s, t)| |A_{\alpha, \beta, \gamma, \delta}(2s)| ds dt,$$

using a divisor-type bound on $B(s)$. By Stirling’s formula applied to $g(s, t)$, we have that this is

$$\ll (hk)^{-1/4+\epsilon} T^s \int_{-\infty}^{\infty} |w(t)| (1 + |\alpha + it|)^{-\frac{1}{4}} (1 + |\beta + it|)^{-\frac{1}{4}} (1 + |\gamma - it|)^{-\frac{1}{4}} (1 + |\delta - it|)^{-\frac{1}{4}}$$

$$\int_{(-1/4+\epsilon)} \frac{|G(s)|}{|s|} |A_{\alpha, \beta, \gamma, \delta}(2s)| ds dt.$$
By the rapid decay of $G(s)$, we bound the second line above by its maximum value (plus an error of size $O(T^{-A})$, which amounts to

\begin{equation}
\ll (hk)^{-1/4+\varepsilon} T^\varepsilon \int_{-\infty}^{\infty} |w(t)| (1 + |\alpha + it|)^{-\frac{1}{2}} (1 + |\beta + it|)^{-\frac{1}{2}} (1 + |\gamma - it|)^{-\frac{1}{2}} (1 + |\delta - it|)^{-\frac{1}{2}} dt
\times \max_{|u| \leq T^\varepsilon} |\zeta\left(\frac{1}{2} + \alpha + \gamma + iu\right) \zeta\left(\frac{1}{2} + \alpha + \delta + iu\right) \zeta\left(\frac{1}{2} + \beta + \gamma + iu\right) \zeta\left(\frac{1}{2} + \beta + \delta + iu\right)|.
\end{equation}

Observe that switching the parameters $\alpha \leftrightarrow -\gamma$, $\beta \leftrightarrow -\delta$, and multiplying by $X_{\alpha,\beta,\gamma,\delta,t}$ (which has absolute value $O(1)$) does not alter the form of this bound, so it is valid also for the “second part” of the approximate functional equation.

Next we look at the off-diagonal terms, which take the form

\begin{equation}
I_{O}^{(1)}(h, k) = \sum_{r \neq 0} \sum_{hm - kn = r} \frac{\sigma_{\alpha,\beta}(m)\sigma_{\gamma,\delta}(n)}{\sqrt{mn}} f(hm, kn),
\end{equation}

with

\begin{equation}
f(x, y) = \int_{-\infty}^{\infty} \left(\frac{x}{y}\right)^{-it} V_{\alpha,\beta,\gamma,\delta,t}(\frac{\pi^2 xy}{hk}) w(t) dt.
\end{equation}

As a working hypothesis, we suppose that we have an asymptotic formula for the shifted divisor sum

\begin{equation}
\sum_{hm - kn = r} \sigma_{\alpha,\beta}(m)\sigma_{\gamma,\delta}(n) F(hm, kn) = N_{\alpha,\beta,\gamma,\delta}(h, k; r; F) + N_{\beta,\alpha,\gamma,\delta}(h, k; r; F) + N_{\alpha,\beta,\gamma,\delta}(h, k; r; F) + N_{\beta,\alpha,\gamma,\delta}(h, k; r; F) + \mathcal{E}_{\alpha,\beta,\gamma,\delta}(h, k; r, F),
\end{equation}

where

\begin{equation}
N_{\alpha,\beta,\gamma,\delta}(h, k; r; F) = \frac{\zeta(1 - \alpha + \beta) \zeta(1 - \gamma + \delta)}{h^{1-\alpha} k^{1-\gamma}} \int_{\text{max}(0, r)}^{\infty} x^{-\alpha}(x-r)^{-\gamma} F(x, x-r) dx
\times \sum_{l=1}^{\infty} \frac{S(r, 0; l)(h, l)^{1-\alpha+\beta}(k, l)^{1-\gamma+\delta}}{l^{2-\alpha-\beta-\gamma+\delta}},
\end{equation}

and $\mathcal{E}$ is a presumed error term, and it is not our purpose here to prove a bound on this $\mathcal{E}$. Here $F(x, y) = (xy)^{-1/2} f(x, y)$. Formally speaking, we have then

\begin{equation}
E_2 = \sum_{r \neq 0} \mathcal{E}_{\alpha,\beta,\gamma,\delta}(h, k; r; (xy)^{-1/2} f(x, y)) + \mathcal{E}_{-\gamma, -\delta, -\alpha, -\beta}(h, k; r; (xy)^{-1/2} f^*(x, y)),
\end{equation}

where $f(x, y) = f_{\alpha,\beta,\gamma,\delta}(x, y)$ is defined by (6.15) and $f^*$ is identical to $f$ but with $\alpha \leftrightarrow -\gamma$, $\beta \leftrightarrow -\delta$, and multiplied by $X_{\alpha,\beta,\gamma,\delta,t}$.

Now define $I_{O,\alpha,\beta,\gamma,\delta}^{(1)} = \sum_{r > 0} N_{\alpha,\beta,\gamma,\delta}(h, k; r, F)$. The arguments of (83)-(87) from [HY] carry over almost without change. There is a small error in (89) which should read

\begin{equation}
K^- = r^{-\alpha - \gamma} \int_{0}^{\infty} x^{-\frac{1}{2} - \alpha}(x + 1)^{-\frac{1}{2} - \gamma}
\int_{0}^{\infty} x^{-ut}(1 + x)^{ut} g(s, t) w(t) dt ds dx,
\end{equation}
where the difference is that we previously had $x^{-\frac{1}{2}}\gamma(1 + x)^{-\frac{1}{2} - \alpha}$ appearing in $K^-$. In addition, there is an error in (91); the ± signs need to be switched on the right hand side. It turns out that these two errors compensate for each other and the resulting calculations of [HY] remain valid.

We cannot use (94) (which is Stirling’s formula) because it assumes the shifts are small. However, we can evaluate the $x$-integrals in terms of gamma functions, and use (98) to evaluate the arithmetical sum. In this way, we arrive at

\[(6.20)\quad I_{\alpha,\beta,\gamma,\delta}^{(1+)} = \frac{\zeta(1 - \alpha + \beta - \gamma - \delta)}{\zeta(2 - \alpha + \beta - \gamma - \delta)} \frac{1}{h_{1/2-\alpha k_{1/2-\gamma}}} \int_{-\infty}^{\infty} \frac{w(t)}{2\pi i} \int_{(e)} \frac{G(s)}{s} g_{\alpha,\beta,\gamma,\delta}(s, t) \]

\[\left(\frac{hk}{\pi^2}\right)^s \frac{\Gamma\left(\frac{1}{2} - \gamma - s + it\right)\Gamma(\alpha + \gamma + 2s)}{\Gamma\left(\frac{1}{2} + \alpha + s + it\right)} C_{\alpha,\beta,\gamma,\delta,h,k}(s) \zeta(\alpha + \gamma + 2s) \zeta(1 + \beta + \delta + 2s) ds dt.\]

