Max-Planck-Institut für Mathematik
Bonn

Asymptotic trace formula for the Hecke operators

by

Junehyuk Jung
Naser T. Sardari
(With an appendix by Simon Marshall)

Max-Planck-Institut für Mathematik
Preprint Series 2019 (54)

Asymptotic trace formula for the Hecke operators

Junehyuk Jung
Naser T. Sardari
(With an appendix by Simon Marshall)

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

Department of Mathematics
Texas A&M University
College Station, TX 77845
USA

Department of Mathematics
University of Wisconsin-Madison
Madison, WI 53706
USA

MPIM 19-54

Asymptotic trace formula for the Hecke operators

Junehyuk Jung Naser T. Sardari
(With an appendix by Simon Marshall)

Abstract

Given integers m, n and k, we give an explicit formula with an optimal error term (with square root cancellation) for the Petersson trace formula involving the m-th and n-th Fourier coefficients of an orthonormal basis of $S_k(N)^*$ (the weight k newforms with fixed square-free level N) provided that $|4\pi\sqrt{mn} - k| = o(k^{\frac{1}{3}})$. Moreover, we establish an explicit formula with a power saving error term for the trace of the Hecke operator $\mathcal{T}_n^*$ on $S_k(N)^*$ averaged over k in a short interval. By bounding the second moment of the trace of $\mathcal{T}_n$ over a larger interval, we show that the trace of $\mathcal{T}_n$ is unusually large in the range $|4\pi\sqrt{n} - k| = o(n^{\frac{1}{6}})$. As an application, for any fixed prime p with $\gcd(p, N) = 1$, we show that there exists a sequence $\{k_n\}$ of weights such that the error term of Weyl's law for $\mathcal{T}_p$ is unusually large and violates the prediction of arithmetic quantum chaos. In particular, this generalizes the result of Gamburd, Jakobson and Sarnak [GJS99, Theorem 1.4] with an improved exponent.

Acknowledgments. J.J. thanks S.M. and Department of Mathematics of UW-Madison for invitation and support. J.J. also thanks Sug Woo Shin, Peter Jaehyun Cho, and Matthew Young for many helpful comments. S.M. was supported by NSF grant DMS-1501230. N.T.S. was supported by NSF grant DMS-1902185 and is grateful to Max Planck Institute for Mathematics in Bonn for its hospitality and financial support. N.T.S. thanks his Ph.D. advisor Peter Sarnak for several insightful and inspiring conversations regarding the error term of the Weyl law while he was a graduate student at Princeton University.

1 Introduction

1.1 Background

We begin by explaining Weyl's law, and bounds on its error term, in some arithmetic examples which reduce to deep problems in Number Theory. Let $X \subset \mathbb{R}^d$ be a bounded

J. Jung: Department of Mathematics, Texas A&M University, College Station, TX 77845; e-mail: junehyuk@math.tamu.edu

N. T.Sardari: Max Planck Institut for Mathematics, Vivatsgasse 7 53111 Bonn Germany; e-mail: ntalebiz@mpim-bonn.mpg.de

S. Marshall: Department of Mathematics, UW-Madison, Madison, WI 53706; e-mail: marshall@math.wisc.edu)

Mathematics Subject Classification (2010): Primary 11F25; Secondary 11F72

domain with smooth boundary. Let T be a positive real number, and let $N(T)$ be the number of Dirichlet Laplacian eigenvalues of X less than T^2 (counted with multiplicity). It was conjectured independently by Sommerfeld and Lorentz, based on the work of Rayleigh on the theory of sound, and proved by Weyl [Wey11] shortly after, that

$$N(T) = c_d \mathrm{vol}(X) T^d (1 + o(1)) \quad \text{as } \lambda \to \infty,$$

where c_d is a constant depending only on d and $\mathrm{vol}(X)$ is the volume of X in $\mathbb{R}^d$. More generally, let (M^d, g) be a compact smooth Riemannian manifold of dimension d with Laplace-Beltrami operator Δ. Then Hörmander [Hör68] proved that

$$N(T) = c_d \mathrm{vol}(M) T^d + R_M(T),$$

where $R_M(T) = O(T^{d-1})$. In fact, this general estimate is sharp for the round sphere $M = S^d$. However, given a manifold M the question of finding the optimal bound for the error term $R_M(T)$ is a very difficult problem.

We now restrict to the case $d = 2$, and discuss the relation between the size of $R_M(T)$, and the geodesic flow on the unit cotangent bundle S^*M, predicted by the correspondence principle. The two extreme behaviors that the geodesic flow can have are being chaotic or completely integrable, and in these two cases the correspondence principle predicts the distribution of eigenvalues to be modeled by a large random matrix, and a Poisson process, respectively [Ber85, Ber86].

In particular, we expect that for a generic 2 dimensional flat torus, or a compact arithmetic hyperbolic surface [Sar95, Figure 1.3 and Section 3][3], the set of eigenvalues inside the universal interval $\left[T^2, \left(T + \frac{1}{L}\right)^2\right]$ where $\log T \ll L = o(T)$ is modeled by Poisson process; see the very interesting work of Rudnick [Rud05] and Sarnak's letter [Sar02] explaining the critical window $\log(T) \ll L = o(T)$ using Kuznetsov's trace formula. This suggests that these surfaces satisfy $R_M(T) = O(T^{\frac{1}{2}+\varepsilon})$. In fact, Petridis and Toth proved that the average order of the error term in Weyl's law for a random torus chosen in a compact part of the moduli space of two dimensional tori is $R(T) = O_\varepsilon(T^{\frac{1}{2}+\varepsilon})$; see [PT02]. Moreover, for compact arithmetic surfaces it was proved by Selberg [Hej76, p.315] that $R(T) = \Omega(T^{\frac{1}{2}} / \log T)$.

For the rational torus $\mathbb{T} = \mathbb{R}^2/\mathbb{Z}^2$, bounding $R_{\mathbb{T}}(T)$ is equivalent to the classical Gauss circle problem. It was conjectured by Hardy that $R_{\mathbb{T}}(T) = O_\varepsilon(T^{\frac{1}{2}+\varepsilon})$, and it is known by Hardy and Landau [HL24] that $R_{\mathbb{T}}(T) = \Omega(T^{\frac{1}{2}} \log^{\frac{1}{4}} T)$. Note that the eigenvalue distribution here is known not to be Poisson [Sar97].

As mentioned above, for generic compact hyperbolic surfaces, we expect the set of eigenvalues inside the interval $[T^2, (T+1)^2]$ to follow the eigenvalue distribution of a

[3]The geodesic flow in this case is chaotic, but Sarnak explains that one expects to see Poisson behavior due to the high multiplicity of the geodesic length spectrum.

large symmetric matrix, which has a rigid structure. As a result, it is conjectured that these surfaces satisfy $R_M(T) = O(T^{\varepsilon})$.

Proving an optimal upper bound on $R_M(T)$ is extremely difficult, and we don't have any explicit example of M other than the sphere where the optimal bound is known! The best known upper bound for hyperbolic manifolds is $R_M(T) = O(T^{d-1}/\log(T))$, due to Bérard [Bér77]. As pointed out by Sarnak [Sar02, Page 2], even improving the constant and showing that $R(T) = o(T/\log(T))$ for the cuspidal spectrum of $SL_2(\mathbb{Z})\backslash\mathbb{H}$ (after removing the contribution of the Eisenstein series) is very difficult; see Remark 1.2.1.

In this paper, we give bounds on the error term of Weyl's law for the Hecke eigenvalues of the family of classical holomorphic modular forms with a fixed level. We briefly describe this family, its Weyl's law, and known bounds and predictions on its error term. Next, we explain our results and compare them with the previous results and predictions.

Let $\Gamma_0(N) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z}, N|c \right\}$ be the Hecke congruence subgroup of level N. Let $S_k(N)$ be the space of even weight $k \in \mathbb{Z}$ modular forms of level N. It is the space of the holomorphic functions f such that

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \tag{1.1}$$

for every $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(N)$, and f converges to zero as it approaches each cusp (we have finitely many cusps for $\Gamma_0(N)$ that are associated to the orbits of $\Gamma_0(N)$ acting by Möbius transformations on $\mathbb{P}^1(\mathbb{Q})$); see [Sar90]. It is well-known that $S_k(N)$ is a finite dimensional vector space over $\mathbb{C}$, and is equipped with the Petersson inner product $\langle f, g\rangle := \int_{\Gamma_0(N)\backslash\mathbb{H}} f(z)\bar{g}(z)y^k dxdy/y^2$ which makes it into a Hilbert space. Assume that p is a fixed prime number where $p \nmid N$. Then one can define a self-adjoint Hecke operator $\mathcal{T}_p$ on $S_k(N)$:

$$\mathcal{T}_p(f)(z) := p^{-\frac{k-1}{2}} \sum_{n=1}^{\infty} a_{np} e(nz) + p^{\frac{k-1}{2}} \sum_{n=1}^{\infty} a_n e(pnz), \tag{1.2}$$

where $f(z) = \sum_{n=1}^{\infty} a_n e(nz)$ is the Fourier expansion of f at the cusp ∞. In particular, if f is an eigenfunction of $\mathcal{T}_p$ with eigenvalues $\lambda_p(f)$ then $a_p = a_1\lambda_p(f)p^{\frac{k-1}{2}}$. By Deligne's result [Del74] the Ramanujan-Petersson conjecture holds for f and we have $|\lambda_p(f)| \leq 2$. Under Langlands' philosophy, the Hecke operator $\mathcal{T}_p$ is the p-adic analogue of the Laplace operator (the eigenvalues of $\mathcal{T}_p$ determine the Satake parameters of the associated local representation π_p of $GL_2(\mathbb{Q}_p)$ just as the Laplace eigenvalue of the Maass form determines the associated local representation π_∞ of $GL_2(\mathbb{R})$). Let $B_{k,N}$ be a basis for the eigenfunctions of $\mathcal{T}_p$ acting on $S_k(N)$. Let $\mu_{k,N} := \frac{1}{\dim(S_k(N))} \sum_{f \in B_{k,N}} \delta_{\lambda_p(f)}$ be the spectral probability measure associated to $\mathcal{T}_p$ acting on $S_k(N)$ which is supported in $[-2, 2]$.

Using the Eichler-Selberg trace formula, Serre [Ser97] proved that $\mu_{k,N}$ converges weakly to μ_p as $k+N\to\infty$, where μ_p is the Plancherel measure of $GL_2(\mathbb{Q}_p)$ given by

$$\mu_p(x) := \frac{p+1}{\pi}\frac{(1-\frac{x^2}{4})^{\frac{1}{2}}}{(p^{\frac{1}{2}}+p^{-\frac{1}{2}})^2-x^2}dx.$$

Moreover, let $\nu_{k,N} := \frac{1}{\dim(S_k(N))}\sum_{f\in B_{k,N}}^h \delta_{\lambda_p(f)}$, where the superscript h means the expression in the sum is multiplied by the harmonic weights $\frac{\Gamma(k-1)}{(4\pi)^{k-1}\langle f,f\rangle}$. It follows from the Petersson trace formula (see Section 2) that $\nu_{k,N}$ converges weakly to the semi-circle law

$$\mu_\infty(x) := \frac{1}{\pi}\sqrt{1-\frac{x^2}{4}}dx,$$

as $k+N\to\infty$. These are the analogues of Weyl's law for this family of classical modular forms. In fact, Weyl's law is formulated and expected to hold in great generality for other families of automorphic forms; see [SST16, Conjecture 1]. In [GJS99], Gamburd, Jakobson and Sarnak studied the spectrum of the elements in the group ring of $SU(2)$. In particular, they proved the analogue of Selberg's lower bound and Bérard's upper bound on the error term of Weyl's law in that context. By the Jacquet-Langlands correspondence one can interpret their results in our context as follows. Given two probability measures μ_1 and μ_2 on $\mathbb{R}$, we denote the discrepancy between them by $D(\mu_1,\mu_2)$, where

$$D(\mu_1,\mu_2) := \sup\{|\mu_1(I)-\mu_2(I)| : I=[a,b]\subset\mathbb{R}\}.$$

Then [GJS99, Theorem 1.3] is equivalent to $D(\mu_{k,2},\mu_p) = O(1/\log(k))$, which is the analogue of Bérard's upper bound. Moreover, [GJS99, Theorem 1.4] is equivalent to the existence of a sequence of integers $k_n\to\infty$ such that

$$D(\mu_{k_n,2},\mu_p) \gg \frac{1}{k_n^{\frac{1}{2}}\log^2 k_n}, \tag{1.3}$$

which is the analogue of Selberg's lower bound. This is a corollary of their lower bound on the variance of the trace of the Hecke operators by varying the weight k; see Theorem 1.5.

1.2 Main results

1.2.1 Large discrepancy for $\mu_{k,N}^*$

Let $S_k(N)^*$ be the space of newforms of weight k and fixed level N. Let $\mathcal{T}_p^*$ be the restriction of $\mathcal{T}_p$ from $S_k(N)$ to its subspace $S_k(N)^*$. We denote by $\mu_{k,N}^*$ and $\nu_{k,N}^*$ the corresponding measures associated to $\mathcal{T}_p^*$. The main theorem of this paper is a generalization of (1.3) to $\mu_{k,N}^*$ with any squarefree level N and an improved exponent of k in the lower bound:

Main Theorem 1.1. *Let $N>1$ be a fixed square-free integer. Then there exists an infinite sequence of weights $\{k_n\}$ with $k_n\to\infty$ such that*

$$D(\mu^*_{k_n,N},\mu_p)\gg\frac{1}{k_n^{\frac{1}{3}}\log^2 k_n}.$$

Remark. As mentioned in the introduction the best known upper bound for $D(\mu^*_{k,N},\mu_p)$ is

$$D(\mu^*_{k,N},\mu_p)=O\left(\log(k)^{-1}\right), \tag{1.4}$$

see [MS09]. The standard method for giving an upper bound on the discrepancy of a sequence of points is the Erdös-Turán inequality [ET48]. Even to improve the implied constant in (1.4) using the Erdös-Turán inequality, one needs to obtain a nontrivial upper bound on the trace of the Hecke operator $\mathcal{T}_n$ for $n\gg k^A$, where $A>0$ is an arbitrarily large constant. But the error term in the Selberg trace formula is large in this range and makes the problem very difficult by this approach.

