Algebra & Number Theory

Volume 12

2018

No. 10

Bounds for traces of Hecke operators and applications to modular and elliptic curves over a finite field

Ian Petrow

msp

ALGEBRA AND NUMBER THEORY 12:10 (2018)
dx.doi.org/10.2140/ant.2018.12.2471

Bounds for traces of Hecke operators and applications to modular and elliptic curves over a finite field

Ian Petrow

We give an upper bound for the trace of a Hecke operator acting on the space of holomorphic cusp forms with respect to certain congruence subgroups. Such an estimate has applications to the analytic theory of elliptic curves over a finite field, going beyond the Riemann hypothesis over finite fields. As the main tool to prove our bound on traces of Hecke operators, we develop a Petersson formula for newforms for general nebentype characters.

1. Introduction

1A. ***Statement of results.*** Let $S_\kappa(\Gamma, \epsilon)$ be the space of holomorphic cusp forms of weight κ, for a subgroup Γ of a Hecke congruence group, and of nebentype character ϵ. We write $\mathrm{Tr}(T\,|\,S_\kappa(\Gamma,\epsilon))$ for the trace of a linear operator T acting on $S_\kappa(\Gamma,\epsilon)$. The aim of this paper is to give estimates for $\mathrm{Tr}(T_m|S_\kappa(\Gamma,\epsilon))$, where T_m is the m-th Hecke operator, as the parameters m, κ, Γ, and ϵ vary simultaneously.

Consider first the case that $\Gamma = \Gamma_0(N)$ and ϵ is any Dirichlet character modulo N. Let $d(m)$ denote the number of divisors of m, let $\sigma(m)$ denote the sum of the divisors of m, and let $\psi(N) = [\Gamma_0(N) : \mathrm{SL}_2(\mathbb{Z})] = N\prod_{p|N}\bigl(1+\frac{1}{p}\bigr)$. We assume that $\kappa \geqslant 2$ is an integer throughout the paper. Deligne's theorem tells us that each eigenvalue $\lambda(m)$ of T_m satisfies $|\lambda(m)| \leqslant d(m)m^{(\kappa-1)/2}$. Therefore we have the "trivial" estimate on the trace

$$\mathrm{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon)) \leqslant \dim S_\kappa(\Gamma_0(N),\epsilon)d(m)m^{\frac{\kappa-1}{2}} \leqslant \frac{(\kappa-1)\psi(N)}{12}d(m)m^{\frac{\kappa-1}{2}}. \tag{1-1}$$

For the bound on $\dim S_\kappa(\Gamma_0(N),\epsilon)$, see, e.g., [Ross 1992, Corollary 8]. The power of m in (1-1) is sharp by the Sato–Tate distribution for Hecke eigenvalues. On the other hand, by a careful analysis using the Eichler–Selberg trace formula, Conrey, Duke and Farmer [Conrey et al. 1997] and in more generality Serre [1997, Proposition 4] showed that if $\epsilon(-1) = (-1)^\kappa$ then

$$\mathrm{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon)) = \frac{\kappa-1}{12}\epsilon(m^{\frac12})m^{\frac{\kappa}{2}-1}\psi(N) + O\Bigl(\bigl(\sigma(m)\max_{f^2<4m}\psi(f) + d(m)N^{\frac12}\bigr)m^{\frac{\kappa-1}{2}}d(N)\Bigr), \tag{1-2}$$

This work was supported by Swiss National Science Foundation grant PZ00P2_168164.
MSC2010: primary 11F25; secondary 11F11, 11F72, 11G20, 14G15.
Keywords: traces of Hecke operators, modular curves over a finite field, elliptic curves over a finite field, Petersson formula for newforms, Tsfasman–Vlăduţ–Zink theorem.

where $\epsilon(m^{1/2})$ is understood to be 0 unless m is a perfect square. If $\epsilon(-1)\neq(-1)^\kappa$ then $S_\kappa(\Gamma_0(N),\epsilon)=\{0\}$, so the left-hand side vanishes identically. We expect the estimate (1-2) to be sharp if m is fixed and $\kappa+N\to\infty$.

Write $\boldsymbol{c}(\epsilon)$ for the conductor of the Dirichlet character ϵ, and $\boldsymbol{c}^*(\epsilon)=\prod_{p\mid \boldsymbol{c}(\epsilon)} p$ for its square-free part. In this paper we prove:

Theorem 1.1. *Suppose that* $\epsilon(-1)=(-1)^\kappa$, $(N,m)=1$, *and* $m\boldsymbol{c}(\epsilon)\boldsymbol{c}^*(\epsilon)\ll(N^4\kappa^{10/3})^{1-\eta}$ *for some* $\eta>0$. *Then we have*

$$\operatorname{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon)) = \frac{\kappa-1}{12}\epsilon(m^{\frac12})m^{\frac{\kappa}{2}-1}\psi(N)+O_{\eta,\varepsilon}(N^{\frac{10}{11}}m^{\frac{\kappa-1}{2}+\frac{1}{44}}\kappa^{\frac{61}{66}}\boldsymbol{c}(\epsilon)^{\frac{1}{44}}\boldsymbol{c}^*(\epsilon)^{\frac{1}{44}}(Nm\kappa)^\varepsilon). \quad (1\text{-}3)$$

We remark that the hypothesis that $m\boldsymbol{c}(\epsilon)\boldsymbol{c}^*(\epsilon)\ll(N^4\kappa^{10/3})^{1-\eta}$ for some $\eta>0$ in Theorem 1.1 is no restriction in practice, since if the hypothesis fails then (1-1) is a superior bound anyway. Indeed, the error term in (1-3) is smaller than that in both (1-1) and (1-2) when

$$N^{\frac{8}{13}}\kappa^{\frac{122}{195}}(N\kappa)^\varepsilon\boldsymbol{c}(\epsilon)^{\frac{1}{65}}\boldsymbol{c}^*(\epsilon)^{\frac{1}{65}}\ll m\ll\frac{(N^4\kappa^{\frac{10}{3}})^{1-\eta}}{\boldsymbol{c}(\epsilon)\boldsymbol{c}^*(\epsilon)}.$$

For example, if ϵ is trivial and the weight κ is fixed, then (1-3) is better than (1-1) and (1-2) for

$$N^{\frac{8}{13}+\varepsilon}\ll m\ll N^{4-\varepsilon}.$$

Note that our result requires the hypothesis $(N,m)=1$, whereas the estimates (1-1) and (1-2) do not. We discuss the source of this condition in the sketch of the proof, below.

We are also interested in spaces of modular forms for groups other than $\Gamma_0(N)$. In particular, for positive integers $M\mid N$ let

$$\Gamma(M,N)=\left\{\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbb{Z}): a,d\equiv1\ (\mathrm{mod}\ N),\ c\equiv0\ (\mathrm{mod}\ NM)\right\}. \quad (1\text{-}4)$$

These congruence groups interpolate between $\Gamma_1(N)=\Gamma(1,N)$ and $\Gamma(N)\simeq\Gamma(N,N)$. We write $S_\kappa(M,N)$ for the space of modular forms of weight κ for the group $\Gamma(M,N)$ (without nebentype character). Let $\delta(a,b)$ be the indicator function of $a=b$ and $\delta_c(a,b)$ be the indicator function of $a\equiv b\ (\mathrm{mod}\ c)$. Let T_m be the m-th Hecke operator acting on $S_\kappa(M,N)$ and for $(d,N)=1$ let $\langle d\rangle$ be the d-th diamond operator. These operators commute and $T_1=\langle1\rangle=\mathrm{id}$; for definitions see [Diamond and Shurman 2005, §5.1, 5.2] or [Kaplan and Petrow 2017, §4]. In particular, we have

$$\operatorname{Tr}(\langle d\rangle T_m|S_\kappa(\Gamma(M,N)))=\sum_{\epsilon\ (\mathrm{mod}\ N)}\epsilon(d)\operatorname{Tr}(T_m|S_\kappa(\Gamma_0(NM),\epsilon)). \quad (1\text{-}5)$$

Applying (1-1) to (1-5) we have

$$\operatorname{Tr}(\langle d\rangle T_m|S_\kappa(\Gamma(M,N)))\leqslant\frac{\kappa-1}{12}\varphi(N)\psi(NM)d(m)m^{\frac{\kappa-1}{2}}. \quad (1\text{-}6)$$

Meanwhile, summing (1-2) over characters $\epsilon \pmod N$ such that $\epsilon(-1) = (-1)^\kappa$ we find

$$\operatorname{Tr}(\langle d\rangle T_m | S_\kappa(\Gamma(M,N))) = \frac{\kappa-1}{24} m^{\frac{\kappa}{2}-1}\varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1) + (-1)^\kappa \delta_N(m^{\frac{1}{2}}d,-1)\big) + O\Big((\sigma(m)\max_{f^2<4m}\psi(f) + d(m)(MN)^{\frac{1}{2}})m^{\frac{\kappa-1}{2}} d(MN)N\Big). \quad (1\text{-}7)$$

The following result improves on both (1-6) and (1-7) in an intermediate range of parameters.

Theorem 1.2. *Suppose that* $M\,|\,N$, $(N,m) = 1$, *and* $m \ll (N^6\kappa^{10/3})^{1-\eta}$ *for some* $\eta > 0$. *We have*

$$\operatorname{Tr}(\langle d\rangle T_m | S_\kappa(\Gamma(M,N))) = \frac{\kappa-1}{24} m^{\frac{\kappa}{2}-1}\varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1) + (-1)^\kappa \delta_N(m^{\frac{1}{2}}d,-1)\big) + O_{\eta,\varepsilon}(MN^{\frac{41}{22}} m^{\frac{\kappa-1}{2}+\frac{1}{44}} \kappa^{\frac{61}{66}} (Nm\kappa)^\varepsilon).$$

1B. *Applications to modular and elliptic curves over a finite field.* Hecke operators appear throughout number theory, and estimates for their traces are especially relevant to equidistribution problems. See for example [Serre 1997, §5–§8] and [Murty and Sinha 2009]. We mention here a few consequences in the analytic theory of modular and elliptic curves over a finite field.

Let C be a nonsingular projective curve of genus g over a finite field $\mathbb{F}_q$ with q elements. Then we have (see, e.g., [Milne 2017, Chapter 11])

$$|C(\mathbb{F}_{q^n})| = q^n + 1 - \sum_{i=1}^{2g} \alpha_i^n,$$

where $\{\alpha_i\}$ are the inverse zeros of the zeta function of C

$$Z(C,T) = \frac{(1-\alpha_1 T)\cdots(1-\alpha_{2g}T)}{(1-T)(1-qT)}.$$

The Riemann hypothesis for curves over finite fields asserts that $|\alpha_i| = \sqrt{q}$ for all i. Igusa [1959] showed that there exists a nonsingular projective model for $X_0(N)$ over $\mathbb{Q}$ whose reductions modulo primes p, $p \nmid N$, are also nonsingular (see also the survey [Diamond and Im 1995, §9]), and so the preceding discussion applies to $X_0(N)$ when $p \nmid N$. Since $g \sim \psi(N)/12$ as $N \to \infty$ we have

$$|X_0(N)(\mathbb{F}_q)| = q + 1 + O(\psi(N)q^{1/2}). \quad (1\text{-}8)$$

In particular, $|X_0(N)(\mathbb{F}_q)| \sim q$ as $q \to \infty$ as soon as $q \gg N^{2+\delta}$ for some $\delta > 0$. On the other hand, the Eichler–Shimura correspondence (see, e.g., [Milne 2017, Theorem 11.14]) asserts that

$$Z(X_0(N),T) = \frac{\prod_{f\in H_2(N)}(1-\lambda_f(p)T + pT^2)}{(1-T)(1-pT)},$$

where $H_2(N)$ is a basis for $S_2(\Gamma_0(N))$ consisting of eigenforms of $\{T_p : p \nmid N\}$ and $\lambda_f(p)$ is the T_p eigenvalue of f. We therefore have

$$|X_0(N)(\mathbb{F}_q)| = q + 1 - \operatorname{Tr}(T_q | S_2(\Gamma_0(N))) + p\operatorname{Tr}(T_{q/p^2} | S_2(\Gamma_0(N))),$$

where we set $T_{p^{-1}} = 0$. Applying (1-1), (1-2), and Theorem 1.1 we get:

Corollary 1.3. *Suppose $q = p^v$ is a prime power such that $p \nmid N$. We have*

$$|X_0(N)(\mathbb{F}_q)| = q + (p-1)\frac{\psi(N)}{12}\delta_2(v,0) + O_\varepsilon\big(\min(\psi(N), q^{\frac{1}{44}}N^{\frac{10}{11}}(qN)^\varepsilon, (q^{\frac{3}{2}} + N^{\frac{1}{2}})d(N)q^\varepsilon)q^{\frac{1}{2}}\big).$$

In particular, the main term is larger than the error term as soon as $q \gg N^{\frac{40}{21}+\delta}$ for some fixed $\delta > 0$.