Here $C(s)$ is a finite Euler product defined by (99)–(103) of [HY]. Similarly, for the case of $r < 0$, we find

\[(6.21)\quad I_{\alpha,\beta,\gamma,\delta}^{(1-)} = \frac{\zeta(1 - \alpha + \beta - \gamma - \delta)}{\zeta(2 - \alpha + \beta - \gamma - \delta)} \frac{1}{h_{1/2-\alpha k_{1/2-\gamma}}} \int_{-\infty}^{\infty} \frac{w(t)}{2\pi i} \int_{(e)} \frac{G(s)}{s} g_{\alpha,\beta,\gamma,\delta}(s, t) \]

\[\left(\frac{hk}{\pi^2}\right)^s \frac{\Gamma\left(\frac{1}{2} - \alpha - s - it\right)\Gamma(\alpha + \gamma + 2s)}{\Gamma\left(\frac{1}{2} + \gamma + s - it\right)} C_{\alpha,\beta,\gamma,\delta,h,k}(s) \zeta(\alpha + \gamma + 2s) \zeta(1 + \beta + \delta + 2s) ds dt.\]

Note the following simple gamma function identity

\[(6.22)\quad \frac{\Gamma\left(\frac{1}{2} - z\right)}{\Gamma\left(\frac{1}{2} + z\right)} + \frac{\Gamma\left(\frac{1}{2} - w\right)}{\Gamma\left(\frac{1}{2} + w\right)} = \pi^{-1} \Gamma(1/2 - z) \Gamma(1/2 - w) 2 \cos(\pi/2(z + w)) \cos(\pi/2(z - w)).\]

We apply this with $z = \alpha + s + it$, $w = \gamma + s - it$, whence

\[(6.23)\quad \sum_{\pm} I_{\alpha,\beta,\gamma,\delta}^{(1\pm)} = \frac{\zeta(1 - \alpha + \beta - \gamma - \delta)}{\zeta(2 - \alpha + \beta - \gamma - \delta)} \frac{1}{h_{1/2-\alpha k_{1/2-\gamma}}} \int_{-\infty}^{\infty} \frac{w(t)}{2\pi i} \int_{(e)} \frac{G(s)}{s} g_{\alpha,\beta,\gamma,\delta}(s, t) \]

\[\left(\frac{hk}{\pi^2}\right)^s \frac{\Gamma\left(\frac{1}{2} - \alpha - s - it\right)\Gamma\left(\frac{1}{2} - \gamma - s + it\right)}{\Gamma\left(\frac{1}{2} + \gamma + s - it\right)} C_{\alpha,\beta,\gamma,\delta,h,k}(s) \cos\left(\frac{\pi}{2}(\alpha - \gamma + 2it)\right) \]

\[2 \cos\left(\frac{\pi}{2}(\alpha + \gamma + 2s)\right) \Gamma(\alpha + \gamma + 2s) \zeta(\alpha + \gamma + 2s) \zeta(1 + \beta + \delta + 2s) ds dt.\]

Next we use the functional equation of the zeta function in the form

\[(6.24)\quad \pi^{-2s} \Gamma(\alpha + \gamma + 2s) \zeta(\alpha + \gamma + 2s) = \frac{\pi^{\alpha + \gamma} 2^{\alpha + \gamma + 2s}}{2 \cos(\frac{\pi}{2}(\alpha + \gamma + 2s))} \zeta(1 - \alpha - \gamma - 2s).\]

Thus we obtain (we suppress the subscripts on $C(s)$ as they match $g(s, t)$)

\[(6.25)\quad \sum_{\pm} I_{\alpha,\beta,\gamma,\delta}^{(1\pm)} = \frac{\zeta(1 - \alpha + \beta - \gamma - \delta)}{\zeta(2 - \alpha + \beta - \gamma - \delta)} \frac{1}{h k^{s+\alpha} k^{s+\gamma} \Gamma\left(\frac{1}{2} - \alpha - s - it\right) \Gamma(\frac{1}{2} - \gamma + s - it)} \pi^{\alpha + \gamma - 2s} \]

\[\cos\left(\frac{\pi}{2}(\alpha - \gamma + 2it)\right) \zeta(1 - \alpha - \gamma - 2s) \zeta(1 + \beta + \delta + 2s) ds dt.\]
Anticipating some future simplifications, we apply another gamma function identity. We claim
\[
\pi^{-1} 2^{1-z-w} \Gamma(z) \Gamma(w) \cos \left( \frac{\pi}{2} (w - z) \right) = \frac{1}{2} \frac{\Gamma \left( \frac{\alpha}{2} \right)}{\Gamma \left( \frac{1-\alpha}{2} \right)} \left( 1 + \tan \left( \frac{\pi}{2} z \right) \tan \left( \frac{\pi}{2} w \right) \right),
\]
which can quickly be checked using the duplication and reflection formulas for the gamma function, and the addition formula for cosine. We apply this with \( z = 1/2 - \alpha - s - it \) and \( w = 1/2 - \gamma - s + it \). Our basic assumption \( |\alpha + it| \gg T^s \) (and similarly for the other shifts) implies that \( \tan \left( \frac{\pi}{2} z \right) = \pm i + O_A(T^{-A}) \) (the choice of sign depending on the sign of \( t \)) and similarly \( \tan \left( \frac{\pi}{2} w \right) = \mp i + O_A(T^{-A}) \). Thus
\[
\sum_{\pm} I_{\alpha,\beta,\gamma,\delta}^{(1\pm)} = \frac{\zeta(1 - \alpha + \beta) \zeta(1 - \gamma + \delta)}{\zeta(2 - \alpha + \beta - \gamma + \delta) \sqrt{hk}} \int_{-\infty}^{\infty} \frac{w(t)}{2\pi i} \int_{(-\varepsilon)} G(s) g_{\alpha,\beta,\gamma,\delta}(s, t) C(s) h^{s+\alpha} k^{s+\gamma} \pi^{\alpha+\gamma} \zeta(1 - \alpha - \gamma - 2s) \zeta(1 + \beta + \delta + 2s) ds dt + O(T^{-A}).
\]
We need to match this term with another one arising from the second part of the approximate functional equation. Define \( I_{\alpha,\beta,\gamma,\delta}^{(2\pm)} \) to be the term arising from taking \( I_{\beta,\alpha,\delta,\gamma}^{(1\pm)} \) and then applying the symmetries \( \alpha \leftrightarrow -\gamma, \beta \leftrightarrow -\delta \), and multiplying by \( X_{\alpha,\beta,\gamma,\delta} \) (so in all we applied \( \alpha \leftrightarrow -\delta, \beta \leftrightarrow -\gamma \), and multiplied by \( X \)). Inserting these changes into (6.25) and changing variables \( s \to -s \), we obtain
\[
\sum_{\pm} I_{\alpha,\beta,\gamma,\delta}^{(2\pm)} = -\frac{\zeta(1 - \alpha + \beta) \zeta(1 - \gamma + \delta)}{\zeta(2 - \alpha + \beta - \gamma + \delta) \sqrt{hk}} \int_{-\infty}^{\infty} \frac{w(t)}{2\pi i} \int_{(-\varepsilon)} G(s) g_{-\delta, -\gamma, -\beta, -\alpha}(-s, t) C(-s) h^{s+\delta} k^{s+\beta} X_{\alpha,\beta,\gamma,\delta} t \frac{\Gamma \left( \frac{1+\delta+s-it}{2} \right)}{\Gamma \left( \frac{1-\delta-s-it}{2} \right)} \pi^{-\beta-\delta} \zeta(1 - \alpha - \gamma - 2s) \zeta(1 + \beta + \delta + 2s) ds dt + O(T^{-A}).
\]
By Corollary 6.5 of [HY], we have
\[
I_{\alpha,\beta,\gamma,\delta}^{s+\alpha} k^{s+\gamma} C_{\alpha,\beta,\gamma,\delta,\alpha,\beta,\alpha}(s) = h^{s-\delta} k^{-s-\beta} C_{-\delta, -\gamma, -\beta, -\alpha, \alpha, \beta, \alpha}(-s).
\]
We also claim that
\[
g_{\alpha,\beta,\gamma,\delta}(s, t) \frac{\Gamma \left( \frac{1-\alpha-s-it}{2} \right)}{\Gamma \left( \frac{1+\alpha+s+it}{2} \right)} \pi^{\alpha+\gamma} = g_{-\delta, -\gamma, -\beta, -\alpha}(-s, t) X_{\alpha,\beta,\gamma,\delta} \frac{\Gamma \left( \frac{1+\delta+s-it}{2} \right)}{\Gamma \left( \frac{1-\delta-s-it}{2} \right)} \pi^{-\beta-\delta},
\]
which immediately follows upon writing out the definitions of both sides.