Theorem 1.1 follows from an explicit asymptotic formula for the weighted average of the trace of the Hecke operator in a short interval. More precisely, let ψ be a positive smooth function supported in $[-1,1]$, and satisfying $\int_{-1}^{1}\psi(t)dt=1$. Let $\mathrm{Tr}\ \mathcal{T}_n(N,k)^*$ be the trace of the Hecke operator $\mathcal{T}_n^*$ on $S_k(N)^*$. Let K be a number satisfying $K=4\pi\sqrt{n}+o(n^{\frac{1}{6}})$.

Theorem 1.2. *Let $\frac{1}{6}<\delta<\frac{1}{3}$ be any real number. We have*

$$\frac{1}{K^\delta}\sum_{k>0,k\in2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)\mathrm{Tr}\ \mathcal{T}_n(N,k)^*=J_K(4\pi\sqrt{n})\frac{\mu(N)K}{12}\zeta^{-1}(2)\frac{\sigma(n)}{n}(1+O(K^{-\varepsilon})),$$

where J_K is the J-Bessel function, μ is the Möbius function, σ is the sum of the divisors of n and ε is some small fixed constant depending on δ. Moreover, the implicit constant in O depends only on the fixed variables N and ε.

Remark. By the asymptotic of the J-Bessel function in the transition range, we have $|J_K(4\pi\sqrt{n})|\gg K^{-\frac{1}{3}}$; see [DLMF, 10.19.8]. Hence, we have $|\mathrm{Tr}\ \mathcal{T}_n(N,k)^*|\gg k^{\frac{2}{3}}$ for some $k\in[K-K^\delta,K+K^\delta]$. This lower bound violates the naive expected square root cancelation for the eigenvalues of the Hecke operator $\mathcal{T}_n(N,l)^*$.

We give a brief description of the proof. We give the proof of the above theorem in Section 3. The proof is based on the Petersson trace formula and the proof of Theorem 1.7 that we give in Section 2. The main term of the above formula comes from the J-Bessel function in the transition range. Next, we simplify the error term by using bounds on the J-Bessel function outside the transition range. For the remaining error terms, we average over weights and apply the Poisson summation formula and obtain a sum of Kloosterman sums twisted by oscillatory integrals. The Theorem follows from Weil's bound for Kloosterman sums, and bounds on the oscillatory integrals that we prove by the stationary phase method in Section 3.1. There are some similarities between our method and the circle method, specially the version developed by Heath-Brown [HB96].

1.2.2 Variance of the trace

If we consider the variance of the trace of the Hecke operator over $k \sim \sqrt{n}$, the largeness of the trace in Theorem 1.2 is no longer present. To be precise, we have the following results:

Theorem 1.3. *Let $N > 1$ be a squarefree integer. For any n, we have*

$$\sum_{\substack{k\in 2\mathbb{Z} \\ 3\pi\sqrt{n}<k<5\pi\sqrt{n}}} \left| \sum_{f\in B^*_{k,N}} \lambda_n(f) - \frac{k-1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}} \right|^2 \ll_N n(\log n)^2(\log\log n)^4.$$

In particular, almost all k in the range $[3\pi\sqrt{n}, 5\pi\sqrt{n}]$ satisfy

$$\sum_{f\in B^*_{k,N}} \lambda_n(f) = O_\epsilon\left(k^{\frac{1}{2}+\epsilon}\right).$$

We also prove a lower bound for the variance of the trace of the Hecke operator:

Theorem 1.4. *Let $N > 1$ be a squarefree integer and let $n = p^m$ where p is an odd prime. There exists a sufficiently large fixed constant $A > 0$ such that for any $K > A\sqrt{n}$, we have*

$$\frac{1}{\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right)} \sum_{k>0, k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right) \left| \sum_{f\in B^*_{k,N}} \lambda_n(f) - \frac{k-1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}} \right|^2 \gg_N n^{\frac{1}{2}}, \quad (1.5)$$

where $\delta_{\sqrt{n}} = 1$ if n is a square, and 0 otherwise.

This immediately implies the following weaker version of Theorem 1.1.

Corollary 1.5. *Let $N > 1$ be a fixed square-free integer and let p be an odd prime. Then we have*

$$D(\mu^*_{k,N}, \mu_p) = \Omega\left(\frac{1}{k^{\frac{1}{2}}\log^2 k}\right).$$

Remark. Note that this generalizes [GJS99] to any square-free level $N > 1$.

Theorem 1.3 and 1.4 are consequences of the following asymptotic formula, which we derive from the Eichler–Selberg trace formula for $T \geq \sqrt{n}$ (Lemma 4.6):

$$\sum_{k>0, k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \left| \sum_{f\in B^*_{k,N}} \lambda_n(f) - \frac{k-1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}} \right|^2$$
$$= 2\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \sum_{t^2<4n} |D(t,n)|^2 - \phi\left(\frac{1}{T}\right)\frac{\sigma_1(n)^2}{n} + O\left(n^{\frac{1}{2}+\epsilon}\right). \quad (1.6)$$

Here $D(t,n)$ is a weighted sum of class numbers:

$$D(t,n) = \frac{i}{2\sqrt{4n-t^2}} \sum_f h_w\left(\frac{t^2-4n}{f^2}\right) \tilde{\mu}(t,f,n,N),$$

with weights $|\tilde{\mu}(t,f,n,N)| = O_N(1)$ (for the precise definition, see Lemma 4.2).

The upper bound (Theorem 1.3) then follows by applying a standard upper bound for the class numbers of imaginary quadratic fields.

Note that inputting the sharp lower bound for the class numbers of imaginary quadratic fields,

$$h_w(-d) \gg_\epsilon d^{\frac{1}{2}-\epsilon},$$

to (1.6) is not sufficient to prove the lower bound in Theorem 1.4. Therefore we relate the problem of estimating the sparse sum of sums of class numbers

$$\sum_{t^2<4n} |D(t,n)|^2$$

to the problem of counting integral lattice points on 3-spheres, under certain congruence conditions on the coordinates. This can be done by following the circle method developed by Kloosterman [Klo27], and we have

$$\sum_{t^2<4n} |D(t,n)|^2 \gg_N \sqrt{n},$$

under the assumption that n is odd (Theorem 4.7). Now if $n = p^m$ with a fixed odd prime p, and if $T > A\sqrt{n}$ for some large A, we see that

$$2\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \sum_{t^2<4n} |D(t,n)|^2$$

is larger than $\phi\left(\frac{1}{T}\right)\frac{\sigma_1(n)^2}{n} = O(n)$, from which Theorem 1.4 follows. These steps are carried out in Section 4.

1.2.3 Large discrepancy for the measure with harmonic weights

Next, we give our results on the error term of the Weyl law associated to the measures $\nu_{k,N}^*$ as $k \to \infty$.

Theorem 1.6. *There exists an infinite sequence of weights $\{k_n\}$ with $k_n \to \infty$ such that*

$$D(\nu_{k_n,N}^*, \mu_\infty) \gg \frac{1}{k_n^{\frac{1}{3}} \log^2 k_n}. \tag{1.7}$$

Remark. The above exceptional sequence of weights is very explicit and is given by $k_n = \lfloor 4\pi p^n \rfloor$. Based on arithmetic quantum chaos, numerical evidence [GJS99, Figure 5 and Figure 6], and the random model described in the introduction for the eigenvalues of the Hecke operator, it is expected that

$$D(\mu_{k,N}^*, \mu_p) = O_{\varepsilon,N}\left(k^{\frac{1}{2}+\varepsilon}\right) \text{ and } D(\nu_{k,N}^*, \mu_\infty) = O_{\varepsilon,N}\left(k^{\frac{1}{2}+\varepsilon}\right), \tag{1.8}$$

for a density 1 set of k. In this context, the exponent $\frac{1}{3}$ in Theorem 1.6 (and Theorem 1.1) shows that one can not achieve (1.8) for every weight k.

Theorem 1.6 is an immediate consequence of an explicit asymptotic formula for the Petersson trace formula. More precisely, let

$$\Delta_{k,N}^*(m,n) := \sum\nolimits_{f \in B_{k,N}^*}^{h} \lambda_m(f)\lambda_n(f).$$

Theorem 1.7. *Assume that $|4\pi\sqrt{mn} - k| = O(k^{\frac{1}{3}})$ and $gcd(mn, N) = 1$. Then*

$$\Delta_{k,N}^*(m,n) = \frac{\varphi(N)}{N}\delta(m,n) + J_{k-1}(4\pi\sqrt{mn})\frac{\mu(N)}{N}\prod_{p|N}(1-1/p^2) + O(k^{-\frac{1}{2}}).$$

where $\delta(m,n) = 1$ if $m = n$ and $\delta(m,n) = 0$ otherwise, J_{k-1} is the J-Bessel function and the implicit constant in O depends only on the fixed variables N and ε.

Remark. Since $|4\pi\sqrt{mn} - k| = O(k^{\frac{1}{3}})$, by the asymptotic of the J-Bessel function in the transition range [DLMF, 10.19.8], we have $|J_{k-1}(4\pi\sqrt{mn})| \gg 1/k^{\frac{1}{3}}$. It follows that $J_{k-1}(4\pi\sqrt{mn})\frac{\mu(N)}{N}\prod_{p|N}(1-1/p^2)$ is the main term, and

$$|\Delta_{k,N}^*(m,n) - \delta(m,n)| \gg 1/k^{\frac{1}{3}}.$$

The above lower bound violates the naive expected square root cancelation in the sum of the normalized Fourier coefficients of the newforms in this range. More generally, one can generalize Theorem 1.7 if $|4\pi\sqrt{mn} - \alpha k| = O(k^{\frac{1}{3}})$ for any fixed integer α. In the appendix by Simon Marshall, the existence of this asymptotic trace formula is explained via the geometric side of the Petersson trace formula.

We prove Theorem 1.7 in Section 2 by applying the Petersson trace formula and partitioning the geometric side of this formula into three parts according to the various behavior of the J-Bessel function in different ranges. This partition is explained in the appendix according to the incidence of the associated pairs of horocycles. The main term comes from the J-Bessel function in the transition range where the associated horocycles are tangent to each other, and the error term stays in the ranges where the J-Bessel function decays rapidly.

Theorem 1.2 follows form Theorem 1.7 and averaging over m and k parameters. In fact, we expect that a stronger version of of Theorem 1.2 to be true, namely:

$$\text{Tr } \mathcal{T}_n(N,k)^* = J_k(4\pi\sqrt{n})\frac{\mu(N)k}{12}\zeta^{-1}(2)\frac{\sigma(n)}{n}(1+O(k^{-\varepsilon})),$$

where $k = 4\pi\sqrt{n} + o(n^{\frac{1}{6}})$. However, removing the harmonic weights in Kuznetsov's formula by only averaging over m in our context is equivalent to a very strong unproven bound on L-functions, namely:

Hypothesis 1.8. *Let $n = O(k^2)$ and N be a fixed square free integer. Then*

$$\sum\nolimits^h_{f\in B^*_{k,N}} \lambda_n(f)L(\frac{1}{2}+it, sym^2 f) = O(k^{-\frac{1}{6}-\epsilon}), \tag{1.9}$$

where $t = O(\log(k)^A)$ for some $A > 0$.

We are overcoming this problem by taking average over k on a very short interval. Hence unlike standard averaging techniques, it is only good for estimating lower bound.

1.3 Notations

We let $S_k(N)$ and its subspace $S^*_k(N)$ denote the space of holomorphic cusp forms and the subspace of newforms of weight k on $\Gamma_0(N)\backslash\mathbb{H}$. If $\gcd(n,N)=1$, we let $\mathcal{T}_n = \mathcal{T}_n(N,k)$ be the n-th Hecke operator acting on $S_k(N)$. For a joint eigenfunction $f \in S_k(N)$ of $\mathcal{T}_n$, let $\lambda_n(f)$ be the eigenvalues of $\mathcal{T}_n,\ n \geq 1$. We normalize $\mathcal{T}_n$ so that $|\lambda_n(f)| \leq 2$ is the Ramanujan bound. We use the divisor function parameterized by t: $\sigma_t(n) = \sum_{d|n} d^t$. We write $B_{k,N}$ and $B^*_{k,N}$ for an orthonormal basis of $S_k(N)$ and $S_k(N)^*$ respectively. If $f \in S_k(N)$, we write $\left(\frac{\Gamma(k-1)}{(4n\pi)^{k-1}}\right)^{\frac{1}{2}}\frac{f}{|f|_2} = \sum_n \rho_f(n)n^{k-\frac{1}{2}}e(nz)$ for the L^2-normalized Fourier coefficients of f. The sum $\sum^h_{f\in B_{k,N}}$ means the expression in the sum is multiplied by the harmonic weights $\frac{\Gamma(k-1)}{(4\pi)^{k-1}\langle f,f\rangle}$. We write $\nu(N) = [Sl_2(Z):\Gamma_0(N)]$, and when N is square free we have $\nu(N) = N\prod_{p|N}(1+1/p)$.

