NON-VANISHING OF HIGH DERIVATIVES OF AUTOMORPHIC L-FUNCTIONS AT THE CENTER OF THE CRITICAL STRIP

E. KOWALSKI, P. MICHEL, AND J. VANDERKAM

Abstract. We prove non-vanishing results for the central value of high derivatives of the complete L-function $\Lambda(f,s)$ attached to primitive forms of weight 2 and prime level q. For fixed $k\geqslant 0$ the proportion of primitive forms f such that $\Lambda^{(k)}(f,1/2)\neq 0$ is $\geqslant p_k+o(1)$ with $p_k>0$ and $p_k=1/2+O(k^{-2})$, as the level q goes to infinity. This result is (asymptotically in k) optimal and analogous to a result of Conrey on the zeros of high derivatives of Riemann's ξ function lying on the critical line. As an application we obtain new strong unconditional bounds for the average order of vanishing of the forms f (i.e. the analytic rank of the Jacobian variety $J_0(q)$).

Contents

1. Introduction

In recent years, non-vanishing results for central values of families of L functions and their derivatives have received considerable attention ([Du, Lu, MM, BFH, P-P, Iw, I-S, KM1, KM2, VdK1] and others) mainly because of their implications in various areas such as the Birch-Swinnerton-Dyer conjecture, spectral deformation theory, and classical analytic number theory. In this paper, we consider a similar question for the central value of higher derivatives of one such family.

Given a prime number q, let $S_2(q)^*$ denote the set of primitive Hecke eigenforms of weight 2 relative to the subgroup $\Gamma_0(q)$. Any $f\in S_2(q)^*$ admits a Fourier expansion of the form

$$f(z)=\sum_{n\geqslant 1} n^{1/2}\lambda_f(n)e(nz),$$

P.M. is partially supported by the Institut Universitaire de France.
J. V. supported by NSF grant DMS-9804517 and by the Hertz Foundation.

normalized so that $\lambda_f(1) = 1$, and so $\lambda_f(n) \in \mathbf{R} \cap \overline{\mathbf{Q}}$. To f is associated an L-function with Euler product,

$$L(f,s) = \sum_{n\geqslant 1} \lambda_f(n) n^{-s} = \prod_p \Bigl(1 - \frac{\lambda_f(p)}{p^s} + \frac{\varepsilon_q(p)}{p^{2s}}\Bigr)^{-1}, \tag{1.1}$$

where ε_q is the trivial character mod q. This Dirichlet series is absolutely convergent when $\Re e(s)$ is large, and admits analytic continuation to all of $\mathbf{C}$. The completed L-function

$$\Lambda(f,s) = \widehat{q}^s \Gamma(s + \frac{1}{2}) L(f,s), \text{ where } \widehat{q} = \frac{\sqrt{q}}{2\pi}$$

satisfies the functional equation

$$\Lambda(f,s) = \varepsilon_f \Lambda(f, 1-s), \text{ where } \varepsilon_f = -q^{1/2}\lambda_f(q) = \pm 1.$$

We say that f is even (resp. odd) iff $\varepsilon_f = +1$ (resp. -1) and will use $S_2^+(q)$ (resp. $S_2^-(q)$) to denote the corresponding subset of $S_2(q)^*$. It is known that

$$|S_2^+(q)| \sim |S_2^-(q)| \sim \frac{1}{2}|S_2(q)^*| \sim \frac{q}{24}$$

as $q \to +\infty$ (among prime values).

We denote by r_f the "analytic rank" of $L(f,s)$, defined as the order of vanishing of $L(f,s)$ at $s = \frac{1}{2}$, which is the same as that of $\Lambda(f,s)$, and has the same parity as f.

For $k \geqslant 0$, let p_k be "the proportion of f with k-th derivative $\Lambda^{(k)}(f, \frac{1}{2}) \neq 0$", i.e.

$$p_k = \liminf_{q\to+\infty} \frac{|\{f \in S_2(q)^*,\ \Lambda(f,1/2) \neq 0\}|}{|S_2(q)^*|}. \tag{1.2}$$

Some of the results of [KM1, KM2, VdK1, I-S] can be summarized as follows:

Theorem 1.1. *We have*

$$p_0 > 0, \text{ and } p_1 > 0,$$

in other words, a positive proportion of even forms f are such that $L(f, \frac{1}{2}) \neq 0$, and a positive proportion of odd forms are such that $L'(f, \frac{1}{2}) \neq 0$.

The currently best bounds for these constants are $p_0 \geqslant 1/4$ and $p_1 \geqslant 7/16$ (see [I-S, KM2]). It is not a coincidence that the best bound for p_1 available is larger than the best one for p_0, as we show in this paper by considering higher derivatives.

Theorem 1.2. *For all $k \geqslant 0$, we have $p_k > 0$. In fact,*

$$p_k \geqslant \pi_k$$

where π_k is a function satisfying

$$\pi_k = \frac{1}{2} - \frac{1}{32}k^{-2} + O(k^{-3}). \tag{1.3}$$

In particular

$$p_2 \geqslant 0.48254,\ p_3 \geqslant 0.49478,\ p_4 \geqslant 0.49758, p_5 \geqslant 0.49856.$$

In fact, $p_k > 0$ was already proved in [VdK1] but the lower bound obtained there, although positive, approached 0 as $k \to +\infty$.

Since the set of forms f such that $\Lambda^{(k)}(f, 1/2) \neq 0$ is contained in either $S_2^+(q)$ or $S_2^-(q)$, we must have $p_k \leqslant \frac{1}{2}$. This is conjectured to be the an equality for $k = 0$ and 1 (so that p_k should be 0 for $k \geqslant 2$) by Brumer and Murty [Br, M]. While our result does not imply this conjecture, it at least shows that forms with high order of vanishing are rare.

As an immediate corollary of these computations, we obtain for example the following:

Corollary 1.3. *If q is prime and large enough, then at least 99% of the forms $f \in S_2(q)^*$ have analytic rank $\leqslant 4$.*

This is clear from the values of π_3 and π_4 given in the theorem.

In Section 8 we apply Theorem 1.2 to further investigate the analytic rank of the Jacobian variety $J_0(q)$ of $X_0(q)$; this application arose from discussions with B. Conrey and H. Iwaniec. Recall that by Eichler-Shimura theory [Shi], we have

$$L(J_0(q), s) = \prod_{f \in S_2(q)^*} L(f, s).$$

Using the ideas of [KM1] in addition to the results proved here, we are able to show:

Theorem 1.4. *Let α, $0 < \alpha < 2$ be a fixed real number. For q prime large enough, we have*

$$\frac{1}{|S_2(q)^*|} \sum_{f \in S_2(q)^*} r_f^\alpha \leqslant \tfrac{1}{2} + \sum_{k=0}^{+\infty} ((k+2)^\alpha - k^\alpha)(\tfrac{1}{2} - p_k) + o_\alpha(1).$$

For example[1]

$$\begin{aligned} \tfrac{1}{2} + \sum_{k=0}^{\infty}((k+2)^\alpha - k^\alpha)(\tfrac{1}{2} - p_k) \leqslant \tfrac{1}{2} + \sum_{k=0}^{\infty}((k+2)^\alpha - k^\alpha)(\tfrac{1}{2} - \pi_k) &\leqslant 1.1891 \quad \text{if } \alpha = 1 \\ &\leqslant 3.2191 \quad \text{if } \alpha = 1.9. \end{aligned}$$

In the case $\alpha = 2$, the series diverges but we can still prove that $\sum r_f^2 \leqslant C$, for some absolute constant C (see Theorem 8.1).

Corollary 1.5. *Assume the Birch and Swinnerton-Dyer conjecture for $J_0(q)$. Then for q prime large enough we have*

$$\operatorname{rank} J_0(q) \leqslant (c + o(1)) \dim J_0(q)$$

where $c = 1.1891$.

In [KM1], this was proved for some absolute constant c, and in [KM3], it had previously been proved that one could take $c = 6.5$. Corollary 1.5 is thus much stronger. This is explained by the fact that we are looking directly at the order of $L(f, s)$ at $s = \frac{1}{2}$, and not overcounting zeros in a small neighborhood, as was inherently the case for the methods of [KM1], [KM3]. However, because of lack of uniformity in k in the limit involved, we still have to use ideas similar to those of [KM1] (see Section 8).

Maybe the most striking feature of this inequality is that $c < \frac{3}{2}$, which was the first admissible value obtained by Brumer *assuming the Generalized Riemann Hypothesis*.

There are many circumstances in which the normalization $\lambda_f(1) = 1$ is not the most natural, and it is often more convenient to use the "harmonic" normalization (see [Du, KM1, KM2, M]).

Let $\omega_f = 1/4\pi(f, f)$,where (f, g) is the Petersson inner product on $\Gamma_0(q)\backslash\mathbf{H}$. For any set of complex numbers α_f, $f \in S_2(q)^*$, we write

$$\sum_f{}^h \alpha_f = \sum_f \omega_f \alpha_f.$$

[1] In view of (1.3) the above series is convergent for $\alpha < 2$

The weights ω_f define (asymptotically) a probability measure on $S_2(q)^*$ since

$$\sum_f{}^h 1 = 1 + O(q^{-3/2}), \tag{1.4}$$

and for this measure we prove the following variant of Theorem 1.2.

Theorem 1.6. *For all $k \geqslant 0$*

$$\liminf_{q\to+\infty} \sum_{f,\ \Lambda^{(k)}(f,1/2)\neq 0}{}^h 1 \geqslant \pi_k,$$

with the same π_k's as in Theorem 1.2.

Since $\displaystyle\sum_{f\in S_2^{\pm}(q)}{}^h 1 \sim \frac{1}{2}$, the result is still asymptotically optimal.

The advantage of these weights in our setting is that they make it possible to use the Petersson formula (see 3.2), which shows very clearly the very strong cancellation in the average of the product $\lambda_f(n)\lambda_f(m)$ of Hecke eigenvalues when $n \neq m$.

In this paper we will only prove Theorem 1.6, and the harmonic counterpart of Theorem 1.4. One may then deduce Theorem 1.2 and Theorem 1.4 using the technique for "removing" the weight ω_f developped and implemented in [KM1, Section 3] (see also [I-S, KM2]).

It is also possible to prove Theorem 1.2 directly by applying the Eichler-Selberg trace formula in the form of Lemma 3.4 below, instead of Lemma 3.2, but this leads to a smaller lower-bound $p_k \geqslant \pi_k$, although still satisfying $\pi_k = 1/2 + O(k^{-2})$.

Theorems 1.2 and 1.6 are in perfect analogy with the results of Conrey [Co], where it is proven that the proportion of the zeros of $\xi^{(k)}(s)$ (the k-th derivative of Riemann's ξ function) lying on the critical line tends to 1 as $k \to +\infty$. We thank B. Conrey and D. Farmer for pointing out the relevance of Conrey's paper to our problem (which ultimately leads to simplifications of our initial approach) and for explaining why these methods lead to higher proportions for higher k.

Although this is somewhat obscured by the fact that we are dealing with Euler products of degree 2, our proof is easier than Conrey's since we are looking for zeros at a specific point rather than throughout the critical line. Another aspect which simplifies our job is that since $\lambda_f(n)$ is real the functional equation is symmetric (the sign ε_f excepted), so that the $\Lambda^{(k)}(f, 1/2)$ are all real. This allows the oscillations of the sign of $\Lambda^{(k)}(f, 1/2)$ to be well controlled by the mollifier, which in turn minimizes the loss of information when we apply the Cauchy-Schwarz inequality.

The situation is slightly more difficult when complex values are allowed: for example, considering the same problem for L-functions associated to modular forms twisted by a non real character (or for Dirichlet L-functions of primitive characters) our method also gives proportions π_k approaching 1/2, but in that case there is no parity issue and the expected proportion should be 1. It turns out that a less obvious choice of mollifier leads to proportions approaching 2/3. We hope to come back to this topic in another paper.

As a final remark, we note that the above results are special cases of a general non-vanishing result for arbitrary linear combinations of derivatives of $\Lambda(f, s)$ (at $s = 1/2$): see Section 6 below for the general statement. In particular the numerical values found in Corollary 1.3 and in Corollary 1.5 are not the best possible. We refer the reader interested in numerical experiments and optimization questions to Sections 6, 7 and 8.3 were we describe the process required to get the optimal bounds and its limitations.

Note also that some authors (see [P] for instance) may find it more natural to look directly at derivatives of $L(f,s)$ rather than $\Lambda(f,s)$. Our method can be adapted to cover this problem also (see Section 6).