We can recognize now that \( \sum_{\pm} I_{\alpha,\beta,\gamma,\delta}^{(1\pm)} \) and \( \sum_{\pm} I_{\alpha,\beta,\gamma,\delta}^{(2\pm)} \) have identical integrands (up to the tiny error \( O(T^{-A}) \)), but different contours that when combined give a closed contour around the finitely many poles with real part near 0. Thus by Cauchy’s theorem this inner \( s \)-integral can be expressed as the sum of residues with small real parts, plus a very small error term.
from $O_A(T^{-A})$. The residue at $s = 0$ gives

\begin{equation}
(6.31) \quad \zeta(1 - \alpha + \beta)\zeta(1 - \gamma + \delta)\zeta(1 - \alpha - \gamma)\zeta(1 + \beta + \delta) / \zeta(2 - \alpha + \beta - \gamma + \delta) \sqrt{hk} \int_{-\infty}^{\infty} w(t)X_{\alpha,\gamma,t}C(0)h^\alpha k^\gamma dt + O(T^{-A}).
\end{equation}

From Lemma 6.10 of [HY], we have $h^\alpha k^\gamma C_{\alpha,\beta,\gamma,\delta,h,k}(0) = B_{-\gamma,\beta,-\alpha,\delta,h,k}(0)$. Hence this residue gives the first term on the second line of (6.8). The other three permutations of the off-diagonal terms give the remaining three terms of (6.8).

Finally, we briefly indicate how the “junk terms” cancel. This is luckily almost identical to the calculations of [HY]. Looking at (6.27), the residue at $2s = -\alpha - \gamma$ has the pleasant feature that the ratio of two gamma factors on the second line evaluates to 1. This term exactly matches the analogous residue of the diagonal term at $2s = -\alpha - \gamma$, using Lemma 6.11 of [HY] to match the arithmetical factors. Similarly, Lemma 6.12 of [HY] matches the other residue at $2s = -\beta - \delta$ (in this case it is easier to compare via (6.28)). □

7. Unconditional shifted divisor sum

Theorem 7.1. Suppose that $w(x)$ is a smooth function on the positive reals supported on $Y \leq x \leq 2Y$ and satisfying $w^{(j)}(x) \ll (P/Y)^j$ for some parameters $1 \leq P \leq Y$. Let $\theta = 7/64$, and set $R = P + \frac{\tau|m|}{Y}$. Then for $m \neq 0, R \ll T/(TY)^{\delta}$, we have

\begin{equation}
(7.1) \quad \sum_{n \in \mathbb{Z}} \tau_{iT}(n)\tau_{iT}(n + m)w(n) = M.T. + E.T.,
\end{equation}

where

\begin{equation}
(7.2) \quad M.T. = \sum_{\pm} |\frac{\zeta(1 + 2iT)^2}{\zeta(2)}|_{\sigma = -1}(m) \int_{\max(0,-m)}^{\infty} (x + m)^{\mp iT}x^{\mp iT}w(x)dx
\end{equation}

\begin{equation}
+ \sum_{\pm} \frac{\zeta(1 \mp 2iT)^2}{\zeta(2 \mp 4iT)}|_{\sigma = -1 \pm iT}(m) \int_{\max(0,-m)}^{\infty} (x + m)^{\pm iT}x^{\pm iT}w(x)dx.
\end{equation}

and

\begin{equation}
(7.3) \quad E.T. \ll (|m|^{\theta T^{\frac{1}{2}}Y^{\frac{1}{2}}R^2 + T^{\frac{1}{2}}Y^{\frac{3}{4}}R^2})TY^\varepsilon.
\end{equation}

Furthermore, with $R = P + \frac{\tau M}{Y}$, we have

\begin{equation}
(7.4) \quad \sum_{1 \leq |m| \leq M} |E.T.| \ll (MT^{\frac{1}{2}}Y^{\frac{1}{2}}R^2 + MT^{\frac{1}{2}}Y^{\frac{3}{4}}R^2)TY^\varepsilon.
\end{equation}

The crucial qualitative feature of the error term is that when $Y = T$, $m \asymp P \asymp 1$ (so $R \asymp 1$), it is a power saving in $Y$.

The main term here is identical to the main term in (6.16), because when $h = k = 1$ the sum over $l$ in (6.17) can be explicitly evaluated in the above form. Actually, the second line of (7.2) is $O(T^{-100})$ and thereby may be omitted; it is retained here to match the structure of (6.17).

In his work on the $L^4$ norm of Eisenstein series, Spinu [Sp] was also led to estimating a shifted divisor sum with large spectral parameter, but of the form $\sum_{n \in \mathbb{Z}} \tau_{iT}(n)\tau_0(n + m)$; see his (1.11).
For the proof, we proceed in an analogous way to Motohashi [Mo1] who found a spectral expansion for shifted convolution sums. We would not save a lot of effort by directly appealing to the work in [Mo1] because the exact form we need does not appear there, and anyway, with any approach it is necessary to make estimations which forms the bulk of our work (by comparison, examine Chapter 5 of [Mo2]). Therefore we chose to give a mostly self-contained exposition.

The proof requires the full spectral theory of automorphic forms. The older Estermann-type approach using Weil’s bound for Kloosterman sums is far from satisfactory; see (7.31) for this estimate.

By a symmetry argument, we may assume that $m \geq 1$; this uses the fact that $w_m(x) := w(x - m)$ has support on $x \asymp Y$ too, since $m = O(Y)$ by assumption.

7.1. Separation of variables. As a starting point, we find it convenient to appeal to an approximate functional equation for the divisor function, as in Lemma 5.4 of [Y]. This is a particularly simple alternative to the circle method or the delta symbol method for the divisor function. We have $\tau_{IT}(m) = m^{-iT} \sigma_{2IT}(m)$. The formula is

\begin{equation}
\tau_{IT}(m) = m^{-iT} \sum_{l=1}^{\infty} S(m, 0; l) f_{2IT} \left( \frac{l}{\sqrt{m}} \right) + m^{iT} \sum_{l=1}^{\infty} S(m, 0; l) f_{-2IT} \left( \frac{l}{\sqrt{m}} \right),
\end{equation}

where

\begin{equation}
f_\lambda(x) = \frac{1}{2\pi i} \int_{(\varepsilon)} x^{-w} \zeta(1 + w, \lambda + w) \frac{G(w)}{w} dw,
\end{equation}

and $G(w)$ is any fixed even holomorphic function with rapid decay in vertical strips satisfying $G(0) = 0$ (e.g., $\exp(u^2)$). The terms are symmetric under $T \to -T$. By shifting the contour to the right, we see that $f_{\pm 2IT}(x) \ll x^{-A}$ with $A > 0$ arbitrary (uniformly in $T$). One can verify this identity by reversing the orders of integration and summation, evaluating the sum over $l$, and using symmetry to express the sum of the two contour integrals as the residue.