2 Petersson trace formula

In this section, we give a proof of Theorem 1.7. First, we explain the Petersson trace formula. Recall that $B_{k,N}$ is any orthonormal basis of $S_k(N)$. Let $\frac{f}{|f|_2} = \left(\frac{\Gamma(k-1)}{(4\pi)^{k-1}}\right)^{-\frac{1}{2}} \sum_n \rho_n(f) n^{k-\frac{1}{2}} e(nz)$. Then Petersson proved [Pet32]; (see also [ILS00, Proposition 2.1])

$$\Delta_{k,N}(m,n) := \sum_{f\in B_{k,N}} \rho_m(f)\rho_n(f) = \delta(m,n) + 2\pi i^{-k} \sum_{N|c} \frac{S(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right), \tag{2.1}$$

where J_{k-1} is the J-Bessel function and $S(m,n;c)$ is the Kloosterman sum. We have the well-known Weil's bound

$$S(m,n;c) \leq \tau(c)\sqrt{\gcd(m,n,c)}\sqrt{c}. \tag{2.2}$$

Each new from f of level M gives rise to $\tau(M)$ old forms in $S_k(N)$; see [AL70]. By choosing a special orthonormal basis of Hecke eigenfunctions, it is possible to write the Petersson formula only for the new forms of level N; see [ILS00, Proposition 2.9]

$$\Delta^*_{k,N}(m,n) := \sum^h_{f\in B^*_{k,N}} \bar{\lambda}_m(f)\lambda_n(f) = \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} l^{-1}\Delta_{k,M}(ml^2,n). \tag{2.3}$$

We assume that $\gcd(mn,N)=1$ and

$$|4\pi\sqrt{mn}-k| = O(k^{\frac{1}{3}}). \tag{2.4}$$

2.1 Proof of Theorem 1.7.

Proof. We apply the identity (2.3) and obtain

$$\Delta^*_{k,N}(m,n) = \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} l^{-1}\Delta_{k,M}(ml^2,n).$$

First, we analyze the contribution of $\delta(ml^2,n)$ by applying the Petersson formula (2.1). Since, $l|N^\infty$ and $\gcd(N,mn)=1$, then the only possibility for $ml^2=n$ is that $l=1$ and $m=n$. By summing over l, we obtain

$$\sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} l^{-1}\delta(ml^2,n) = \sum_{LM=N} \frac{\mu(L)}{L}\delta(m,n) = \frac{\varphi(N)}{N}\delta(m,n).$$

Therefore,

$$\Delta^*_{k,N}(m,n) := \frac{\varphi(N)}{N}\delta(m,n) + S_1 + S_2,$$

where

$$S_1 := \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} l^{-1} \sum_{M|c \text{ and } c=l} \frac{S(ml^2,n;c)}{c} J_{k-1}(\frac{4\pi l\sqrt{mn}}{c}),$$

$$S_2 := \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} l^{-1} \sum_{M|c \text{ and } c\neq l} \frac{S(ml^2,n;c)}{c} J_{k-1}(\frac{4\pi l\sqrt{mn}}{c}).$$

In what follows, we assume that $l = c$ and give an explicit formula for S_1. Since $M|c$, $l|L$ and $\gcd(L,M) = 1$ then $M = 1$ and we have

$$S_1 = J_{k-1}(4\pi\sqrt{mn})\frac{\mu(N)}{N} \sum_{l|N^\infty} l^{-1} \frac{S(ml^2,n;l)}{l}. \tag{2.5}$$

By using the Ramanujan identity $S(0,n;l) = \mu(l)$, we obtain

$$S_1 = J_{k-1}(4\pi\sqrt{mn})\frac{\mu(N)}{N} \prod_{p|N}(1-1/p^2). \tag{2.6}$$

Note that we have the following asymptotic for the J Bessel function in the transition range where $a = O(1)$; see [DLMF, 10.19.8]

$$J_\nu(\nu + a\nu^{\frac{1}{3}}) = \frac{2^{\frac{1}{3}}}{\nu^{\frac{1}{3}}} Ai(-2^{\frac{1}{3}}a) + O(\frac{1}{\nu}). \tag{2.7}$$

By the inequality (2.7) and the assumption (2.4), we have $|S_1| \gg \frac{1}{k^{\frac{1}{3}}}$, where the constant involved in $\gg$ only depends on N which is fixed. Next, we give an upper bound on S_2. Let $\delta > 0$ be some positive real number and $S_{2,\delta}$ be the same sum as S_2 but subjected to $k^\delta < l$,

$$S_{2,\delta} := \sum_{LM=N} \frac{\mu(L)}{L} \sum_{k^\delta<l|L^\infty} l^{-1} \sum_{M|c \text{ and } c\neq l} \frac{S(ml^2,n;c)}{c} J_{k-1}(\frac{4\pi l\sqrt{mn}}{c}).$$

Since, N is fixed and S_1 is supported on $l|N^\infty$ and $\mu(l) \neq 0$, it follows from (2.3) that for sufficiently large k ;.e.g., $k^\delta > N$

$$S_{2,\delta} = \sum_{LM=N} \frac{\mu(L)}{L} \sum_{k^\delta<l|L^\infty} l^{-1}\big(\Delta_{k,M}(ml^2,n) - \delta(ml^2,n)\big).$$

By [ILS00, Corollary 2.2], we have

$$\Delta_{k,M}(ml^2,n) - \delta(ml^2,n) = O(\frac{(mn)^{\frac{1}{4}+\varepsilon} l^{\frac{1}{2}+\varepsilon}}{k^{5/6}}).$$

where the implied constant in O only depends on the fix number N and ε. Therefore,

$$S_{2,\delta} \ll \sum_{k^\delta<l|N^\infty} l^{-1} \frac{(mn)^{\frac{1}{4}+\varepsilon} l^{\frac{1}{2}+\varepsilon}}{k^{5/6}}.$$

By (2.4), we have

$$S_{2,\delta} \ll k^{-\frac{1}{3}+\varepsilon} \sum_{k^\delta < l | N^\infty} l^{-\frac{1}{2}+\varepsilon} = O(k^{-\frac{1}{3}-\delta/2+2\varepsilon}). \tag{2.8}$$

Finally, we give an upper bound on $S(\delta) := S_2 - S_{2,\delta}$. We split $S(\delta)$ into three ranges:

1. $2l < c$
2. $l < c < 2l < 2k^\delta$
3. $c < l < k^\delta$

and we write $S_i(\delta)$ for the sum $S(\delta)$ subjected to the i-th condition listed above. We give an upper bound on $S_1(\delta)$ by using the following upper bound for J_v when the order ν is large; see [DLMF, 10.14.7]

$$1 \leq \frac{J_\nu(\nu x)}{x^\nu J_\nu(\nu)} \leq e^{\nu(1-x)}, \tag{2.9}$$

where $\nu \geq 0$ and $0 < x \leq 1$. By (2.9), (2.7) and Weil's bound (2.2) on Kloosterman's sum, we have

$$\begin{aligned}
|S_1(\delta)| &\leq \left| \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty, l<k^\delta} l^{-1} \sum_{M|c, 2l<c} \frac{S(ml^2,n;c)}{c} J_{k-1}(\frac{4\pi l\sqrt{mn}}{c}) \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} l^{-1} \sum_{2l<c} \left| \frac{S(ml^2,n;c)}{c} J_{k-1}(\frac{4\pi l\sqrt{mn}}{c}) \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} l^{-1} \sum_{2l<c} \left| \frac{e^{k(1-l/c+\log(l/c))}}{k^{\frac{1}{3}}} \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} \frac{e^{k(1-\frac{1}{2}-\log(2))}}{k^{\frac{1}{3}}} \\
&\ll e^{-(0.19)k}.
\end{aligned} \tag{2.10}$$

Next, we give an upper bound on $S_2(\delta)$ and $S_3(\delta)$. From [DLMF, (10.20.4)] of NIST functions and the following upper bound on the Airy function for real nonpositive x; see [DLMF, (9.8.1) and (9.8.20)]

$$\mathrm{Ai}(x) \ll |x|^{-\frac{1}{4}},$$

we have for $1 > z \geq \frac{1}{2}$,

$$J_\nu(\nu z) \ll \frac{1}{(1-z^2)^{\frac{1}{4}} \nu^{\frac{1}{2}}}, \tag{2.11}$$

and for $z \geq 1$,

$$J_\nu(\nu z) \ll \frac{1}{(z^2-1)^{\frac{1}{4}}\nu^{\frac{1}{2}}}. \tag{2.12}$$

Assume that $l < c < 2l < 2k^\delta$. By the inequality (2.11), (2.4) and Weil's bound (2.2)

$$\begin{aligned}
|S_2(\delta)| &= \left| \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty, l<k^\delta} l^{-1} \sum_{M|c, c<2l} \frac{S(ml^2,n;c)}{c} J_{k-1}\left(\frac{4\pi l\sqrt{mn}}{c}\right) \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} l^{-1} \sum_{l<c<2l} \sqrt{\gcd(m,n,c)}\, c^{-\frac{1}{2}+\varepsilon} \left| J_{k-1}\left(\frac{4\pi l\sqrt{mn}}{c}\right) \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} l^{-1} \sum_{l<c<2l} \sqrt{\gcd(m,n,c)}\, c^{-\frac{1}{2}+\varepsilon} k^{-\frac{1}{2}} \frac{1}{\left(1-\frac{l^2}{c^2}\right)^{\frac{1}{4}}} \qquad (2.13) \\
&\ll k^{-\frac{1}{2}} \sum_{l|N^\infty, l<k^\delta} l^{-5/4+\varepsilon} \sum_{l<c<2l} \frac{\sqrt{\gcd(m,n,c)}}{(c-l)^{\frac{1}{4}}} \\
&\ll k^{-\frac{1}{2}} \sum_{l|N^\infty, l<k^\delta} l^{-\frac{1}{2}+\varepsilon} \ll k^{-\frac{1}{2}}.
\end{aligned}$$

where the implied constant only depends on the fixed number N. Finally, assume that $c < l < k^\delta$ then by (2.12) and Weil's bound (2.2)

$$\begin{aligned}
|S_3(\delta)| &= \left| \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty, l<k^\delta} l^{-1} \sum_{M|c, c<l} \frac{S(ml^2,n;c)}{c} J_{k-1}\left(\frac{4\pi l\sqrt{mn}}{c}\right) \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} l^{-1} \sum_{c<l} \sqrt{\gcd(m,n,c)}\, c^{-\frac{1}{2}+\varepsilon} \left| J_{k-1}\left(\frac{4\pi l\sqrt{mn}}{c}\right) \right| \\
&\ll \sum_{l|N^\infty, l<k^\delta} l^{-1} \sum_{c<l} \sqrt{\gcd(m,n,c)}\, c^{-\frac{1}{2}+\varepsilon} k^{-\frac{1}{2}} \frac{1}{\left(\frac{l^2}{c^2}-1\right)^{\frac{1}{4}}} \qquad (2.14) \\
&\ll k^{-\frac{1}{2}} \sum_{l|N^\infty, l<k^\delta} l^{-5/4} \sum_{c<l} \sqrt{\gcd(m,n,c)}\, c^{-\frac{1}{2}+\varepsilon} \frac{c^{\frac{1}{2}}}{(l-c)^{\frac{1}{4}}} \\
&\ll k^{-\frac{1}{2}} \sum_{l|N^\infty, l<k^\delta} l^{-\frac{1}{2}+\varepsilon} \ll k^{-\frac{1}{2}}.
\end{aligned}$$

Let $\delta = \frac{1}{3} + \varepsilon$ and apply (2.8), (2.8), (2.10), (2.13) and (2.14), to obtain

$$\Delta_{k,N}^*(m,n) := \frac{\varphi(N)}{N}\delta(m,n) + J_{k-1}(4\pi\sqrt{mn})\frac{\mu(N)}{N}\prod_{p|N}\left(1-\frac{1}{p^2}\right) + O\left(k_n^{-\frac{1}{2}}\right).$$

This concludes the proof of Theorem 1.7. ∎

2.2 Proof of Theorem 1.6.

Proof. Recall that $\nu^*_{k,N} := \sum^h_{f\in B_{k,N}} \delta_{\lambda_p(f)}$. Since $|\lambda_p(f)| \leq 2$, we can write $\lambda_p(f) = 2\cos(\theta_p(f))$ for a unique $0 \leq \theta_p(f) \leq \pi$. Let $U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta}$ for $n \geq 0$ be the $n-$th Chebyshev polynomial of the second kind. It is well-known that $\lambda_{p^n}(f) = U_n(\lambda_p(f)/2)$. In order to give a lower bound on the discrepancy of $\nu^*_{k_n,N}$ and μ_∞ for $k_n := \lfloor 4\pi\sqrt{p^n} \rfloor$, we compute the difference between the expected value of $U_n(x)$ with respect to these measures. It is well-known that $U_n(x)$ are orthogonal set of polynomials with respect to μ_∞. Hence,

$$\int_{-2}^{2} U_n(x) d\mu_\infty(x) = 0$$

On the other hand, by Theorem 1.7, since $|k_n - 4\pi\sqrt{p^n}| < 1$ we have

$$\int_{-2}^{2} U_n(x) d\nu^*_{k,N} = \Delta^*_{k_n,N}(1,p^n) = J_{k_n-1}(4\pi\sqrt{p^n}) \frac{\mu(N)}{N} \prod_{p|N}(1-1/p^2) + O(k_n^{-\frac{1}{2}}).$$

As pointed out in Remark 1.2, since $|k_n - 4\pi\sqrt{p^n}| < 1$ then by the known lower bound in the transition range of the J-Bessel function, we have

$$\int_{-2}^{2} U_n(x) d\nu^*_{k,N} \gg_N k_n^{-\frac{1}{3}}.$$

By integration by parts and upper bound $|U_n'(x)| \ll n^2$, it follows that

$$D(\nu^*_{k_n,N}, \mu_\infty) \gg \frac{1}{n^2} k_n^{-\frac{1}{3}}. \tag{2.15}$$

Since $k_n = \lfloor 4\pi\sqrt{p^n} \rfloor$, it follows that

$$D(\nu^*_{k_n,N}, \mu_\infty) \gg \frac{1}{k_n^{\frac{1}{3}} \log^2 k_n}.$$

This concludes the proof of our theorem. □

3 Removing the weights

In this section we give the proof of Theorem 1.2 and then Theorem 1.1 follows immediately from it. We give a brief outline of the proof of Theorem 1.2. Our proof is built on the proof of Theorem 1.7 and we assume that the reader is familiar with that proof. Note that the trace of the Hecke operator $\mathcal{T}^*_n(N,k)$ is obtained by removing the arithmetic weights $\frac{1}{Z(1,f)}$ from the Petersson trace formula (2.3) at $m = 1$. The usual trick for removing these weights is to average the Petersson trace formula (2.3) smoothly over m^2 where $\gcd(m,N) = 1$. Unfortunately, the error associated to the $S_2(\delta)$ and $S_3(\delta)$ sums defined in (2.8) are larger than the main term after averaging over m^2. In order to bound the error term associated to these terms, we sum the trace formula as k varies inside a short interval of size $\sim k^{\delta}$ for some $1/6 < \delta < \frac{1}{3}$ ($\delta < 1/6$ is not large enough to bound the error term and $\delta > \frac{1}{3}$ makes the main term smaller than the error term!) and then apply the Poisson summation formula on the k sum and obtain some oscillatory integrals. We give bounds on these oscillatory integral in Lemma 3.1. Finally, Theorem 1.2 follows form Weil's bound on the Klossterman's and Lemma 3.1.