Corollary 1.3 shows that there is significant cancellation between the zeros α_i of $Z(X_0(N), T)$, and in this sense it goes beyond the Riemann hypothesis for $Z(X_0(N), T)$. Assuming square-root cancellation between the zeros, one might conjecture an error term of size $(qN)^{1/2+\varepsilon}$ in Corollary 1.3, which would imply that the main term is larger than the error term whenever $q \gg N^{1+\delta}$ for some δ. If one assumes the generalized Lindelöf hypothesis for adjoint square L-functions, then the method in this paper produces an error term of size $q^{1/8+\varepsilon}N^{1/2+\varepsilon}$ in Corollary 1.3 (see Lemma 6.1). In a much more speculative direction, if under the assumption $(mn, W) = 1$ the upper bound $\ll_{\kappa,\varepsilon} (mnW)^\varepsilon W^{-1/2}$ for the sum appearing in Lemma 4.1 holds (cf. the Linnik–Selberg conjecture), then the error term $(qN)^{1/2+\varepsilon}$ in Corollary 1.3 is admissible.

If q is a square then we can compare the second main term in Corollary 1.3 to the error term coming from (1-2) in the range where q is small compared to N. For example, in the special case that p is a prime and $q = p^2$ we have:

Corollary 1.4. *If $p, N \to \infty$ where p runs through primes $p \nmid N$ then for any fixed $\delta > 0$ we have*

$$|X_0(N)(\mathbb{F}_{p^2})| = \begin{cases} p^2 + O(p\psi(N)) & \text{if } p^2 \gg N^{4-\delta}, \\ p^2 + p\frac{1}{12}\psi(N) + O_\varepsilon(p^{\frac{23}{22}}N^{\frac{10}{11}}(qN)^\varepsilon) & \text{if } N^{\frac{40}{21}-\delta} \ll p^2 \ll N^{4-\delta}, \\ p\frac{1}{12}\psi(N) + O_\varepsilon(p^{\frac{23}{22}}N^{\frac{10}{11}}(qN)^\varepsilon) & \text{if } N^{\frac{8}{13}+\delta} \ll p^2 \ll N^{\frac{40}{21}-\delta}, \\ (p-1)\frac{1}{12}\psi(N) + O_\varepsilon((p^4 + N^{\frac{1}{2}}p)d(N)p^\varepsilon) & \text{if } p^2 \ll N^{\frac{2}{3}-\delta}. \end{cases}$$

The first of these cases is just (1-8), and the last is the Tsfasman–Vlăduţ–Zink theorem [Tsfasman et al. 1982], which has important applications to algebraic coding theory; see [Moreno 1991, Chapter 5].

Using Theorem 1.2 we can make more explicit statements about elliptic curves themselves. Let E be an elliptic curve defined over $\mathbb{F}_q$ and let $t_E = q + 1 - \#E(\mathbb{F}_q)$ be the trace of the associated Frobenius endomorphism. Hasse's theorem tells us that $|t_E| \leqslant 2\sqrt{q}$. The set of $\mathbb{F}_q$-isomorphism classes of elliptic curves defined over $\mathbb{F}_q$ is naturally a probability space where the probability of a singleton is given by

$$\boldsymbol{P}_q(\{E\}) = \frac{1}{q|\operatorname{Aut}_{\mathbb{F}_q}(E)|}.$$

We would like to study the expectations as $q \to \infty$ of various random variables associated to t_E or the structure of the group of $\mathbb{F}_q$-rational points of E. To be precise: let A be a finite abelian group with at most two generators, and let Φ_A denote the indicator function of the event that there exists an injective group homomorphism $A \hookrightarrow E(\mathbb{F}_q)$. Let $U_j(x)$ for $j \geqslant 0$ be the Chebyshev polynomials of the second kind. The Chebyshev polynomials form an orthonormal basis for the Hilbert space $L^2\big([-1,1], \frac{2}{\pi}\sqrt{1-x^2}dx\big)$.

N. Kaplan and the author [Kaplan and Petrow 2017, Theorem 2] gave explicit formulas for the expectations

$$\boldsymbol{E}_q(U_j(t_E/2\sqrt{q})\Phi_A) = \frac{1}{q} \sum_{\substack{E/\mathbb{F}_q \\ A\hookrightarrow E(\mathbb{F}_q)}} \frac{U_j(t_E/2\sqrt{q})}{|\operatorname{Aut}_{\mathbb{F}_q}(E)|}$$

in terms of $\operatorname{Tr}(\langle d\rangle T_m|S_\kappa(\Gamma(M,N)))$ and elementary arithmetic functions of m, M, N, and j.

Theorem 1.2 yields the following refinement of the error term in the main corollary of [Kaplan and Petrow 2017]. Let

$$v(n_1,n_2) = \frac{n_1}{\psi(n_1)\varphi(n_1)n_2^2} \prod_{\ell \mid \frac{n_1}{(q-1,n_1)}} (1+\ell^{-1-2v_\ell(\frac{(q-1,n_1)}{n_2})}).$$

Corollary 1.5. *Let $n_1 = n_1(A)$ and $n_2 = n_2(A)$ be the first and second invariant factors of A (i.e., we have $n_2|n_1$). Suppose that $(|A|,q)=1$ and $q\equiv 1 \pmod{n_2}$. Then*

$$\boldsymbol{E}_q(U_j(t_E/2\sqrt{q})\Phi_A) = v(n_1,n_2)\big(\delta(j,0) + O_{j,\varepsilon}(\min(n_1, q^{\frac{1}{44}}n_1^{\frac{19}{22}})n_1n_2q^{-\frac{1}{2}}(qn_1)^{\varepsilon})\big).$$

If $q\not\equiv 1 \pmod{n_2}$, then $\boldsymbol{E}_q(U_j\Phi_A)$ vanishes identically.

In particular, the traces of the Frobenius t_E for $\{E/\mathbb{F}_q : A\hookrightarrow E(\mathbb{F}_q)\}$ become equidistributed with respect to the Sato–Tate measure as $q\to\infty$ through prime powers $q\equiv 1 \pmod{n_2}$. The equidistribution is uniform in A as soon as $q \gg n_2^2 n_1^{41/11+\delta}$ for any fixed $\delta>0$.

In [Kaplan and Petrow 2017] Kaplan and the author showed that the equidistribution of t_E for $\{E/\mathbb{F}_q : A\hookrightarrow E(\mathbb{F}_q)\}$ is uniform as soon as $q\gg n_2^2n_1^{4+\delta}$ by applying (1-6) to bound the trace. In this sense, Corollary 1.5 goes beyond what one can conclude using the Riemann hypothesis of Deligne alone. All of the error terms in the theorems and corollaries found in Section 2 of [loc. cit.] are similarly improved by applying Theorem 1.2 in addition to (1-6).

1C. *Outline of proof.* Thanks to (1-5), the structural steps of the proof of Theorem 1.2 reduce to those of Theorem 1.1. The details of the analytic arguments differ however (see Section 5). For these reasons, we only discuss the proof of Theorem 1.1 in this outline.

By Atkin–Lehner theory, to estimate $\operatorname{Tr}(T_m|S_\kappa(\Gamma_0(N),\epsilon))$ it suffices to estimate

$$\sum_{f\in H_\kappa^\star(N,\epsilon)} \lambda_f(m), \tag{1-9}$$

where $H_\kappa^\star(N,\epsilon)$ is set of Hecke-normalized newforms of level N and character ϵ, and $\lambda_f(m)$ is the m-th Hecke eigenvalue of f, normalized so that $|\lambda_f(n)|\leqslant d(n)$. Whereas Serre and Conrey, Duke, and Farmer used the Eichler–Selberg trace formula to access the trace of T_m, we take a different path and use the Petersson trace formula.

Let $\mathcal{B}_\kappa(\Gamma_0(N),\epsilon)$ be an orthonormal basis for $S_\kappa(\Gamma_0(N),\epsilon)$. Let $g\in\mathcal{B}_\kappa(\Gamma_0(N),\epsilon)$ and write its Fourier coefficients as $\{b_g(n)\}_{n\geqslant 1}$. Then the Petersson formula [Iwaniec and Kowalski 2004, Proposition 14.5] says that

$$\frac{\Gamma(\kappa-1)}{(4\pi\sqrt{mn})^{\kappa-1}}\sum_{f\in\mathcal{B}_\kappa(\Gamma_0(N),\epsilon)} b_f(n)\overline{b_f(m)} = \delta(m,n)+2\pi i^{-\kappa}\sum_{\substack{c>0\\ c\equiv 0 \ (\mathrm{mod}\ N)}}\frac{S_\epsilon(m,n,c)}{c}J_{\kappa-1}\biggl(\frac{4\pi\sqrt{mn}}{c}\biggr), \tag{1-10}$$

where J_α is the J-Bessel function, $S_\epsilon(m,n,c)$ is the twisted Kloosterman sum

$$S_\epsilon(m,n,c)=\sideset{}{^*}\sum_{d\ (\mathrm{mod}\ c)}\epsilon(d)e\biggl(\frac{dm+\bar{d}n}{c}\biggr),$$

and the $*$ indicates we run over invertible $d\ (\mathrm{mod}\ c)$.

Our goal is to apply the Petersson formula to (1-9), and so we are faced with two technical difficulties:

(1) Only the newforms in $S_\kappa(\Gamma_0(N),\epsilon)$ have Fourier coefficients proportional to the Hecke eigenvalues appearing in (1-9).

(2) If f is a newform, the constant of proportionality between Fourier coefficients $b_f(n)$ and the Hecke eigenvalues $\lambda_f(n)$ is $\approx \|f\|_{L^2}$, which is not constant across $H_\kappa^\star(N,\epsilon)$.

We overcome (1) in Theorem 3.1 by developing a Petersson formula for newforms for $S_\kappa(\Gamma_0(N),\epsilon)$. There has been much recent interest in such formulas; see for example [Barrett et al. 2017; Nelson 2017; Petrow and Young 2018; Young 2018]. Theorem 3.1 is a generalization of [Barrett et al. 2017, Proposition 4.1] to nontrivial central characters, which itself is a generalization of work of Iwaniec, Luo and Sarnak [Iwaniec et al. 2000], Rouymi [2011] and Ng [2012]. Peter Humphries has also shared a preprint with the author in which he independently obtains Theorem 3.1, and uses it to study low-lying zeros of the L-functions associated to $f\in H_\kappa^\star(N,\epsilon)$. Theorem 3.1 is the only place in the proof where we have used the hypothesis $(N,m)=1$, in an essential way, and so is the source of the relatively prime conditions in Theorems 1.1 and 1.2.

We deal with (2) by appealing to the special value formula

$$L(1,\mathrm{Ad}^2 f)=\frac{\zeta^{(N)}(2)(4\pi)^\kappa}{\Gamma(\kappa)}\frac{\|f\|_{L^2}^2}{\mathrm{Vol}\,X_0(N)},$$

where $L(s,\mathrm{Ad}^2 f)$ is a certain Dirichlet series whose coefficients involve $\lambda_f(n^2)$, and which we discuss in more detail in Section 2. One may then swap the sum over f and this Dirichlet series, and apply our Petersson formula for newforms (Theorem 3.1). Estimating the resulting sums directly using the Weil bound for $S_\epsilon(a,b,c)$ (see Lemma 4.2), one recovers that the trace of T_m is $\ll_m (N\kappa)^{1+\varepsilon}$ (compare with (1-1)).

To save a bit more and obtain Theorem 1.1 we remove the weights $\|f\|_{L^2}^2$ more efficiently using a method due to Kowalski and Michel [1999, Proposition 2]. Their method is based on Hölder's inequality and a large sieve inequality due to Duke and Kowalski [2000, Theorem 4] for subfamilies of automorphic forms on GL_3. There are other notable large sieve inequalities for GL_3 in the literature; see, e.g., [Blomer et al. 2017, Theorem 3] and [Venkatesh 2006, Theorem 1]. However, these two are not useful to us since we need a large sieve inequality which is efficient for the proper subfamily of GL_3 forms cut out by the image of the adjoint square lift from GL_2. The inequality of Duke and Kowalski is superior to the results [Blomer et al. 2017, Theorem 3] and [Venkatesh 2006, Theorem 1] in the case of a thin subfamily and a long summation variable, which is the situation of interest to us.