Letting $B_w(T)$ denote the left hand side of (7.1), we then have

\begin{equation}
B_w(T) = \sum_{\pm} \sum_{l=1}^{\infty} \frac{1}{l^{1+2iT}} \sum_{n=1}^{\infty} \tau_{IT}(n) S(n + m, 0; l) \left( 1 + \frac{m}{n} \right)^{\mp iT} f_{\pm 2IT} \left( \frac{l}{\sqrt{n + m}} \right) w(n).
\end{equation}

For later use, we wish to have the variables $l$ and $n$ separated, so we rearrange this as follows:

\begin{equation}
B_w(T) = \sum_{\pm} \frac{1}{2\pi i} \int_{(1)} G(u) u \zeta(1 \mp 2iT + u) B_{w, \pm, u, m}(T) du,
\end{equation}

with

\begin{equation}
B_{w, \pm, u, m}(T) = \sum_{l=1}^{\infty} \sum_{n=1}^{\infty} \tau_{IT}(n) S(n + m, 0; l) \frac{g_{m,u,T}(n)}{l^{1+2iT + u + \frac{\pm iT + u}{2}}} w(x),
\end{equation}

$g_{m,u,T}(x) = \left( 1 + \frac{m}{x} \right)^{\mp iT + \frac{u}{2}} w(x)$.

The changing of orders of integration and summation is justified by absolute convergence. This function $g$ has support for $x \asymp Y$ and satisfies

\begin{equation}
\frac{d^j}{dx^j} g_{m,u,T}(x) \ll_{u,j} \left( \frac{Tm}{Y^2} + \frac{P}{Y} \right)^j = Y^{-j} R^j,
\end{equation}

with an implied constant depending polynomially on $u$. In practice, we think of $u$ as almost bounded because the rapid decay of $G(u)$ overcomes any $u$-dependence of $B$. Notice that in
estimating \( B_{u}(T) \), we are free to move the \( u \)-contour to any fixed line with \( \text{Re}(u) > 0 \) since the integrand is holomorphic in this region.

We record a bound on the Mellin transform of \( g \). Integration by parts shows

\[
\tilde{g}(s) = \frac{(-1)^j}{s(s+1)\ldots(s+j-1)} \int_{0}^{\infty} g^{(j)}(x) x^{s+j} \frac{dx}{x}, \tag{7.11}
\]

so that using (7.10) we derive the bound

\[
\tilde{g}(s) \ll_{\sigma,j} Y^{\sigma} \frac{R^j}{|s(s+1)\ldots(s+j-1)|}, \tag{7.12}
\]

so that \( \tilde{g}(s) \) is very small for \( |\text{Im}(s)| \gg (TY)^{\varepsilon}R \). Since \( g \) has compact support, \( \tilde{g} \) is entire, and the apparent singularities at \( s = 0, -1, -2, \ldots \) in (7.11) are necessarily removable.

### 7.2. Relationships with the Estermann function

Next we open the Kloosterman sum, and use the Mellin transform of \( g = g_{m,u,T} \) to write (7.9) as

\[
B_{w,\pm,u,m}(T) = \sum_{l=1}^{\infty} \sum_{n \equiv a \pmod{l}}^{*} e\left(\frac{am}{l^{1+2iT+u}}\right) \frac{1}{2\pi i} \int_{(2)} \tilde{g}(s + \frac{u}{2}) D(s, \mp 2iT, \frac{a}{l}) ds, \tag{7.13}
\]

where we have used \( \tau_{iT}(n)/n^{\pm iT} = \sigma_{\mp 2iT}(n) \), and \( D \) is the Estermann function defined for \( \text{Re}(s) > \max(1, 1 + \text{Re}(\xi)) \)

\[
D(s, \xi, \frac{a}{l}) = \sum_{n=1}^{\infty} \sigma_{\xi}(n) e\left(\frac{an}{l}\right) n^{-s}. \tag{7.14}
\]

See Lemma 3.7 of [Mo2] for proofs of the following properties. The Estermann function has a meromorphic continuation and functional equation

\[
D(s, \xi, a/l) = 2(2\pi)^{2s-\xi-2}\Gamma(1-s)\Gamma(1+\xi-s) \Gamma(1-s, -\xi, -\pi/a/l) \cos(\pi(s - \frac{1}{2}\xi)) \cos(\pi(s + \frac{1}{2}\xi)). \tag{7.15}
\]

Furthermore, the pole at \( s = 1 \) has residue \( \ell^{-1}\xi(1 - \xi) \) and the one at \( 1 + \xi \) has residue \( l^{-\xi-1}\zeta(1 + \xi) \).

We analyze \( B \) by moving the contour to the line \(-1 - \varepsilon\), applying the functional equation, and reversing the order of summation and integration. There are poles at \( s = 1 \) and \( s = 1 \mp 2iT \) which lead to the main terms that we shall examine shortly in Section 7.3. Thus we obtain, with shorthand \( \alpha_{T} = 2(2\pi)^{2T-2} \),

\[
B_{w,\pm,u,m} - M.T. = \alpha_{T} \sum_{l=1}^{\infty} \sum_{n \equiv a \pmod{l}}^{*} e\left(\frac{an}{l}\right) \sum_{n=1}^{\infty} \sigma_{\pm 2iT}(n) \frac{1}{2\pi i} \int_{(-1-\varepsilon)} \tilde{g}(s + \frac{u}{2}) \left(2\pi\right)^{2s} \left(\frac{n}{l^{2}}\right)^{s} \Gamma(1-s)\Gamma(1 \mp 2iT-s) \left(e\left(\frac{\pi n}{l}\right) \cosh(\pi T) - e\left(-\frac{\pi n}{l}\right) \cos(\pi(s \pm iT))\right) ds, \tag{7.16}
\]

where \( M.T. \) stands for the main terms.
7.3. Computing the main terms. The pole at $s = 1$ gives to $B_w(T)$

$$
(7.17) \sum_{\pm} \frac{1}{2\pi i} \int_{(1)} \frac{G(u)}{u} \zeta(1 \mp 2iT + u) \sum_{l=1}^{\infty} \frac{e^{\frac{am}{l}}}{1 + 2i\pi l + u} g(1 + \frac{u}{2}) \zeta(1 \pm 2iT) \ell^{-1+2iT} du.
$$

Using $S(m,0;l) = \sum d_l(m,l) d\mu(l/d)$, this simplifies as

$$
(7.18) \sum_{\pm} \frac{1}{2\pi i} \int_{(1)} \frac{G(u)}{u} g(1 + \frac{u}{2}) \zeta(1 \mp 2iT + u) \frac{\sigma_{-1-u}(m)}{\zeta(2 + u)} \zeta(1 \pm 2iT) du.
$$

By shifting the contour left to $\text{Re}(u) = -1/2$, we cross a pole at $u = 0$ which gives

$$
(7.19) \sum_{\pm} g(1) \frac{\zeta(1 + 2iT)^2}{\zeta(2)} \sigma_{-1}(m).
$$

The pole at $u = \pm 2iT$ is very small but it also will cancel a forthcoming term (this should not be surprising because one could choose a different function $G$ and change the value of $G(\pm 2iT)$). Using (7.12) and Weyl’s bound, we see that the new integral satisfies the bound

$$
(7.20) (TY)^{\varepsilon} \int_{-\infty}^{\infty} \left| \frac{G(-\frac{1}{2} + it)}{-\frac{1}{2} + it} g(3/4 + it/2) \zeta(1/2 + it \mp 2iT) \right| dt \ll T^{1/6} Y^{3/4} (YT)^{\varepsilon},
$$

and this is absorbed by the second error term appearing in (7.3).