3.1 Averaging over the weight

In Lemma 3.1, we prove a lower bound on the average of the J-Bessel function in the transition range and also a non-trivial upper bound on this outside the transition range. We use this lemma in the proof of Theorem 1.2 where we bound the average of $S_2(\delta)$ and $S_3(\delta)$ over k.

Recall that ψ is a positive smooth function supported in $[-1, 1]$ and $\int_{-1}^{1} \psi(t)dt = 1$. Let $K > 0$ be a positive real number.

Lemma 3.1. *Let* $0 < \delta < \frac{1}{3}$ *and* $x > 0$. *If* $\frac{x-K}{K^{\delta}} > \max(\frac{x}{K^{3\delta}}, K^{\epsilon})$, *then*

$$\sum_{l \equiv 1 \mod 2} \psi(\frac{l-K}{K^{\delta}}) J_l(x) \ll_{A,\psi} K^{-A} \tag{3.1}$$

otherwise

$$\frac{1}{K^{\delta}} \sum_{l \equiv 1 \mod 2} \psi(\frac{l-K}{K^{\delta}}) J_l(x) \ll_{\psi} K^{-\frac{1}{3}} \tag{3.2}$$

for any $A > 0$ *where* $\ll_{A,\psi}$ *means the implicit constant is independent of* x *and* K *and only depends on the smooth weight function* ψ *and the exponent* A. *Moreover, if* $x = K + o(K^{\frac{1}{3}})$ *then*

$$\frac{1}{K^{\delta}} \sum_{l \equiv 1 \mod 2} \psi(\frac{l-K}{K^{\delta}}) J_l(x) = J_K(x)(1 + O(K^{-\epsilon})) \gg_{\psi} K^{-\frac{1}{3}}. \tag{3.3}$$

Proof. It is well-known that

$$J_l(x) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-2\pi i l t} e^{-ix \sin 2\pi t} dt.$$

By the Poisson summation formula, it follows that

$$\sum_{l \equiv 1 \mod 2} g(l) J_l(x) = \int_{-\infty}^{\infty} \hat{\psi}(u) e^{-2\pi i u K^{1-\delta}} \left(e^{-ix \sin(2\pi u/K^{\delta})} - e^{ix \sin(2\pi u/K^{\delta})}\right) du.$$

By writing the Taylor expansion of the sin function at zero, we obtain

$$-2\pi i u K^{1-\delta} \pm ix \sin(2\pi u/K^{\delta}) = 2\pi i \frac{\pm x - K}{K^{\delta}} u \mp \frac{ix(2\pi u)^3}{6K^{3\delta}} \pm c(u),$$

where $|c'(u)| \leq \frac{xk^{4\epsilon}}{K^{5\delta}}$ for $u \in [-K^{\epsilon}, K^{\epsilon}]$. Assume that $\frac{x-K}{K^{\delta}} > \max(\frac{x}{K^{3\delta}}, K^{\epsilon})$ then it follows that

$$\frac{d}{du}\Big(-2\pi i u K^{1-\delta} \pm ix\sin(2\pi u/K^{\delta})\Big) \gg \frac{\pm x - K}{K^{\frac{1}{3}}},$$

where $u \in [-K^{\epsilon}, K^{\epsilon}]$. Therefore, by the stationary phase theorem

$$\int_{-K^{\epsilon}}^{K^{\epsilon}} \hat{\psi}(u) e^{-2\pi i u K^{1-\delta}} \big(e^{-ix\sin(2\pi u/K^{\delta})} - e^{ix\sin(2\pi u/K^{\delta})}\big) du \ll_{A,\psi} |K|^{-A},$$

for any $A > 0$. We note that for $|u| > K^{\epsilon}$ the Fourier transform of ψ decays faster than any polynomial and we have

$$\int_{|u|>K^{\epsilon}} |\hat{\psi}(u)| \ll_{A,\psi} |K|^{-A},$$

This completes the proof of (3.1). The inequality (3.2) follows, from the well-known upper bound $J_K(x) \ll K^{-\frac{1}{3}}$ and the fact that ψ is supported in $[-1, 1]$. Finally, (3.3) follows from the asymptotic of the J-Bessel function in the transition range (2.7). This concludes the proof of our lemma.

□

Finally, we give the proof of Theorem 1.2. Recall that $\mathcal{T}_n^*(N, k) := \sum_{f \in B_{k,N}^*} \lambda_n(f)$ and $K := 4\pi\sqrt{n} + o(n^{\frac{1}{6}})$. First, we cite some identities from [ILS00] that we use in the proof. Let f be a newform of $S_k(N)$ of level M, then by [ILS00, Lemma 2.5], we have

$$\rho_m(f)\rho_n(f) = \frac{12\lambda_m(f)\lambda_n(f)M}{(k-1)\nu(N)Z(1,f)\varphi(M)}, \tag{3.4}$$

where $Z(s, f) := \sum_n \lambda_f(n^2)n^{-s}$. Note that $Z(s, f)$ is related to $L(s, \text{sym}^2(f))$ by; see [ILS00, (3.14)]

$$L(s, \text{sym}^2(f)) = \zeta(2s)\zeta_N(2s)^{-1}Z(s, f),$$

where $\zeta_N(2s) = \prod_{p|N}(1-p^{-2s})^{-1}$. Let $Z^N(s, f) := \sum_{\gcd(m,N)=1} \frac{\lambda_{m^2}(f)}{m^s}$, then by [ILS00, (3.16)]

$$Z^N(s, f) = L(s, sym^2(f))\zeta(2s)^{-1}\zeta_N(2s)\zeta_N(s+1)^{-1}. \tag{3.5}$$

By the celebrated result of Shimura [Shi75] $L(s, sym^2(f))$ is an entire function, so $Z^N(s, f)$ is holomorphic for $\Re(s) > \frac{1}{2}$ and has meromorphic continuation to the complex plane. Let $w(x) = \exp(-x)$. Note that the Mellin transform of w is the Gamma function

$$\hat{w}(s) := \int_0^{\infty} x^{s-1}w(x)dx = \Gamma(s).$$

3.2 Proof of Theorem 1.2

Proof. Assume that $k \in [K - K^\delta, K + K^\delta]$ where $\delta < \frac{1}{3}$. By the Petersson formula (2.3)

$$\sum_{f \in B^*_{k,N}} \rho_{m^2}(f)\rho_f(n) = \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} l^{-1}\Delta_{k,M}(m^2l^2, n). \tag{3.6}$$

Let $T := k^\alpha$ for some $0 < \alpha < 1$ that we choose at the end of the proof. We average the LHS of the above by the smooth function $w(x/T)/x$ and use (3.4) to obtain

$$\begin{aligned}\sum_{\gcd(m,N)=1} w(m/T)/m \sum_{f \in B^*_{k,N}} \rho_{m^2}(f)\rho_f(n) &= \sum_{f \in B^*_{k,N}} \sum_{\gcd(m,N)=1} w(m/T)\frac{12\lambda_n(f)\lambda_f(m^2)\zeta_N(2)}{m(k-1)NZ(1,f)}. \\ &= \frac{12}{(k-1)N}\sum_{f \in B^*_{k,N}} \lambda_n(f)\frac{\zeta_N(2)}{Z(1,f)} \sum_{\gcd(m,N)=1} w(m/T)\frac{\lambda_f(m^2)}{m}.\end{aligned} \tag{3.7}$$

By the inverse of the Mellin transform, we have $w(x/T) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty}\Gamma(s)T^s x^{-s}ds$ and this implies

$$\sum_{\gcd(m,N)=1} w(m/T)\frac{\lambda_{f_i}(m^2)}{m} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} Z^N(s+1,f)T^s\Gamma(s)ds.$$

We change the contour integral to the $\Re(s) = -\frac{1}{2}$ and pick up the pole of $\Gamma(s)$ at $s = 0$ with residue $Z^N(1,f) = \frac{Z(1,f)}{\zeta_N(2)}$, hence

$$\sum_{\gcd(m,N)=1} w(m/T)\frac{\lambda_f(m^2)}{m} = \frac{Z(1,f)}{\zeta_N(2)} + \frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} Z^N(s+1,f)T^s\Gamma(s)ds. \tag{3.8}$$

By (3.5),

$$\begin{aligned}&\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} Z^N(s+1,f)T^s\Gamma(s)ds \\ &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} L(\frac{1}{2}+it, sym^2(f))\zeta(1+2it)^{-1}\zeta_N(1+2it)\zeta_N(\frac{3}{2}+it)^{-1}T^{-\frac{1}{2}+it}\Gamma(-\frac{1}{2}+it)dt.\end{aligned}$$

First, we assume that $|t| > \log(k)^2$. By Stirling's formula; see [DLMF, 5.11.9]

$$\Gamma(-\frac{1}{2}+it) = O((1+|t|)^{-1}e^{-\pi|t|/2}).$$

By using the above bound, the convexity bound on $L(\frac{1}{2}+it, sym^2 f)$, the well-known bound $\zeta(1+2it)^{-1} = O(\log(t)^7)$, the fact that $\zeta_N(2s)\zeta_N(s+1)^{-1}$ is bounded on $\Re(s) = \frac{1}{2}$ and $|T^{-\frac{1}{2}+it}| \leq T^{-\frac{1}{2}} \leq k^{-\alpha/2}$, it follows that

$$\int_{-\infty}^{\log(k)^2} + \int_{\log(k)^2}^{\infty} L(\frac{1}{2}+it, sym^2(f))\zeta(1+2it)^{-1}\zeta_N(1+2it)\zeta_N(\frac{3}{2}+it)^{-1}T^{-\frac{1}{2}+it}\Gamma(-\frac{1}{2}+it)dt = O(k^{-A}),$$

for any $A > 0$ where the implicit constant in O depends on A. By the above, (3.7) and (3.8), we obtain

$$\sum_{\gcd(m,N)=1} w(m/T)/m \sum_{f\in B^*_{k,N}} \rho_{m^2}(f)\rho_n(f) = \frac{12}{(k-1)N}\mathcal{T}^*_n(N,k) + O(k^{-A})$$
$$+\int_{-\log(k)^2}^{\log(k)^2} \Big(\sum\nolimits^h_{f\in B^*_{k,N}} \lambda_n(f) L\left(\frac{1}{2}+it, sym^2 f\right)\Big)\frac{\zeta_N(1+2it)}{\zeta(1+2it)\zeta_N\left(\frac{3}{2}+it\right)} T^{-\frac{1}{2}+it}\Gamma\left(-\frac{1}{2}+it\right)dt \tag{3.9}$$

By the Ramanujan bound on the holomorphic cusp forms $|\lambda_n(f)| \ll n^{\epsilon}$. Hence,

$$\sum\nolimits^h_{f\in B^*_{k,N}} \lambda_n(f) L(\frac{1}{2}+it, sym^2 f) \ll n^{\epsilon}\sum\nolimits^h_{f\in B^*_{k,N}} |L(\frac{1}{2}+it, sym^2 f)| \ll n^{\epsilon}.$$

Therefore,

$$\int_{-\log(k)^2}^{\log(k)^2} \Big(\sum\nolimits^h_{f\in B^*_{k,N}} \lambda_n(f) L(\frac{1}{2}+it, sym^2 f)\Big)\zeta(1+2it)^{-1}\zeta_N(1+2it)\zeta_N(\frac{3}{2}+it)^{-1}T^{-\frac{1}{2}+it}\Gamma(-\frac{1}{2}+it)dt$$
$$= O(T^{-\frac{1}{2}}k^{\epsilon}).$$

By the above and (3.9), we have

$$\sum_{\gcd(m,N)=1} w(m/T)/m \sum_{f\in B^*_{k,N}} \rho_{m^2}(f)\rho_f(n) = \frac{12}{(k-1)N}\mathcal{T}^*_n(N,k) + O(T^{-\frac{1}{2}}k^{\varepsilon}). \tag{3.10}$$

Finally, we average the RHS of (3.6) with similar weights $w(m/T)/m$. Our method is very similar to our argument in the proof of Theorem 1.7. Let

$$S := \sum_{\gcd(m,N)=1} w(m/T)/m \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^{\infty},} l^{-1}\Delta_{k,M}(m^2l^2, n).$$