Acknowledgment. It is a pleasure to thank B. Conrey, D. Farmer, R. Heath-Brown, H. Iwaniec, P. Sarnak, R. de la Bretèche for discussions related to this work. This project was begun while the second author visited Rutgers University in April 98 and he would like to thank H. Iwaniec for his generous hospitality.

The computations were performed on MAPLE and on the PARI System of Batut, Bernardi, Cohen. The code of the programs used can be obtained from the authors.

2. The mollification

In this section, and until Section 6, we fix the integer $k \geqslant 0$. The dependency on k will become important later on and will be emphasized accordingly.

As is classical, our non-vanishing results are an application of the Cauchy-Schwarz inequality supplemented by the method of mollification of moments of the special values being investigated. We recall the motivation of the mollification technique.

Suppose that we were to consider the first and second (unmollified) moments

$$\mathcal{L}^h = \sum_{f\in S_2(q)^*}^h \Lambda^{(k)}(f,1/2),$$

$$\mathcal{Q}^h = \sum_{f\in S_2(q)^*}^h \Lambda^{(k)}(f,1/2)^2.$$

Using Lemma 3.2, one can show that, as $q\to+\infty$,

$$\mathcal{L}^h \sim c_k(\log\widehat{q})^k,\ \mathcal{Q}^h\sim c_k'(\log\widehat{q})^{2k+1} \tag{2.1}$$

for some $c_k, c_k' > 0$ (see in particular [Du] for this proof in the case $k=0$). This means that although the "typical" value of $\Lambda^{(k)}(f,1/2)$ is of size $c_k(\log\widehat{q})^k$ there is a small (but not negligible) subset of forms f for which $\Lambda^{(k)}(f,1/2)$ may be as large as $(\log\widehat{q})^{k+1/2}$ (in other words, the variance of the "random variable" $f\mapsto \frac{\Lambda^{(k)}(f,1/2)}{(\log\hat{q})^k}$ is large). By Cauchy-Schwarz, (2.1) implies that

$$\sum_{f,\ \Lambda^{(k)}(f,1/2)\neq 0}^h 1 \geqslant \frac{(\mathcal{L}^h)^2}{\mathcal{Q}^h} \gg \frac{1}{\log q}.$$

To save the $\log q$ factor (that is, to control the abnormally large oscillations of $\Lambda^{(k)}(f,1/2)$), we multiply it by a well chosen mollifier

$$M(f)=\sum_{m<M}\lambda_f(m)\frac{x_m}{\sqrt{m}},$$

with M of size $\widehat{q}^\Delta$ for some $\Delta>0$, such that $\Lambda^{(k)}(f,1/2)M(f)$ is more nearly constant over all f. We consider the "mollified" moments

$$\mathcal{L}^h = \sum_f^h \Lambda^{(k)}(f,1/2)M(f), \tag{2.2}$$

$$\mathcal{Q}^h := \sum_f^h |\Lambda^{(k)}(f,1/2)M(f)|^2. \tag{2.3}$$

While those sums become significantly more difficult to deal with, it is still possible to estimate them asymptotically for Δ small enough, and we prove that for M as before we have

$$\mathcal{L}^h \sim \mathcal{L}(x)(\log \widehat{q})^k$$
$$\mathcal{Q}^h \sim \mathcal{Q}(x)(\log \widehat{q})^{2k}$$

as $q \to +\infty$, where $\mathcal{L}$ and $\mathcal{Q}$ are linear and quadratic forms, respectively, in the coefficients $x = (x_m)_{m<M}$ of the mollifier (and do not depend on q).

Hence

$$\sum_{f,\ \Lambda^{(k)}(f,1/2)\neq 0}^{h} 1 \geqslant \frac{\mathcal{L}(x)^2}{\mathcal{Q}(x)}.$$

For some choices of (x_m), it is possible to make this ratio approach a constant as q (and thus M) goes to infinity.

The problem is now to choose (x_m) so that this ratio becomes as large as possible. Intuitively, the task of the mollifier becomes easier as k growths, since it will have to dampen oscillations of amplitude up to $(\log \widehat{q})^{1/2}$ of a variable of typical size $(\log \widehat{q})^k$. Thus even a "lazy" mollifier can be adequate, and in particular its length M can be rather short, i.e. Δ can be rather small. As we will see, for well chosen (x_m), any $\Delta > 0$ suffices to insure that

$$p_k \geqslant \tfrac{1}{2} + O(k^{-2}) \text{ as } k \to +\infty.$$

In the special cases $k = 0, 1$ the proportions p_0, p_1 have a deep arithmetical significance, so the authors of [I-S, KM1, KM2] considered a more general class of mollifiers, picking (x_m) only at the end of a delicate optimization process. Due to the nature of the expressions involved for more general k, it is not possible to do this here, but since the demand on the mollifier becomes less stringent for large k as well, it is sufficient (and much simpler) to start with a particular type of mollifier, which mimick those of [Co].

Let $M = \widehat{q}^{\Delta}$ for a fixed $0 < \Delta < 1$ to be chosen later, with the proviso that M is not an integer. Let P be a polynomial which satisfies $P(0) = P'(0) = 0$, and let ψ be the arithmetic function

$$\psi(m) = \prod_{p|m}(1 + \frac{1}{p}). \tag{2.4}$$

We choose the following as our mollifier for the kth derivative:

$$M(f) = M_P(f) = \sum_{m<M} \frac{\lambda_f(m)\mu(m)}{\psi(m)m^{1/2}} P(\frac{\log M/m}{\log M}). \tag{2.5}$$

The optimal choices of mollifiers for $k = 0, 1$ were of this form[2], so we expect that this is not a particularly harsh restriction on the mollifiers. This expectation is borne out by the strength of our results, since we do obtain the asymptotically best possible result $p_k \to 1/2$.

As usual, the specific coefficients were chosen to make $M_P(f)$ a reasonable approximation to the inverse of $\Lambda^{(k)}(f, 1/2)$ (the Möbius function is the "universal inverse" for Euler products, and the function ψ accounts for the degree two portion of the Euler factors). The choice of P will be made at the end of the proof. For now we prefer to keep the argument as general as possible. Since optimization will now be performed on the space of polynomials, rather than vectors, we henceforth denote the moments by $\mathcal{L}^h(P)$ and $\mathcal{Q}^h(P)$.

[2] It turns out that the choices $P = x^2$, for $k = 0$, $P = x^2 - x^3/6$, for $k = 1$, give again the non-vanishing proportion $p_0 \geqslant 1/4$ and $p_1 \geqslant 7/16$ for the critical values and their first derivative, if we let $\Delta \to 1$. However the study of the general quadratic form has other applications

One advantage of this type of mollifier is that the polynomial has a convenient contour integral expression. Given a polynomial $P(x) = \sum_k a_k x^k$, we define the auxiliary function

$$\widehat{P_M}(s) = \sum_k k! a_k (s \log M)^{-k}.$$

Then we have

Lemma 2.1. *If $P(0) = 0$ and M is not an integer,*

$$\delta_{m<M} P(\frac{\log M/m}{\log M}) = \frac{1}{2\pi i} \int_{(3)} \frac{M^s}{m^s} \widehat{P_M}(s) \frac{ds}{s},$$

where $\delta_{m<M}$ is 1 if $m < M$ and 0 if $m > M$.

Proof. To evaluate the integral, we shift the contour. If $m > M$, we shift to $\Re\mathrm{e}(s) = S$, where S is large, so that we have a convergent integral (since $P(0) = 0$) dominated by $(M/m)^S \to 0$. If $m < M$, we shift to $\Re\mathrm{e}(s) = -S$, with S large. The resulting integral again tends to zero, and as explained in the appendix, Lemma 9.1, the residue at $s = 0$ is exactly that given above. The condition $P(0) = 0$ was needed to make the integrals absolutely convergent, since all denominators will thus have s^2 or higher powers. □

3. Auxiliary Lemmas

Here we recall some lemmas about modular forms used in the sequel. More technical lemmas are presented in an Appendix at the end of the paper.

We begin with Hecke's recursion formula for primitive forms, which is equivalent to the Euler product (1.1).

Lemma 3.1. *For $m, n \geqslant 1$ and $f \in S_2(q)^*$ one has*

$$\lambda_f(m)\lambda_f(n) = \sum_{d|(m,n)} \varepsilon_q(d) \lambda_f\Big(\frac{mn}{d^2}\Big) \tag{3.1}$$

where ε_q is the trivial character modulo q.

The next lemma is the special case of the Petersson formula for prime level q and weight 2 (in this case $S_2(q)^*$ is an orthogonal basis of $S_2(q)$).

Lemma 3.2. *For $m, n \geqslant 1$ one has*

$$\sum_{f \in S_2(q)^*}^{h} \lambda_f(m)\lambda_f(n) = \delta_{m,n} - 2\pi \sum_{c \geqslant 1} \frac{S(m, n; cq)}{cq} J_1\Big(\frac{4\pi\sqrt{mn}}{cq}\Big) \tag{3.2}$$

where $\delta_{m,n}$ is the Kronecker symbol,

$$S(m, n; c) = \sum_{\substack{x \bmod c \\ (x,c)=1}} e\Big(\frac{mx + n\overline{x}}{c}\Big)$$

is the classical Kloosterman sum and $J_1(x)$ is the Bessel function of order 1. Moreover one has the estimation

$$\sum_{f \in S_2(q)^*}^{h} \lambda_f(m)\lambda_f(n) = \delta_{m,n} + O((m, n, q)^{1/2}(mn)^{1/2} q^{-3/2}). \tag{3.3}$$

This last bound turns out to be sufficient to allow us to to take a mollifier of length $M = \hat{q}^{\Delta}$ for any $\Delta < 1/2$. To take a longer mollifier requires greater cancellation, and to get this one must "open" the Kloosterman sums and take advantage of the cancellation inherent in geometric sums. This allows better bounds on average. One can show (see [VdK2] for details) the following

Lemma 3.3. *Let N_1, N_2, m_1, m_2 be such that*

$$N_1N_2 \ll q(\log q)^2, \quad m_1m_2 \ll q^{1-\delta}$$

for some $\delta > 0$. Then for all $\varepsilon > 0$

$$\sum_{\substack{n_1\sim N_1\\ n_2\sim N_2}} \sum_{c\geqslant 1} \frac{S(m_1n_1, m_2n_2; cq)}{cq} J_1\Big(\frac{4\pi\sqrt{m_1n_1m_2n_2}}{cq}\Big) \ll_{\varepsilon,\delta} q^{\varepsilon} \frac{(m_1m_2N_1N_2)^{1/2}}{q}.$$

For comparison, using the Weil bound alone gives $q^{-1/2+\varepsilon}(m_1m_2N_1N_2)^{1/2}$ here. Lemma 3.3 allows one to take any $\Delta < 1$ and this leads to the optimal values of p_k stated in the introduction.

Finally, we mention a lemma which is at the core of the argument in [VdK1]. It is not used in this paper, but as we have mentioned before, it could replace Lemma (3.3) to prove Theorem 1.2 without recourse to the harmonic average.

Lemma 3.4. *Given (a_n) any bounded sequence of numbers and $N \leqslant q^{2-\delta}$ for some $\delta > 0$, where q is prime.*

$$\frac{1}{|S_2(q)^*|} \sum_{n\sim N} a_n \sum_{f\in S_2(q)^*} \lambda_f(n) = \sum_{n\sim N} \frac{a_n}{n^{1/2}} \delta_{n=\square} + O_{\delta}(N^{7/4}q^{-3/2} + N^{39/32}q^{-3/4}), \tag{3.4}$$

where $\delta_{n=\square}$ is the characteristic function of squares of integers.

4. The first moment

We now begin the proof of Theorem 1.6. The first step is to evaluate the first moment, $\mathcal{L}^h(P)$ (see (2.2)) with $M_P(f)$ defined in (2.5).