The other residue at $s = 1 \mp 2iT$ equals

$$
(7.21) \frac{1}{2\pi i} \int_{(1)} \frac{G(u)}{u} g(1 \mp 2iT + \frac{u}{2}) \zeta(1 \mp 2iT + u) \frac{\sigma_{-1\pm 4iT}(m)}{\zeta(2 \mp 4iT + u)} \zeta(1 \pm 2iT) du.
$$

Again shifting contours to the left to $\text{Re}(u) = -1/2$, we cross a pole at $u = 0$ which gives

$$
(7.22) g(1 \mp 2iT) \frac{\zeta(1 \pm 2iT)^2}{\zeta(2 \mp 4iT)} \sigma_{-1\pm 4iT}(m).
$$

The pole at $u = \pm 2iT$ is small, but could also be checked to cancel a previous pole. The new contour is much smaller than the bound appearing in (7.20).

Note that

$$
(7.23) g(1) = \int_{0}^{\infty} (x + m)^{\mp iT} x^{\pm iT} w(x) dx, \quad g(1 \mp 2iT) = \int_{0}^{\infty} (x + m)^{\mp iT} x^{\mp iT} w(x) dx.
$$

Thus (7.19) and (7.22) form the main term in (7.2). This combination of main terms precisely agrees with the main terms of (6.16).

Additionally, one may see that (7.21) is $O(T^{-100})$ using that $g(s)$ is small for $|\text{Im}(s)| \gg R(TY)^{\varepsilon}$, and recalling $R \ll T(TY)^{-\delta}$. This explains why the main term on the second line of (7.2) is small.

7.4. Analysis of the weight functions. We return to (7.16). By changing variables $s \to s + \frac{1}{2} - \frac{u}{2}$, it becomes

$$
(7.24) B_{w,\pm,u,m} - M.T. = \alpha_T \sum_{\delta = \pm 1} \delta \sum_{n=1}^{\infty} \frac{\sigma_{-2iT}(n)}{n^{\frac{1}{2} + \frac{u}{2}}} \sum_{l=1}^{\infty} S(n, \delta m; l) \frac{\varphi_{\delta}(\frac{4\pi \sqrt{mn}}{l})}{l},
$$
where $\alpha_T' = 2(2\pi)^{\pm 2i T - 1 - u}$ and

$$\varphi_\delta(x) = \frac{1}{2\pi i} \int_{(\sigma)} \tilde{g}(s + \frac{1}{2})(x/2)^{2s} m^{-s} \Gamma(s + \frac{3}{2} - s - \frac{u}{2}) \Gamma(\frac{1}{2} - s + \frac{u}{2} + 2i T) c_\delta(s) ds,$$

with $c_+(s) = \cosh(\pi T)$ and $c_-(s) = \cos(\pi(s - \frac{1}{2} - \frac{u}{2} \pm iT))$. Here $\sigma < 1/2$ (recall $\text{Re}(u) > 0$).

We shall derive some bounds on $\varphi_\delta$. Let $\text{Re}(s) = \sigma$ and $\text{Re}(u) = \nu > 0$. We claim that for $\sigma < 1/2$ fixed, and each choice of $\delta = \pm$, that

$$\Gamma(\frac{1}{2} - \sigma - it) \Gamma(\frac{1}{2} - \sigma - it \mp 2iT) c_\delta(\sigma + it) \ll (1 + |t|)^{-\sigma} (1 + |t \pm 2T|)^{-\sigma}. $$

To prove this, initially in the case $\delta = +$, observe that Stirling’s formula gives a bound of this form, except multiplied by $\exp(\frac{\pi}{2} Q_1(t))$, where $Q_1(t) = 2T - |t \pm 2T| - |t|$. Notice $Q_1(t \mp T) = Q_1(t, T/2)$ defined by (4.5), so $Q_1(t) \leq 0$ for all $t, T \in \mathbb{R}$ (this can also be checked directly).

The case $\delta = -$ is similar, except with $Q_1(t)$ replaced by $Q_2(t) = 2|t \pm T| - |t| - |t \pm 2T|$, which we again claim is $\leq 0$ for all $t$. It takes the value 0 for $|t|$ large, and so its maximum value must occur at $t = 0$ or $t \pm 2T = 0$, and $Q_2(t) = 0$ at both of these points.

By (7.12) and (7.26), we have

$$\varphi_\delta(x) \ll \left(\frac{x^2}{m}\right)^{\sigma} \int_{-\infty}^{\infty} (1 + |t|)^{-\sigma + \frac{\nu}{2}} (1 + |t \pm 2T|)^{-\sigma + \frac{\nu}{2}} |g(\sigma + \frac{1}{2} + it)| dt,$$

the implied constant depends polynomially on $\text{Im}(u)$. Using the assumption $R \ll T/(TY)^{\delta}$, we deduce for fixed $\sigma < 1/2$ that

$$\varphi_\delta(x) \ll_{\sigma, \nu} x^{2\sigma} Y^{1 + \frac{\nu}{2}} m^{-\sigma} T^{-\sigma + \frac{\nu}{2}} R^{1 - \sigma + \frac{\nu}{2}}.$$

By taking $\sigma = 1/2 - \varepsilon$ or $-\sigma > 0$ very large, we obtain

$$\varphi_\delta(x) \ll x Y m^{-1/2} T^{-1/2} R^{1/2} (TY/x)^{\varepsilon} (1 + \frac{x^2}{mTR/Y})^{-A} (TR)^{\nu/2}.$$

Since it is straightforward at this point, we record the effect of bounding $B$ trivially with the Weil bound, even though this is not our final objective. The sum over $l$ can be truncated at $Y^{1/2 + \varepsilon}$. Taking $\nu = \varepsilon$ and $\sigma = 1/2 - \varepsilon$, we then obtain that the error term in $B$ is

$$\ll (TY)^{\varepsilon} \sum_{l \in \mathbb{Z}^{1/2 + \varepsilon}} \sum_{n \in (TY)^{1/2} R T / Y} |S(n, \pm m, l)| \frac{Y R^{1/2}}{l^2 T^{1/2}}.$$

A short computation shows that this error term is

$$\ll Y^{3/4} T^{1/2} R^{3/2} (TY)^{\varepsilon},$$

which is only nontrivial for $T \leq Y^{1/2 - \delta}$ when $R \approx 1$.