We analyze the contribution of $\delta(m^2l^2, n)$ by applying the Petersson formula (2.1). Since, $l|N^{\infty}$ and $\gcd(N, mn) = 1$, then the only possibility for $m^2l^2 = n$ is that $l = 1$ and $m^2 = n$. Therefore,

$$\sum_{\gcd(m,N)=1} w(m/T)/m \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^{\infty}} l^{-1}\delta(m^2l^2, n)$$
$$= w(\sqrt{n}/T)/\sqrt{n} \sum_{LM=N} \frac{\mu(L)}{L}\delta(\sqrt{n}) = \frac{\varphi(N)w(\frac{\sqrt{n}}{T})}{N\sqrt{n}}\delta(\sqrt{n}).$$

where $\delta(\sqrt{n}) = 1$ if n is a perfect square and $\delta(\sqrt{n}) = 0$ otherwise. Note that by our choice of w if $T \ll n^{\frac{1}{2}-\varepsilon}$, then

$$\frac{\varphi(N)w(\frac{\sqrt{n}}{T})}{N\sqrt{n}}\delta(\sqrt{n}) = O(k^{-A}), \tag{3.11}$$

for any $A > 0$. Let

$$S^T := \sum_{\gcd(m,N)=1, T^{1+\varepsilon}<m} w(m/T)/m \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty,} l^{-1}\big(\Delta_{k,M}(m^2l^2,n) - \delta(m^2l^2,n)\big).$$

By [ILS00, Corollary 2.2], we have

$$\Delta_{k,M}(m^2l^2,n) - \delta(m^2l^2,n) = O(\frac{n^{\frac{1}{4}+\varepsilon}(ml)^{\frac{1}{2}+\varepsilon}}{k^{5/6}}).$$

where the implied constant in O only depends on the fix number N and ε. It follows from the above and the choice of w and T that $S^T = O(k^{-A})$. Hence,

$$S = S_1 + S_2 + O(k^{-A}), \tag{3.12}$$

where

$$S_1 := \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} \sum_{\gcd(m,N)=1, m<T^{1+\varepsilon}} w(m/T)/ml \sum_{c|M, c=ml} \frac{S(m^2l^2,n;c)}{c} J_{k-1}(\frac{4\pi ml\sqrt{n}}{c}),$$

$$S_2 := \sum_{LM=N} \frac{\mu(L)}{L} \sum_{l|L^\infty} \sum_{\gcd(m,N)=1, m<T^{1+\varepsilon}} w(m/T)/ml \sum_{c|M, c\neq ml} \frac{S(m^2l^2,n;c)}{c} J_{k-1}(\frac{4\pi ml\sqrt{n}}{c}).$$

In what follows, we give an asymptotic formula for S_1 which is the sum over the diagonal terms $ml = c$ where $\gcd(m,N) = 1$ and $l|L^\infty$. Similarly, $ml = c$ happens when $M = 1$ and $L = N$ and we have

$$S(m^2l^2,n;c) = S(0,n;c) = \sum_{d|\gcd(c,n)} \mu(\frac{c}{d})d.$$

Hence,

$$\begin{aligned}
S_1 &= J_{k-1}(4\pi\sqrt{n})\frac{\mu(N)}{N} \sum_{l|N^\infty} \sum_{\gcd(m,N)=1, m<T^{1+\varepsilon}} w(m/T)/(ml)^2 \sum_{d|\gcd(ml,n)} \mu(\frac{ml}{d})d \\
&= J_{k-1}(4\pi\sqrt{n})\frac{\mu(N)}{N}\Big(\sum_{l|N^\infty} \frac{\mu(l)}{l^2}\Big)\Big(\sum_{\gcd(m,N)=1, m<T^{1+\varepsilon}} w(m/T)/m^2 \sum_{d|\gcd(m,n)} \mu(\frac{m}{d})d\Big) \\
&= J_{k-1}(4\pi\sqrt{n})\frac{\mu(N)}{N}\zeta_N(2)^{-1}\Big(\sum_{d|n} 1/d \sum_{\gcd(h,N)=1, h<T^{1+\varepsilon}/d} w(hd/T)\mu(h)/h^2\Big) \\
&= J_{k-1}(4\pi\sqrt{n})\frac{\mu(N)}{N}\zeta^{-1}(2)\sigma(n)/n(1+O(T^{-1})).
\end{aligned} \tag{3.13}$$

Next, we give an upper bound on S_2. Let $\beta > 0$ be some positive real number and $S_{2,\beta}$ be the same sum as S_2 but subjected to $K^\beta < l$,

$$S_{2,\beta} := \sum_{LM=N} \frac{\mu(L)}{L} \sum_{K^\beta<l|L^\infty} \sum_{\gcd(m,N)=1, m<T^{1+\varepsilon}} w(m/T)/ml \sum_{c|M, c\neq ml} \frac{S(m^2l^2,n;c)}{c} J_{k-1}(\frac{4\pi ml\sqrt{n}}{c}).$$

Since, N is fixed and S_1 is supported on $l|N^\infty$ and $\mu(l)\neq 0$, it follows from (2.3) that for sufficiently large k ;.e.g., $K^\beta > N$

$$S_{2,\beta}=\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}} w(m/T)/m\sum_{LM=N}\frac{\mu(L)}{L}\sum_{K^\beta<l|L^\infty} l^{-1}\big(\Delta_{k,M}(m^2l^2,n)-\delta(ml^2,n)\big).$$

By [ILS00, Corollary 2.2], we have

$$\Delta_{k,M}(m^2l^2,n)-\delta(m^2l^2,n)=O(\frac{n^{\frac{1}{4}+\varepsilon}(ml)^{\frac{1}{2}+\varepsilon}}{k^{5/6}}).$$

where the implied constant in O only depends on the fix number N and ε. Therefore,

$$S_{2,\beta}\ll\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}} w(m/T)/m\sum_{K^\beta<l|N^\infty} l^{-1}\frac{n^{\frac{1}{4}+\varepsilon}(ml)^{\frac{1}{2}+\varepsilon}}{k^{5/6}}.$$

By (2.4), we have

$$S_{2,\beta}\ll k^{-\frac{1}{3}+\varepsilon}\sum_{m<T^{1+\varepsilon}}\sum_{K^\beta<l|N^\infty}(ml)^{-\frac{1}{2}+\varepsilon}=O(T^{\frac{1}{2}}k^{-\frac{1}{3}-\beta/2+\varepsilon}). \tag{3.14}$$

Finally, we give an upper bound on $S(\beta):=S_2-S_{2,\beta}$. We split $S(\beta)$ into two ranges:

1. $2ml<c$,
2. $c<2ml$ and $c\neq ml$

and we write $S_i(\beta)$ for the sum $S(\beta)$ subjected to the i-th condition listed above. First, we give an upper bound on $S_1(\beta)$. Assume that $2ml<c$ then by (2.9), (2.7) and Weil's bound (2.2) on Kloosterman's sum, we have

$$\begin{aligned}|S_1(\beta)|&=\Big|\sum_{LM=N}\frac{\mu(L)}{L}\sum_{l|L^\infty,l<K^\beta}\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}} w(m/T)/(ml)\sum_{M|c,2ml<c}\frac{S(m^2l^2,n;c)}{c}J_{k-1}(\frac{4\pi ml\sqrt{n}}{c})\Big|\\&\ll\sum_{l|N^\infty,l<K^\beta}\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}} w(m/T)/(ml)\sum_{2ml<c}\Big|\frac{S(m^2l^2,n;c)}{c}\Big|\Big|J_{k-1}(\frac{4\pi ml\sqrt{n}}{c})\Big|\\&\ll\sum_{h<K^\beta M^{1+\varepsilon}} h^{-1}\sum_{2h<c}\Big|\frac{e^{k(1-h/c+\log(h/c))}}{k^{\frac{1}{3}}}\Big|\ll e^{-(0.19)k}.\end{aligned}\tag{3.15}$$

By inequalities (3.10), (3.11), (3.12), (3.13), (3.14), (3.15), we have

$$\mathcal{T}_n^*(N,k)=J_{k-1}(4\pi\sqrt{n})\frac{\mu(N)k}{12}\zeta^{-1}(2)\sigma(n)/n+S_2(\beta)+O(T^{-\frac{1}{2}}k^{\varepsilon}+T^{\frac{1}{2}}k^{-\frac{1}{3}-\beta/2+\varepsilon}).$$

We average the above identity by $\frac{1}{K^\delta}\sum_{k>0,k\in 2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)$ and apply inequality (3.3) in Lemma 3.1

$$\frac{1}{K^\delta}\sum_{k>0,k\in 2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)\mathcal{T}_n^*(N,k)=J_K(4\pi\sqrt{n})\frac{\mu(N)K}{12}\zeta^{-1}(2)\frac{\sigma(n)}{n}(1+K^{-\epsilon})$$
$$+\frac{1}{K^\delta}\sum_{k>0,k\in 2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)S_2(\beta)+O(T^{-\frac{1}{2}}k^{\varepsilon}+T^{\frac{1}{2}}k^{-\frac{1}{3}-\beta/2+\varepsilon}). \quad (3.16)$$

Next, we give an upper bound on the average of $S_2(\beta)$.

$$\Big|\frac{1}{K^\delta}\sum_{k>0,k\in 2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)S_2(\beta)\Big|=\Big|\sum_{LM=N}\frac{\mu(L)}{L}\sum_{l|L^\infty,l<K^\beta}\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}}\frac{w(m/T)}{ml}$$
$$\times\sum_{M|c,c<2ml}\frac{S(m^2l^2,n;c)}{c}\frac{1}{K^\delta}\sum_{k>0,k\in 2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)J_{k-1}\left(\frac{4\pi ml\sqrt{n}}{c}\right)\Big| \quad (3.17)$$

For the summation $S_2(\beta)$, we have $c<2ml<2T^{1+\epsilon}K^\beta$. Let $x:=\frac{4\pi ml\sqrt{n}}{c}$. First, we check the condition of inequality (3.1) in Lemma 3.1, that is if $\frac{x-K}{K^\delta}>\max(\frac{x}{K^{3\delta}},K^\epsilon)$. We assumed that $|K-4\pi\sqrt{n}|<n^{\frac{1}{6}}$, $\delta<\frac{1}{3}$ and $c<2ml$, hence $\frac{x}{K^{3\delta}}>K^\epsilon$. So, it is enough to check if $\frac{x-K}{K^\delta}>\frac{x}{K^{3\delta}}$. In particular, if $|\frac{ml}{c}-1|>K^{-2\delta}$ then we can apply inequality (3.1). Hence we consider two cases:

1. $c<2ml$ and $|\frac{ml}{c}-1|>K^{-2\delta}$
2. $c<2ml$ and $|\frac{ml}{c}-1|<K^{-2\delta}$

We denote the above sums by $S_{2,1}$ and $S_{2,2}$ respectively where $S_2(\beta)=S_{2,1}+S_{2,2}$. By Lemma inequality (3.1), identity (3.17) and Weil's bound (2.2) on Kloosterman's sum, we have

$$\Big|\frac{1}{K^\delta}\sum_{k>0,k\in 2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)S_{2,1}\Big|\ll\sum_{l|N^\infty,l<K^\beta}\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}}\frac{1}{ml}\sum_{c<2ml}\Big|\frac{S(m^2l^2,n;c)}{c}\Big|K^{-A}$$
$$\ll K^{-A}\sum_{l|N^\infty,l<K^\beta}\sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}}\frac{1}{ml}\sum_{c<2ml}\sqrt{\gcd(m,n,c)}c^{-\frac{1}{2}+\varepsilon}$$
$$=O(T^{\frac{1}{2}+\epsilon}K^{-A}). \quad (3.18)$$

Finally, we bound the $S_{2,2}$ sum. We apply inequality (3.2) and Weil's bound (2.2) on Kloosterman's sum:

$$\begin{aligned}
\Big|\frac{1}{K^\delta}\sum_{k>0,k\in2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)S_{2,2}\Big| &\ll \sum_{l|N^\infty,l<K^\beta}\ \sum_{\gcd(m,N)=1,m<T^{1+\varepsilon}}\frac{1}{ml}\sum_{|\frac{ml}{c}-1|<K^{-2\delta}}\Big|\frac{S(m^2l^2,n;c)}{c}\Big|K^{-\frac{1}{3}}\\
&\ll K^{-\frac{1}{3}}\sum_{l|N^\infty,l<K^\beta}\ \sum_{m<T^{1+\varepsilon}}\frac{1}{ml}\sqrt{\gcd(m,n)}\sum_{|\frac{ml}{c}-1|<K^{-2\delta}}c^{-\frac{1}{2}+\varepsilon}\\
&\ll K^{-\frac{1}{3}}\sum_{l|N^\infty,l<K^\beta}\ \sum_{m<T^{1+\varepsilon}}\frac{1}{ml}\sqrt{\gcd(m,n)}\frac{(ml)^{\frac{1}{2}+\epsilon}}{K^{2\delta}}\\
&\ll K^{-\frac{1}{3}}\sum_{l|N^\infty,l<K^\beta}\ \sum_{m<T^{1+\varepsilon}}\sqrt{\gcd(m,n)}\frac{(ml)^{-\frac{1}{2}+\epsilon}}{K^{2\delta}}\\
&= O(T^{\frac{1}{2}+\epsilon}K^{-\frac{1}{3}-2\delta}).
\end{aligned} \tag{3.19}$$

Therefore, by inequalities (3.16), (3.18), (3.19), we have

$$\begin{aligned}
&\frac{1}{K^\delta}\sum_{k>0,k\in2\mathbb{Z}}\psi\left(\frac{k-K}{K^\delta}\right)\mathcal{T}_n^*(N,k)\\
&= J_K(4\pi\sqrt{n})\frac{\mu(N)K\sigma(n)}{12\zeta(2)n}(1+K^{-\epsilon})+O(T^{-\frac{1}{2}}k^{\varepsilon}+T^{\frac{1}{2}}k^{-\frac{1}{3}-\frac{\beta}{2}+\varepsilon}+T^{\frac{1}{2}+\epsilon}K^{-\frac{1}{3}-2\delta})
\end{aligned}$$

By choosing β large enough, $T\sim K^{\frac{2}{3}+\varepsilon}$ and $\frac{1}{6}<\delta<\frac{1}{3}$ we conclude our theorem.