Since $1/2$ is outside the region where the Dirichlet series expression for $L(f,s)$ converges, we must first find a way to express $\Lambda^{(k)}(f,1/2)$ as a rapidly convergent series. We note that when $\Re e(s) > 1/2$, so that all of the relevant sums are absolutely convergent, we have

$$\Lambda_f^{(k)}(\frac{1}{2}+s) = \sum_n \frac{\lambda_f(n)\widehat{q}^{1/2}}{n^{1/2}} \left(\frac{\widehat{q}^s}{n^s}\Gamma(1+s)\right)^{(k)}$$

A standard contour shift, along with the functional equation, gives

$$\Lambda^{(k)}(f,1/2) = (1+(-1)^k\varepsilon_f)\widehat{q}^{1/2} \sum_n \frac{\lambda_f(n)}{n^{1/2}} V_{total}(\frac{n}{\widehat{q}}), \tag{4.1}$$

where

$$V_{total}(y) = \frac{1}{2\pi i}\int_{(3)} \left(\frac{\Gamma(1+t)}{y^t}\right)^{(k)} \frac{dt}{t} = \sum_{l=0}^{k} \binom{k}{l} (-\log y)^l \frac{1}{2\pi i}\int_{(3)} (\Gamma(1+t))^{(k-l)} y^{-t}\frac{dt}{t} \tag{4.2}$$

satisfies

$$V_{total}(y) = O_N(y^{-N})$$

for all $N \geqslant 1$ when y is large. In addition, when y is small,

$$V_{total}(y) = (-\log y)^k + R(-\log y) + O(y), \text{ as } y \to 0$$

where R is a polynomial of degree $\leqslant k-1$. Over the course of this paper, we will only be interested in the leading term of the expressions occuring, which give the asymptotic behavior, and we will drop terms of lower order of magnitude in $\log \widehat{q}$ and $\log M$ whenever feasible. In order to simplify nearly every formula yet to come, we restrict ourselves to evaluating the single term corresponding to $(-\log y)^k$ in (4.2), whose contribution to (4.1) we denote by $\Lambda^{(k)}_{main}(f, 1/2)$. All other terms can be evaluated in exactly the same way and will be smaller by a factor of at least $(\log \widehat{q})^{-1}$. As $\widehat{q}$ goes to infinity, we may ignore them to get the asymptotic formula.

Thus, multiplying by $M_P(f)$, the main term of the first harmonic moment is

$$\mathcal{L}_{main}(P) = \sum_f^h \Lambda^{(k)}_{main}(f, 1/2) M_P(f),$$

where

$$\Lambda^{(k)}_{main}(f, 1/2) M_P(f) = (1 + \epsilon_f(-1)^k)\widehat{q}^{1/2} \sum_{m,n} \frac{\lambda_f(m)\lambda_f(n)\mu(m)}{\psi(m)\sqrt{mn}} (\log \frac{\widehat{q}}{n})^k P(\frac{\log(M/m)}{\log M}) V(\frac{n}{\widehat{q}})$$

and now

$$V(y) = \frac{1}{2\pi i} \int_{(3)} \frac{\Gamma(1+t)}{y^t} \frac{dt}{t}.$$

Applying the Petersson formula (3.3), we infer that when $M = \widehat{q}^{\Delta}$ for a fixed Δ, $0 < \Delta < 1$, we have

$$\mathcal{L}_{main}(P) = \widehat{q}^{1/2} \sum_m \frac{\mu(m)}{\psi(m)\sqrt{m}} (\log \frac{\widehat{q}}{m})^k P(\frac{\log M/m}{\log M}) V(\frac{m}{\widehat{q}}) + O(\widehat{q}^{1/2-\gamma}). \tag{4.3}$$

for some $\gamma = \gamma(\Delta) > 0$ depending on Δ only (see [I-S, KM1] for details). In particular, the terms involving the sign of the functional equation ε_f only contribute part of the remainder term.

We now focus on the main term of (4.3), which can be evaluated in any of several ways. We could replace $V(m/\widehat{q})$ by 1 (since $\Delta < 1$, $m/\widehat{q} \to 0$) and then evaluate arithmetic sums involving P and the Möbius function in the spirit of [Co]. For diversity we choose instead to apply directly the method of residues: we use the integral expressions for P and V, and introduce the integral expression (again by Lemma 9.1 in the Appendix)

$$(\log \frac{\widehat{q}}{m})^k = \frac{k!}{2\pi i} \int_{C_\delta} \frac{\widehat{q}^z dz}{m^z z^{k+1}},$$

where C_δ is a circle of radius δ around the origin in the complex plane. We will be adjusting δ as becomes necessary through the course of the proof, the exact radius is unimportant. Thus $\mathcal{L}_{main}(P)$ is given by the integral

$$\frac{k!\widehat{q}^{1/2}}{(2\pi i)^3} \int_{C_\delta} \int_{(3)} \int_{(3)} \sum_m \frac{\mu(m)}{\psi(m)m} \frac{M^s \widehat{q}^t \Gamma(1+t)\widehat{P_M}(s)}{m^s m^t} \frac{\widehat{q}^z}{z^{k+1} m^z} \frac{ds}{s} \frac{dt}{t} dz. \tag{4.4}$$

Executing the summation in m, which is absolutely convergent for the variables in the given range, we have

$$\frac{k!\widehat{q}^{1/2}}{(2\pi i)^3} \int_{C_\delta} \int_{(3)} \int_{(3)} \frac{M^s \widehat{q}^{t+z} \Gamma(1+t)\widehat{P_M}(s)}{z^{k+1}} \frac{\eta_1(s,t,z)}{\zeta(1+s+t+z)} \frac{ds}{s} \frac{dt}{t} dz, \tag{4.5}$$

where

$$\eta_1(s,t,z) = \prod_p \left(\frac{1 - p^{-1-s-t-z}(1+p^{-1})^{-1}}{1 - p^{-1-s-t-z}} \right)$$

is absolutely convergent and holomorphic on $\Re\mathrm{e}(s+t+z) > -1/2$, with $\eta_1(0,0,0) = \zeta(2)$.

It remains to evaluate (4.5). We will first evaluate the residues coming from the poles in t, s, and z at the origin, using the lemmas from the appendix. Then we will show that, other than these residues, (4.5) contributes a negligible amount to the first moment.

4.1. **Evaluating the residues.** As far as the residues at the origin are concerned, we may replace $\zeta(1+s+t+z)$ by $(s+t+z)^{-1}$. In addition, the $\Gamma(1+t)$ and $\eta_1(s,t,z)$ terms, which have neither poles nor zeros at the point in question, and which are independent of M and $\widehat{q}$, can be treated as constants, since any part of the residue coming ¿from their derivatives will have fewer powers of $\log M$ and $\log \widehat{q}$ than the main term. Thus we are interested in evaluating

$$k!\zeta(2)\widehat{q}^{1/2}\mathrm{Res}_{s=z=t=0}\frac{M^s\widehat{q}^{z+t}\widehat{P_M}(s)(s+t+z)}{stz^{k+1}}. \tag{4.6}$$

The t residue comes from a simple pole, leaving us with

$$k!\zeta(2)\widehat{q}^{1/2}\mathrm{Res}_{s=z=0}\frac{M^s\widehat{q}^{z}\widehat{P_M}(s)(s+z)}{sz^{k+1}}.$$

The remaining terms can be evaluated by the lemmas in the appendix, giving a total residue of

$$\zeta(2)\widehat{q}^{1/2}\Big(\log\widehat{q}\Big)^k\Big(\frac{1}{\log M}P'(1)+\frac{1}{\log\widehat{q}}kP(1)\Big). \tag{4.7}$$

4.2. **Shifting the contours.** In evaluating (4.5), we start by shifting the t and s contours to $\Re\mathrm{e}(s) = 1/2$, passing no poles in the process. Then we shift the t contour almost to $-1/2$, far enough to make $\widehat{q}^{\Re\mathrm{e}(t)+\delta}M^{1/2}$ negligible (this is possible so long as $\log_{\widehat{q}} M = \Delta < 1$). The resulting integral is too small to contribute to the main term, since the argument of the zeta function is > 1. Thus only the residue at $t = 0$ matters, and we have to estimate (up to constants) the integral

$$\frac{k!}{(2\pi i)^2}\int_{(1/2)}\int_{C_\delta}\frac{\widehat{q}^zM^s\widehat{P_M}(s)\eta(s,0,z)}{sz^{k+1}\zeta(1+s+z)}dsdz.$$

Next we evaluate the z integral, which merely involves taking k derivatives of the inner expression. The exact result does not matter, merely note that the only terms in the denominator will be powers of $\zeta(1+s)$, and there will only be powers of M remaining.

Finally we shift the s contour to the left, past $\Re\mathrm{e}(s) = 0$ but before crossing any zeros of $\zeta(1+s)$, so as not to introduce any extra poles. Such a contour exists by the prime number theorem (were the Riemann Hypothesis known, we could shift s all the way to $\Re\mathrm{e}(s) = -1/2+\delta$ for any $\delta > 0$, but fortunately this is not needed). It is known that there exists a constant $c > 0$ such that we can shift the contour this way to the one given by

$$\Re\mathrm{e}(s) = -c/\log(1+|\mathrm{Im}(s)|),$$

and this will be enough. The pole we acquire from $s = 0$ has already been discussed in the previous section. As for the integral, note that on the new contour, $1/\zeta(1+s)$ has no pole, and in fact

$$\zeta^{(j)}(1+s) \ll (1+|\mathrm{Im}(s)|)^\delta$$

for any $\delta > 0$ and any $j \geqslant 0$. So the various derivatives of $1/\zeta(1+s)$ acquired from the z integration can be replaced by $(1+|\mathrm{Im}(s)|)^\delta$ for any small δ. Thus the integral is bounded (since $P(0) = 0$) by

$$\int M^{-c/\log(1+|t|)}(1+|t|)^{-2+\delta}dt \ll e^{-c'(\log M)^{1/2-\delta}}$$

which is small enough to overcome any $\log\widehat{q}$ terms that might arise from the z derivatives. Thus the contributions from these contours are bounded, and they can be ignored in the calculation of the first moment, as desired.

Hence we have proved the following proposition:

Proposition 4.1. *For fixed* $0<\Delta<1$ *and* $M=\widehat{q}^{\Delta}$

$$\mathcal{L}^h(P)=(1+O_k((\log q)^{-1}))\zeta(2)\widehat{q}^{1/2}(\log\widehat{q})^k\Big(\frac{1}{\log M}P'(1)+\frac{1}{\log\widehat{q}}kP(1)\Big).$$

5. The second moment

The principle is the same as before. We start again with a rapidly convergent series which gives $\Lambda^{(k)}(f,1/2)^2$. The functional equation for $\Lambda(f,1/2+s)$ has always sign $+1$ so manipulations similar as those performed in the previous section yield

$$\Lambda^{(k)}(f,1/2)^2=2\widehat{q}\sum_{n_1,n_2}\frac{\lambda_f(n_1)\lambda_f(n_2)}{(n_1n_2)^{1/2}}\times\frac{1}{2\pi i}\int_{(3)}\Big(\frac{\widehat{q}^t\Gamma(1+t)}{n_1^t}\Big)^{(k)}\Big(\frac{\widehat{q}^t\Gamma(1+t)}{n_2^t}\Big)^{(k)}\frac{dt}{t}.$$

Again, derivatives of $\Gamma(1+t)$ will contribute lower orders of magnitude in $\log\widehat{q}$, and can thus be ignored (for the exact value of their contributions, simply carry through this proof with $\Gamma^{(\ell)}$ replacing one or both of the Γ's, and $k-\ell$ replacing either k). Thus we need only consider

$$\Lambda^{(k)}_{main}(f,1/2)^2=2\widehat{q}\sum_{n_1,n_2}\frac{\lambda_f(n_1)\lambda_f(n_2)}{(n_1n_2)^{1/2}}(\log\frac{\widehat{q}}{n_1})^k(\log\frac{\widehat{q}}{n_2})^kW(\frac{n_1n_2}{\widehat{q}^2}),\tag{5.1}$$

where

$$W(y)=\frac{1}{2\pi i}\int_{(3)}\frac{\Gamma(1+t)^2}{y^t}\frac{dt}{t}\tag{5.2}$$

decays faster than any negative power of y.

Multiplying by the square of the mollifier we evaluate

$$\begin{aligned}\mathcal{Q}_{main}(P)&=\sum_f{}^h M_P(f)^2\Lambda^{(k)}_{main}(f,1/2)^2\\&=2\widehat{q}\sum_{n_1,n_2,m_1,m_2}\frac{\mu(m_1)\mu(m_2)}{\psi(m_1)\psi(m_2)\sqrt{m_1m_2n_1n_2}}P(\frac{\log M/m_1}{\log M})P(\frac{\log M/m_2}{\log M})\\&\quad\times(\log\frac{\widehat{q}}{n_1})^k(\log\frac{\widehat{q}}{n_2})^kW(\frac{n_1n_2}{\widehat{q}^2})\sum_f{}^h\lambda_f(m_1)\lambda_f(n_1)\lambda_f(m_2)\lambda_f(n_2).\end{aligned}$$

As in the previous section, we wish to use the Petersson formula for the inner sum over f. Since, it only applies to products of two eigenvalues, we first appeal to the Hecke recursion (3.1).