2. Preliminaries on L-series

If $L(s)$ is a meromorphic function defined in $\mathrm{Re}(s) \gg 1$ by an infinite product over primes p of local factors $L_p(s)$, then for any integer N we write

$$L^{(N)}(s) = \prod_{p \nmid N} L_p(s) \quad \text{and} \quad L_N(s) = \prod_{p \mid N} L_p(s),$$

so that $L(s) = L_N(s) L^{(N)}(s)$ for any $N \in \mathbb{N}$. To deal with the $\|f\|_{L^2}^2$-normalization alluded to in Section 1C, we introduce the "naive" adjoint square L-function. For $f \in H_\kappa^\star(N, \epsilon)$, let

$$L(s, \mathrm{Ad}^2 f) = \frac{\zeta^{(N)}(2s)}{\zeta(s)} \sum_{n \geqslant 1} \frac{|\lambda_f(n)|^2}{n^s} = \prod_p L_p(s, \mathrm{Ad}^2 f),$$

where $\zeta(s)$ is the Riemann zeta function, and where

$$L_p(s, \mathrm{Ad}^2 f) = \begin{cases} \left(1 - \dfrac{1}{p^{2s}}\right)^{-1} \displaystyle\sum_{\alpha \geqslant 0} \frac{\bar{\epsilon}(p^\alpha) \lambda_f(p^{2\alpha})}{p^{\alpha s}} & \text{if } p \nmid N, \\ \left(1 - \dfrac{1}{p^s}\right)\left(1 - \dfrac{|\lambda_f(p)|^2}{p^s}\right)^{-1} & \text{if } p \mid N. \end{cases} \tag{2-1}$$

Warning: the $L(s, \mathrm{Ad}^2 f)$ is *not* the true adjoint square L-function of f as defined by functoriality (see [Iwaniec and Kowalski 2004, p. 133] and the online errata). But if $p \nmid N$, then $L_p(s, \mathrm{Ad}^2 f)$ *does* match the local L-factor at p of the true adjoint square L-function. Our "naive" adjoint square L function $L(s, \mathrm{Ad}^2 f)$ is chosen to be the Dirichlet series for which the following lemma is true.

Lemma 2.1. *The series $L(s, \mathrm{Ad}^2 f)$ defined above is holomorphic for $\mathrm{Re}(s) > 0$ and*

$$L(1, \mathrm{Ad}^2 f) = \frac{\zeta^{(N)}(2)(4\pi)^\kappa}{\Gamma(\kappa)} \frac{\langle f, f \rangle_N}{\mathrm{Vol}\, X_0(N)}, \tag{2-2}$$

where

$$\langle f, f \rangle_N = \int_{\Gamma_0(N) \backslash \mathcal{H}} |f(z)|^2 y^\kappa \frac{dx\, dy}{y^2}$$

and

$$\mathrm{Vol}\, X_0(N) = \int_{\Gamma_0(N) \backslash \mathcal{H}} \frac{dx\, dy}{y^2} = \frac{\pi}{3} \psi(N).$$

Proof. For the first statement, let π denote the irreducible admissible cuspidal automorphic representation of GL_2 generated by f, and denote by $L(s, \mathrm{Ad}^2\,\pi)$ the L-function of its adjoint square lift. We have by [Gelbart and Jacquet 1978] that $L(s, \mathrm{Ad}^2\,\pi)$ is an entire function of s. Therefore, the prime-to-N part of the naive L-function $L^{(N)}(s, \mathrm{Ad}^2\,f)$ is holomorphic for $\mathrm{Re}(s) > 0$.

For the second statement, take the standard nonholomorphic Eisenstein series for $\Gamma_0(N)$ at the cusp ∞ given by

$$E(z,s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \mathrm{Im}(\gamma z)^s.$$

Then we have by the classical Rankin–Selberg unfolding argument

$$\int_{\Gamma_0(N)\backslash \mathcal{H}} |f(z)|^2 E(z,s) y^\kappa \,\frac{dx\,dy}{y^2} = \frac{\Gamma(s+\kappa-1)}{(4\pi)^{s+\kappa-1}} \sum_{n \geqslant 1} \frac{|\lambda_f(n)|^2}{n^s}.$$

We deduce the lemma by taking residues on both sides and recalling [Iwaniec 1997, Theorem 13.2] that

$$\mathrm{Res}_{s=1}\, E(z,s) = \mathrm{Vol}\, X_0(N)^{-1}.$$ □

Let $\varrho_f(n)$ be the Dirichlet series coefficients of $L^{(N)}(s, \mathrm{Ad}^2\,f)$. Explicitly,

$$\varrho_f(n) = \begin{cases} \sum_{n=m^2\ell} \bar{\epsilon}(\ell)\lambda_f(\ell^2) & \text{if } (n,N)=1, \\ 0 & \text{if } (n,N)>1. \end{cases} \tag{2-3}$$

Inverting, we also have

$$\bar{\epsilon}(n)\lambda_f(n^2) = \sum_{m^2\ell=n} \mu(m)\varrho_f(\ell). \tag{2-4}$$

For future reference, we write the partial sums of $L^{(N)}(1, \mathrm{Ad}^2\,f)$ compactly as

$$\omega_f(x) = \sum_{n \leqslant x} \frac{\varrho_f(n)}{n}. \tag{2-5}$$

By contrast, when $p \mid N$ we have that $L_p(s, \mathrm{Ad}^2\,f)$ is constant along $f \in H_\kappa^\star(N,\epsilon)$ by the following lemma.

Lemma 2.2 [Ogg 1969, Theorems 2 and 3]. *Let $p \mid N$ be a prime, and ϵ a Dirichlet character mod N. Write*

$$a_{N,\epsilon}(p) = \begin{cases} 1 & \text{if } \epsilon \text{ is } \textbf{not} \text{ a character mod } N/p, \\ \frac{1}{p} & \text{if } \epsilon \text{ is a character mod } N/p \text{ and } p^2 \nmid N, \\ 0 & \text{if } \epsilon \text{ is a character mod } N/p \text{ and } p^2 \mid N. \end{cases}$$

Then we have $|\lambda_f(p)|^2 = a_{N,\epsilon}(p)$.

3. Structural steps

We study the operator $T'_m = T_m/m^{(\kappa-1)/2}$ on $S_\kappa(\Gamma_0(N),\epsilon)$ so that each eigenvalue $\lambda_f(m)$ of the T'_m operator is normalized by Deligne's theorem to have $|\lambda_f(m)| \leqslant d(m)$. We write $H_\kappa^\star(N,\epsilon)$ for the set of

Hecke-normalized newforms in $S_\kappa(N, \epsilon)$ in the sense of Atkin–Lehner theory [1970]; see also [Li 1975]. Also by Atkin–Lehner theory we have when $(m, N) = 1$ that

$$\operatorname{Tr}(T'_m | S_\kappa(\Gamma_0(N), \epsilon)) = \sum_{LM=N} d(L) \sum_{f \in H^\star_\kappa(M,\epsilon)} \lambda_f(m), \tag{3-1}$$

where we consider the interior sum to be empty if ϵ is not a character mod M. Thanks to (1-5), we can reduce the structural steps for traces on $S_\kappa(\Gamma(M, N))$ to the case of $S_\kappa(\Gamma_0(N), \epsilon)$.

Recall the notation from Section 1C and write $c_\kappa = \Gamma(\kappa - 1)/(4\pi)^{\kappa-1}$. Let

$$\Delta_{\kappa,N,\epsilon}(m, n) = \frac{c_\kappa}{(\sqrt{mn})^{\kappa-1}} \sum_{f \in \mathcal{B}_\kappa(\Gamma_0(N),\epsilon)} b_f(n)\overline{b_f(m)},$$

so that the Petersson formula (1-10) is

$$\Delta_{\kappa,N,\epsilon}(m, n) = \delta(m, n) + 2\pi i^{-\kappa} \sum_{\substack{c>0 \\ c\equiv 0 \pmod N}} \frac{S_\epsilon(m, n, c)}{c} J_{\kappa-1}\left(\frac{4\pi\sqrt{mn}}{c}\right). \tag{3-2}$$

The following theorem is our main tool for computing sums over the set of newforms $H^\star_\kappa(N, \epsilon)$.

Theorem 3.1. *If $(mn, N) = 1$ then we have*

$$c_\kappa \sum_{f \in H^\star_\kappa(N,\epsilon)} \frac{\overline{\lambda_f(m)}\lambda_f(n)}{\langle f, f\rangle_N} = \sum_{LM=N} \mu(L) R(M, L, \epsilon) \sum_{\substack{\ell | L^\infty \\ (\ell,M)=1}} \frac{\bar\epsilon(\ell)}{\ell} \Delta_{\kappa,M,\epsilon}(m, n\ell^2),$$

where

$$R(M, L, \epsilon) := \frac{1}{L} \prod_{\substack{p^2 | L \\ p\nmid M}} \left(1 - \frac{1}{p^2}\right)^{-1} \prod_{p | (M,L)} \left(1 - \frac{a_{M,\epsilon}(p)}{p}\right)^{-1},$$

and $a_{M,\epsilon}(p)$ was defined in Lemma 2.2.

Proof. See Section 7. □

Theorem 3.1 does not directly apply to (3-1) because of the normalization by $\langle f, f\rangle_N$.

We present a technique for removing the weights $\langle f, f\rangle_N$, which is a slight generalization of [Kowalski and Michel 1999, §3]. The idea for removing such weights first appeared in [Ram Murty 1995]. Let $\alpha = (\alpha_f)$ be a sequence of complex numbers indexed by

$$f \in \bigcup_{N\geq 1} \bigcup_{\epsilon \pmod N} H^\star_\kappa(N, \epsilon).$$

Define the natural averaging operator

$$A[\alpha] = A_{N,\epsilon}[\alpha] = \sum_{f \in H^\star_\kappa(N,\epsilon)} \alpha_f.$$

Let

$$\omega_f = c_\kappa \frac{L_N(1, \mathrm{Ad}^2 f)}{\langle f, f\rangle_N}.$$

Then we define the *harmonic* averaging operator

$$A^h[\alpha] = A^h_{N,\epsilon}[\alpha] = \sum_{f \in H_\kappa^\star(N,\epsilon)} \omega_f \alpha_f.$$

The following proposition is a minor generalization of Proposition 2 of [Kowalski and Michel 1999]. It allows us to pass from natural averages of newforms to harmonic averages of newforms.

Proposition 3.2. *Let $\alpha = (\alpha_f)$ be a sequence of complex numbers indexed by $f \in H_\kappa^\star(N,\epsilon)$ running over all N and all ϵ. Suppose that for all $\varepsilon > 0$*

$$A^h[|\alpha_f|] \ll_\varepsilon (N\kappa)^\varepsilon \tag{3-3}$$

and

$$\max_{f \in H_\kappa^\star(N,\epsilon)} |\omega_f \alpha_f| \ll (N\kappa)^{-\delta+\varepsilon} \tag{3-4}$$

for some absolute $\delta > 0$. For any integer $r \geqslant 1$ write $x = (N\kappa)^{10/r}$. Then we have

$$A[\alpha_f] = \frac{\kappa-1}{4\pi} \frac{\mathrm{Vol}\, X_0(N)}{\zeta^{(N)}(2)} \big(A^h[\omega_f(x)\alpha_f] + O_{\varepsilon,r}(x^{-\frac{\delta}{20}+\varepsilon} + (N\kappa)^{-1})\big).$$

Proof. See Section 6. □

One of the main ingredients in the proof of Proposition 3.2 is a large sieve inequality for the Dirichlet series coefficients of the automorphic adjoint square L-function $L(s, \mathrm{Ad}^2\,\pi)$, see Proposition 6.2, which is a quotation of [Duke and Kowalski 2000, Corollary 6]. This inequality is only valid when the length of summation X satisfies $X \gg (N\kappa)^8$, which is far from the expected truth. Nonetheless, as of now it is the best available such inequality in the range of parameters of interest to us. The exponent $-\delta/20$ in Proposition 3.2 is optimized given the exponent 8 above, and any improvement over the result of Duke and Kowalski would lead to a corresponding improvement to the value $20 = 2(8+2)$.

We apply Proposition 3.2 with $\alpha_f = \overline{\lambda_f(m)}$ to (3-1) to get

$$\begin{aligned}\overline{\mathrm{Tr}(T'_m|S_\kappa(\Gamma_0(N),\epsilon))} &= \sum_{LM=N} d(L) \frac{\kappa-1}{4\pi} \frac{\mathrm{Vol}\, X_0(M)}{\zeta^{(M)}(2)} A^h_{M,\epsilon}[\omega_f(x)\overline{\lambda_f(m)}] + O(\kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon) \\ &= \frac{\kappa-1}{12} \sum_{LM=N} \frac{d(L)\psi(M)}{\zeta^{(M)}(2)} \sum_{\substack{n \leqslant x \\ (n,M)=1}} \frac{1}{n} \sum_{n=k^2\ell} \bar{\epsilon}(\ell) A^h_{M,\epsilon}[\overline{\lambda_f(m)}\lambda_f(\ell^2)] \\ &\quad + O(\kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon).\end{aligned} \tag{3-5}$$

We are now ready to apply Theorem 3.1. We deduce a version of the newform formula for the harmonic averages $A^h[\overline{\lambda_f(m)}\lambda_f(n)]$ appearing in (3-5).

Lemma 3.3. *Let $\boldsymbol{c}_p(\epsilon)$ denote the exponent of the p-part of $\boldsymbol{c}(\epsilon)$. If $(mn, N) = 1$ then we have*

$$A^h_{N,\epsilon}[\overline{\lambda_f(m)}\lambda_f(n)] = \frac{1}{\psi(N)} \sum_{LM=N} \mu(L) M F(M,\epsilon) \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \Delta_{\kappa,M,\epsilon}(m, n\ell^2), \tag{3-6}$$

where

$$F(M,\epsilon) = \prod_{\substack{p \| M \\ \boldsymbol{c}_p(\epsilon)=1}} \left(1 + \frac{1}{p}\right) \prod_{\substack{p^\alpha \| M \\ \alpha \geqslant 2 \\ \boldsymbol{c}_p(\epsilon)=\alpha}} \left(1 - \frac{1}{p}\right)^{-1}.$$

In particular, if $\epsilon = \epsilon_0$ is trivial we have

$$A^h_{N,\epsilon_0}[\overline{\lambda_f(m)}\lambda_f(n)] = \frac{1}{\psi(N)} \sum_{LM=N} \mu(L) M \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{1}{\ell} \Delta_{\kappa,M,\epsilon_0}(m, n\ell^2). \tag{3-7}$$

Note that formula (3-7) resembles closely the formula found in [Barrett et al. 2017, Proposition 4.1].