□

3.3 Proof of Theorem 1.1

Proof. The method of the proof is similar to the proof of Theorem 1.6. Let $U_n(x)$ be the $n-$th Chebyshev polynomial of the second kind. It is well known that

$$\int_{-2}^{2}U_n(x)d\mu_p(x)=\begin{cases}\frac{1}{p^{n/2}} & \text{if } n \text{ is a even}\\ 0 & \text{otherwise.}\end{cases}$$

By Theorem 1.2, there exists $k_n\in[\lfloor 4\pi\sqrt{p^n}\rfloor-p^{n/6},\lfloor 4\pi\sqrt{p^n}\rfloor+p^{n/6}]$ such that

$$\int_{-2}^{2}U_n(x)d\mu_p(x)-\int_{-2}^{2}U_n(x)d\mu^*_{k_n,N}\gg k_n^{-\frac{1}{3}}$$

By the above inequality and a similar argument as in Theorem 1.6, we have

$$D(\mu^*_{k_n,N},\mu_p)\gg\frac{1}{k_n^{\frac{1}{3}}\log^2 k_n}.$$

This concludes the proof of our theorem. □

4 Selberg's trace formula

The main purpose of this section is to prove Theorem 1.3 and Theorem 1.4. We first recall Eichler–Selberg trace formula. We use the version from [MS09] (see also [Ser97]).

Theorem 4.1 (Eichler–Selberg trace formula, Theorem 10 [MS09])**.** *For every positive integer $n \geq 1$, the trace* Tr *of $\mathcal{T}_n = \mathcal{T}_n(N,k)$ acting on $S_k(N)$ is given by*

$$\mathrm{Tr}\ \mathcal{T}_n = A_1(n,k,N) + A_2(n,k,N) + A_3(n,k,N) + A_4(n,k),$$

where $A_i(n,k)$'s are as follows:

$$A_1(n,k,N) = \begin{cases} \frac{k-1}{12}\psi(N)\frac{1}{\sqrt{n}} & \textit{if } n \textit{ is a square} \\ 0 & \textit{otherwise} \end{cases} \quad \textit{where } \psi(N) = N\prod_{p|N}\left(1+\frac{1}{p}\right).$$

$$A_2(n,k,N) = -\frac{1}{2}n^{-\frac{k-1}{2}} \sum_{t\in\mathbb{Z},\ t^2<4n} \frac{\rho_{t,n}^{k-1} - \bar{\rho}_{t,n}^{k-1}}{\rho_{t,n} - \bar{\rho}_{t,n}} \sum_f h_w\left(\frac{t^2-4n}{f^2}\right)\mu(t,f,n,N),$$

where $\rho_{t,n}$ and $\bar{\rho}_{t,n}$ are zeros of x^2-tx+n, and the inner sum runs over all positive divisors of t^2-4n such that $(t^2-4n)/f^2 \in \mathbb{Z}$ is congruent to 0 or $1 \pmod 4$. $\mu(t,f,n,N)$ is given by

$$\mu(t,f,n,N) = \frac{\psi(N)}{\psi(N/N_f)}M(t,n,NN_f)$$

where $N_f = \gcd(N,f)$ and $M(t,n,K)$ denotes the number of solutions of the congruence $x^2 - tx + n \equiv 0 \pmod K$.

$$A_3(n,k,N) = -n^{-\frac{k-1}{2}} \sum_{d|n,\ 0<d\leq\sqrt{n}} d^{k-1} \sum_{c|N, \gcd\left(c,\frac{N}{c}\right)|\gcd\left(N,\frac{n}{d}-d\right)} \varphi\left(\gcd\left(c,\frac{N}{c}\right)\right).$$

Here, φ is Euler's totient function, and in the first summation, if there is a contribution from the term $d=\sqrt{n}$, it should be multiplied by $\frac{1}{2}$.

$$A_4(n,k) = \begin{cases} n^{-\frac{1}{2}}\sum_{t|n} t & \textit{if } k=2, \\ 0 & \textit{otherwise.} \end{cases}$$

To relate the trace of $\mathcal{T}_n$ acting on $S_k(N)$ and the trace of its restriction $\mathcal{T}_n^*$ to $S_k(N)^*$, one may use Atkin–Lehner decomposition for squarefree integers N to derive (see for instance, [Ham98])

$$\mathrm{Tr}\ \mathcal{T}_n(N,k) = \sum_{d|N} \sigma_0(N/d)\mathrm{Tr}\ \mathcal{T}_n^*(d,k),$$

and by Möbius inversion, this implies that

$$\mathrm{Tr}\ \mathcal{T}_n^*(N,k) = \sum_{d|N} \sigma_0(N/d)\mu(N/d)\mathrm{Tr}\ \mathcal{T}_n(d,k). \tag{4.1}$$

Therefore we have:

Lemma 4.2. *Assume that N is a squarefree integer. For every positive integer $n \geq 1$, the trace* Tr *of $\mathcal{T}_n = \mathcal{T}_n(N,k)$ restricted to $S_k(N)^*$, which we denote by $\mathcal{T}_n^* = \mathcal{T}_n^*(N,k)$ is given by*

$$\mathrm{Tr}\,\mathcal{T}_n^* = B_1(n,k,N) + B_2(n,k,N) + B_3(n,k,N) + B_4(n,k,N),$$

where $B_i(n,k)$'s are as follows:

$$B_1(n,k,N) = \begin{cases} \frac{k-1}{12}\varphi(N)\frac{1}{\sqrt{n}} & \textit{if } n \textit{ is a square} \\ 0 & \textit{otherwise} \end{cases}$$

$$B_2(n,k,N) = -\frac{1}{2}n^{-\frac{k-1}{2}} \sum_{t\in\mathbb{Z},\ t^2<4n} \frac{\rho_{t,n}^{k-1} - \bar{\rho}_{t,n}^{k-1}}{\rho_{t,n} - \bar{\rho}_{t,n}} \sum_f h_w\left(\frac{t^2-4n}{f^2}\right)\tilde{\mu}(t,f,n,N).$$

where $\rho_{t,n}$ and $\bar{\rho}_{t,n}$ are zeros of $x^2 - tx + n$, and the inner sum runs over all positive divisors of $t^2 - 4n$ such that $(t^2-4n)/f^2 \in \mathbb{Z}$ is congruent to 0 or $1 \pmod 4$. $\tilde{\mu}(t,f,n,N)$ is given by

$$\tilde{\mu}(t,f,n,N) = \sum_{d|N} \sigma_0(N/d)\mu(N/d)\mu(t,f,n,d).$$

$$B_3(n,k,N) = \begin{cases} -n^{-\frac{k-1}{2}} \sum_{d|n,\ 0<d\leq\sqrt{n}} d^{k-1} & \textit{if } N = 1 \\ 0 & \textit{otherwise} \end{cases}$$

In the first summation, if there is a contribution from the term $d = \sqrt{n}$, it should be multiplied by $\frac{1}{2}$.

$$B_4(n,k,N) = \begin{cases} \mu(N)n^{-\frac{1}{2}} \sum_{t|n} t & \textit{if } k = 2, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. From Theorem 4.1 and (4.1), we have

$$\mathrm{Tr}\,\mathcal{T}_n^* = B_1(n,k,N) + B_2(n,k,N) + B_3(n,k,N) + B_4(n,k,N),$$

where

$$B_i(n,k,N) = \sum_{d|N} \sigma_0(N/d)\mu(N/d)A_i(n,k,d).$$

Note that when N is squarefree, $\gcd\left(c, \frac{N}{c}\right) = 1$, so the inner sum of $A_3(n,k,N)$ becomes $\sigma_0(N)$.

To prove the lemma, it is sufficient to compute for $i = 1, 3, 4$

$$B_i(n,k,p) = A_i(n,k,p) - 2A_i(n,k,1)$$

by multiplicity of Dirichlet convolution, and the assumption that N is squarefree:

$$\psi(p) - 2\psi(1) = p - 1 = \varphi(p),$$

when $i = 1$,

$$\sigma_0(p) - 2\sigma_0(1) = 2 - 2 = 0,$$

when $i = 3$, and

$$1 - 2 = -1 = \mu(p),$$

when $i = 4$. □

4.1 Analytic setup

Let ϕ be a positive even rapidly decaying function whose Fourier transform $\hat{\phi}$ is supported in $[-1/100, 1/100]$. In this section, we study the second moment of B_2:

$$\sum_{k>0, k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) |B_2(n,k,N)|^2 = \frac{1}{2} \sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) |B_2(n,k,N)|^2, \tag{4.2}$$

where we used $B_2(n,k,N) = -B_2(n, 2-k, N)$.

We first collect some preliminary estimates.

Lemma 4.3. *We have*

$$|S_k(N)^*| = \frac{k-1}{12}\varphi(N) + O_N(1), \tag{4.3}$$

and

$$B_2(n,k,N) \ll_N \sigma_1(n). \tag{4.4}$$

Proof. (4.3) follows from Theorem 13 of [MS09], and (4.1).

To prove (4.4), note that

$$\left| n^{-\frac{k-1}{2}} \frac{\rho_{t,n}^{k-1} - \bar{\rho}_{t,n}^{k-1}}{\rho_{t,n} - \bar{\rho}_{t,n}} \right| \leq \frac{2}{|\rho_{t,n} - \bar{\rho}_{t,n}|} = \frac{2}{\sqrt{4n - t^2}} \leq 2.$$

Therefore

$$|B_2(n,k,N)| \leq 2 \sum_{t^2 < 4n} \sum_{f} h_w\left(\frac{t^2 - 4n}{f^2}\right) \tilde{\mu}(t,f,n,N) \ll_N \sigma_1(n),$$

where we combined Lemma 16 [MS09] and a trivial upper bound $\tilde{\mu}(t,f,n,N) \ll_N 1$ in the last estimate. □

For $t \in \mathbb{Z}$ such that $t^2 < 4n$, define $0 < \theta_{t,n} < \pi$ by

$$\sqrt{n} e^{i\theta_{t,n}} = \frac{1}{2}(t + i\sqrt{4n - t^2}).$$

We record some trivial estimates regarding $\theta_{t,n}$'s

Lemma 4.4. *For integer t such that $t^2 < n$, we have*

$$\pi - \frac{1}{2\sqrt{n}}\theta_{t,n} \geq \frac{1}{2\sqrt{n}}$$

and

$$\theta_{t,n} - \theta_{t+1,n} \geq \frac{1}{2\sqrt{n}}.$$

Proof. We have

$$\sin\theta_{t,n} = \frac{\sqrt{4n-t^2}}{2\sqrt{n}} \geq \frac{1}{2\sqrt{n}}.$$

Also,

$$e^{i(\theta_{t,n}-\theta_{t+1,n})} = \frac{1}{4n}(t+i\sqrt{4n-t^2})(t+1-i\sqrt{4n-(t+1)^2})$$

so

$$\begin{aligned}\sin(\theta_{t,n}-\theta_{t+1,n}) &= \frac{1}{4n}((t+1)\sqrt{4n-t^2} - t\sqrt{4n-(t+1)^2}) \\ &= \frac{1}{4n}\frac{(t+1)^2(4n-t^2) - t^2(4n-(t+1)^2)}{(t+1)\sqrt{4n-t^2}+t\sqrt{4n-(t+1)^2}} \\ &= \frac{2t+1}{(t+1)\sqrt{4n-t^2}+t\sqrt{4n-(t+1)^2}} \\ &\geq \frac{1}{\sqrt{4n}}.\end{aligned}$$

□

We introduce $D(t,n)$ as follows:

$$B_2(n,k,N) = \sum_{t\in\mathbb{Z},\ t^2<4n} \left(e^{i(k-1)\theta_{t,n}} - e^{-i(k-1)\theta_{t,n}}\right) D(t,n).$$

Then expanding (4.2) and using $D(t,n) = -D(-t,n)$, we get

$$\begin{aligned}&\sum_{k\in2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right)|B_2(n,k,N)|^2 \\ =&4\sum_{k\in2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right)\sum_{t^2<4n}|D(t,n)|^2 \\ &+\sum_{t_1\neq t_2}\sum_{k\in2\mathbb{Z}}\phi\left(\frac{k-1}{T}\right)e^{\pm i(k-1)(\theta_{t_1,n}-\theta_{t_2,n})}D(t_1,n)D(t_2,n) \\ &-\sum_{t_1\neq -t_2}\sum_{k\in2\mathbb{Z}}\phi\left(\frac{k-1}{T}\right)e^{\pm i(k-1)(\theta_{t_1,n}+\theta_{t_2,n})}D(t_1,n)D(t_2,n) \\ =&D+OD,\end{aligned}$$

where the diagonal part D comes from $\theta_{t_1,n} + \theta_{t_2,n} = \pi$ and from $\theta_{t_1,n} = \theta_{t_2,n}$. Note from Lemma 4.4 that, unless it is an integer multiple of π, $\theta_{t_1,n} \pm \theta_{t_2,n}$ are contained in $\left[\frac{1}{2\sqrt{n}}, \pi - \frac{1}{2\sqrt{n}}\right]$ modulo π. Therefore we have

$$\begin{aligned} OD &\ll \sup_{\theta \in \left[\frac{1}{2\sqrt{n}}, \pi - \frac{1}{2\sqrt{n}}\right]} \left| \sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) e^{i(k-1)\theta} \right| \sum_{t_1,t_2} |D(t_1,n)D(t_2,n)| \\ &\ll_N \sup_{\theta \in \left[\frac{1}{2\sqrt{n}}, \pi - \frac{1}{2\sqrt{n}}\right]} \left| \sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) e^{i(k-1)\theta} \right| \sigma_1(n)^2. \end{aligned}$$