7.5. **Application of the Kuznetsov formulas.** To get a stronger result, we apply the spectral theory of automorphic forms in the guise of the Kuznetsov formula. Let

$$K_\delta(m, n, \varphi_\delta) = \sum_{l=1}^{\infty} \frac{S(n, \delta m; l)}{l} \varphi_\delta \left( \frac{4\pi \sqrt{mn}}{l} \right).$$
The Kuznetsov formula (in one particularly useful way for us) states

\( K_+(m, n, \varphi_+) = \sum_j \frac{|\rho_j(1)|^2}{\cosh(\pi t_j)} \lambda_j(m) \lambda_j(n) \frac{1}{2\pi i} \int_{(\sigma)} \varphi_+^*(s) h_+(s, t_j) ds \)

\[ + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau_{iv}(m) \tau_{iv}(n)}{\zeta(1 + 2ir)^2} \frac{1}{2\pi i} \int_{(\sigma)} \varphi_+^*(s) h_+(s, r) ds dr \]

\[ + \sum_k (2k - 1) q_{m,n}(k) \frac{1}{2\pi i} \int_{(\sigma)} \varphi_+^*(s) \frac{\Gamma(k - 1 + s)}{\Gamma(k + 1 - s)} ds - \frac{1}{2\pi i} \delta_{m=n} \frac{1}{2\pi i} \int_{(\sigma)} \varphi_+^*(s) \frac{\Gamma(s)}{\Gamma(1 - s)} ds, \]

with notation as follows: the \( s \)-integrals are over a vertical line with \( \sigma = 3/4 + \varepsilon \),

\[ h_+(s, r) = \frac{1}{2} \sin(\pi s) \Gamma(s - \frac{1}{2} + ir) \Gamma(s - \frac{1}{2} - ir), \]

\[ \varphi_+^*(s) = \int_{0}^{\infty} \varphi_0(x)(x/2)^{-2s} dx, \]

and \( q_{m,n}(k) \) is the sum over a Hecke eigenbasis of weight \( 2k \) (even) holomorphic forms, with scaling such that

\[ q_{m,n}(k) = k^{-1+o(1)} \sum_{f \in B_{2k}} \lambda_f(m) \lambda_f(n). \]

We also recall that \( \rho_j(1) \) is scaled so that \( \frac{|\rho_j(1)|^2}{\cosh(\pi t_j)} = t_j^\alpha(1) \). One can find this form of Kuznetsov’s formula as (2.4.13) of [Mo2]. It is easy to check that \( \varphi_0(x) \) satisfies the required bounds to apply the Kuznetsov formula, using (7.29) as well as simple variants for the derivatives of \( \varphi_0^* \).

By the Mellin inversion formula and (7.25), we have

\[ \varphi_0^*(s) = 2^{2s} \tilde{\varphi}_0(1 - 2s) = \tilde{g}(s)m^{\frac{1}{2}-s} \Gamma(1 - s + \frac{u}{2}) \Gamma(1 - s + \frac{u}{2} + 2iT) c_\delta(s - \frac{1}{2}). \]

It follows from (7.12) and (7.26) that for \( |t| \leq R(TY)^\varepsilon \) (recall \( R \ll T(TY)^{-\delta} \) so \( t = o(T) \)):

\[ \varphi_0^*(t) \ll_{\sigma, u} Y^\sigma m^{\frac{1}{2}-\sigma}(1 + |t|)^{\frac{1}{2}-\sigma+\frac{u}{2}} T^{\frac{1}{2}-\sigma+\frac{u}{2}}. \]

Also, \( \varphi_0^*(s) \) is very small for \( |t| \geq R(TY)^{\varepsilon} \), and the implied constant depends polynomially on \( \text{Im}(u) \).

We also need the opposite sign case where the Kuznetsov formula has a different shape than (7.33). For this, we claim

\[ K_-(m, n, \varphi_-) = \sum_j \frac{|\rho_j(1)|^2}{\cosh(\pi t_j)} \lambda_j(-m) \lambda_j(n) \frac{1}{2\pi i} \int_{(\sigma)} \varphi_-^*(s) h_-(s, t_j) ds \]

\[ + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau_{iv}(m) \tau_{iv}(n)}{\zeta(1 + 2ir)^2} \frac{1}{2\pi i} \int_{(\sigma)} \varphi_-^*(s) h_-(s, r) ds, \]

where \( 1/2 < \sigma < 1 \), and

\[ h_-(s, r) = \frac{1}{2} \cosh(\pi r) \Gamma(s - \frac{1}{2} + ir) \Gamma(s - \frac{1}{2} - ir). \]

Towards this, we first quote Theorem 2.5 of [Mo2], stating

\[ K_-(m, n, \varphi) = \sum_j \frac{|\rho_j(1)|^2}{\cosh(\pi t_j)} \lambda_j(-m) \lambda_j(n) \tilde{\varphi}(t_j) + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau_{iv}(m) \tau_{iv}(n)}{\zeta(1 + 2ir)^2} \tilde{\varphi}(r) dr \]
Then inserting (7.33) and (7.39) into (7.24), and then in turn into (7.8), we get, say
\[ (7.47) \]
valid for \( 0 < \sigma < 1 \) (the upper bound arises to ensure the holomorphy of \( \tilde{\varphi}(s) \)), so by (6.561.16) of [GR] and a change of variables, we have for \( 1/2 < \sigma < 1 \) that
\[ (7.44) \]
Using \( \varphi_\delta^*(s) = 2^{2s}\tilde{\varphi}_\delta(1-2s) \), we have
\[ (7.45) \]
Hence we derive (7.39).
Define
\[ (7.46) \]
\[ (7.47) \]
Then inserting (7.33) and (7.39) into (7.24), and then in turn into (7.8), we get, say
\[ (7.48) \]
Here this error term comes from the calculation of the residues of the Estermann function as in Section 7.3, and \( E_{\text{holo}} \) and \( E_{\text{diag}} \) are given by analogous formulas to \( E_{\text{Maass}} \) and \( E_{\text{Eis}} \). See (7.58) below for \( E_{\text{holo}} \). In the forthcoming sections, we show
\[ (7.49) \]
and
\[ (7.50) \]
which together account for (7.3). Note the second term in (7.50) is smaller than (7.49), as is the error term in (7.48).
7.6. **Cusp form contributions.** We pick up with (7.46). By absolute convergence, we can bring the sum over $n$ to the inside. By an exercise with the Hecke relations, one can check that for $\Re(v) > 1$,

$$
(7.51) \quad \sum_{n=1}^{\infty} \frac{\sigma_{\pm 2iT}(n)\lambda_j(n)}{n^v} = \frac{L(v \mp 2iT, u_j)L(v, u_j)}{\zeta(2v \mp 2iT)}.
$$

We evaluate the sum over $n$, and then shift the $u$-contour back to $\nu = \varepsilon$ (crossing no poles), and obtain

$$
(7.52) \quad E_{\text{Maass}}^\delta = \sum \frac{1}{2\pi i} \int (\varepsilon) \frac{G(u)}{u} \alpha'_T \sum_j \frac{|\rho_j(1)|^2}{\cosh(\pi t_j)} \lambda_j(\delta m)L(\frac{1}{2} + \frac{u}{2} \mp 2iT, u_j)L(\frac{1}{2} + \frac{u}{2}, u_j) \frac{1}{2\pi i} \int_{(3/4 + \varepsilon)} \varphi^*_s(s)h_\delta(s, t_j)dsdtu.
$$