Note that since q, the level, is prime, and we have the restriction $m_1,m_2<M<\widehat{q}^{\Delta}\leqslant q^{\Delta/2}$, the trivial character ε_q won't appear in the recursion formula.

Replacing m_1,n_1,m_2,n_2 by $m_1d_1,n_1d_1,m_2d_2,n_2d_2$, we can rewrite the main term of the second moment as

$$\begin{aligned}\mathcal{Q}_{main}=2\widehat{q}\sum_{d_1,d_2}\frac{1}{d_1d_2}\sum_{n_1,n_2,m_1,m_2}&\frac{\mu(m_1d_1)\mu(m_2d_2)}{\psi(m_1d_1)\psi(m_2d_2)\sqrt{m_1m_2n_1n_2}}P(\frac{\log M/(m_1d_1)}{\log M})P(\frac{\log M/(m_2d_2)}{\log M})\\&\times(\log\frac{\widehat{q}}{n_1d_1})^k(\log\frac{\widehat{q}}{n_2d_2})^kW(\frac{n_1n_2d_1d_2}{\widehat{q}^2})\sum_f{}^h\lambda_f(m_1n_1)\lambda_f(m_2n_2).\end{aligned}$$

The inner sum can now be evaluated by the Petersson formula. Integrating by parts and using Lemma 3.3, one shows (see [I-S] or [VdK2] for details, another possibility could be to use the more robust but much less elementary large sieve type inequality for sums of Kloosterman sums given by Proposition 1 of [DFI]) that the terms coming from the Kloosterman sums have a total contribution which is

$$\ll \widehat{q}^{1-\gamma}$$

for some $\gamma=\gamma(\Delta)>0$ if $\Delta<1$ (if one uses the Weil bound on the individual Kloosterman sums, this follows only in the range $\Delta<1/2$).

Thus we can replace the inner average over f by the Kronecker symbol $\delta_{m_1n_1,m_2n_2}$. The most convenient way to do this is to let $c=(m_1,m_2)$, so that n_1 and n_2 must be proportional to m_2/c and m_1/c, respectively. We make the substitutions $n_1=nm_2/c$ and $n_2=nm_1/c$, then replace m_1,m_2 by m_1c,m_2c, so that the main term of the second moment becomes

$$2\widehat{q}\sum_{d_1,d_2}\frac{1}{d_1d_2}\sum_{c}\frac{1}{c}\sum_{n}\frac{1}{n}\sum_{(m_1,m_2)=1}\frac{\mu(m_1d_1c)\mu(m_2d_2c)}{\psi(m_1d_1c)\psi(m_2d_2c)m_1m_2}P(\frac{\log M/(m_1d_1c)}{\log M})P(\frac{\log M/(m_2d_2c)}{\log M})$$

$$\times(\log\frac{\widehat{q}}{m_2d_1n})^k(\log\frac{\widehat{q}}{m_1d_2n})^kW(\frac{m_1m_2d_1d_2n^2}{\widehat{q}^2}). \tag{5.3}$$

As before, we evaluate this by converting to integral expressions and using residues and contour shifts.

By the formula

$$(\log\frac{\widehat{q}}{n_i})^k=\frac{k!}{2\pi i}\int_{C_{\delta_i}}\frac{\widehat{q}^{z_i}dz_i}{z_i^{k+1}n^{z_i}},$$

our main term is expressed as

$$\frac{2\widehat{q}(k!)^2}{(2\pi i)^5}\int_{C_{\delta_2}}\int_{C_{\delta_1}}\int_{(3)}\int_{(3)}\int_{(3)}\frac{\widehat{q}^{2t+z_1+z_2}M^{s_1+s_2}\widehat{P_M}(s_1)\widehat{P_M}(s_2)}{\Gamma(1+t)^{-2}s_1s_2tz_1^{k+1}z_2^{k+1}}\eta(s_1,s_2,t,z_1,z_2)ds_1ds_2dtdz_1dz_2,$$

with

$$\eta(s_1,s_2,t,z_1,z_2)=\sum_{\substack{m_1,m_2,d_1,d_2,c,n\\(m_1,m_2)=1}}\frac{\mu(m_1d_1c)\mu(m_2d_2c)}{\psi(m_1d_1c)\psi(m_2d_2c)} \tag{5.4}$$

$$\times\frac{1}{d_1^{1+s_1+t+z_2}d_2^{1+s_2+t+z_1}c^{1+s_1+s_2}n^{1+2t+z_1+z_2}m_1^{1+s_1+t+z_1}m_2^{1+s_2+t+z_2}}.$$

The sums are rather intricate, but the ψ functions will only lead to second-order corrections, as will the relative primality restrictions. Thus, up to Euler products which converge to the left of $s_1=s_2=t=z_1=z_2=0$, the sums must lead to two factors of ζ and four factors of ζ^{-1}. Thus we write the second moment as

$$\frac{2\widehat{q}(k!)^2}{(2\pi i)^5}\int_{C_{\delta_2}}\int_{C_{\delta_1}}\int_{(3)}\int_{(3)}\int_{(3)}\frac{\Gamma(1+t)^2\widehat{q}^{2t+z_1+z_2}M^{s_1+s_2}\widehat{P_M}(s_1)\widehat{P_M}(s_2)}{s_1s_2tz_1^{k+1}z_2^{k+1}}$$

$$\times\frac{\zeta(1+2t+z_1+z_2)\zeta(1+s_1+s_2)\eta_2(s_1,s_2,t,z_1,z_2)ds_1ds_2dtdz_1dz_2}{\zeta(1+s_1+t+z_1)\zeta(1+s_2+t+z_1)\zeta(1+s_1+t+z_2)\zeta(1+s_2+t+z_2)}, \tag{5.5}$$

where η_2 is an Euler product which is absolutely convergent even when all five variables are slightly to the left of the origin (in particular, $\eta_2(0,0,0,0,0)=\prod_p(1+O(p^{-2}))$).

Once again, we will evaluate the moment through evaluating the residues when all the variables are zero, and then we will show that the contour integrals which result after shifting all of the contours to the left of zero give negligible amounts. Lest the main theme be lost in the mass of calculations which are to follow, note that the total powers of $\log \widehat{q}$ and $\log M$ which come from an integral of this sort depend entirely on the degree of pole at the origin. The ζ's increase that degree, the $1/\zeta$'s decrease that degree, and the net effect is that the degree of the second moment is twice that of the first moment, as one would want. Had we not mollified, there would be no ζ's in the numerator, one fewer in the denominator, and no s poles, leading to a net increase of a single factor of $\log \widehat{q}$ from that which we will now find.

5.1. **Evaluating the residues.** We can again replace $\zeta(1+z)$ by z^{-1} in (5.5) as far as calculations at the origin are concerned, so the relevant expression is now

$$2\widehat{q}(k!)^2 \mathrm{Res}_{(0,0,0,0,0)} \frac{\Gamma(1+t)^2 M^{s_1+s_2} \widehat{q}^{2t+z_1+z_2} \widehat{P_M}(s_1)\widehat{P_M}(s_2)\eta_2(s_1,s_2,z_1,z_2,t)}{ts_1s_2z_1^{k+1}z_2^{k+1}}$$

$$\times \frac{(s_1+z_1+t)(s_2+z_1+t)(s_1+z_2+t)(s_2+z_2+t)}{(z_1+z_2+2t)(s_1+s_2)}. \tag{5.6}$$

As with η_1 in the first moment, derivatives of η_2 and Γ can be ignored, since they lead to lower powers of $\log \widehat{q}$ and $\log M$. Once again we have a simple pole in t, so its residue involves merely replacing t with zero everywhere. We thus have

$$2\widehat{q}(k!)^2 \eta_2(0,0,0,0,0) \mathrm{Res}_{(0,0,0,0)} \frac{Q^{s_1+s_2} \widehat{q}^{z_1+z_2} \widehat{P_M}(s_1)\widehat{P_M}(s_2)}{s_1s_2z_1^{k+1}z_2^{k+1}}$$

$$\times \frac{(s_1+z_1)(s_2+z_1)(s_1+z_2)(s_2+z_2)}{(z_1+z_2)(s_1+s_2)}. \tag{5.7}$$

Note that this is now nearly symmetric in $s \leftrightarrow z$. It would be completely so if we were using arbitrary linear combinations of derivatives of Λ, replacing $k!/z_i^k$ by a general $\widehat{Q}(z_i)$. We multiply out

$$(s_1+z_1)(s_2+z_1)(s_1+z_2)(s_2+z_2) = s_1^2s_2^2 + (s_1+s_2)s_1s_2(z_1+z_2) + (s_1^2+s_2^2)z_1z_2$$
$$+ s_1s_2(z_1+z_2)^2 + (s_1+s_2)z_1z_2(z_1+z_2) + z_1^2z_2^2,$$

collecting terms in a manner which is more convenient for the application of the lemmas in the appendix. To evaluate the second and fifth terms, we need only apply Lemma 9.1 and Corollary 9.2, while the other terms require Lemma 9.4 as well. The calculations are now completely straightforward, and we wind up with a residue at the origin which is equal to

$$2\widehat{q}\eta_2(0,0,0,0,0)(\log \widehat{q})^{2k} \left(\frac{\log \widehat{q}}{(\log M)^3} \frac{1}{2k+1} \int_0^1 P''(x)^2 dx + \frac{1}{(\log M)^2} P'(1)^2 \right.$$

$$+ \frac{1}{(\log \widehat{q})(\log M)} (2\frac{k^2}{2k-1} \int_0^1 P(x)P''(x)dx + 2k \int_0^1 P'(x)^2 dx)$$

$$\left. + \frac{k^2}{(\log \widehat{q})^2} P(1)^2 + \frac{\log M}{(\log \widehat{q})^3} \frac{k^2(k-1)^2}{2k-3} \int_0^1 P(x)^2 dx \right). \tag{5.8}$$

It remains to show that $\eta_2(0,0,0,0,0) = \zeta(2)^2$. Returning to the original sum (5.4), we need to factor the zeta functions out of

$$\sum_{d_1,d_2,c,n} \sum_{(m_1,m_2)=1} \frac{\mu(m_1d_1c)\psi(m_1d_1c)^{-1}\mu(m_2d_2c)\psi(m_2d_2c)^{-1}}{d_1^{1+s_1+t+z_2}d_2^{1+s_2+t+z_1}c^{1+s_1+s_2}n^{1+2t+z_1+z_2}m_1^{1+s_1+t+z_1}m_2^{1+s_2+t+z_2}}.$$

The n sum is immediate, giving $\zeta(1+2t+z_1+z_2)$.

We remove the condition $(m_1,m_2)=1$ by Möbius inversion

$$\sum_{(m_1,m_2)=1} f(m_1,m_2) = \sum_{m_1,m_2} f(m_1,m_2) \sum_{b|m_1,m_2} \mu(b) = \sum_b \mu(b) \sum_{m_1,m_2} f(bm_1,bm_2),$$

(for any arithmetic function f).

Replacing bc by c, our sum is

$$\zeta(1+2t+z_1+z_2)\sum_c \frac{\mu(c)^2\psi(c)^{-2}}{c^{1+s_1+s_2}} \sum_{b|c} \frac{\mu(b)}{b^{1+2t+z_1+z_2}}$$

$$\times \sum_{(m_1d_1,c)=1} \frac{\mu(m_1d_1)\psi(m_1d_1)^{-1}}{m_1^{1+s_1+t+z_1}d_1^{1+s_1+t+z_2}} \sum_{(m_2d_2,c)=1} \frac{\mu(m_2d_2)\psi(m_2d_2)^{-1}}{m_2^{1+s_2+t+z_2}d_2^{1+s_2+t+z_1}}.$$

The last two sums have the same form, so we next evaluate

$$\sum_{(md,c)=1} \frac{\mu(md)}{\psi(md)d^{1+u}m^{1+v}}.$$

The sum over m is

$$\prod_{(p,cd)=1} (1-p^{-1-v}\psi(p)^{-1}) = \prod_{(p,cd)=1} (1-p^{-1-v})\eta_p(v),$$

where $\eta_p(0) = (1-p^{-2})^{-1}$. Inserting this into the sum over d, we get

$$\prod_{(p,c)=1} (1-p^{-1-v})\eta_p(v) \sum_{(d,c)=1} \frac{\mu(d)}{\psi(d)d^{-1-u}} \prod_{p|d} \frac{1+p^{-1}}{1+p^{-1}-p^{-1-v}}.$$

We find that this is equal to

$$\prod_{(p,c)=1} \frac{(1+p^{-1}-p^{-1-u}-p^{-1-v})}{(1+p^{-1})} = \frac{\eta^*(u,v)}{\zeta(1+u)\zeta(1+v)} \prod_{p|c} \frac{(1+p^{-1})}{1+p^{-1}-p^{-1-u}-p^{-1-v}}, \tag{5.9}$$

where $\eta^*(0,0) = \zeta(2)$. Thus the m,d sums give the zeta functions desired for the denominator, and also two factors $\zeta(2)$ for the numerator.