Lemma 4.5. *Let $T \geq \sqrt{n}$. Then for any θ that satisfies $\theta \in \left[\frac{1}{2\sqrt{n}}, \pi - \frac{1}{2\sqrt{n}}\right]$, we have*

$$\sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) e^{i(k-1)\theta} = 0,$$

and as a result

$$\sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) |B_2(n,k,N)|^2 = 4 \sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \sum_{t^2 < 4n} |D(t,n)|^2.$$

Proof. From Poisson summation formula we have

$$\sum_{k \in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) e^{i(k-1)\theta} = \sum_{n \in \mathbb{Z}} \phi\left(\frac{2n-1}{T}\right) e^{i(2n-1)\theta} = \sum_{m \in \mathbb{Z}} \Phi(m), \tag{4.5}$$

where

$$\begin{aligned} \Phi(y) &= \int_{-\infty}^{\infty} \phi\left(\frac{2x-1}{T}\right) e^{i(2x-1)\theta} e^{-2\pi i x y} dx \\ &= \frac{1}{2} e^{-\pi i y} \int_{-\infty}^{\infty} \phi\left(\frac{x}{T}\right) e^{ix(\theta - \pi y)} dx \\ &= \frac{T}{2} e^{-\pi i y} \int_{-\infty}^{\infty} \phi(x) e^{ixT(\theta - \pi y)} dx \\ &= \frac{T}{2} e^{-\pi i y} \hat{\phi}\left(\frac{T(\pi y - \theta)}{2\pi}\right). \end{aligned}$$

In the last expression, for any $m \in \mathbb{Z}$, we have

$$\left| \frac{T(\pi m - \theta)}{2\pi} \right| \geq \frac{1}{4\pi},$$

and since $\hat{\phi}$ is assumed to be supported in $[-1/100, 1/100]$, the right hand side of (4.5) vanishes. □

We are ready to prove:

Lemma 4.6. *Let ϕ be a positive even rapidly decaying function whose Fourier transform $\hat{\phi}$ is supported in $[-1/100, 1/100]$. Let $T \geq \sqrt{n}$. Then we have*

$$\sum_{k>0, k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \left|\mathrm{Tr}\ \mathcal{T}_n^* - \frac{k-1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}}\right|^2$$
$$= 2\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \sum_{t^2<4n} |D(t,n)|^2 - \phi\left(\frac{1}{T}\right)\frac{\sigma_1(n)^2}{n} + O\left(n^{\frac{1}{2}+\epsilon}\right), \quad (4.6)$$

where $\delta_{\sqrt{n}} = 1$ if n is square, and 0 otherwise.

Proof. The summand agrees with $B_2(n,k,N)$ unless $k = 2$, so from the computation given above, we have

$$\sum_{k>0, k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \left|\mathrm{Tr}\ \mathcal{T}_n^* - \frac{k-1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}}\right|^2$$
$$= 2\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{T}\right) \sum_{t^2<4n} |D(t,n)|^2 + \phi\left(\frac{1}{T}\right)\left(\left|\mathrm{Tr}\ \mathcal{T}_n^* - \frac{1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}}\right|^2 - |B_2(n,2,N)|^2\right).$$

By Lemma 4.2, for $N > 1$ we have

$$B_2(n,2,N) = \mathrm{Tr}\ \mathcal{T}_n^* - \frac{1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}} - \mu(N)\frac{\sigma_1(n)}{\sqrt{n}}$$

By the Ramanujan bound on the weight 2 modular forms, we have

$$\mathrm{Tr}\ \mathcal{T}_n^* \ll_{\epsilon,N} n^{\epsilon}.$$

Hence,

$$\left|\mathrm{Tr}\ \mathcal{T}_n^* - \frac{1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}}\right|^2 - |B_2(n,2,N)|^2 = -\frac{\sigma_1(n)^2}{n} + O\left(n^{\frac{1}{2}+\epsilon}\right).$$

This concludes our lemma. □

4.2 Arithmetic sum

In this section, we estimate the arithmetic part of (4.6):

$$\sum_{t^2<4n} |D(t,n)|^2.$$

Theorem 4.7. *Assume for simplicity that n is odd. Then we have*

$$\sqrt{n} \ll_N \sum_{t^2<4n} |D(t,n)|^2 \ll_N \sqrt{n}(\log n)^2(\log\log n)^4.$$

Recall that

$$D(t,n) = \frac{i}{2\sqrt{4n-t^2}} \sum_f h_w\left(\frac{t^2-4n}{f^2}\right) \tilde{\mu}(t,f,n,N),$$

where the inner sum runs over all positive divisors of t^2-4n such that $(t^2-4n)/f^2 \in \mathbb{Z}$ is congruent to 0 or $1 \pmod 4$. $\tilde{\mu}(t,f,n,N)$ is given by

$$\tilde{\mu}(t,f,n,N) = \sum_{d|N} \sigma_0(N/d)\mu(N/d)\mu(t,f,n,d),$$

and $\mu(t,f,n,N)$ is given by

$$\mu(t,f,n,N) = \frac{\psi(N)}{\psi(N/N_f)} M(t,n,NN_f),$$

where $N_f = \gcd(N,f)$ and $M(t,n,K)$ denotes the number of solutions of the congruence $x^2 - tx + n \equiv 0 \pmod K$.

Denote by $H(n) = \sum_{f^2|n} h_w(-n/f^2)$ the Hurwitz class number. For the upper bound of the arithmetic sum, we write

$$\sum_{t^2<4n} D(t,n)^2 \ll_N \sum_{t^2<4n} \frac{1}{4n-t^2} H^2\left(t^2-4n\right), \tag{4.7}$$

using the estimate $\mu(t,f,n,N) \ll_N 1$.

For the lower bound, we first prove the following:

Lemma 4.8. *Assume that n is odd. Fix an odd integer $0 < n_0 < 2N$ such that $\left(\frac{n_0^2-4n}{p}\right) = -1$ for all odd primes $p|N$. Then $\tilde{\mu}(t,f,n,N) = \sigma_0(N)\mu(N)$ for any $t \equiv n_0 \pmod{2N}$.*

Proof. For such t, we have $\mu(t,f,n,d) = 0$ unless $d = 1$ or 2. So for an odd N,

$$\tilde{\mu}(t,f,n,N) = \sigma_0(N)\mu(N).$$

When N is even, we have

$$\tilde{\mu}(t,f,n,N) = \sigma_0(N)\mu(N) + \sigma_0(N/2)\mu(N/2)\mu(t,f,n,2) = \sigma_0(N/2)\mu(N/2)(\mu(t,f,n,2)-2),$$

where

$$\mu(t,f,n,2) = M(t,n,2),$$

because $\gcd(N,f)|\gcd(N,t^2-4n) = 1$. Then $M(t,n,2) = 0$ since both n and t are assumed to be odd, and therefore

$$\tilde{\mu}(t,f,n,N) = \sigma_0(N/2)\mu(N/2) \times (-2) = \sigma_0(N)\mu(N). \qquad \square$$

Using this lemma, we bound the arithmetic sum from the below under the assumption that n is odd as follows:

$$\sum_{t^2<4n} D(t,n)^2 \geq \sum_{\substack{t^2<4n \\ t\equiv n_0 \pmod{2N}}} D(t,n)^2 = \sum_{\substack{t^2<4n \\ t\equiv n_0 \pmod{2N}}} \frac{\sigma_0(N)^2}{4n-t^2} H^2\left(t^2-4n\right)$$
$$\geq \sum_{\substack{t^2<4n \\ t\equiv n_0 \pmod{2N}}} \frac{1}{4n-t^2} H^2\left(t^2-4n\right). \quad (4.8)$$

We now handle the right hand sides of (4.7) and (4.8) separately.

4.2.1 Upper bound

We first recall from [Coh75], that for $n = Df^2 < 0$,

$$H(n) = \frac{h(D)}{w(D)} \sum_{d|f} \mu(d)\chi_D(d)\sigma_1\left(\frac{f}{d}\right). \quad (4.9)$$

where $2w(D)$ is the number of units in $\mathbb{Q}(\sqrt{-D})$. Note that

$$\sum_{d|f} \mu(d)\chi_D(d)\sigma_1\left(\frac{f}{d}\right)$$

is multiplicative in f, and

$$\sum_{d|p^k} \mu(d)\chi_D(d)\sigma_1\left(\frac{p^k}{d}\right) = \sigma_1\left(p^k\right) - \chi_D(p)\sigma_1\left(p^{k-1}\right) \leq \sigma_1\left(p^k\right) + \sigma_1\left(p^{k-1}\right) < \left(1+\frac{1}{p}\right)\sigma_1\left(p^k\right).$$

Therefore

$$\sum_{d|f} \mu(d)\chi_D(d)\sigma_1\left(\frac{f}{d}\right) < \sigma_1(f)\prod_{p|f}\left(1+\frac{1}{p}\right) \ll f(\log\log f)^2$$

where we used Grönwall's theorem in the last inequality. Using a standard upper bound $h(D) \ll \sqrt{D}\log D$ yields:

$$H(n) \ll \sqrt{D} f \log D (\log\log f)^2 \ll \sqrt{n}\log n(\log\log n)^2.$$

Now we apply this to (4.7) to conclude that

$$\sum_{t^2<4n} D(t,n)^2 \ll_N \sqrt{n}(\log n)^2(\log\log n)^4.$$

4.2.2 Lower bound

From Cauchy-Schwarz inequality,

$$\sum_{\substack{t^2<4n \\ t\equiv n_0 \pmod{2N}}} \frac{1}{4n-t^2} H^2\left(t^2-4n\right) \sum_{\substack{t^2<4n \\ t\equiv n_0 \pmod{2N}}} (4n-t^2) \geq \left(\sum_{\substack{t^2<4n \\ t\equiv n_0 \pmod{2N}}} H\left(t^2-4n\right)\right)^2,$$

we have

$$\sum_{t^2<4n,\ t\equiv n_0 \pmod{2N}} \frac{1}{4n-t^2} H^2\left(t^2-4n\right) \gg n^{-\frac{3}{2}} \left(\sum_{t^2<4n,\ t\equiv n_0 \pmod{2N}} H\left(t^2-4n\right)\right)^2.$$

Let $r_3(n)$ be the number of ways of representing n as a sum of three squares. Then Gauss' formula (see for instance, [KO99]) asserts that

$$\begin{aligned} r_3(n) &= 12H(-4n) && (n \equiv 1, 2 \pmod 4) \\ &= 24H(-n) && (n \equiv 3 \pmod 8) \\ &= r(n/4) && (n \equiv 0 \pmod 4) \\ &= 0. && (n \equiv 7 \pmod 8) \end{aligned}$$

Observe from (4.9) that if $4 \nmid m$, then

$$H(4^k m) = H(m)\left(\sigma_1(2^k) - \chi_D(2)\sigma_1(2^{k-1})\right)$$

and so

$$2^k H(m) \leq H(4^k m) \leq (2^{k+1} + 2^k - 2)H(m).$$

Combining all these, we conclude that

$$r_3(n) \leq 48H(-n).$$

Therefore we have

$$48 \sum_{t^2<4n,\ t\equiv n_0 \pmod{2N}} H\left(t^2-4n\right) \geq \sum_{t^2<4n,\ t\equiv n_0 \pmod{2N}} r_3(4n-t^2),$$

and observe that the last sum is equal to the number of elements in the following set:

$$A_{2N}(n) := \{4n = t^2 + x^2 + y^2 + z^2 \ :\ x, y, z, t \in \mathbb{Z},\ t \equiv n_0 \pmod{2N}\}. \quad (4.10)$$

Note that we assume that n is odd and N is fixed. Then by the result of Kloosterman [Klo27] who developed a version of the classical circle method with no minor arcs for quadratic forms in four variables we have

$$A_N(n) \gg n,$$

where the implicit constant in $\gg$ only depends on the fixed number N; see also the work of the second author [Sar15, Theorem 11] for the optimal strong approximation for quadratic forms in four and more variables which implies the above lower bound with an explicit dependence on N.

This completes the proof of the lower bound in Theorem 4.7.

4.3 Completion of proofs

In this section, we prove Theorem 1.3, 1.4, and Corollary 1.5.