Proof. By the definition of $L_p(1, \mathrm{Ad}^2 f)$ and Theorem 3.1 we have

$$A^h[\overline{\lambda_f(m)}\lambda_f(n)] = \prod_{p \mid N} \left(1 - \frac{1}{p}\right) \left(1 - \frac{a_{N,\epsilon}(p)}{p}\right)^{-1} \sum_{LM=N} \mu(L) R(M, L, \epsilon) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \Delta_{\kappa,M,\epsilon}(m, n\ell^2).$$

It suffices to show for any L, M that

$$\frac{\psi(LM)}{M} \prod_{p \mid LM} \left(1 - \frac{1}{p}\right) \left(1 - \frac{a_{LM,\epsilon}(p)}{p}\right)^{-1} R(M, L, \epsilon) = \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right) F(M, \epsilon). \tag{3-8}$$

We may also assume that $c(\epsilon) \mid M$, since otherwise $\Delta_{\kappa,M,\epsilon}(m, n\ell^2) = 0$. Both sides of (3-8) are multiplicative, so it suffices to check the case $M = p^\alpha$ and $L = p^\beta$ for an arbitrary prime p. The following cases can be easily verified one-by-one:

- $\alpha \geqslant 2,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) = \alpha,$
- $\alpha \geqslant 2,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) < \alpha,$
- $\alpha \geqslant 2,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) = \alpha,$
- $\alpha \geqslant 2,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) < \alpha,$
- $\alpha = 1,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) = 1,$
- $\alpha = 1,\ \beta \geqslant 1,$ and $\boldsymbol{c}_p(\epsilon) = 0,$
- $\alpha = 1,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) = 1,$
- $\alpha = 1,\ \beta = 0,$ and $\boldsymbol{c}_p(\epsilon) = 0,$
- $\alpha = 0,\ \beta \geqslant 2,$ and $\boldsymbol{c}_p(\epsilon) = 0,$
- $\alpha = 0,\ \beta = 1,$ and $\boldsymbol{c}_p(\epsilon) = 0.$ □

4. Analysis for $\Gamma_0(N)$

Now we put together (3-5), the newform formula (3-6), and the Petersson formula (3-2). By (3-5) and (3-6) we have

$$\overline{\mathrm{Tr}(T'_m | S_\kappa(\Gamma_0(N), \epsilon))} = A + E,$$

where for an integer $r \geqslant 1$ to be chosen later we set $x^r = (N\kappa)^{10}$ and have

$$A = \frac{\kappa-1}{12} \sum_{LM=N} \frac{d(L)}{\zeta^{(M)}(2)} \sum_{\substack{k \leqslant x^{1/2} \\ (k,M)=1}} \frac{1}{k^2} \sum_{\substack{\ell \leqslant x/k^2 \\ (\ell,M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \times \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{\bar{\epsilon}(q)}{q} \Delta_{\kappa,W,\epsilon}(m, q^2\ell^2), \tag{4-1}$$

and E is the error term from (3-5) of size

$$E \ll_{r,\varepsilon} \kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon. \tag{4-2}$$

Applying (3-2) to A we get that

$$A = D + OD,$$

where D and OD are the contributions from the diagonal term and off-diagonal term of (3-2), respectively. We insert $\delta_{m=q^2\ell^2}\delta_{\boldsymbol{c}(\epsilon)\mid W}$ for $\Delta_{\kappa,W,\epsilon}(m, q^2\ell^2)$ in (4-1) to find

$$D = \frac{\kappa-1}{12} \frac{\bar{\epsilon}(m^{\frac12})}{m^{\frac12}} \sum_{LM=N} \frac{d(L)}{\zeta^{(M)}(2)} \sum_{\substack{k \leqslant x^{1/2}/m^{1/4} \\ (k,M)=1}} \frac{1}{k^2} \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \delta_{\boldsymbol{c}(\epsilon)\mid W}.$$

Extending the sum over k to infinity we conclude that

$$D = \frac{\kappa-1}{12} \frac{\bar{\epsilon}(m^{\frac12})}{m^{\frac12}} \sum_{LM=N} d(L) \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \delta_{\boldsymbol{c}(\epsilon)\mid W} + O_\varepsilon\left(\frac{\kappa N^{1+\varepsilon}}{x^{\frac12} m^{\frac14}} |\epsilon(m^{\frac12})|\right).$$

By a tedious case check on prime powers we have

$$\psi(N) \delta_{\boldsymbol{c}(\epsilon)\mid N} = \sum_{LM=N} M F(M,\epsilon) \delta_{\boldsymbol{c}(\epsilon)\mid M} \prod_{p^2 \mid M} \left(1 - \frac{1}{p^2}\right).$$

Therefore the result of the diagonal contribution is

$$D = \frac{\kappa-1}{12} \frac{\bar{\epsilon}(m^{\frac12})}{m^{\frac12}} \psi(N) + O\left(\frac{\kappa N^{1+\varepsilon}}{x^{\frac12} m^{\frac14}} |\epsilon(m^{\frac12})|\right), \tag{4-3}$$

which matches what one finds directly from the identity contribution of the Eichler–Selberg trace formula.

Now we treat the off-diagonal terms. Let

$$B(Y,m,W) = \sum_{\substack{\ell \leqslant Y \\ (\ell,M)=1}} \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{\bar{\epsilon}(q\ell)}{q\ell} \sum_{c \equiv 0 \ (\mathrm{mod}\ W)} \frac{S_\epsilon(m, q^2\ell^2, c)}{c} J_{\kappa-1}\left(\frac{4\pi q\ell\sqrt{m}}{c}\right). \tag{4-4}$$

Then we have

$$OD = \frac{\kappa-1}{12} \sum_{LM=N} \frac{d(L)}{\zeta^{(M)}(2)} \sum_{WQ=M} \mu(Q) W F(W,\epsilon) \prod_{p^2 \mid W} \left(1 - \frac{1}{p^2}\right) \sum_{\substack{k \leqslant x^{1/2} \\ (k,M)=1}} \frac{1}{k^2} B\left(\frac{x}{k^2}, m, W\right).$$

Lemma 4.1. *Let d_3 denote the 3-divisor function. For any $m, n \geqslant 1$ we have*

$$\sum_{c \equiv 0 \ (\mathrm{mod}\ W)} \frac{S_\epsilon(m,n,c)}{c} J_{\kappa-1}\left(\frac{4\pi\sqrt{mn}}{c}\right) \ll \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \prod_{p \mid \boldsymbol{c}(\epsilon)} p^{\frac{1}{4}} \frac{(m,n,W)^{\frac{1}{2}} d_3((m,n)) d(W)}{W\kappa^{\frac{5}{6}}} \left(\frac{mn}{\sqrt{mn}+\kappa W}\right)^{\frac{1}{2}} \log 2mn.$$

Proof. The proof is identical to [Iwaniec et al. 2000, Corollary 2.2] but with the following bound on the Kloosterman sum in lieu of the standard bound without nebentype character.

Lemma 4.2. *For integers $c \in N\mathbb{Z}$ and $a, b \in \mathbb{Z}$ with $c \neq 0$ and $\boldsymbol{c}(\epsilon) \mid N$, we have the estimate*

$$|S_\epsilon(a,b;c)| \le d(c)\,(a,b,c)^{\frac{1}{2}}\, c^{\frac{1}{2}}\, \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}}.$$

Proof. See [Knightly and Li 2013, Theorem 9.2]. □

Applying Lemma 4.1 and estimating sums by integrals we find

$$B(Y,m,W) \ll \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} \frac{d(W) m^{\frac{1}{4}} Y^{\frac{1}{2}}}{W\kappa^{\frac{5}{6}}} \log mY;$$

hence one estimates that

$$OD \ll_\varepsilon \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} x^{\frac{1}{2}} \kappa^{\frac{1}{6}} m^{\frac{1}{4}} N^\varepsilon \log mx.$$

We have $\mathrm{Tr}(T'_m \mid S_\kappa(\Gamma_0(N,\epsilon))) = \overline{D} + \overline{OD} + \overline{E}$, and so collecting error terms we obtain

$$\mathrm{Tr}(T'_m \mid S_\kappa(\Gamma_0(N,\epsilon))) = \frac{\kappa-1}{12} \frac{\epsilon(m^{\frac{1}{2}})}{m^{\frac{1}{2}}} \psi(N) + O_\varepsilon\big(\boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} x^{\frac{1}{2}} \kappa^{\frac{1}{6}} m^{\frac{1}{4}} N^\varepsilon \log mx + \kappa N^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N^\varepsilon\big). \tag{4-5}$$

We now optimize the value of r. By [Goldfeld et al. 1994; Banks 1997], the exponent $\delta = 1$ is admissible. The error in (4-5) is minimized when

$$x^{\frac{11}{20}} = \frac{N\kappa^{\frac{5}{6}}}{m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}}}.$$

Let us assume that there is some $\eta > 0$ such that

$$m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}} \ll (N\kappa^{\frac{5}{6}})^{1-\eta}. \tag{4-6}$$

We choose $r \geqslant 1$ to be the nearest integer to

$$\frac{11}{2}\left(1 + \frac{\log(\kappa^{\frac{1}{6}} m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}})}{\log(N\kappa^{\frac{5}{6}}) - \log(m^{\frac{1}{4}} \boldsymbol{c}(\epsilon)^{\frac{1}{4}} \boldsymbol{c}^*(\epsilon)^{\frac{1}{4}})}\right),$$

which by (4-6) is then bounded above uniformly in terms of $\eta > 0$ only.

5. Analysis for $\Gamma(M,N)$

Recall from (1-5) that

$$\operatorname{Tr}(\langle\bar{d}\rangle T'_m|S_\kappa(\Gamma(M,N))) = \sum_{\epsilon \ (\mathrm{mod}\ N)} \bar{\epsilon}(d)\operatorname{Tr}(T'_m|S_\kappa(\Gamma_0(MN),\epsilon)),$$

and that in Section 4 we decomposed the interior of this as

$$\overline{\operatorname{Tr}(T'_m|S_\kappa(\Gamma_0(MN),\epsilon))} = D + OD + E.$$

Summing the formula (4-3) for D and (4-2) for E trivially over characters $\epsilon \ (\mathrm{mod}\ N)$ we get

$$\begin{aligned}\operatorname{Tr}(\langle\bar{d}\rangle T'_m|S_\kappa(\Gamma(M,N))) = \frac{\kappa-1}{24}m^{-\frac{1}{2}}\varphi(N)\psi(NM)\big(\delta_N(m^{\frac{1}{2}}d,1)+(-1)^\kappa\delta_N(m^{\frac{1}{2}}d,-1)\big)\\ +\overline{OD^*}+O_{\eta,\varepsilon}(\kappa(MN^2)^{1+\varepsilon}x^{-\frac{\delta}{20}+\varepsilon}+N(MN)^\varepsilon), \quad (5\text{-}1)\end{aligned}$$

where $x^r=(MN\kappa)^{10}$, r is a parameter to be chosen later, and

$$OD^* = \sum_{\substack{\epsilon \ (\mathrm{mod}\ N)\\ \epsilon(-1)=(-1)^\kappa}} \epsilon(d)OD.$$

Let

$$\begin{aligned}&B^*(Y,m,W)\\ &\quad= \sum_{\substack{\epsilon \ (\mathrm{mod}\ N)\\ \epsilon(-1)=(-1)^\kappa}} \epsilon(d)F(W,\epsilon) \sum_{\substack{(\ell,K)=1\\ \ell\leqslant Y}} \sum_{\substack{q\,|\,Q^\infty\\ (q,W)=1}} \frac{\overline{\epsilon(q\ell)}}{q\ell} \sum_{c\equiv 0 \ (\mathrm{mod}\ W)} \frac{S_\epsilon(m,q^2\ell^2,c)}{c} J_{\kappa-1}\Big(\frac{4\pi\ell q\sqrt{m}}{c}\Big),\end{aligned}$$

so that we have

$$OD^* = \frac{\kappa-1}{12}\sum_{LK=MN}\frac{d(L)}{\zeta^{(K)}(2)}\sum_{\substack{k\leqslant x^{1/2}\\ (k,K)=1}}\frac{1}{k^2}\sum_{WQ=K}\mu(Q)W\prod_{p^2|W}\Big(1-\frac{1}{p^2}\Big)B^*\Big(\frac{x}{k^2},m,W\Big). \quad (5\text{-}2)$$

We would like to utilize the orthogonality of characters over $\epsilon \ (\mathrm{mod}\ N)$. To implement this, we now refresh the notation. Suppose $W,N\geqslant 1$ are integers such that $W\,|\,N^2$. For $a,b,d,\kappa\in\mathbb{Z}$ and $1\leqslant c\equiv 0 \ (\mathrm{mod}\ W)$ define

$$T_W(a,b,c) := \sum_{\substack{\epsilon \ (\mathrm{mod}\ N)\\ \epsilon(-1)=(-1)^\kappa\\ c(\epsilon)\,|\,W}} \epsilon(d)\bar{\epsilon}(b)F(W,\epsilon)S_\epsilon(a,b,c).$$

With this notation, we have

$$B^*(Y,m,W) = \sum_{\substack{(\ell,K)=1\\ \ell\leqslant Y}} \sum_{\substack{q\,|\,Q^\infty\\ (q,W)=1}} \frac{1}{q\ell} \sum_{c\equiv 0 \ (\mathrm{mod}\ W)} \frac{T_W(m,q^2\ell^2,c)}{c} J_{\kappa-1}\Big(\frac{4\pi\ell q\sqrt{m}}{c}\Big). \quad (5\text{-}3)$$

We can derive a bound on T_W by appealing to the Weil bound for Kloosterman sums.