Finally, we must evaluate the sum over b and c. Since all we want is $\eta_2(0,0,0,0,0)$, we can set $u=v=0$ in the product on the right-hand side of (5.9). Likewise we may set $2t+z_1+z_2=0$ in the exponent of b. Our remaining sum is thus

$$\sum_c \frac{\mu(c)^2}{\psi(c)^2c^{1+s_1+s_2}}(\sum_{b|c}\frac{\mu(b)}{b})\prod_{p|c}\frac{(1+p^{-1})^2}{(1-p^{-1})^2} = \sum_c \frac{\mu(c)^2}{c^{1+s_1+s_2}}\prod_{p|c}(1-p^{-1})^{-1}$$

$$= \prod_p (1+\frac{p^{-1-s_1-s_2}}{1-p^{-1}}) = \zeta(1+s_1+s_2)\prod_p \frac{(1-p^{-1-s_1-s_2})(1-p^{-1}+p^{-1+s_1+s_2})}{(1-p^{-1})}$$

$$= \zeta(1+s_1+s_2)\eta_3(s_1,s_2),$$

where $\eta_3(0,0)=1$. Thus $\eta_2(0,0,0,0,0)=\eta^*(0,0)^2\eta_3(0,0)=\zeta(2)^2$, and we have completed the evaluation of the main term of the second moment.

5.2. **Shifting the contours.** The goal is to bound all contributions to the integral

$$\frac{2(k!)^2}{(2\pi i)^5}\int_{C_{\delta_2}}\int_{C_{\delta_1}}\int_{(3)}\int_{(3)}\int_{(3)}\frac{\Gamma(1+t)^2\widehat{q}^{2t+z_1+z_2}M^{s_1+s_2}}{s_1s_2tz_1^{k+1}z_2^{k+1}}$$

$$\times\frac{\zeta(1+2t+z_1+z_2)\zeta(1+s_1+s_2)\eta_2(s_1,s_2,t,z_1,z_2)ds_1ds_2dtdz_1dz_2}{\zeta(1+s_1+t+z_1)\zeta(1+s_2+t+z_1)\zeta(1+s_1+t+z_2)\zeta(1+s_2+t+z_2)},$$

other than that arising from the pole at $s_1=s_2=t=z_1=z_2=0$, by something which is smaller than $(\log\widehat{q})^{2k-2}$. The general approach is the same as with the first moment, but the presence of extra zeta functions, especially in the numerator, makes things slightly more delicate.

Since the radii of the z contours were arbitrary, we set $\delta_2>\delta_1$ so that $z_1+z_2\neq 0$. As with the first moment, we shift the s_1,s_2,t contours to $1/2$, then shift the t contour to $\Re(t)=-1/2+\delta$, with $\delta>\delta_1,\delta_2$, but δ small enough that $\widehat{q}^{2\delta+\delta_1+\delta_2-1}M$ is still negligible. Provided that $\Delta<1$, we can always choose δ,δ_1,δ_2 so that this is the case. Since everything in the integral other than $\widehat{q}^{2t+z_1+z_2}M^{s_1+s_2}$ is bounded in the range in question, we can thus ignore the resulting t contour. Since s_i+z_j+t has real part at least $\delta-\delta_2>0$ on the new contour, the only poles we pick up in this process come from $t=-\frac{1}{2}(z_1+z_2)$ and $t=0$. We start with the former.

When $t=-\frac{1}{2}(z_1+z_2)$, we are left with the integral over z_1,z_2,s_1,s_2 of (up to constants and irrelevant terms)

$$\frac{M^{s_1+s_2}\zeta(1+s_1+s_2)\widehat{P_M}(s_1)\widehat{P_M}(s_2)}{\prod\zeta(1+s_i\pm(z_1-z_2)/2)s_1s_2z_1^{k+1}z_2^{k+1}(z_1+z_2)}.$$

Note in particular that this integral is independent of $\widehat{q}$. We evaluate the z_1,z_2 integrals first, they simply give linear combinations of derivatives of $\zeta(1+s_i)^{-1}$. We then shift the s_1 and s_2 contours to the contour (say γ) previously described in section 4 lying to the right of all zeros of $\zeta(1+s)$ but to the left of $\Re(s)=0$. Since $\zeta(1+s)^{-1}$ is analytic in the region crossed, its derivatives will not increase the number of poles involved, and may decrease them. In addition, there is a constant depending on k which will enable us to bound the contribution of derivatives of $\zeta(1+s)^{-1}$ by $c_k(1+\log|s|)^k$, which will have no significant impact on the resulting contour integral. Thus we may simply ignore these terms in our consideration of the integral and we are left with the integral of

$$M^{s_1+s_2}\zeta(1+s_1+s_2)\widehat{P_M}(s_1)\widehat{P_M}(s_2)s_1^{-1}s_2^{-1}$$

over $\gamma\times\gamma$, along with the poles at $s_1=s_2=0$ and $s_1=-s_2$. Note that this is now independent of k. The contour integral is bounded in the same fashion as in the previous section, so we focus on the poles. At $s_1=-s_2$, the only dependence of this expression on M or $\hat{q}$ comes ¿from the $\log M$ factors in the denominators of the P's, so this part is clearly bounded. This leaves us with the pole at $s_1=s_2=0$, which simply by counting poles clearly gives a factor of $\log M$, which is dominated by $(\log\widehat{q})^{2k-2}$ so long as $k>1$. This completes the analysis of the pole at $t=-\frac{1}{2}(z_1+z_2)$.

Thus we are left with the pole at $t=0$, which gives, up to constants and irrelevant factors,

$$\frac{\widehat{q}^{z_1+z_2}M^{s_1+s_2}\widehat{P_M}(s_1)\widehat{P_M}(s_2)\zeta(1+s_1+s_2)\zeta(1+z_1+z_2)ds_1ds_2dz_1dz_2}{s_1s_2z_1^{k+1}z_2^{k+1}\prod\zeta(1+s_i+z_j)}.$$

Again, we evaluate the z_2 and z_1 integrals, then shift the s_2,s_1 contours to γ. The integrals over γ are again small enough to cancel any positive power of $\log\widehat{q}$, so the only term of importance remaining (other than the pole at $s_1 = s_2 = z_1 = z_2 = 0$, which was the main term evaluated in the previous section) comes from the pole at $s_2 = -s_1$, integrated over $s_1 \in \gamma$. This integral is independent of M other than the presence of the $1/\log M$ factors in the P_M. Provided that $P(0) = P'(0) = 0$, there will be a factor of $(\log M)^{-4}$ or smaller coming from these coefficients. On the other hand, the z integrals contribute at most $(\log \widehat{q})^{2k+1}$, so the contribution from $t = z_1 = z_2 = 0, s_1 = -s_2$ is no larger than a constant depending on Δ times $(\log \widehat{q})^{2k-3}$. As seen above, the main term of the second moment involves a power of Δ times $(\log \widehat{q})^{2k-2}$, so these terms can be ignored.

This concludes the analysis of the contribution of the contour shifts, and with it the evaluation of the second moment. We thus have shown the following.

Proposition 5.1. *For $0 < \Delta < 1$, and P a fixed polynomial such that $P(0) = P'(0) = 0$ we have*

$$Q^h(P) = 2(1 + O_k((\log q)^{-1}))\widehat{q}\zeta(2)^2 \frac{(\log \widehat{q})^{2k-2}}{\Delta^2}$$
$$\Big(\frac{\Delta^{-1}}{2k+1}\int_0^1 P''(x)^2 dx + P'(1)^2 + 2k\Delta\int_0^1 (\frac{k}{2k-1}P(x)P''(x) + P'(x)^2)dx$$
$$+k^2\Delta^2 P(1)^2 + \frac{k^2(k-1)^2}{2k-3}\Delta^3\int_0^1 P(x)^2 dx\Big).$$

6. A generalization to linear combinations of derivatives

In this section we generalize our estimation of the first and the second moment to rather arbitrary linear combinations of derivatives of $\Lambda(f, s)$ at $s = 1/2$ and derive a very general non-vanishing result for these combinations.

First of all we will rewrite Propositions 4.1 and 5.1 in a slightly more intrinsic form: set $Q(y) = y^k$, then we have ($\Delta = \log M/\log \hat{q}$)

$$\mathcal{L}^h(P) = (1 + O_Q(\frac{1}{\log q}))\zeta(2)\widehat{q}^{1/2}\frac{(\log \widehat{q})^{k-1}}{\Delta}(Q(1)P'(1) + \Delta Q'(1)P(1)). \tag{6.1}$$

Similarly we can write

$$\mathcal{Q}^h(P) = (1 + O_k(\frac{1}{\log q}))2\widehat{q}\zeta(2)^2\frac{(\log \widehat{q})^{2k-2}}{\Delta^2}$$
$$\Big(\Delta^{-1}\int_0^1 Q^2(y)dy\int_0^1 P''(x)^2 dx + Q(1)^2P'(1)^2 + 2\Delta\int_0^1 Q'(y)^2 dy\int_0^1 (PP'')(x)dx +$$
$$+2\Delta Q(1)Q'(1)\int_0^1 P'(x)^2 dx + \Delta Q'(1)^2 P(1)^2 + \Delta^3\int_0^1 Q''^2(y)dy\int_0^1 P(x)^2 dx\Big).$$

Remarking the equalities

$$\int_0^1 (P'(x)^2 + (PP'')(x))dx = P(1)P'(1)$$
$$\int_0^1 (Q'(x)^2 + (QQ'')(x))dx = Q(1)Q'(1)$$

we may express the last factor as

$$(Q(1)P'(1) + \Delta Q'(1)P(1))^2 + \Delta^{-1}\int\int_{[0,1]\times[0,1]}\big(P''(x)Q(y) - \Delta^2 P(x)Q''(y)\big)^2 dxdy.$$

So we have

$$\mathcal{Q}^h(P) = (1 + O_Q(\frac{1}{\log q}))2\widehat{q}\zeta(2)^2 \frac{(\log \widehat{q})^{2k-2}}{\Delta^2} \times \tag{6.2}$$
$$\Big[(Q(1)P'(1) + \Delta Q'(1)P(1))^2 + \Delta^{-1} \int\int_{[0,1]\times[0,1]} \big(P''(x)Q(y) - \Delta^2 P(x)Q''(y)\big)^2 dxdy\Big].$$

Then we can generalize Proposition 4.1 and 5.1 in the following way: for any polynomial $Q(Y) = \sum_{k\geqslant 0} a_k Y^k$, consider the differential operator

$$\tilde{Q} = Q(\frac{1}{\log \hat{q}}\frac{\partial}{\partial s}) = \sum_k a_k \frac{1}{(\log \hat{q})^k}\frac{\partial^k}{\partial^k s}.$$

We consider the generalized moments

$$\mathcal{L}^h(P,Q) := \sum_f{}^h \tilde{Q}(\Lambda(f,s))(1/2)M_P(f)$$

$$\mathcal{Q}^h(P,Q) := \sum_f{}^h \big|\tilde{Q}(\Lambda(f,s))(1/2)M_P(f)\big|^2$$

which are linear and quadratic forms in both variables P, Q respectively. From this discussion it is clear that $\mathcal{L}^h(P,Q)$ is given by

$$\mathcal{L}^h(P,Q) = (1 + O_Q(\frac{1}{\log q}))\zeta(2)\widehat{q}^{1/2}\frac{1}{\Delta \log \widehat{q}}(Q(1)P'(1) + \Delta Q'(1)P(1)).$$

Less clear is the case of the second moment, but returning quickly to the proof of Proposition 5.1, one shows that the equality

$$\mathcal{Q}^h(P,Q) = (1 + O_Q(\frac{1}{\log q}))2\widehat{q}\zeta(2)^2 \frac{1}{\Delta^2(\log \widehat{q})^2} \times$$
$$\Big[(Q(1)P'(1) + \Delta Q'(1)P(1))^2 + \Delta^{-1} \int\int_{[0,1]\times[0,1]} \big(P''(x)Q(y) - \Delta^2 P(x)Q''(y)\big)^2 dxdy\Big].$$

remains true as soon as Q is either an odd or an even polynomial (this condition on Q is necessary to insure a nice functional equation for $\tilde{Q}(\Lambda(f,s))$). An intriguing fact is the symmetry in both the first and the second moment between the variables P and Q.