Next we estimate the $s$-integral above, initially for $\delta = +$. Compared to (7.25), the integral representation is the same except we have $s$ shifted by $1/2$, and we have multiplied by $h_\pm(s, t_j)$. Stirling’s formula shows that

$$
(7.53) \quad h_+(\sigma + it, r) \ll (1 + |t + r|)^{\sigma - 1}(1 + |t - r|)^{\sigma - 1}\exp\left(\frac{\pi}{2}(2|t| - |t + r| - |t - r|)\right).
$$

The exponential part above is 1 for $|r| \leq |t|$, and is $\exp(-\pi|t - r|)$ for $|r| > |t|$ (again we have encountered (4.5)). Thus the $s$-integral in (7.52) is very small unless $t_j \leq R(TY)^{\varepsilon}$. Suppose that $1/2 < \sigma < 1$. Then with $\nu = \varepsilon$, using (7.38) and (7.53), we have

$$
(7.54) \quad \int_{(\sigma)} \varphi^*_+(s)h_+(s, r)ds \ll \int_{|t| \leq R(TY)^{\varepsilon}} (1 + |t|)^{\frac{1}{2} - \sigma}(1 + |t + r|)^{\sigma - 1}(1 + |t - r|)^{\sigma - 1}dt + \ldots,
$$

with the dots representing a very small error term arising from the truncation. Taking $\sigma = 1/2 + \varepsilon$, we obtain the bound

$$
(7.55) \quad \int_{(\sigma)} \varphi^*_+(s)h_+(s, r)ds \ll (TY)^{\varepsilon}Y^{1/2}\left(1 + \frac{|r|}{R(TY)^{\varepsilon}}\right)^{-100}.
$$

Using the uniform subconvexity bound $L(1/2 + u/2 \pm 2iT, u_j) \ll (t_j + T)^{1/3 + \varepsilon}$ of Jutila-Motohashi [JM], the bound $|\lambda_j(m)| \ll m^{\theta + \varepsilon}$ (with $\theta = 7/64$), and the mean value result (with polynomial dependence on $u$) $\sum_{t_j \leq R} |L(1/2 + u/2, u_j)|^2 \ll R^{2 + \varepsilon}$ (see [Mo2], Theorem 3.1, for example), we obtain (7.49) for $E_{\text{Maass}}^\delta$.

Furthermore, if we sum this error term over $m \leq M$, then we can use $\sum_{m \leq M} |\lambda_j(m)|^2 \ll M(t_j M)^{\varepsilon}$ [ItS3], which explains why (7.4) effectively has $\theta = 0$.

We claim that the bound of (7.49) holds for $E_{\text{Maass}}^-\delta$ also. By Stirling’s formula, analogously to (7.53), we have

$$
(7.56) \quad h_-(\sigma + ir, r) \ll (1 + |t + r|)^{\sigma - 1}(1 + |t - r|)^{\sigma - 1}\exp\left(\frac{\pi}{2}(2|t| - |t + r| - |t - r|)\right)
$$

The polynomial factor here is identical to that in (7.53), while the exponential factor is 1 for $|r| \geq |t|$, and is $\exp(-\pi|r - t|)$ for $|r| < |t|$, which simply means that we cannot immediately truncate $|r|$ at $R(TY)^{\varepsilon}$ in this case, in contrast to the $E_{\text{Maass}}^\delta$ case. If $|r| \leq R(TY)^{\varepsilon}$, then in fact (7.54) holds in the case $\delta = -\varepsilon$, since the bound (7.38) is independent of $\delta$, and the
right hand side of (7.54) drops the exponential part of \( h_+(s, r) \) anyway, and so we get the same bound for \( h_-(s, r) \). Since the estimates on the weight functions are identical, we have that the contributions to \( E_{\text{Maass}}^- \) from \( |t_j| \leq R(TY)^\varepsilon \) immediately leads to (7.49).

We claim that \( \int \varphi_+(s) h_-(s, r) ds \) is very small if \( |r| \geq R(TY)^\varepsilon \). We note that on the line \( \Re(s) = \sigma, \Re(u) = \varepsilon \) (which we assume avoids any pole of a gamma function), we have

\[
\int_{(\sigma)} \varphi_+(s) h_-(s, r) ds \ll |r|^{2\sigma - 2Y\sigma} m_{\frac{1}{2} - \sigma} R_{\frac{1}{2} - \sigma} (TY)^\varepsilon \frac{m_1 \Gamma_{\frac{1}{2} + \frac{1}{2}}}{\Gamma_{k + 1 - s}} \left( \frac{r^2 Y}{m RT} \right)^\sigma.
\]

If \( -\sigma > 0 \) is very large, then this bound becomes very small unless \( r^2 \leq \frac{mRT}{Y} (TY)^\varepsilon \leq R^2 (TY)^\varepsilon \). Our original integral representation requires \( \frac{1}{2} < \sigma < 1 \), and moving the contour far to the left crosses poles at \( s - \frac{1}{2} \pm ir = 0, -1, -2, \ldots \). However, these residues are also very small because these occur at \( |t| = |r| \geq R(TY)^\varepsilon \), but \( g_*(\varepsilon + it) \) is very small for such \( t \)'s, by (7.12). Thus \( E_{\text{Maass}}^- \) satisfies the same bounds as \( E_{\text{Maass}}^0 \), as desired.

### 7.7. Holomorphic forms

Define \( E_{\text{holo}} \) analogously to \( E_{\text{Maass}}^0 \). Following the argument of \( E_{\text{Maass}} \), we arrive at

\[
E_{\text{holo}} = \sum_{\pm} \frac{1}{2\pi i} \int_{(\varepsilon)} \frac{G(u)}{u} 2(2\pi)^{\pm 2T - 1 - u} \sum_{k \in B_{2k}} \sum_{f \in B_{2k}} k_1 \lambda_f (m)
L(\frac{1}{2} + \frac{u}{2}, f) L(\frac{1}{2} + \frac{u}{2} + 2iT, f) \frac{1}{2\pi i} \int_{(s)} \varphi_+(s) \frac{\Gamma(k - 1 + s)}{\Gamma(k + 1 - s)} ds du.
\]

Now we estimate this \( s \)-integral. We first show that it is very small unless \( k \ll R(TY)^\varepsilon \). Suppose otherwise. Then by Stirling and a trivial bound,

\[
\int_{(s)} \varphi_+(s) \frac{\Gamma(k - 1 + s)}{\Gamma(k + 1 - s)} ds \ll_m m^{1/2} R^{3/2} T^{1/2} k^{-2} (TY)^\varepsilon \frac{Y k^2}{T R m} \sigma,
\]

so by taking \( \sigma < 0 \) very far to the left (but with \( \sigma > 1 - k \), we see that this bound can be made to be very small since \( TR m / Y \leq R^2 \). Note the similarity with (7.57). Then with the truncation \( k \ll R(TY)^\varepsilon \), setting \( \sigma = 1/2 \), and using \( L(\frac{1}{2} + iT, f) \ll (T + k)^{1/3 + \varepsilon} \), another result of Jutila-Motohashi [JM], the bound on the holomorphic forms becomes equivalent to that of the Maass forms, except we can take \( \theta = 0 \) in this case since we have Deligne's bound. That is, (7.49) holds for \( E_{\text{holo}} \).