Lemma 5.1. *Suppose $W, N \geqslant 1$ such that $W \mid N^2$, $a, b, d, \kappa \in \mathbb{Z}$ such that $(b, W) = 1$, $(d, N) = 1$, and $1 \leqslant c \equiv 0 \pmod{W}$. We factor $c = c_1 c_2$ with $c_1 \mid W^\infty$ and $(c_2, W) = 1$. Then*

$$|T_W(a, b, c)| \leqslant \psi(c_1) d(c_2) (a, b, c_2)^{\frac{1}{2}} c_2^{\frac{1}{2}}.$$

Proof. Consider the sum

$$T'_W(a, b, c) := \sum_{\substack{\epsilon \ (\mathrm{mod}\ N) \\ \boldsymbol{c}(\epsilon) \mid W}} \epsilon(d) \bar{\epsilon}(b) F(W, \epsilon) S_\epsilon(a, b, c),$$

which is a minor variation of $T_W(a, b, c)$ omitting the global condition $\epsilon(-1) = (-1)^\kappa$. We first consider the sum T'_W locally, returning to T_W at the end of the proof. Let $\alpha, \beta, \gamma \geqslant 0$ such that $\alpha \leqslant \gamma$, $\alpha \leqslant 2\beta$, $(d, p^\beta) = (b, p^\alpha) = 1$, and consider $T'_{p^\alpha}(a, b, p^\gamma)$. Let

$$I(\alpha, \beta) := \sum_{\epsilon \ (\mathrm{mod}\ p^\beta)} \epsilon(dx) \bar{\epsilon}(b) \delta_{\boldsymbol{c}_p(\epsilon) \leqslant \alpha} \begin{cases} 1 + \frac{1}{p} & \text{if } \boldsymbol{c}_p(\epsilon) = \alpha = 1, \\ \left(1 - \frac{1}{p}\right)^{-1} & \text{if } \boldsymbol{c}_p(\epsilon) = \alpha \geqslant 2, \\ 1 & \text{else.} \end{cases}$$

By opening the Kloosterman sum and exchanging order of summation we have

$$T'_{p^\alpha}(a, b, p^\gamma) = \sum_{x \ (\mathrm{mod}\ p^\gamma)}^{*} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) I(\alpha, \beta). \tag{5-4}$$

Next we break into four cases:

(1) $\alpha > \beta$.

(2) $0 = \alpha \leqslant \beta$.

(3) $1 = \alpha \leqslant \beta$.

(4) $2 \leqslant \alpha \leqslant \beta$.

Recall the orthogonality relation

$$\sum_{\epsilon \ (\mathrm{mod}\ n)} \epsilon(a) \bar{\epsilon}(b) = \varphi(n) \delta_n(a, b)$$

and the almost-orthogonality relation (see, e.g., [Heath-Brown 1981, Section 2])

$$\sum_{\boldsymbol{c}(\epsilon) = c} \epsilon(a) \bar{\epsilon}(b) = \sum_{\delta \mid (a-b, c)} \varphi(\delta) \mu\left(\frac{c}{\delta}\right).$$

We apply these to evaluate $I(\alpha, \beta)$ in cases (1)–(4). We find

$$I(\alpha, \beta) = \begin{cases} \varphi(p^\beta) \delta_{p^\beta}(xd, b) & \text{if } \alpha > \beta, \\ 1 & \text{if } 0 = \alpha \leqslant \beta, \\ \varphi(p) \delta_p(xd, b) + \frac{1}{p} \sum_{\delta \mid (p, xd-b)} \varphi(\delta) \mu(p^\gamma / \delta) & \text{if } 1 = \alpha \leqslant \beta, \\ \varphi(p^\alpha) \delta_{p^\alpha}(xd, b) + \frac{1}{p-1} \sum_{\delta \mid (p^\alpha, xd-b)} \varphi(\delta) \mu(p^\gamma / \delta) & \text{if } 1 = \alpha \leqslant \beta. \end{cases} \tag{5-5}$$

Recall that $(d, p^\beta) = 1$, so that $d^{-1} \pmod{p^\beta}$ (or $\pmod{p^\gamma}$ in cases (3) and (4)) exists. Inserting (5-5) to (5-4), we find the following.

Case (1): $\beta < \alpha$.

$$T'_{p^\alpha}(a, b, p^\gamma) = \varphi(p^\beta) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{p^\beta}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right).$$

Case (2): $\beta \geqslant \alpha = 0$.

$$T'_1(a, b, p^\gamma) = \sideset{}{^*}\sum_{x \pmod{p^\gamma}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) = S(a, b, p^\gamma).$$

Case (3): $\beta \geqslant \alpha = 1$.

$$T'_p(a, b, p^\gamma) = \varphi(p) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{p}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) + \frac{1}{p} \sum_{\delta \mid p} \varphi(\delta)\mu(p/\delta) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{\delta}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right).$$

Case (4): $\beta \geqslant \alpha \geqslant 2$.

$$T'_{p^\alpha}(a, b, p^\gamma) = \varphi(p^\alpha) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{p^\alpha}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right) + \frac{1}{p-1} \sum_{\delta \mid p^\alpha} \varphi(\delta)\mu(p^\alpha/\delta) \sideset{}{^*}\sum_{\substack{x \pmod{p^\gamma} \\ x \equiv d^{-1} b \pmod{\delta}}} e\left(\frac{ax + b\bar{x}}{p^\gamma}\right).$$

Using the Weil bound for Kloosterman sums and trivial bounds, we find for all integers a, b, and nonzero integers $0 \leqslant i \leqslant j$, and $(y, p) = 1$ we have

$$\left| \sideset{}{^*}\sum_{\substack{x \pmod{p^j} \\ x \equiv y \pmod{p^i}}} e\left(\frac{ax + b^2\bar{x}}{p^\gamma}\right) \right| \leqslant \begin{cases} d(p^j)(a, b^2, p^j)^{\frac{1}{2}} \sqrt{p^j} & \text{if } i = 0, \\ p^{j-i} & \text{else.} \end{cases} \tag{5-6}$$

Applying (5-6) to the various cases above, we find that for cases (1), (3), and (4), i.e., when $\alpha > 0$, we have the bound

$$|T'_{p^\alpha}(a, b, p^\gamma)| \leqslant \psi(p^\gamma). \tag{5-7}$$

In case (2), i.e., when $\alpha = 0$, we have

$$|T'_{p^\alpha}(a, b, p^\gamma)| \leqslant d(p^\gamma)(a, b, p^\gamma)^{\frac{1}{2}} p^{\frac{\gamma}{2}}. \tag{5-8}$$

Thus the estimation of $T'_{p^\alpha}(a, b, p^\gamma)$ is finished.

Now we return to the case of $T_W(a, b, c)$. We have

$$T_W(a, b, c) = \tfrac{1}{2} T'_W(a, b, c) + \tfrac{1}{2}(-1)^\kappa T'_W(a, -b, c),$$

so it suffices to establish the bound stated in the lemma for $T'_W(a,b,c)$. We have that $T'_W(a,b,c)$ is twisted multiplicative, i.e., we have a factorization

$$T'_W(a,b,c) = \prod_{\substack{p^\alpha || W \\ p^\gamma || c}} T'_{p^\alpha}(a\overline{cp^{-\gamma}}, b\overline{cp^{-\gamma}}, p^\gamma). \tag{5-9}$$

Bounding the left-hand side of (5-9) using (5-7) and (5-8), we conclude the proof of the lemma. □

Applying Lemma 5.1 to (5-3) we get

$$B^*(Y,m,W) \leqslant \sum_{\substack{(\ell,K)=1 \\ \ell \leqslant Y}} \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{1}{q\ell} \sum_{W \mid c_1 \mid W^\infty} \frac{\psi(c_1)}{c_1} \sum_{(c_2,W)=1} \frac{d(c_2)(m,q^2\ell^2,c_2)^{\frac12}}{\sqrt{c_2}} \left| J_{\kappa-1}\left(\frac{4\pi q\ell\sqrt{m}}{c_1c_2}\right)\right|.$$

Again following closely the proof of [Iwaniec et al. 2000, Corollary 2.2] we have

$$\sum_{(c_2,W)=1} \frac{d(c_2)(a,b^2,c_2)^{\frac12}}{\sqrt{c_2}} \left| J_{\kappa-1}\left(\frac{4\pi b\sqrt{a}}{c_1c_2}\right)\right| \ll \frac{d_3((a,b^2))}{\kappa^{\frac56}\sqrt{c_1}} \left(\frac{b^2a}{b\sqrt{a}+c_1\kappa}\right)^{\frac12} \log 2b^2a.$$

We have moreover that

$$\sum_{W \mid c_1 \mid W^\infty} \frac{1}{\sqrt{c_1}} \frac{1}{(b\sqrt{a}+c_1\kappa)^{\frac12}} \leqslant \frac{2}{W^{\frac12} b^{\frac12} a^{\frac14}}.$$

These last two estimations lead to

$$\begin{aligned} B^*(Y,m,W) &\ll \frac{m^{\frac14}}{\kappa^{\frac56}} \frac{\psi(W)}{W^{\frac32}} \sum_{\substack{(\ell,K)=1 \\ \ell \leqslant Y}} \sum_{\substack{q \mid Q^\infty \\ (q,W)=1}} \frac{d_3((m,q^2\ell^2))}{\sqrt{q\ell}} \log(2q^2\ell^2 m) \\ &\ll \frac{m^{\frac14} Y^{\frac12} (\log Y)^3 \log 2m}{\kappa^{\frac56}} \frac{\psi(W)}{W^{\frac32}}. \end{aligned}$$

Inserting this into (5-2) we get

$$OD^* \ll \kappa^{\frac16} M N^{\frac12+\varepsilon} x^{\frac12} (\log x)^3 m^{\frac14} \log 2m,$$

and inserting this into (5-1) we conclude that

$$\begin{aligned} &\mathrm{Tr}(\langle \bar d\rangle T'_m | S_\kappa(\Gamma(M,N))) \\ &\quad = \frac{\kappa-1}{24} m^{-\frac12} \varphi(N)\psi(NM)\big(\delta_N(m^{\frac12}d,1) + (-1)^\kappa \delta_N(m^{\frac12}d,-1)\big) \\ &\qquad + O_{\eta,\varepsilon}\big(\kappa^{\frac16} M N^{\frac12} x^{\frac12+\varepsilon} m^{\frac14} \log 2m + \kappa (MN^2)^{1+\varepsilon} x^{-\frac{\delta}{20}+\varepsilon} + N(MN)^\varepsilon\big). \end{aligned} \tag{5-10}$$

Now we optimize the value of r. The error term is minimized when

$$x^{\frac{11}{20}} = \frac{N^{\frac32}\kappa^{\frac56}}{m^{\frac14}}.$$

Let us assume that there is some $\eta > 0$ such that

$$m^{\frac{1}{4}} \ll (N^{\frac{3}{2}}\kappa^{\frac{5}{6}})^{1-\eta}.$$

We choose $r \geqslant 1$ to be the nearest integer to

$$\frac{11}{2}\left(\frac{\log MN\kappa}{\log N^{\frac{3}{2}}\kappa^{\frac{5}{6}} - \log m^{\frac{1}{4}}}\right),$$

which is then bounded above uniformly in terms of $\eta > 0$ only.