The proof of this equality follows the same steps: for example the expression in (5.6) has to be replaced by

$$2\widehat{q}\mathrm{Res}_{(0,0,0,0,0)} \frac{\Gamma(1+t)^2 M^{s_1+s_2}\widehat{P_M}(s_1)\widehat{P_M}(s_2)\widehat{q}^{2t+z_1+z_2}\widehat{Q}_{\hat{q}}(z_1)\widehat{Q}_{\hat{q}}(z_2)}{ts_1s_2z_1z_2}$$
$$\times \eta_2(s_1,s_2,z_1,z_2,t)\frac{(s_1+z_1+t)(s_2+z_1+t)(s_1+z_2+t)(s_2+z_2+t)}{(z_1+z_2+2t)(s_1+s_2)}.$$

At this point we are in position to apply Cauchy's inequality to prove our main result, which gives a lower bound for the proportion of non-vanishing of a general linear combination of the derivatives of $\Lambda(f,s)$ at 1/2.

Theorem 6.1. *Let Q be a fixed polynomial which is either odd or even. Then as $q \to +\infty$ we have*

$$\liminf_{q\to+\infty} \sum^{h}_{\substack{f \\ \tilde{Q}(\Lambda(f,s))(1/2)\neq 0}} 1 \geqslant \operatorname*{Max}_{P,\Delta} \mathcal{R}(P,Q) = \operatorname*{Max}_{P,\Delta} \frac{1}{2(1+\mathcal{R}_2(P,Q))}$$

where $\mathcal{R}_2(P,Q)$ is the ratio

$$\frac{\Delta^{-1}\int\int_{[0,1]\times[0,1]} \big(P''(x)Q(y) - \Delta^2 P(x)Q''(y)\big)^2 dxdy}{(Q(1)P'(1)+\Delta Q'(1)P(1))^2}, \tag{6.3}$$

P ranges over all polynomials such that $P(0) = P'(0) = 0$, and Δ over all real numbers such that $0 < \Delta < 1$.

Remark.— Note that we have always $\mathcal{R}(P,Q) < 1/2$, which is to be expected, since only half of the L-functions in question are of a given parity, and thus at most half can be such that $\tilde{Q}(\Lambda(f,s))(1/2) \neq 0$.

Remark.— In particular one can use these techniques to analyze the behavior of derivatives of $L(f,s)$ rather than $\Lambda(f,s)$, since the former at $s = 1/2$ is just linear combinations of the latter, multiplied by appropriate factors of $\log \widehat{q}$.

Remark.— Note that when looking for the value of the supremum above, $\operatorname*{Max}_{P,\Delta} \mathcal{R}(P,Q)$, one may assume by continuity that $\Delta = 1$ and that P ranges over all power series

$$P(x) = a_2x^2 + a_3x^3 + \ldots$$

which are absolutely convergent (as are the Taylor series for everything up to $P''(x)^2$) on the interval $[0,1]$.

7. Optimizing the ratio

In this section we consider in great detail our original case, that of the k-derivatives, namely $Q(y) = y^k$. In this case, we will find the optimal polynomial $P = P_k$, that is the function which minimizes the ratio $\mathcal{R}_2(P, y^k)$. Then Theorem 1.6 will be true with π_k defined by

$$\pi_k = \frac{1}{2(1+\mathcal{R}_2(P_k, y^k))},$$

and we will show that this satisfies

$$\pi_k = \frac{1}{2} - \frac{1}{32}k^{-2} + O(k^{-3}),$$

as claimed in the statement of Theorem 1.2.

A large part of this optimization process works in greater generality and we will switch to our favorite polynomial only at the very end. For this we denote

$$I(Q) := \int_0^1 Q(y)dy.$$

We need to minimize

$$\mathcal{R}_2(P,Q) = \frac{\int_0^1 \left(\Delta^{-1}I(Q^2)P''(x)^2 - 2\Delta I(QQ'')P''(x)P(x) + \Delta^3 I(Q''^2)P(x)^2\right)dx}{(\Delta Q'(1)P(1) + Q(1)P'(1))^2} \tag{7.1}$$

over all functions $P(x) = a_2x^2 + a_3x^3 + \ldots$ where this Taylor series is absolutely convergent (as are the Taylor series for everything up to $P''(x)^2$) on the interval $[0,1]$.

Assuming for the moment the existence of an optimal such P, we see that the first derivative with respect to ϵ of $\mathcal{R}_2(P(x)+\epsilon f(x))$ must be zero at $\epsilon = 0$ for all admissible functions f. From this it is easy to show that P, up to scaling by a constant, must satisfy

$$\Delta Q'(1)f(1) + Q(1)f'(1) = \int_0^1 f''(x)\left(\Delta^{-1}I(Q^2)P''(x) - \Delta I(QQ'')P(x)\right)dx$$

$$+\int_0^1 f(x)\left(\Delta^3 I(Q''^2)P(x) - \Delta I(QQ'')P''(x)\right)dx \tag{7.2}$$

for all such f. Let $\Pi(x)$ be a function with absolutely convergent Taylor series in $[0,1]$ satisfying $\Pi''(x) = P(x)$. Then two integrations by parts (using $f(0) = f'(0) = 0$) turn (7.2) into

$$\Delta Q'(1)f(1) + Q(1)f'(1) = f(1)\left(\Delta^3 I(Q''^2)\Pi'(1) - \Delta I(QQ'')\Pi'''(1)\right)$$

$$+f'(1)\left(\Delta I(QQ'')\Pi''(1) - \Delta^3 I(Q''^2)\Pi(1)\right)$$

$$+\int_0^1 f''(x)\left(\Delta^{-1}I(Q^2)\Pi''''(x) - 2\Delta I(QQ'')\Pi''(x) + \Delta^3 I(Q''^2)\Pi(x)\right)dx. \tag{7.3}$$

For this to be true for all f with Taylor series as described, the expression in the final integrand must be of the form $a_0 + a_1 x$, which gives us a fourth-order differential equation for Π. When $Q(y) = 1$ or $Q(y) = y$, one finds that Π (and thus P) is a polynomial of low degree, and it is easy to check that, for any Δ, $P_0(x) = x^2$ and $P_1(x) = x^2 - x^3/6$ are the correct choices (this leads to the $p_0 = 1/4$ and $p_1 = 7/16$ mentioned in the introduction). When $\deg Q \geqslant 2$, the differential equation is more intricate. However, its characteristic polynomial can be factored explicitly, and upon differentiating Π twice (which eliminates the $a_0 + a_1 x$ term) to get P, we find that $P(x)$ is a linear combination of the functions $e^{(\pm\alpha\pm i\beta)x}$, where

$$\alpha = \frac{\Delta}{\sqrt{2I(Q^2)}}\sqrt{\sqrt{I(Q^2)I(Q''^2)} + I(QQ'')}, \tag{7.4}$$

$$\beta = \frac{\Delta}{\sqrt{2I(Q^2)}}\sqrt{\sqrt{I(Q^2)I(Q''^2)} - I(QQ'')}.$$

Note that by design $(\alpha \pm i\beta)^2$ are the two roots of the polynomial

$$\Delta^{-1}I(Q^2)X^2 - 2\Delta I(QQ'')X + \Delta^3 I(Q''^2). \tag{7.5}$$

The conditions $P(0) = P'(0)$ imply that, up to scaling,

$$P(x) = \sinh(\alpha x)(\cos(\beta x) - L\sin(\beta x)) - \frac{\alpha}{\beta}e^{-\alpha x}\sin(\beta x). \tag{7.6}$$

One then solves for L by substituting back into the right side of (7.3) and requiring that the ratio of $f(1)$ coefficients to $f'(1)$ coefficients be exactly k for all f (this is a linear constraint with a single solution).

Note that P is now uniquely determined, so that if there exists a single P which optimizes $\mathcal{R}_2(P)$, it is the one we have just found. Suppose instead that the optimal value of $\mathcal{R}_2(.,Q)$ comes from a limiting sequence, and in particular suppose there is a function F with Taylor series of the usual type for which $\mathcal{R}_2(F,Q) < \mathcal{R}_2(P,Q)$. This then implies that the function $\mathcal{R}_2(P+\epsilon(F-P))$

has derivative zero at $\epsilon = 0$ but takes lower values elsewhere. But, taken as a function of ϵ, this is a ratio of non-negative quadratics of the form

$$\mathcal{R}_2(\epsilon)\frac{a_0 + a_1\epsilon + a_2\epsilon^2}{(b_0 + b_1\epsilon)^2}.$$

Provided that $b_0 + b_1\epsilon$ does not divide the numerator, this function has a unique minimum lying below its horizontal asymptote a_2/b_1^2. This minimum lies at the unique solution to $\mathcal{R}_2'(\epsilon) = 0$, which as we have seen is at $\epsilon = 0$. On the other hand, if $b_0 + b_1\epsilon$ divides the numerator (which is positive definite), the function must be a constant in ϵ, implying that there is an infinite family of minimal functions. Again, the uniqueness of our P as a solution to the differential equation rules this possibility out. Thus $\mathcal{R}_2$ is minimized at $P(x)$ of the form given above.

It now remains to calculate the values of $\mathcal{R}_2(P, y^k)$ for $\Delta = 1$. In this particular case we find that

$$\alpha = \Delta\sqrt{\frac{k(k-1)}{2}\left(\sqrt{\frac{2k+1}{2k-3}} + \frac{2k+1}{2k-1}\right)}, \tag{7.7}$$

$$\beta = \Delta\sqrt{\frac{k(k-1)}{2}\left(\sqrt{\frac{2k+1}{2k-3}} - \frac{2k+1}{2k-1}\right)}.$$

As one may easily check, we have asymptotically as $k \to +\infty$ $\alpha = \Delta(k + 1/2) + O(k^{-1})$ and $\beta = \Delta/2 + O(k^{-1})$. One can solve for L by substituting back into (7.3), although the calculations are quite ugly, especially for small k where one cannot ignore $e^{-\alpha}$. Asymptotically in k one has

$$L = \frac{(\alpha + k)\sin\beta + \beta\cos\beta}{-(\alpha + k)\cos\beta + \beta\sin\beta} + O(e^{-\alpha}) \tag{7.8}$$

$$= -\tan\beta + O(k^{-1}) = -\tan(\Delta/2) + O(k^{-1}).$$

Rather than solve for L explicitly, we used the MAPLE computer algebra system to optimize $\mathcal{R}_2(P)$ as a function of L, and this gave the values of π_k listed in Theorem 1.2 for $\Delta = 1$. The optimal values of L for $k = 2, 3, 4, 5$ were -1.407, -0.8827, -0.7634, and -0.7078, respectively.

Finally, we examine the rate of convergence $\mathcal{R}_2(P, y^k)$ to zero, which gives the rate of convergence of $\pi_k := 1/(2 + 2\mathcal{R}_2(P, y^k))$ to $1/2$.