### 7.8. Diagonal term

The diagonal term, say \( E_{\text{diag}} \), is easily checked, with \( \sigma = 1/2 + \varepsilon \), to give the following bound, which is much smaller than (7.49):

\[
E_{\text{diag}} \ll_m m^{-1/2} Y^{1/2} R(TY)^\varepsilon.
\]

### 7.9. Eisenstein contribution

We continue with (7.47). By absolute convergence, we can sum over \( n \) first, getting now

\[
E_{\text{Eis}} = \sum_{\pm} \frac{1}{2\pi i} \int_{(2)} \frac{G(u)}{u} 2(2\pi)^{\pm 2T - 1 - u} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\tau_{ir}(m)}{|\zeta(1 + 2ir)|^2} \frac{1}{2\pi i} \int_{(\sigma)} \int_{(s)} \int_{(r)} |\varphi_+(s) h_+(s, r) - \varphi_-(s) h_-(s, r)| ds dr du,
\]
where we recall the definition (2.11). For ease of reference, we recall that

\[(7.62)\]
\[\varphi^*_h(s, r) = \frac{1}{2} g(s) m^{1 - s} \Gamma(1 - s + \frac{3}{2}) \Gamma(1 - s \mp 2iT) \Gamma(s - \frac{1}{2} + ir) \Gamma(s - \frac{1}{2} - ir) \times \begin{cases} \cosh(\pi T) \sin(\pi s), & \delta = + \\ \cos(\pi(s - \frac{3}{2} + i\tau)) \cosh(\pi\tau), & \delta = -. \end{cases} \]

Now we move the \(u\)-integral to the inside and shift it to the line \(\nu = \varepsilon\). In contrast to the cusp form cases, there are poles at \(u = 1 + 2ir, u = 1 - 2ir, u = 1 + 2ir \pm 2iT\), and \(u = 1 - 2ir \pm 2iT\). The bound on the new line is completely analogous to the bound on \(E_{\text{Maass}}\) (i.e., (7.49)), and in fact the estimates here are slightly better since the spectral measure of the Eisenstein series is smaller (by a factor of \(R\)), and we have the estimate \(|\tau_{\nu}(m)| \leq d(m)\) (“Ramanujan”). This explains the second-listed error term in (7.50).

We examine the residues now. The poles at \(u = 1 + 2ir \pm 2iT\) and \(u = 1 - 2ir \pm 2iT\) give a very small contribution because the inner integral over \(s\) is small unless \(|r| \leq R(TY)^\varepsilon = o(T)\), and \(G(u)\) is small at height \(T\). We work with the residue at \(u = 1 + 2ir\) as the other is similar. By a calculation, we have that this residue, denoted say \(E^r_{\text{Eis}}\) is

\[(7.63)\]
\[E^r_{\text{Eis}} = \sum_{\pm} \frac{1}{\pi} \int_{-\infty}^{\infty} G(1 + 2ir) \frac{G(1 + 2ir)}{1 + 2ir} 2(2\pi)^{2+2iT-2-2ir} \frac{\tau_{ir}(m)}{\zeta(1 - 2ir)} \zeta(1 + 2ir) \zeta(1 + 2iT) \times \begin{cases} \cosh(\pi T) \sin(\pi s) - \cos(\pi(s - \frac{1}{2} - i\tau) \cosh(\pi\tau), & \delta = -. \end{cases} \]

With a cursory examination, it appears that the integrand apparently passes through a pole at \(r = \pm T\) (from the pole of the zeta function), but note that the third line of (7.63) vanishes there. This is the reason we combined the + and - cases.

Now we can fix \(\sigma\) with \(\frac{3}{4} < \sigma \leq 1\) and shift the \(r\)-contour to \(\text{Im}(r) = 1/4\) without crossing any poles. From the decay of \(G\), we can truncate the \(r\)-integral at \((TY)^\varepsilon\), and then we obtain a bound

\[(7.64)\]
\[E^r_{\text{Eis}} \ll m^{1 - \sigma} T^{1/6} Y^\sigma T^{3/4 - \sigma} (TY)^\varepsilon \int_{|t| \leq R(TY)^\varepsilon} (1 + |t|)^{\sigma - \frac{5}{4}} dt \ll T^{1/2 - \sigma} Y^\sigma m^{3/4 - \sigma} R^{1/4} (TY)^\varepsilon.\]

Taking \(\sigma = 3/4 + \varepsilon\), this becomes

\[(7.65)\]
\[E^r_{\text{Eis}} \ll T^{1/6} Y^{3/4} R^{1/2} (TY)^\varepsilon,\]

which is the first-stated bound appearing in (7.50).

This completes the proof of Theorem 7.1.

References

[Be] S. Bettin, The second moment of the Riemann zeta function with unbounded shifts. Int. J. Number Theory 6 (2010), no. 8, 1933–1944.

[BKY] V. Blomer, R. Khan, M. Young, Distribution of Mass of Hecke eigenforms. Duke Math. J. 162 (2013), no. 14, 2609–2644.

[C] V. Chandee, On the correlation of shifted values of the Riemann zeta function. Q. J. Math. 62 (2011), no. 3, 545–572.
[LS3] W. Luo and P. Sarnak, *Quantum variance for Hecke eigenforms*. Ann. Sci. École Norm. Sup. (4) 37 (2004), no. 5, 769–799.

[MV] P. Michel and A. Venkatesh, *The subconvexity problem for GL$_2$*. Publ. Math. Inst. Hautes Études Sci. No. 111 (2010), 171–271.

[Mo1] Y. Motohashi, *The binary additive divisor problem*. Ann. Sci. École Norm. Sup. (4) 27 (1994), no. 5, 529–572.

[Mo2] Y. Motohashi, *Spectral theory of the Riemann zeta-function*. Cambridge Tracts in Mathematics, 127. Cambridge University Press, Cambridge, 1997.

[R] K. Ramachandra, *Application of a theorem of Montgomery and Vaughan to the zeta-function*. J. London Math. Soc. (2) 10 (1975), no. 4, 482–486.

[RS] Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*. Comm. Math. Phys. 161 (1994), no. 1, 195–213.

[Sol] K. Soundararajan, *Moments of the Riemann zeta function*. Ann. of Math. (2) 170 (2009), no. 2, 981–993.

[So2] K. Soundararajan, *Quantum unique ergodicity for SL$_2$(Z)\H*. Ann. of Math. (2) 172 (2010), no. 2, 1529–1538.

[Sp] F. Spinu, *The $L^4$-norm of Eisenstein series*. Princeton PhD thesis 2003.

[TZ1] J. Toth, and S. Zelditch, *Quantum ergodic restriction theorems. I: Interior hypersurfaces in domains with ergodic billiards*. Ann. Henri Poincaré 13 (2012), no. 4, 599–670.

[TZ2] J. Toth, and S. Zelditch, *Quantum ergodic restriction theorems, II: manifolds without boundary*. Geom. Funct. Anal. 23 (2013), no. 2, 715–775.

[W] T. Watson, *Rankin triple products and quantum chaos*. http://arxiv.org/abs/0810.0425.

[Y] M. P. Young, *The fourth moment of Dirichlet L-functions*. Ann. of Math. (2) 173 (2011), no. 1, 1–50.

[Za] 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 (1982).

[Ze1] S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*. Duke Math. J. 55 (1987), no. 4, 919–941.

[Ze2] S. Zelditch, *Mean Lindelöf hypothesis and equidistribution of cusp forms and Eisenstein series*. J. Funct. Anal. 97 (1991), no. 1, 1–49.

Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, U.S.A.

E-mail address: myoung@math.tamu.edu