6. Proof of Proposition 3.2

Proof. We have by Lemma 2.1 that

$$A[\alpha_f] = \frac{\kappa-1}{4\pi}\frac{\operatorname{Vol} X_0(N)}{\zeta^{(N)}(2)} \sum_{f \in H_\kappa^\star(N,\epsilon)} \omega_f \alpha_f L^{(N)}(1, \operatorname{Ad}^2 f) = \frac{\kappa-1}{4\pi}\frac{\operatorname{Vol} X_0(N)}{\zeta^{(N)}(2)} A^h[\alpha_f L^{(N)}(1, \operatorname{Ad}^2 f)].$$

Recall that we have set $\varrho_f(n)$ to be the Dirichlet series coefficients of $L^{(N)}(s, \operatorname{Ad}^2 f)$, along with

$$\omega_f(x) = \sum_{n \leqslant x} \frac{\varrho_f(n)}{n} \quad \text{and} \quad \omega_f(x, y) = \sum_{x < n \leqslant y} \frac{\varrho_f(n)}{n}.$$

Lemma 6.1. *We have*

$$L^{(N)}(1, \operatorname{Ad}^2 f) = \omega_f(x) + \omega_f(x, y) + O_\varepsilon((N\kappa)^{\frac{1}{2}} y^{-\frac{1}{2}+\varepsilon}).$$

Assuming the generalized Lindelöf hypothesis, the $(N\kappa)^{1/2}$ *can be reduced to* $(N\kappa)^\varepsilon$.

Proof (*sketch*). For $c, T, y > 0$, we apply Perron's formula (see, e.g., [Davenport 2000, p. 105]) to calculate $\omega_f(y)$, finding

$$\omega_f(y) = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} L^{(N)}(1+s, \operatorname{Ad}^2 f)\frac{y^s}{s}\,ds + O\left(y^c \sum_{n \geqslant 1} \frac{\varrho_f(n)}{n^{1+c}} \min(1, T^{-1}|\log y/n|^{-1})\right).$$

We shift the contour to $\operatorname{Re}(s) = -2$ to get

$$\omega_f(y) = L^{(N)}(1, \operatorname{Ad}^2 f) + \frac{1}{2\pi i}\left(\int_{c-iT}^{-2-iT} + \int_{-2-iT}^{-2+iT} + \int_{-2+iT}^{c+iT}\right) L^{(N)}(1+s, \operatorname{Ad}^2 f)\frac{y^s}{s}\,ds \\ + O\left(y^c \sum_{n \geqslant 1} \frac{\varrho_f(n)}{n^{1+c}} \min(1, T^{-1}|\log y/n|^{-1})\right). \quad (6\text{-}1)$$

By an inspection of the functional equation for $L(s, f \otimes \bar{f})$ found in [Li 1979, Example 1], we have the convexity bound (see, e.g., [Iwaniec and Kowalski 2004, (5.20)])

$$L^{(N)}(s, \operatorname{Ad}^2 f) \ll [(\kappa N)^2(1+|t|)^3]^{\frac{1-\sigma}{2}+\varepsilon}, \quad (6\text{-}2)$$

where $s=\sigma+it$, valid for $\sigma\leqslant 1$. Choosing $c=\varepsilon$, $T=(N\kappa)^{-\frac{1}{2}}y^{\frac{1}{2}+\varepsilon}$, and estimating all of the terms in (6-1) directly, one finds the estimate in the statement of the lemma.

If one assumes the generalized Lindelöf hypothesis in place of (6-2), then we shift the contour to $\mathrm{Re}(s)=-\frac{1}{2}$ instead of -2 and follow the same steps. □

By Lemma 6.1 we have

$$A[\alpha_f]=\frac{\kappa-1}{4\pi}\frac{\mathrm{Vol}\,X_0(N)}{\zeta^{(N)}(2)}\big(A^h[\omega_f(x)\alpha_f]+A^h[\omega_f(x,y)\alpha_f]+O((N\kappa)^{\frac{1}{2}}y^{-\frac{1}{2}+\varepsilon}A^h[|\alpha_f|])\big). \tag{6-3}$$

By the hypothesis (3-3) we have $A^h[|\alpha_f|]\ll_\varepsilon (N\kappa)^\varepsilon$, and so taking $y=(N\kappa)^{3+\varepsilon}$, we find that the O term in (6-3) is $\ll (N\kappa)^{-1}$.

Next we consider the second term and treat it using the following large sieve inequality. This is a slight variation on Corollary 6 of [Duke and Kowalski 2000]; see also [Kowalski and Michel 1999, Proposition 1]. Let $\lambda_f^{(2)}(n)$ be the Dirichlet series coefficients of the automorphic adjoint-square L-function $L(s,\mathrm{Ad}^2\,\pi)$, where f is a newform for the representation π. If $(n,N)=1$ then we have $\lambda_f^{(2)}(n)=\varrho_f(n)$.

Proposition 6.2. *Let $X\geqslant (N\kappa)^8$. We have for all $\varepsilon>0$ that*

$$\sum_{f\in H_\kappa^\star(N,\epsilon)}\Big|\sum_{n\leqslant X}a_n\lambda_f^{(2)}(n)\Big|^2\ll_\varepsilon X^{1+\varepsilon}\sum_{n\leqslant X}|a_n|^2 \tag{6-4}$$

for any finite family $(a_n)_{1\leqslant n\leqslant X}$ of complex numbers, where the constant depends only on ε.

By following closely [Kowalski and Michel 1999, §3.3] one deduces from Proposition 6.2 the following lemma.

Lemma 6.3. *Let $r\geqslant 1$ be an integer such that $x^r\geqslant (N\kappa)^{10}$. Then for all $\varepsilon>0$ we have*

$$A[\omega_f(x,y)^{2r}]\ll_{r,\varepsilon}(N\kappa)^\varepsilon,$$

where the implied constant depends only on r and ε.

Proof. It suffices to replace instances of $\lambda_f(n^2)$ in [Kowalski and Michel 1999, Lemma 3] by $\bar\epsilon(n)\lambda_f(n^2)$ and to use (2-3) and (2-4) in the place of (15) and (16) of [loc. cit.]. □

We now can give an estimate for the second term of (6-3). We use Hölder's inequality to separate $\omega_f(x,y)$ from α_f, and Lemma 6.3 to handle the the sum involving $\omega_f(x,y)$. Precisely, let s be defined by $(2r)^{-1}+s^{-1}=1$. Applying Hölder's inequality we find for any integer $r\geqslant 1$ that

$$\begin{aligned}A^h[\omega_f(x,y)\alpha_f]=\sum_{f\in H_\kappa^\star}\omega_f\omega_f(x,y)\alpha_f&\leqslant A[\omega_f(x,y)^{2r}]^{\frac{1}{2r}}\Big(\sum_{f\in H_\kappa^\star(N,\epsilon)}(\omega_f|\alpha_f|)^s\Big)^{\frac{1}{s}}\\&\leqslant A^{\frac{1}{2r}}A[\omega_f(x,y)^{2r}]^{\frac{1}{2r}}A^h[|\alpha_f|]^{\frac{1}{s}},\end{aligned}$$

where

$$A=\max_{f\in H_\kappa^\star(N,\epsilon)}\omega_f|\alpha_f|\ll_\varepsilon (N\kappa)^{-\delta+\varepsilon}$$

by hypothesis (3-4). Suppose now that r is sufficiently large so that $x^r \geqslant (N\kappa)^{10}$. Then Lemma 6.3 applies, and we have

$$A[\omega_f(x,y)^{2r}]^{\frac{1}{2r}} \ll_{r,\varepsilon} (N\kappa)^{\varepsilon}.$$

Lastly, by hypothesis (3-3) we have

$$A^h[|\alpha_f|]^{\frac{1}{s}} \ll_{\varepsilon} (N\kappa)^{\varepsilon}.$$

Putting these estimates together, we find that $A^h[\omega_f(x,y)\alpha_f] \ll_{r,\varepsilon} (N\kappa)^{-\frac{\delta}{2r}+\varepsilon}$, and so derive the bound claimed in Proposition 3.2. □

7. Proof of Theorem 3.1

Proof. The strategy of the proof is to pick an orthogonal basis for $S_\kappa(\Gamma_0(N),\epsilon)$ and compute the Fourier coefficients of basis elements explicitly. For f a modular function of weight κ, we define $f_{|d}(z) = d^{\frac{\kappa}{2}} f(dz)$. Atkin–Lehner theory gives an orthogonal direct sum decomposition

$$S_\kappa(\Gamma_0(N),\epsilon) = \bigoplus_{LM=N} \bigoplus_{f\in H_\kappa^\star(M,\epsilon)} S_\kappa(L; f,\epsilon),$$

where $S_\kappa(L; f,\epsilon) = \operatorname{span}\{f_{|\ell} : \ell \mid L\}$ is called an oldclass. Note that the inner sum is $\{0\}$ unless $\boldsymbol{c}(\epsilon) \mid M$, so we may assume this for the remainder of the proof.

To pick an orthogonal basis for $S_\kappa(\Gamma_0(N),\epsilon)$ it then suffices to pick a orthonormal basis for each oldclass $S_\kappa(L; f,\epsilon)$. We use a basis for the oldclasses first due to [Schulze-Pillot and Yenirce 2018, Theorem 8]. The basis constructed by Schulze-Pillot and Yenirce is the same as the one found by Rouymi [2011] in the case of prime power level and trivial nebentypus and Ng [2012] in the case of arbitrary level and trivial nebentypus; see also [Blomer and Milićević 2015, Chapter 5] and [Humphries 2018, Lemma 3.15]. Each of these preceding works used the Rankin–Selberg method to compute inner products and orthonormalize the oldclasses. Schulze-Pillot and Yenirce however took a different and simpler path, using the trace operator to compute the inner products.

Let $f \in H^\star(M,\epsilon)$. For integers $d \mid g$ one defines a joint multiplicative function $\xi_g(d)$. On prime powers, $\xi_g(d)$ is given for $\nu \geqslant 2$ by

$$\xi_1(1) = 1, \quad \xi_{p^\nu}(p^\nu) = \left(1 - \frac{|\lambda_f(p)|^2}{p(1+\frac{\varepsilon_{0,M}(p)}{p})^2}\right)^{-\frac{1}{2}} \left(1 - \frac{\varepsilon_{0,M}(p)^2}{p^2}\right)^{-\frac{1}{2}},$$

$$\xi_p(p) = \left(1 - \frac{|\lambda_f(p)|^2}{p(1+\frac{\epsilon_{0,M}(p)}{p})^2}\right)^{-\frac{1}{2}}, \quad \xi_{p^\nu}(p^{\nu-1}) = \frac{-\overline{\lambda_f(p)}}{\sqrt{p}} \xi_{p^\nu}(p^\nu),$$

$$\xi_p(1) = \frac{-\overline{\lambda_f(p)}}{\sqrt{p}(1+\epsilon_{0,M}(p)/p)} \xi_p(p), \quad \xi_{p^\nu}(p^{\nu-2}) = \frac{\overline{\epsilon(p)}}{p} \xi_{p^\nu}(p^\nu),$$

and $\xi_{p^a}(p^b) = 0$ in all other cases.

Proposition 7.1 [Schulze-Pillot and Yenirce 2018, Theorem 9]. *Let $M \mid N$ and let $f \in H_\kappa^\star(M, \epsilon)$. The set of functions*

$$\{f^{(g)}(z) = \textstyle\sum_{d \mid g} \xi_g(d) d^{\frac{\kappa}{2}} f(dz) : g \mid L\}$$

is an orthogonal basis for $S_k(L; f, \epsilon)$. In fact, if f is $L^2(\Gamma_0(N)\backslash\mathcal{H})$-normalized, then the above set is in fact orthonormal.

Now that we have an orthonormal basis for $S_\kappa(\Gamma_0(N), \epsilon)$, we follow Barrett, Burkhardt, DeWitt, Dorward, and Miller [Barrett et al. 2017] to derive the Petersson formula for newforms Theorem 3.1.

Let $f \in H_\kappa^\star(M, \epsilon)$ have Fourier coefficients $a_f(n)$ and be normalized so that $a_f(1) = 1$. Of course $f(z)/\|f\|_N$ is $L^2(\Gamma_0(N)\backslash\mathcal{H})$-normalized, so using the basis in Proposition 7.1 we have

$$\begin{aligned}\Delta_{\kappa,N,\epsilon}(m,n) &= \frac{c_\kappa}{(mn)^{\frac{\kappa-1}{2}}} \sum_{g \in \mathcal{B}_\kappa(\Gamma_0(N),\epsilon)} b_g(n)\overline{b_g(m)} \\ &= \frac{c_\kappa}{(mn)^{\frac{\kappa-1}{2}}} \sum_{LM=N} \sum_{f \in H_\kappa^\star(m,\epsilon)} \frac{1}{\langle f, f\rangle_N} \sum_{g \mid L} \overline{a_{f^{(g)}}(m)} a_{f^{(g)}}(n). \end{aligned} \tag{7-1}$$

By the definition of $f^{(g)}$ we have

$$a_{f^{(g)}}(n) = \sum_{d \mid (g,n)} \xi_g(d) d^{\frac{\kappa}{2}} a_f\left(\frac{n}{d}\right),$$

which are now expressible in terms of Hecke eigenvalues $\lambda_f(n)$ normalized so that $|\lambda_f(n)| \leqslant d(n)$. We have then that

$$\begin{aligned}&\Delta_{\kappa,N,\epsilon}(m,n) \\ &= \frac{c_\kappa}{(mn)^{\frac{\kappa-1}{2}}} \sum_{LM=N} \sum_{f \in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_N^2} \sum_{g \mid L} \overline{\left(\sum_{d \mid (g,m)} \xi_g(d) d^{\frac{\kappa}{2}} a_f\left(\frac{m}{d}\right)\right)} \left(\sum_{d \mid (g,n)} \xi_g(d) d^{\frac{\kappa}{2}} a_f\left(\frac{n}{d}\right)\right) \\ &= c_\kappa \sum_{LM=N} \sum_{f \in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_N^2} \sum_{g \mid L} \overline{\left(\sum_{d \mid (g,m)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{m}{d}\right)\right)} \left(\sum_{d \mid (g,n)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{n}{d}\right)\right) \\ &= c_\kappa \sum_{LM=N} \sum_{f \in H_\kappa^\star(N,\epsilon)} \frac{1}{\|f\|_N^2} \sum_{g \mid L} \Xi_g(m,n,f), \end{aligned}$$

where we have set

$$\Xi_g(m,n,f) = \overline{\left(\sum_{d \mid (g,m)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{m}{d}\right)\right)} \left(\sum_{d \mid (g,n)} \xi_g(d) d^{\frac{1}{2}} \lambda_f\left(\frac{n}{d}\right)\right)$$

for $g \mid L \mid N$.