Lemma 7.1. *For the optimal P_k as chosen above and $0 < \Delta \leqslant 1$,*

$$\mathcal{R}_2(P_k, y^k) = \frac{1}{16k^2} + O(k^{-3}),$$

and therefore

$$\pi_k = \frac{1}{2} - \frac{1}{32}k^{-2} + O(k^{-3}).$$

Proof. We examine the denominator of $\mathcal{R}_2$ first. When evaluating $P(x)$ or its derivatives at $x = 1$ for large k, it is clearly enough to evaluate the parts coming ¿from $e^{\alpha x}$, since those from $e^{-\alpha x}$ will be exponentially small in k. Using the expression for L in (7.8), we find that

$$\Delta kP(1) + P'(1) = \frac{e^\alpha}{2}\left(\frac{\Delta k + \alpha}{\cos\beta} + O(1)\right) + O(ke^{-\alpha}) = \frac{\Delta ke^\alpha}{\cos\beta} + O(e^\alpha). \tag{7.9}$$

Thus the only terms which will matter in the numerator of $\mathcal{R}_2$ are those which include a factor of $e^{2\alpha}$, so we may again ignore the $e^{-\alpha x}$ terms of P, and also only evaluate the integrals at $x = 1$. Thus we are only interested in the contribution from

$$P^{main}(x) = \frac{1}{4}\left((1+iL)e^{(\alpha+i\beta)x} + (1-iL)e^{(\alpha-i\beta)x}\right)$$

to

$$\int_0^1 \left(\Delta^{-1}I(Q^2)(P^{main}(x)'')^2 - 2\Delta I(QQ'')P^{main}(x)''P^{main}(x) + \Delta^3 I(Q''^2)P^{main}(x)^2\right) dx.$$

By (7.5), each term in P^{main} gives zero when put into this expression by itself. Thus the only non-zero terms are those involving the products of the two terms in P^{main}, which give

$$\frac{1+L^2}{16}\left(2\Delta^{-1}I(Q^2)(\alpha^2+\beta^2)^2 - 2\Delta I(QQ'')((\alpha+i\beta)^2+(\alpha-i\beta)^2) + 2\Delta^3 I(Q''^2)\right)\int_0^1 e^{2\alpha x}dx.$$

Making the substitution

$$2\Delta I(QQ'')(\alpha\pm i\beta)^2 = \Delta^{-1}I(Q^2)(\alpha\pm i\beta)^4 + \Delta^3 I(Q''^2)$$

and evaluating the integral leaves us with

$$\frac{e^{2\alpha}}{2\alpha}\frac{1+L^2}{16\Delta(2k+1)}\left(2(\alpha^2+\beta^2)^2 - (\alpha+i\beta)^4 - (\alpha-i\beta)^4\right) = \frac{e^{2\alpha}}{2\alpha}\frac{1+L^2}{\Delta(2k+1)}(\alpha\beta)^2.$$

Using $1+L^2 = (\cos\beta)^{-2} + O(1/k)$, $\alpha = \Delta(k+1/2) + O(1/k)$, and $\beta = \Delta/2 + O(1/k)$ turns this into

$$\frac{e^{2\alpha}\Delta^2}{16\cos^2\beta}(1+O(\frac{1}{k})).$$

Comparing with the square of (7.9) completes the proof. □

8. Application to the analytic rank of $J_0(q)$

8.1. **Proof of Theorem 1.4.** Here we will prove Theorem 1.4, the statement of which we recall: let α, $0<\alpha<2$, be a fixed real number, then for q prime large enough we have

$$\sum_{f\in S_2(q)^*} r_f^\alpha \leqslant \left(\tfrac{1}{2} + \sum_{k=0}^{+\infty}((k+2)^\alpha - k^\alpha)(\tfrac{1}{2}-p_k) + o_\alpha(1)\right)|S_2(q)^*|. \tag{8.1}$$

The proof is relatively straightforward; indeed, if we had

$$\sum_{\Lambda^{(k)}(f,\frac{1}{2})\neq 0} 1 \geqslant p_k|S_2(q)^*|$$

instead of the limit formula

$$\sum_{\Lambda^{(k)}(f,\frac{1}{2})\neq 0} 1 \geqslant (p_k + o_k(1))|S_2(q)^*|,$$

it would be immediate by summing by parts over k. However, we do not know how large the $o_k(1)$ is as function of k (in the case of π_k, we have actually shown that the $o_k(1)$ is of size (roughly) $(k!)^2(\log q)^{-1}$, which is too large).

To avoid this problem, we will need the following theorem, of independent interest:

Theorem 8.1. *There exists an absolute constant $C > 0$ such that for all q prime*

$$\sum_{f\in S_2(q)^*} r_f^2 \leqslant C|S_2(q)^*| \text{ and } \sum_{f\in S_2(q)^*}^h r_f^2 \leqslant C.$$

Using this, which is proved below in Section 8.2, we now prove (8.1).

Lemma 8.2. *Let $k \geqslant 0$. We have*

$$\sum_{r_f\geqslant k} 1 \ll (\tfrac{1}{2} - p_{k-1} + \tfrac{1}{2} - p_{k-2} + o_k(1))|S_2(q)^*|$$

where the implied constant is absolute, and we put $p_{-1} = 0$.

Proof. If the analytic rank of f is $\geqslant k$, at least we must have

$$\Lambda^{(k-1)}(f, \tfrac{1}{2}) = \Lambda^{(k-2)}(f, \tfrac{1}{2}) = 0.$$

By parity considerations, the proportion of forms satisfying those two conditions is

$$\ll \tfrac{1}{2} - p_{k-1} + \tfrac{1}{2} - p_{k-2} + o_k(1).$$

□

Let α, $0 < \alpha < 2$ be fixed. We consider the average

$$\sum_{f\in S_2(q)^*} r_f^\alpha.$$

Introducing a fixed $n \geqslant 1$ we write

$$\sum_{f\in S_2(q)^*} r_f^\alpha = \sum_{r_f>n} r_f^\alpha + \sum_{r_f\leqslant n} r_f^\alpha = N_1 + N_2, \text{ say.}$$

By Hölder's inequality,

$$\begin{aligned} N_1 &\leqslant \Bigl(\sum_f r_f^2\Bigr)^{\alpha/2}\Bigl(\sum_{r_f>n} 1\Bigr)^{1-\alpha/2} \\ &\ll |S_2(q)^*|^{\alpha/2}\Bigl(\sum_{r_f>n} 1\Bigr)^{1-\alpha/2} \text{ (Proposition 8.1)} \\ &\ll (\tfrac{1}{2} - p_n + \tfrac{1}{2} - p_{n-1} + o_n(1))^{1-\alpha/2}|S_2(q)^*|^{1-\alpha/2} \text{ (Lemma 8.2)} \\ &\ll (\tfrac{1}{2} - \pi_n + \tfrac{1}{2} - \pi_{n-1} + o_n(1))^{1-\alpha/2}|S_2(q)^*|^{1-\alpha/2} \\ &\ll \Bigl(\frac{1}{n^{2-\alpha}} + o_{\alpha,n}(1)\Bigr)|S_2(q)^*|^{1-\alpha/2} \end{aligned}$$

and thus

$$N_1 \ll \Bigl(\frac{1}{n^{2-\alpha}} + o_{\alpha,n}(1)\Bigr)|S_2(q)^*|.$$

Turning to the other term, we have by partial summation

$$N_2 = \sum_{k=1}^n k^\alpha\Bigl(\sum_{r_f=k} 1\Bigr) \leqslant \sum_{k=1}^n (k^\alpha - (k-1)^\alpha)\Bigl(\sum_{r_f\geqslant k} 1\Bigr)$$

since the difference between those two expressions comes from forms whose L-function has order $> n$. By the lemma

$$\sum_{r_f \geqslant k} 1 \leqslant (\tfrac{1}{2} - p_{k-1} + \tfrac{1}{2} - p_{k-2} + o_k(1))|S_2(q)^*|$$

(where we put $p_{-1} = 0$). Therefore, after some manipulation

$$N_2 \leqslant |S_2(q)^*|\Big(\tfrac{1}{2} + \sum_{k \leqslant n-2} \big((k+2)^\alpha - k^\alpha\big)(\tfrac{1}{2} - p_k) + (n^\alpha - (n-1)^\alpha)(\tfrac{1}{2} - p_n) + o_n(1)\Big).$$

We extend the sum to the infinite series, which has non-negative terms, so a limit, finite or $+\infty$. In fact by Theorem 1.2, we have

$$\tfrac{1}{2} - p_k \ll \tfrac{1}{2} - \pi_k \ll k^{-2},$$

so it converges for $\alpha < 2$, and we obtain

$$N_2 \leqslant \Big(\tfrac{1}{2} + \sum_{k=0}^{+\infty} ((k+2)^\alpha - k^\alpha)(\tfrac{1}{2} - p_k) + o_n(1)\Big)|S_2(q)^*|.$$

Adding the estimate for N_1, and letting then n go to $+\infty$, this proves Theorem 1.4.

Specific values were calculated with the MAPLE computer algebra system. Thus we find for example that

$$\sum_{f \in S_2(q)^*} r_f \leqslant (1.1891 + o(1))|S_2(q)^*|, \qquad \sum_{f \in S_2(q)^*} r_f^{1.9} \leqslant (3.2191 + o(1))|S_2(q)^*|.$$

If α approaches 2, it is eventually better to use the (unspecified but computable) constant of Proposition 8.1, as the expression above tends to $+\infty$ for $\alpha \to 2$.

8.2. **The average rank squared.** We now prove Theorem 8.1. The method is based on that used in [KM1], to which we refer for complete details of the steps only briefly sketched below.

Applying the explicit formula and proceeding as in [KM1, 4.1] with some help from Cauchy-Schwarz, we reduce the proof to a density theorem for zeros of automorphic L-functions.

For any $\sigma \geqslant \frac{1}{2}$, t_1 and t_2 real, we denote by $N(f;\sigma,t_1,t_2)$ the number of zeros $\rho = \beta + i\gamma$ of $L(f,s)$ which satisfy $\beta \geqslant \sigma$ and $t_1 \leqslant \gamma \leqslant t_2$. Then it is enough to prove the

Proposition 8.3. *There exist absolute constants $B > 0$, $c > 0$, such that for any $\sigma \geqslant \frac{1}{2} + (\log q)^{-1}$ and any real numbers $t_1 < t_2$ such that $t_2 - t_1 \geqslant (\log q)^{-1}$, it holds*

$$\sum_{f \in S_2(q)^*} N(f,\sigma,t_1,t_2)^2 \leqslant (1 + |t_1| + |t_2|)^B q^{1-c(\sigma - \frac{1}{2})}(t_2 - t_1)(\log q).$$

Remark. About the transition from this density theorem to Proposition 8.1: if we follow [KM1] closely, we see that we need an estimate for the term denoted $S_1(f,\lambda)$ on average, namely we require

$$\sum_f S_1(f,\lambda)^2 \ll |S_2(q)^*|(\log q)^2.$$

The harmonic analogue of this is proved in the course of proving [KM1, Lemma 7]), and the weight is removed as usual (very easily in this case).

To prove Proposition 8.3, we will appeal to the following fundamental result of [KM1], which estimates a mollified second moment of $L(f,s)$ on average.

Let $M = \hat{q}^{\Delta}$ for some parameter Δ and let $g = g_M$ be the function

$$g(x) = \begin{cases} 1 \text{ if } x \leqslant \sqrt{M} \\ \frac{\log M/x}{\log \sqrt{M}} \text{ if } \sqrt{M} \leqslant x \leqslant M \\ 0 \text{ if } x > M. \end{cases}$$

We then put

$$x_m(s) = \mu(m) m^{-s} \sum_{n \geqslant 1} \frac{\varepsilon(n)\mu(mn)^2}{n^{2s}} g(mn)$$

and finally

$$M(f,s) = \sum_m \lambda_f(m) x_m(s) = \sum_m \frac{\mu(m)\lambda_f(m)}{m^s} \sum_n \frac{\varepsilon(n)\mu(mn)^2}{n^{2s}} g(mn).$$

Proposition 8.4. *Let $M = \hat{q}^{\Delta}$ with $0 < \Delta < \frac{1}{2}$, let $c < \Delta$ be any positive real number. There exists an absolute constant $B > 0$ such that for all q large enough we have*

$$\sum_{f \in S_2(q)^*} |L(f, \sigma + it) M(f, \sigma + it) - 1|^2 \ll (1 + |t|)^B q^{1 - \frac{c}{2}(\sigma - \frac{1}{2})},$$

uniformly for $\sigma \geqslant \frac{1}{2} + (\log q)^{-1}$ and $t \in \mathbf{R}$, the implied constant depending only on c and Δ.

We can now prove the density theorem. The argument is similar to that of [KM1] (based on Lemma 14 of [S], see also [Kow]), with a simple trick to get to the square.