Now suppose that $(d_1, d_2) = 1$ and $d_1 d_2 \mid m$. Then by Hecke multiplicativity we have

$$\lambda_f\left(\frac{m}{d_1}\right)\lambda_f\left(\frac{m}{d_2}\right) = \lambda_f(m)\lambda_f\left(\frac{m}{d_1 d_2}\right),$$

so that for $(g_1, g_2) = 1$ we have

$$\Xi_{g_1}(m,n,f)\Xi_{g_2}(m,n,f) = \overline{\lambda_f(m)}\lambda_f(n)\Xi_{g_1 g_2}(m,n,f).$$

Therefore

$$\Delta_{\kappa,N,\epsilon}(m,n) = c_\kappa \sum_{LM=N} \sum_{f \in H_\kappa^\star(M,\epsilon)} \frac{1}{\|f\|_N^2} (\overline{\lambda_f(m)}\lambda_f(n))^{1-\omega(L)} \prod_{p^\alpha \| L} \Bigg(\sum_{d \mid p^\alpha} \Xi_d(m,n,f) \Bigg),$$

where $\omega(n)$ is the number of distinct prime factors of n. Let

$$V_{p^\alpha}(m,n,f) = \sum_{d \mid p^\alpha} \Xi_d(m,n,f) = (1 * \Xi)_{p^\alpha}(m,n,f),$$

where $*$ denotes Dirichlet convolution. We suppose now that $(m,n,N) = 1$ and calculate.

Lemma 7.2 [Barrett et al. 2017, Appendix A]. *If $(m,n,N) = 1$ then we have*

$$\begin{aligned} V_{p^\alpha}(m,n,f) = &\overline{\lambda_f(m)}\lambda_f(n)(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2) \\ &+ \delta_{p|m}\overline{\lambda_f(m/p)}\lambda_f(n)p^{\frac12}(\overline{\xi_p(p)}\xi_p(1)+\overline{\xi_{p^2}(p)}\xi_{p^2}(1)) \\ &+ \delta_{p|n}\overline{\lambda_f(m)}\lambda_f(n/p)p^{\frac12}(\overline{\xi_p(1)}\xi_p(p)+\overline{\xi_{p^2}(1)}\xi_{p^2}(p)) \\ &+ \delta_{p^2|m}\overline{\lambda_f(m/p^2)}\lambda_f(n)p\overline{\xi_{p^2}(p^2)}\xi_{p^2}(1) + \delta_{p^2|n}\overline{\lambda_f(m)}\lambda_f(n/p^2)p\overline{\xi_{p^2}(1)}\xi_{p^2}(p^2), \end{aligned}$$

if $\alpha \geqslant 2$ and

$$\begin{aligned} V_{p^\alpha}(m,n,f) = \overline{\lambda_f(m)}\lambda_f(n)(1+|\xi_p(1)|^2) &\\ + \delta_{p|m}\overline{\lambda_f(m/p)}\lambda_f(n)p^{\frac12}\overline{\xi_p(p)}\xi_p(1) &\\ + \delta_{p|n}\overline{\lambda_f(m)}\lambda_f(n/p)p^{\frac12}\overline{\xi_p(1)}\xi_p(p), & \end{aligned}$$

if $\alpha = 1$.

Proof. We actually have if $\alpha \geqslant 2$ that

$$V_{p^\alpha}(m,n,f) = \Xi_1(m,n,f) + \Xi_p(m,n,f) + \Xi_{p^2}(m,n,f).$$

The other summands vanish because by our assumption $(m,n,N) = 1$, since if $p \mid m$ then $p \nmid n$ because $p \mid N$. So each p divides either m or n but never both. Then, we have $\xi_{p^\beta}(1) = 0$ for $\beta \geqslant 3$. In fact, even more terms vanish. We have

$$\begin{aligned} V_{p^\alpha}(m,n,f) = &\overline{\lambda_f(m)}\lambda_f(n)(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2) \\ &+ \delta_{p|m}\overline{\lambda_f(m/p)}\lambda_f(n)p^{\frac12}(\overline{\xi_p(p)}\xi_p(1)+\overline{\xi_{p^2}(p)}\xi_{p^2}(1)) \\ &+ \delta_{p|n}\overline{\lambda_f(m)}\lambda_f(n/p)p^{\frac12}(\overline{\xi_p(1)}\xi_p(p)+\overline{\xi_{p^2}(1)}\xi_{p^2}(p)) \\ &+ \delta_{p^2|m}\overline{\lambda_f(m/p^2)}\lambda_f(n)p\overline{\xi_{p^2}(p^2)}\xi_{p^2}(1) + \delta_{p^2|n}\overline{\lambda_f(m)}\lambda_f(n/p^2)p\overline{\xi_{p^2}(1)}\xi_{p^2}(p^2). \end{aligned}$$

Inserting the formulas for ξ, we complete the proof. The formula for the $\alpha = 1$ case is even simpler as we can drop the p^2 terms. □

Recall we write $ML = N$ and $f \in H^\star_\kappa(M,\epsilon)$.

Lemma 7.3. *If* $(m,N) = 1$ *and* $(n,N) = 1$ *then we have*

$$(\overline{\lambda_f(m)}\lambda_f(n))^{1-\omega(L)} \prod_{p^\alpha \| L} V_{p^\alpha}(m,n,f) = \overline{\lambda_f(m)}\lambda_f(n) \prod_{p\|L}(1+|\xi_p(1)|^2) \prod_{p^2|L}(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2).$$

Proof. Note that the conditions $(m,N) = 1$ and $(n,N) = 1$ imply that $p \nmid m$ and $p \nmid n$. So the formula above follows immediately from the formulas in Lemma 7.2. □

One has that

$$\|f\|_N^2 = \frac{\psi(N)}{\psi(M)}\|f\|_M^2$$

since $f \in H^\star_\kappa(M,\epsilon)$. Thus

$$\Delta_{\kappa,N,\epsilon}(m,n) = c_\kappa \sum_{LM=N} \frac{\psi(M)}{\psi(N)} \sum_{f\in H^\star_\kappa(M,\epsilon)} \frac{1}{\|f\|_M^2}\overline{\lambda_f(m)}\lambda_f(n) \prod_{p\|L}(1+|\xi_p(1)|^2) \prod_{p^2|L}(1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2).$$

Next we insert the definitions of the ξ functions. Let

$$r_f(p) = 1 - \frac{|\lambda_f(p)|^2}{p\big(1+\frac{\epsilon_{0,M}(p)}{p}\big)^2},$$

so

$$r_f(p)^{-1} = 1 + \frac{|\lambda_f(p)|^2}{p\big(1+\frac{\epsilon_{0,M}(p)}{p}\big)^2} + \left(\frac{|\lambda_f(p)|^2}{p\big(1+\frac{\epsilon_{0,M}(p)}{p}\big)^2}\right)^2 + \cdots,$$

where $\epsilon_{0,M}$ denotes the trivial character modulo M. Observe that

$$1+|\xi_p(1)|^2 = r_f(p)^{-1}$$

and

$$1+|\xi_p(1)|^2+|\xi_{p^2}(1)|^2 = r_f(p)^{-1}\left(1-\frac{\epsilon_{0,M}(p)}{p^2}\right)^{-1}.$$

Then we get

$$\Delta_{\kappa,N,\epsilon}(m,n) = c_\kappa \sum_{LM=N} \frac{\psi(M)}{\psi(N)} \prod_{p^2|L}\left(1-\frac{\epsilon_{0,M}(p)}{p^2}\right)^{-1} \sum_{f\in H^\star_\kappa(M,\epsilon)} \frac{\overline{\lambda_f(m)}\lambda_f(n)}{\|f\|_M^2} \prod_{p|L}\frac{1}{r_f(p)}. \tag{7-2}$$

Next we need a formula for $r_f(p)^{-1}$. Recall from (2-1) that at a prime $p \nmid M$ the local adjoint square L function is given by

$$L_p(1,\mathrm{Ad}^2 f) = \frac{1}{1-p^{-2}}\sum_{\alpha\geqslant 0}\frac{\bar\epsilon(p^\alpha)\lambda_f(p^{2\alpha})}{p^\alpha} = \frac{1}{\big(1-\frac{\alpha(p)/\beta(p)}{p}\big)\big(1-\frac{1}{p}\big)\big(1-\frac{\beta(p)/\alpha(p)}{p}\big)}$$

so that

$$\sum_{\alpha\geqslant 0}\frac{\bar\epsilon(p^\alpha)\lambda_f(p^{2\alpha})}{p^\alpha}=\frac{1+\frac{1}{p}}{\left(1-\frac{\alpha(p)/\beta(p)}{p}\right)\left(1-\frac{\beta(p)/\alpha(p)}{p}\right)}=\frac{1+\frac{1}{p}}{\left(1+\frac{1}{p}\right)^2-\frac{|\lambda_f(p)|^2}{p}}=\frac{1}{\left(1+\frac{1}{p}\right)r_f(p)},$$

where the second equals sign follows from the formulas

$$|\lambda_f(p)|^2=\bar\epsilon(p)\lambda_f(p)^2,\quad \lambda_f(p)=\alpha(p)+\beta(p),\quad \alpha(p)\beta(p)=\epsilon(p),$$

which are valid when $p\nmid M$. We can summarize the above calculation and Lemma 2.2 as

$$r_f(p)^{-1}=\begin{cases}\left(1+\frac{1}{p}\right)\sum_{\alpha\geqslant 0}\frac{\bar\epsilon(p^\alpha)\lambda_f(p^{2\alpha})}{p^\alpha} & \text{if } p\nmid M,\\ \left(1-\frac{a_{M,\epsilon}(p)}{p}\right)^{-1} & \text{if } p\mid M.\end{cases}$$

Let

$$\Delta^\star_{\kappa,N,\epsilon}(m,n)=c_\kappa\sum_{f\in H^\star_\kappa(N,\epsilon)}\frac{\overline{\lambda_f(m)}\lambda_f(n)}{\|f\|_N^2}.$$

Recall the definition of $R(M,L,\epsilon)$ from the statement of Theorem 3.1, which we rearrange to

$$R(M,L,\epsilon)=\frac{\psi(M)}{\psi(ML)}\prod_{\substack{p^2\mid L\\ p\nmid M}}\left(1-\frac{1}{p^2}\right)^{-1}\prod_{\substack{p\mid L\\ p\nmid M}}\left(1+\frac{1}{p}\right)\prod_{p\mid(M,L)}\left(1-\frac{a_{M,\epsilon}(p)}{p}\right)^{-1}.$$

We have then that

$$\Delta_{\kappa,N,\epsilon}(m,n)=\sum_{LM=N}R(M,L,\epsilon)\sum_{\substack{\ell\mid L^\infty\\ (\ell,M)=1}}\frac{\bar\epsilon(\ell)}{\ell}\Delta^\star_{\kappa,M,\epsilon}(m,n\ell^2).\tag{7-3}$$

This is analogous to the first half of [Barrett et al. 2017, Proposition 4.1]. Now we would like to invert this formula, and we prepare for this with two lemmas.

Lemma 7.4. *Let $\alpha,\beta\geqslant 0$ and $0\leqslant\gamma\leqslant\beta$ and $\boldsymbol{c}_p(\epsilon)\leqslant\beta-1$. Then*

$$R(p^\beta,p^\alpha,\epsilon)R(p^\gamma,p^{\beta-\gamma},\epsilon)=R(p^\gamma,p^{\alpha+\beta-\gamma},\epsilon).\tag{7-4}$$

Proof. We check cases.

Case $\alpha\geqslant 0$ and $\beta=\gamma$. Note that $R(p^\gamma,1,\epsilon)=1$ for any $\gamma\geqslant 0$.

Case $\alpha=0$. Note that $R(p^\beta,1,\epsilon)=1$ for any $\beta\geqslant 0$.