We may assume that $t_2 - t_1 = (\log q)^{-1}$. We set

$$\sigma' = \sigma - \frac{1}{2\log q}, \quad t_1' = t_1 - \frac{\eta}{\log q}, \quad t_2' = t_2 + \frac{\eta}{\log q}$$

where $\eta > 0$ is some parameter, large enough so that

$$\frac{\pi\eta}{2\eta + 1} \geqslant \frac{\pi}{6}, \quad \frac{4\pi}{2\eta + 1} < c.$$

If we let $h_f(s) = L(f,s)M(f,s)$, which vanishes at zeros of $L(f,s)$, using Selberg's Lemma 14 ([S]), we find the zero-detecting inequality

$$\begin{aligned} N(f,\sigma,t_1,t_2) \quad &\leqslant \quad \frac{2}{\pi}(\log q) \int_{t_1'}^{t_2'} \sin\Big(\pi \frac{t - t_1'}{t_2' - t_1'}\Big) \log |h_f(\sigma' + it)| dt \\ &+ \quad \frac{2}{\pi}(\log q) \int_{\sigma'}^{+\infty} \sinh\Big(\pi \frac{x - \sigma'}{t_2' - t_1'}\Big) \{\log |h_f(x + it_1')| + \log |h_f(x + it_2')|\} dx. \end{aligned}$$

Since $\log |1 + x| \leqslant \log(1 + |x|) \leqslant |x|$ and $\sinh(x) \geqslant 0$ for $x > 0$, writing

$$h_f(s) = 1 + (L(f,s)M(f,s) - 1),$$

we obtain

$$N(f,\sigma,t_1,t_2) \leqslant \frac{2}{\pi}(\log q) \int_{t_1'}^{t_2'} \sin\Big(\pi \frac{t - t_1'}{t_2' - t_1'}\Big) |LM(f, \sigma' + it) - 1| dt$$

$$+ \frac{2}{\pi}(\log q) \int_{\sigma'}^{+\infty} \sinh\Big(\pi \frac{x - \sigma'}{t_2' - t_1'}\Big) \{> |LM(f, x + it_1') - 1| + |LM(x + it_2') - 1|\} dx.$$

We now square this last inequality (now and not before because $\log |h_f(\sigma' + it)|$ might be negative and very large in absolute value at some point, which would be difficult to handle), and average over f. The average of $N(f,\sigma,t_1,t_2)^2$ is bounded, up to some absolute multiplicative constant, by a sum of three terms, which are double integrals. All are treated similarly, so we pick only one for example, namely

$$(\log q)^2 \int_{\sigma'}^{+\infty}\int_{\sigma'}^{+\infty} \sinh\Big(\pi\frac{x-\sigma'}{t_2'-t_1'}\Big)\sinh\Big(\pi\frac{y-\sigma'}{t_2'-t_1'}\Big)\mathcal{M}_1(x,y)dxdy$$

where the mixed moment $\mathcal{M}_1$ is

$$\mathcal{M}_1(x,y)=\sum_f |LM(f,x+it_1')-1||LM(f,y+it_1')-1|.$$

By Cauchy's inequality, we have

$$\begin{aligned}\mathcal{M}_1(x,y) &\leqslant \Big(\sum_f |LM(x+it_1')-1|^2\Big)^{\frac12}\Big(\sum_f |LM(y+it_1')-1|^2\Big)^{\frac12}\\ &\ll (1+|t_1'|)^B q^{1-\frac{c}{4}(x-\frac12)-\frac{c}{4}(y-\frac12)}.\end{aligned}$$

Now the double integral splits as a product. One is

$$\int_{\sigma'}^{+\infty}\sinh\Big(\pi\frac{x-\sigma'}{t_2'-t_1'}\Big)q^{-\frac{c}{4}(x-\frac12)}dx \ll \frac{1}{\log q}q^{-\frac{c}{4}(\sigma'-\frac12)}$$

because $q^{-cx/4}=e^{-cx(\log q)/4}$ and η has been chosen so that $\pi/(2\eta+1)<c/4$ (making the integral converging, and as small as $(\log q)^{-1}$). The second integral is handled in the same way, and so this first term is seen to be

$$\ll (1+|t_1'|)^B q^{1-\frac{c}{2}(\sigma-\frac12)}.$$

The other two terms are estimated in the same way.

8.3. **Final remarks.** We wish to emphasize here that the value 1.1891 is certainly not the best possible constant within reach of our method. In fact the arguments of the preceding section work without change if we consider, instead of the polynomials y^k, $k\geqslant 0$, any other family Q_k of polynomials such that $\deg Q_k = k$ and Q_k is of the same parity as k.

The best possible bound our results can get will be achieved if, for any k, Q_k is chosen to minimize $\mathcal{R}_2(P_Q,Q)$ among all the polynomials of degree k and of the same parity as k, where P_Q is the optimal function corresponding to Q described in Section 7. Moreover R. Heath-Brown showed us how a refinment of the arguments given in section 8.1 could be used to improve further the value of the constant.

Nevertheless, for $k=0$ or 1, there is no other choice than $Q_0(y)=1$, $Q_1(y)=y$. The optimal P is known in those cases, so the bound for $r(J_0(q))$ will be always

$$>1/2+2((1/2-1/4)+(1/2-7/16))=1.125,$$

and this is the absolute limit of our method, barring any improvement in the (logarithmic) length of the mollifier Δ beyond 1 (which would be of great significance independently of this, and which is feasible assuming GRH for Dirichlet L functions as shown by Iwaniec and Sarnak [I-S]).

Similarly the 99% result cannot be reduced below 3 since $1/2+7/16<99/100$.

9. Appendix

In this section we provide the calculations behind the various residue calculations used throughout the paper. All are in the spirit of lemmas 9 through 11 of Conrey's work on high derivatives of the zeta function, although in order to maintain more flexibility, we have kept ourselves to calculations of residues rather than number-theoretic sums.

As mentioned earlier in the text, given a polynomial

$$P(x) = \sum a_k x^k$$

and a large number M, we define a new polynomial

$$\widehat{P_M}(s) = \sum a_k \frac{k!}{(s\log M)^k}$$

for use in contour integrals.

Lemma 9.1.

$$\mathrm{Res}_{s=0}\frac{M^s\widehat{P_M}(s)}{s} = P(1).$$

Proof. This is a straightforward calculation, one has

$$\sum \frac{a_k k!}{(\log M)^k}\mathrm{Res}_{s=0}\frac{M^s}{s^{k+1}} = \sum a_k,$$

as desired. □

Corollary 9.2.

$$\mathrm{Res}_{s=0}\frac{M^s s^\ell\widehat{P_M}(s)}{s} = \frac{1}{(\log M)^\ell}P^{(\ell)}(1).$$

Proof. Multiplying $\widehat{P_M}(s)$ by s^ℓ gives

$$\sum \frac{a_k k!}{(\log M)^k}s^{\ell-k} = (\log M)^{-\ell}\sum(\frac{a_k k!}{(k-\ell)!})\frac{(k-\ell)!}{(s\log M)^{k-\ell}} = (\log M)^{-\ell}(\widehat{P^{(\ell)})_M}(s).$$

Now use Lemma 9.1. Note that if ℓ is larger than the degree of the polynomial, then the residue is zero, as is the derivative, so that this corollary still applies. □

Corollary 9.3.

$$\mathrm{Res}_{s=0}\frac{M^s\widehat{P_M}(s)}{s^{1+\ell}} = (\log M)^\ell(\int^{(\ell)} P)]_0^1,$$

where $\int^{(\ell)}$ means to take ℓ antiderivatives, without including constants of integration.

Proof. Exactly the same as Corollary 9.2. □

Lemma 9.4.

$$\mathrm{Res}_{s_1,s_2=0}\frac{M^{s_1+s_2}\widehat{P_M}(s_1)\widehat{Q_M}(s_2)}{s_1 s_2(s_1+s_2)} = (\log M)(\int_0^1 P_M(x)Q_M(x)dx).$$

Proof. Again, we break the polynomials up by coefficients, letting b_k represent the coefficients of Q, and calculate the residues one at a time.

$$\mathrm{Res}_{s_1=0}\sum\frac{a_{k_1}k_1!}{(\log M)^{k_1}}\frac{M^{s_1}}{s_1^{k_1+1}}\mathrm{Res}_{s_2=0}\sum\frac{b_{k_2}k_2!}{(\log M)^{k_2}}\frac{M^{s_2}}{s_2^{k_2+1}(s_1+s_2)}$$

$$= \mathrm{Res}_{s_1=0} \sum_{k_1} \frac{a_{k_1}}{(\log M)^{k_1}} \frac{k_1! M^{s_1}}{s_1^{k_1+1}} \sum_{k_2} \frac{b_{k_2}}{(\log M)^{k_2}} \sum_{\ell=0}^{k_2} \binom{k_2}{\ell} \frac{(-1)^\ell \ell!}{s_1^{\ell+1}} (\log M)^{k_2-\ell}$$

$$= \sum_{k_1,k_2} \frac{a_{k_1} a_{k_2}}{(\log M)^{k_1+k_2}} \sum_{\ell=0}^{k_2} \frac{(-1)^\ell k_1! k_2!}{(k_2-\ell)!(k_1+\ell+1)!} (\log M)^{k_1+k_2+1}.$$

We now use the combinatorial identity

$$\sum_{A=0}^{B} (-1)^A \binom{C}{B-A} = \binom{C-1}{B},$$

which is most easily seen by comparing the Bth coefficients of the identity

$$(1+x)^C(1-x+x^2-x^3+\cdots) = (1+x)^{C-1}.$$

Using this to evaluate the sum over ℓ, we have

$$(\log M) \sum_{k_1,k_2} a_{k_1} b_{k_2} \frac{k_1! k_2!}{(k_1+k_2+1)!} \binom{k_1+k_2}{k_2}$$

$$= (\log M) \sum_{k_1,k_2} a_{k_1} b_{k_2} \frac{1}{k_1+k_2+1},$$

as desired. □

References

[Br] A. Brumer: *The rank of $J_0(N)$,* Astérisque, Vol 228 (1995), 41–68. 1

[BFH] D. Bump ,S. Friedberg ,J. Hoffstein : *Nonvanishing theorems for L-functions of modular forms and their derivatives,* Invent. Math. 102, No.3, 543–618 (1990). 1

[Co] B. Conrey : *Zeros of derivatives of Riemann's ξ-function on the critical line,* J. Number Theory 16, 49–74 (1983). 1, 2, 4

[Du] W. Duke : *The critical order of vanishing of automorphic L-functions with high level,* Invent. Math. 119, (1995), 165–174. 1, 1, 2

[DFI] W. Duke, J. Friedlander and H. Iwaniec : *Bounds for automorphic L functions II,* Invent. Math. 115, (1994), 219–239. 5

[HBM] R. Heath-Brown and P. Michel: *Exponential decay in the frequency of analytic ranks of automorphic L-functions*, preprint.

[Iw] H. Iwaniec : *On the order of vanishing of modular L-functions at the critical point,* Sém. Théor. Nombres Bordeaux 2, (1990), 365–376. 1

[I-S] H. Iwaniec and P. Sarnak: *The non-vanishing of central values of automorphic L-functions and Siegel-Landau zeros,* (preprint 1998). 1, 1, 1, 1, 2, 4, 5, 8.3

[KM1] P. Michel and E. Kowalski: *The analytic rank of $J_0(q)$ and zeros of automorphic L-functions*, to appear in Duke Math. Journal. 1, 1, 1, 1, 1, 2, 4, 8.2, 8.2, 8.2

[KM2] P. Michel and E. Kowalski: *A lower bound for the rank of $J_0(q)$*, to appear in Acta Arithmetica. 1, 1, 1, 1, 1, 2

[KM3] P. Michel and E. Kowalski: *An explicit upper bound for the rank of $J_0(q)$*, to appear in Israel Journal of Math. 1

[Kow] E. Kowalski: *The rank of the Jacobian of modular curves: analytic methods,* Ph.D. Thesis, Rutgers University, 1998. 8.2

[Lu] W. Luo: *On the nonvanishing of Rankin-Selberg L-functions,* Duke Math. J. 69, No.2, (1993) 411–425. 1

[M] R. Murty *The analytic rang of $J_0(N)$,* (K. Dilcher Edt.) CMS Conf. Proc. 15, 263–277 (1995). 1, 1

[MM] M. R. Murty and V. K. Murty: *Mean values of derivatives of modular L-series,* Ann. Math., II. Ser. 133, No.3, (1991) 447–475. 1

[P-P] A. Perelli and J. Pomykala: *Averages over twisted elliptic L-functions,* Acta Arith. 80, No 2, 149–163 (1997). 1

[P] J. Pomykala : *Non vanishing of n-th derivatives of twisted elliptic L-functions in the critical point,* J. Théor. Nombres Bordeaux, 9 (1997), 1–10. 1

[S] A. Selberg: *Contributions to the theory of Dirichlet's L functions*, Skr. utg. av Det Norske Videnskaps-Akademi, Mat. Nat. Kl. (1946) No 3, 1–62. 8.2

[Shi] G. Shimura, Introduction to the arithmetic theory of automorphic functions, Iwanomi Shoten, Tokyo, Princeton Univ. Press, Princeton, NJ, 1971. 1

[VdK1] J. VanderKam: *The rank of quotients of $J_0(N)$,* to appear in Duke Math Journal. 1, 1, 1, 3

[VdK2] J. VanderKam: *Linear independance in the homology of $X_0(N)$*, to appear. 3, 5

Dept. of Math. - Fine Hall, Princeton University, Princeton, NJ 08544-1000, USA.
E-mail address: ekowalsk@math.princeton.edu

Mathématiques, Université Montpellier II CC 051, 34095 Montpellier Cedex 05, FRANCE.
E-mail address: michel@darboux.math.univ-montp2.fr

Center for Communications Research, Thanet Road, Princeton, NJ 08540, USA.
E-mail address: vanderkm@idaccr.org