Case $\alpha\geqslant 1$, $\beta=1$ and $\gamma=0$. We have by hypothesis $\boldsymbol{c}_p(\epsilon)=0$, so

$$R(p,p^\alpha)R(1,p)=\frac{\psi(p)}{\psi(p^{\alpha+1})}\left(1-\frac{1}{p^2}\right)^{-1}\frac{1}{\psi(p)}\left(1+\frac{1}{p}\right).$$

On the other hand, we also have

$$R(1, p^{\alpha+1}) = \frac{1}{\psi(p^{\alpha+1})}\left(1-\frac{1}{p^2}\right)^{-1}\left(1+\frac{1}{p}\right).$$

Case $\alpha \geqslant 1,\ \beta \geqslant 2$ and $\gamma = 0$. We have $p \mid (p^\beta, p^\alpha)$ and $a_{p^\beta,\epsilon}(p) = 0$, so $R(p^\beta, p^\alpha, \epsilon) = p^{-\alpha}$ and

$$R(1, p^\beta, \epsilon) = \frac{1}{\psi(p^\beta)}\left(1-\frac{1}{p^2}\right)^{-1}\left(1+\frac{1}{p}\right),$$

$$R(1, p^{\alpha+\beta}, \epsilon) = \frac{1}{\psi(p^{\alpha+\beta})}\left(1-\frac{1}{p^2}\right)^{-1}\left(1+\frac{1}{p}\right).$$

Generic case $\alpha \geqslant 1,\ \beta \geqslant 2,\ 1 \leqslant \gamma \leqslant \beta - 1$, and $\boldsymbol{c}_p(\epsilon) \leqslant \beta - 1$. We have

$$R(p^\beta, p^\alpha, \epsilon) = \frac{\psi(p^\beta)}{\psi(p^{\alpha+\beta})},$$

$$R(p^\gamma, p^{\beta-\alpha}, \epsilon) = \frac{\psi(p^\gamma)}{\psi(p^\beta)}\left(1-\frac{a_{p^\gamma,\epsilon}(p)}{p}\right)^{-1},$$

$$R(p^\gamma, p^{\beta+\alpha-\gamma}, \epsilon) = \frac{\psi(p^\gamma)}{\psi(p^{\alpha+\beta})}\left(1-\frac{a_{p^\gamma,\epsilon}(p)}{p}\right)^{-1}.$$

The above cover all the cases in the lemma. □

Lemma 7.5. *Let $N \in \mathbb{N}$, $N = LM$, and $M = WQ$. Then*

$$R(M, L, \epsilon)R(W, Q, \epsilon)\delta_{\boldsymbol{c}(\epsilon)\mid W} = R(W, LQ, \epsilon)\delta_{\boldsymbol{c}(\epsilon)\mid W}.$$

Proof. Both sides of the desired formula are multiplicative. Let $\alpha = v_p(L)$, $\beta = v_p(M)$, and $\gamma = v_p(W)$. It then suffices to check that

$$R(p^\beta, p^\alpha, \epsilon)R(p^\gamma, p^{\beta-\gamma}, \epsilon)\delta_{\gamma \geqslant \boldsymbol{c}_p(\epsilon)} = R(p^\gamma, p^{\alpha+\beta-\gamma}, \epsilon)\delta_{\gamma \geqslant \boldsymbol{c}_p(\epsilon)}. \tag{7-5}$$

If $\boldsymbol{c}_p(\epsilon) \leqslant \beta - 1$ then (7-5) is true by Lemma 7.4. So, suppose not. Then $\beta \leqslant \boldsymbol{c}_p(\epsilon) \leqslant \gamma$, but $W \mid M$ so $\gamma \leqslant \beta$ and so $\beta = \gamma$. In the case $\beta = \gamma$, (7-5) is true because $R(p^\beta, 1, \epsilon) = 1$. □

We are now prepared to invert (7-3) using Lemma 7.5. We calculate

$$\sum_{LM=N} \mu(L)R(M, L, \epsilon) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell}\Delta_{\kappa,M,\epsilon}(m, n\ell^2)$$

$$= \sum_{LM=N} \mu(L)R(M, L, \epsilon) \sum_{\substack{\ell \mid L^\infty \\ (\ell, M)=1}} \frac{\bar{\epsilon}(\ell)}{\ell} \sum_{QW=M} R(W, Q, \epsilon) \sum_{\substack{q \mid Q^\infty \\ (q, W)=1}} \frac{\bar{\epsilon}(q)}{q}\Delta^\star_{\kappa,W,\epsilon}(m, n\ell^2 q^2)$$

$$= \sum_{LM=N} \mu(L) \sum_{QW=M} R(M, L, \epsilon)R(W, Q, \epsilon) \sum_{\substack{b \mid (LQ)^\infty \\ (b, W)=1}} \frac{\bar{\epsilon}(b)}{b}\Delta^\star_{\kappa,W,\epsilon}(m, nb^2)$$

$$= \sum_{WX=N} R(W,X,\epsilon) \sum_{\substack{b\mid X^\infty \\ (b,W)=1}} \frac{\bar{\epsilon}(b)}{b} \Delta^\star_{\kappa,W,\epsilon}(m,nb^2) \sum_{LQ=X} \mu(L)$$

$$= R(N,1,\epsilon)\Delta^\star_{\kappa,N,\epsilon}(m,n)$$

$$= \Delta^\star_{\kappa,N,\epsilon}(m,n),$$

where the first equality follows by (7-3), the third by Lemma 7.5, and the fourth by Möbius inversion. □

Acknowledgements

I would like to thank Nathan Kaplan for careful reading and pointing out the Tsfasman–Vlăduţ–Zink theorem to me, Corentin Perret-Gentil for some helpful discussions, and the anonymous referee for a thorough and detailed report on the first version of this paper.

References

[Atkin and Lehner 1970] A. O. L. Atkin and J. Lehner, "Hecke operators on $\Gamma_0(m)$", *Math. Ann.* **185** (1970), 134–160. MR Zbl

[Banks 1997] W. D. Banks, "Twisted symmetric-square L-functions and the nonexistence of Siegel zeros on GL(3)", *Duke Math. J.* **87**:2 (1997), 343–353. MR Zbl

[Barrett et al. 2017] O. Barrett, P. Burkhardt, J. DeWitt, R. Dorward, and S. J. Miller, "One-level density for holomorphic cusp forms of arbitrary level", *Res. Number Theory* **3** (2017), art. id. 25. MR Zbl

[Blomer and Milićević 2015] V. Blomer and D. Milićević, "The second moment of twisted modular L-functions", *Geom. Funct. Anal.* **25**:2 (2015), 453–516. MR Zbl

[Blomer et al. 2017] V. Blomer, J. Buttcane, and P. Maga, "Applications of the Kuznetsov formula on GL(3) II: the level aspect", *Math. Ann.* **369**:1-2 (2017), 723–759. MR Zbl

[Conrey et al. 1997] J. B. Conrey, W. Duke, and D. W. Farmer, "The distribution of the eigenvalues of Hecke operators", *Acta Arith.* **78**:4 (1997), 405–409. MR Zbl

[Davenport 2000] H. Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Mathematics **74**, Springer, 2000. MR Zbl

[Diamond and Im 1995] F. Diamond and J. Im, "Modular forms and modular curves", pp. 39–133 in *Seminar on Fermat's Last Theorem* (Toronto, 1993–1994), CMS Conf. Proc. **17**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Diamond and Shurman 2005] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Mathematics **228**, Springer, 2005. MR Zbl

[Duke and Kowalski 2000] W. Duke and E. Kowalski, "A problem of Linnik for elliptic curves and mean-value estimates for automorphic representations", *Invent. Math.* **139**:1 (2000), 1–39. MR Zbl

[Gelbart and Jacquet 1978] S. Gelbart and H. Jacquet, "A relation between automorphic representations of GL(2) and GL(3)", *Ann. Sci. École Norm. Sup.* (4) **11**:4 (1978), 471–542. MR Zbl

[Goldfeld et al. 1994] D. Goldfeld, J. Hoffstein, and D. Lieman, "Appendix: an effective zero-free region", (1994). Appendix to J. Hoffstein and P. Lockhart, "Coefficients of Maass forms and the Siegel zero", *Ann. of Math.* (2) **140**:1 (1994), 161–181. MR Zbl

[Heath-Brown 1981] D. R. Heath-Brown, "The fourth power mean of Dirichlet's L-functions", *Analysis* **1**:1 (1981), 25–32. MR Zbl

[Humphries 2018] P. Humphries, "Density theorems for exceptional eigenvalues for congruence subgroups", *Algebra Number Theory* **12**:7 (2018), 1581–1610. MR Zbl

[Igusa 1959] J. Igusa, "Kroneckerian model of fields of elliptic modular functions", *Amer. J. Math.* **81** (1959), 561–577. MR Zbl

[Iwaniec 1997] H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics **17**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications **53**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Iwaniec et al. 2000] H. Iwaniec, W. Luo, and P. Sarnak, "Low lying zeros of families of L-functions", *Inst. Hautes Études Sci. Publ. Math.* **91** (2000), 55–131. MR Zbl

[Kaplan and Petrow 2017] N. Kaplan and I. Petrow, "Elliptic curves over a finite field and the trace formula", *Proc. Lond. Math. Soc.* (3) **115**:6 (2017), 1317–1372. MR Zbl

[Knightly and Li 2013] A. Knightly and C. Li, *Kuznetsov's trace formula and the Hecke eigenvalues of Maass forms*, Mem. Amer. Math. Soc. **1055**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Kowalski and Michel 1999] E. Kowalski and P. Michel, "The analytic rank of $J_0(q)$ and zeros of automorphic L-functions", *Duke Math. J.* **100**:3 (1999), 503–542. MR Zbl

[Li 1975] W. C. W. Li, "Newforms and functional equations", *Math. Ann.* **212** (1975), 285–315. MR Zbl

[Li 1979] W. C. W. Li, "L-series of Rankin type and their functional equations", *Math. Ann.* **244**:2 (1979), 135–166. MR Zbl

[Milne 2017] J. S. Milne, "Modular functions and modular forms", course notes, 2017, available at https://www.jmilne.org/math/CourseNotes/mf.html. version 1.31.

[Moreno 1991] C. Moreno, *Algebraic curves over finite fields*, Cambridge Tracts in Mathematics **97**, Cambridge University Press, 1991. MR Zbl

[Murty and Sinha 2009] M. R. Murty and K. Sinha, "Effective equidistribution of eigenvalues of Hecke operators", *J. Number Theory* **129**:3 (2009), 681–714. MR Zbl

[Nelson 2017] P. D. Nelson, "Analytic isolation of newforms of given level", *Arch. Math.* (*Basel*) **108**:6 (2017), 555–568. MR Zbl

[Ng 2012] M. Ng, *The basis for space of cusp forms and Petersson trace formula*, master's thesis, University of Hong Kong, 2012.

[Ogg 1969] A. P. Ogg, "On the eigenvalues of Hecke operators", *Math. Ann.* **179** (1969), 101–108. MR Zbl

[Petrow and Young 2018] I. Petrow and M. P. Young, "A generalized cubic moment and the Petersson formula for newforms", *Math. Ann.* (online publication August 2018).

[Ram Murty 1995] M. Ram Murty, "The analytic rank of $J_0(N)(\mathbb{Q})$", pp. 263–277 in *Number theory* (Halifax, 1994), edited by K. Dilcher, CMS Conf. Proc. **15**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Ross 1992] S. L. Ross, II, "A simplified trace formula for Hecke operators for $\Gamma_0(N)$", *Trans. Amer. Math. Soc.* **331**:1 (1992), 425–447. MR Zbl

[Rouymi 2011] D. Rouymi, "Formules de trace et non-annulation de fonctions L automorphes au niveau $\mathfrak{p}^{\nu}$", *Acta Arith.* **147**:1 (2011), 1–32. MR Zbl

[Schulze-Pillot and Yenirce 2018] R. Schulze-Pillot and A. Yenirce, "Petersson products of bases of spaces of cusp forms and estimates for Fourier coefficients", *Int. J. Number Theory* **14**:8 (2018), 2277–2290. MR Zbl

[Serre 1997] J.-P. Serre, "Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p", *J. Amer. Math. Soc.* **10**:1 (1997), 75–102. MR Zbl

[Tsfasman et al. 1982] M. A. Tsfasman, S. G. Vlăduţ, and T. Zink, "Modular curves, Shimura curves, and Goppa codes, better than Varshamov–Gilbert bound", *Math. Nachr.* **109** (1982), 21–28. MR Zbl

[Venkatesh 2006] A. Venkatesh, "Large sieve inequalities for GL(n)-forms in the conductor aspect", *Adv. Math.* **200**:2 (2006), 336–356. MR Zbl

[Young 2018] M. P. Young, "Explicit calculations with Eisenstein series", *J. Number Theory* (online publication December 2018).

Communicated by Roger Heath-Brown
Received 2018-05-06 Revised 2018-07-22 Accepted 2018-08-23

ian.petrow@math.ethz.ch Departement Mathematik, ETH Zürich, Zürich, Switzerland

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2018 is US $340/year for the electronic version, and $535/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 12 No. 10 2018