Proof of Theorem 1.3. This is simple consequence of combining Lemma 4.6 and Theorem 4.7. □

Proof of Theorem 1.4. From Lemma 4.6 and Theorem 4.7, we see that LHS of (1.5) is

$$> c_N\sqrt{n} - \frac{\sigma_1(n)^2}{An\sqrt{n}},$$

for some constant $c_N > 0$ depending only on N. If $n = p^m$, then $\sigma_1(n) = \frac{p^{m+1}-1}{p-1} < 2p^m = 2n$, which implies that

$$c_N\sqrt{n} - \frac{\sigma_1(n)^2}{An\sqrt{n}} > \left(c_N - \frac{4}{A}\right)\sqrt{n}.$$

□

Proof of Corollary 1.5. We first note that from (61), [GJS99] that for $n = p^m$,

$$\left| \sum_{f\in B^*_{k,N}} \lambda_n(f) - |B^*_{k,N}|\frac{\delta_{\sqrt{n}}}{\sqrt{n}} \right| \le 2m^2|B^*_{k,N}|D(\mu^*_{k,N},\mu_p).$$

By (4.3), we have by $2x^2 + 2y^2 \ge (x+y)^2$,

$$2\left| \sum_{f\in B^*_{k,N}} \lambda_n(f) - |B^*_{k,N}|\frac{\delta_{\sqrt{n}}}{\sqrt{n}} \right|^2 \ge \left| \sum_{f\in B^*_{k,N}} \lambda_n(f) - \frac{k-1}{12}\varphi(N)\frac{\delta_{\sqrt{n}}}{\sqrt{n}} \right|^2 + O(n^{-1}).$$

Now from Theorem 1.4, we have

$$\frac{1}{\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right)} \sum_{k>0, k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right) m^4 |B^*_{k,N}|^2 D(\mu^*_{k,N},\mu_p)^2 \gg_N n^{\frac{1}{2}}, \tag{4.11}$$

where $K = A\sqrt{n}$ for some fixed sufficiently large A. Assume for contradiction that

$$D(\mu^*_{k,N},\mu_p) = o\left(\frac{1}{k^{\frac{1}{2}}\log^2 k}\right). \tag{4.12}$$

Then from (4.11), we have

$$n^{\frac{1}{2}} \ll \frac{1}{\sum_{k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right)} \sum_{k>0, k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right) m^4 |B^*_{k,N}|^2 D(\mu^*_{k,N}, \mu_p)^2$$
$$= o\left(\frac{1}{K} \sum_{k>0,k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right) m^4 \frac{k}{\log^4 k}\right).$$

However,

$$\frac{1}{K} \sum_{k>0,k\in 2\mathbb{Z}} \phi\left(\frac{k-1}{K}\right) m^4 \frac{k}{\log^4 k} \ll m^4 \frac{K}{\log^4 K} \ll \sqrt{n}$$

contradicting the assumption (4.12).

5 Appendix: By Simon Marshall

The purpose of this appendix is to illustrate the geometric origin of the transition behavior of the J-Bessel function, by recalling the derivation of the Petersson trace formula as a relative trace formula following [KL06]. Let $G = PSL(2,\mathbb{R})$, and $\Gamma = PSL(2,\mathbb{Z})$. Let $k \geq 2$ be even, and define $f \in C^\infty(G)$ by

$$f(g) = \frac{k-1}{4\pi} \frac{(2i)^k}{(-b+c+(a+d)i)^k}, \qquad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

This is the L^2-normalized matrix coefficient of the lowest weight vector in the weight k discrete series, see e.g. [KL06, Section 3.1]. We form the function

$$K_\Gamma(x,y) = \sum_{\gamma\in\Gamma} f(x^{-1}\gamma y)$$

on $(\Gamma\backslash G)^2$. The Petersson trace formula can be proved by integrating $K_\Gamma(x,y)$ against characters over two horocycles on $\Gamma\backslash G$, and comparing the geometric and spectral expansions of K_Γ. More precisely, if $m, n \geq 1$ and we define

$$\sigma_n = \begin{pmatrix} k/4\pi n & \\ & 1 \end{pmatrix},$$

and likewise for σ_m, then the integral we wish to expand is

$$\int_0^1 \int_0^1 K_\Gamma\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}\sigma_n, \begin{pmatrix} 1 & y \\ & 1 \end{pmatrix}\sigma_m\right) e(-nx+my)dxdy.$$

Note that the heights we have chosen for our horocycles are optimal for picking up the n and mth Fourier coefficients on the spectral side.

We shall analyze the geometric side of this integral, which is

$$\int_0^1\int_0^1\sum_{\gamma\in\Gamma} f\left(\sigma_n^{-1}\begin{pmatrix}1 & -x\\ & 1\end{pmatrix}\gamma\begin{pmatrix}1 & y\\ & 1\end{pmatrix}\sigma_m\right)e(-nx+my)dxdy.$$

We break the sum over γ into double cosets $N\eta N$, which gives

$$\sum_{\eta\in N\backslash\Gamma/N}\int_0^1\int_0^1\sum_{\gamma\in N\eta N} f\left(\sigma_n^{-1}\begin{pmatrix}1 & -x\\ & 1\end{pmatrix}\gamma\begin{pmatrix}1 & y\\ & 1\end{pmatrix}\sigma_m\right)e(-nx+my)dxdy.$$

The contribution from the identity coset is

$$\int_0^1\int_0^1\sum_{\gamma\in N} f\left(\sigma_n^{-1}\begin{pmatrix}1 & -x\\ & 1\end{pmatrix}\gamma\begin{pmatrix}1 & y\\ & 1\end{pmatrix}\sigma_m\right)e(-nx+my)dxdy.$$

This vanishes unless $m=n$, in which case it is

$$\frac{4\pi n}{k}\int_{-\infty}^{\infty} f\left(\begin{pmatrix}1 & x\\ & 1\end{pmatrix}\right)dx,$$

i.e. the integral of f over the horocycle of height 1. If $\eta\neq 1$, there is no repetition among the elements $n_1\gamma n_2$, and so we may unfold the two integrals to obtain

$$I_\eta=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\left(\sigma_n^{-1}\begin{pmatrix}1 & -x\\ & 1\end{pmatrix}\eta\begin{pmatrix}1 & y\\ & 1\end{pmatrix}\sigma_m\right)e(-nx+my)dxdy. \tag{5.1}$$

This integral has a simple geometric meaning, as the integral of the kernel $K(x,y)=f(x^{-1}y)$ against characters over the two horocycles $N\sigma_n$ and $\eta N\sigma_m$. If we write $\eta=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ with $c>0$, then c corresponds to the index of summation on the geometric side of the Petersson formula. Moreover, the ranges $c<4\pi\sqrt{mn}/k$, $c=4\pi\sqrt{mn}/k$, and $c>4\pi\sqrt{mn}/k$ correspond to the oscillation, transition, and decay range of the J-Bessel function in the following way. We shall use the fact that the kernel K concentrates near the diagonal in $\mathbb{H}^2\times\mathbb{H}^2$. If $c<4\pi\sqrt{mn}/k$, then the two horocycles intersect transversally. The integrand is roughly supported on two balls of radius $k^{-\frac{1}{2}}$ and has magnitude k, and we have $I_\eta\sim 1$ as expected. If $c>4\pi\sqrt{mn}/k$ then the horocycles do not intersect, and $I_\eta\ll_N k^{-N}$. The case $c=4\pi\sqrt{mn}/k$ is where the horocycles are tangent, and so the integral is roughly supported on a ball of radius $k^{-\frac{1}{4}}$. One might expect $I_\eta\sim k^{\frac{1}{2}}$ from this, but in fact it is of size $k^{\frac{1}{6}}$. As we shall see below, the point is that the phase in (5.1) has a cubic degeneracy, and this (rather than the support) determines the size of I_η.

We now explicate the relation between I_η and the geometric side of the Petersson formula, and analyze the phase of the integral in the transition range. Writing $\eta=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$

with $c > 0$, the double coset $N\eta N$ is determined by c and the residue class of a mod c. Moreover, we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & a/c \\ & 1 \end{pmatrix} \begin{pmatrix} & -1/c \\ c & \end{pmatrix} \begin{pmatrix} 1 & d/c \\ & 1 \end{pmatrix}.$$

Changing variable in x and y by a translation, we have

$$I_\eta = e(-(na+md)/c) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f \left(\sigma_n^{-1} \begin{pmatrix} 1 & -x \\ & 1 \end{pmatrix} \begin{pmatrix} & -1/c \\ c & \end{pmatrix} \begin{pmatrix} 1 & y \\ & 1 \end{pmatrix} \sigma_m \right) e(-nx+my) dx dy.$$

Conjugating the matrices σ_n and σ_m though to the middle and changing variable gives

$$I_\eta = e(-(na+md)/c) \frac{k^2}{(4\pi)^2 mn} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f \left(\begin{pmatrix} 1 & -x \\ & 1 \end{pmatrix} \begin{pmatrix} & -4\pi n/kc \\ kc/4\pi m & \end{pmatrix} \begin{pmatrix} 1 & y \\ & 1 \end{pmatrix} \right)$$
$$e(k(-x+y)/4\pi) dx dy.$$

If we define

$$A(t,k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f \left(\begin{pmatrix} 1 & -x \\ & 1 \end{pmatrix} \begin{pmatrix} & -1/t \\ t & \end{pmatrix} \begin{pmatrix} 1 & y \\ & 1 \end{pmatrix} \right) e(k(-x+y)/4\pi) dx dy,$$

then the contribution from all η with a given value of c is

$$\frac{k^2}{(4\pi)^2 mn} S(m,n,c) A(kc/4\pi\sqrt{mn}, k).$$

In [KL06, Prop. 3.6], Knightly and Li calculate

$$A(t,k) = \frac{e^{-k} i^k 4\pi k^{k-1}}{2t(k-2)!} J_{k-1}(k/t) \sim \frac{k^{\frac{1}{2}}}{t} J_{k-1}(k/t),$$

which gives the required appearance of J_{k-1} on the geometric side.

One again sees the geometric meaning of $A(t,k)$. It is an integral of $K(x,y)$ against characters over a horocycle of height 1, and a horocycle corresponding to the point $0 \in \partial\mathbb{H}^2$ and whose highest point is at i/t^2. One therefore expects a transition of $A(t,k)$ at $t = 1$, and this corresponds to $c = 4\pi\sqrt{mn}/k$ as claimed above. We now write $A(1,k)$ as an oscillatory integral (with non-imaginary phase function), and examine its critical point. Using our formula for f gives

$$\begin{aligned} f \left(\begin{pmatrix} 1 & -x \\ & 1 \end{pmatrix} \begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \begin{pmatrix} 1 & y \\ & 1 \end{pmatrix} \right) &= f \left(\begin{pmatrix} -x & -1-xy \\ 1 & y \end{pmatrix} \right) \\ &= \frac{k-1}{4\pi} i^k (1 + xy/2 + i(y-x)/2)^{-k} \\ &= \frac{k-1}{4\pi} i^k \exp(-k \log(1 + xy/2 + i(y-x)/2)). \end{aligned}$$

Computing the Taylor expansion of $\log(1+xy/2+i(y-x)/2)$ gives

$$\begin{aligned}\log(1+xy/2+i(y-x)/2) &= xy/2+i(y-x)/2-\frac{1}{2}\left(-(y-x)^2/4+ixy(y-x)/2\right)\\ &\quad -i(y-x)^{\frac{3}{2}}4+O(x^4+y^4)\\ &=(x+y)^2/8+i\left((y-x)/2-xy(y-x)/4-(y-x)^{\frac{3}{2}}4\right)+O(x^4+y^4).\end{aligned}$$

Substituting this into $A(1,k)$ gives

$$A(1,k)=\frac{k-1}{4\pi}i^k\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(-k(x+y)^2/8+ik\left(xy(y-x)/4+(y-x)^3/24\right)+kO(x^4+y^4))dxdy.$$

The leading term $-k(x+y)^2/8$ in the phase truncates the integral to the line $x+y=0$ at scale $k^{-\frac{1}{2}}$, and along this line the leading term in the phase is imaginary with a cubic degeneracy. This is why one has $A(1,k)\sim k^{\frac{1}{6}}$ compared to $A(t,k)\sim 1$ for $t<1$. □

References

[AL70] A. O. L. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[Bér77] Pierre H. Bérard. On the wave equation on a compact Riemannian manifold without conjugate points. *Math. Z.*, 155(3):249–276, 1977.

[Ber85] M. V. Berry. Semiclassical theory of spectral rigidity. *Proc. Roy. Soc. London Ser. A*, 400(1819):229–251, 1985.

[Ber86] M. V. Berry. Fluctuations in numbers of energy levels. In *Stochastic processes in classical and quantum systems (Ascona, 1985)*, volume 262 of *Lecture Notes in Phys.*, pages 47–53. Springer, Berlin, 1986.

[Coh75] Henri Cohen. Sums involving the values at negative integers of L-functions of quadratic characters. *Math. Ann.*, 217(3):271–285, 1975.

[Del74] Pierre Deligne. La conjecture de Weil. I. *Inst. Hautes Études Sci. Publ. Math.*, (43):273–307, 1974.

[DLMF] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.18 of 2018-03-27. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[ET48] P. Erdös and P. Turán. On a problem in the theory of uniform distribution. I. *Nederl. Akad. Wetensch., Proc.*, 51:1146–1154 = Indagationes Math. 10, 370–378 (1948), 1948.

[GJS99] Alex Gamburd, Dmitry Jakobson, and Peter Sarnak. Spectra of elements in the group ring of $\mathrm{SU}(2)$. *J. Eur. Math. Soc. (JEMS)*, 1(1):51–85, 1999.

[Ham98] Carol Hamer. A formula for the traces of the Hecke operators on certain spaces of newforms. *Arch. Math. (Basel)*, 70(3):204–210, 1998.

[HB96] D. R. Heath-Brown. A new form of the circle method, and its application to quadratic forms. *J. Reine Angew. Math.*, 481:149–206, 1996.

[Hej76] Dennis A. Hejhal. *The Selberg trace formula for* $\mathrm{PSL}(2, R)$. *Vol. I.* Lecture Notes in Mathematics, Vol. 548. Springer-Verlag, Berlin-New York, 1976.

[HL24] G. H. Hardy and E. Landau. The lattice points of a circle. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 105(731):244–258, 1924.

[Hör68] Lars Hörmander. The spectral function of an elliptic operator. *Acta Math.*, 121:193–218, 1968.

[ILS00] Henryk Iwaniec, Wenzhi Luo, and Peter Sarnak. Low lying zeros of families of L-functions. *Inst. Hautes Études Sci. Publ. Math.*, (91):55–131 (2001), 2000.

[KL06] A. Knightly and C. Li. A relative trace formula proof of the Petersson trace formula. *Acta Arithmetica*, 122:297–313, 2006.

[Klo27] H. D. Kloosterman. On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$. *Acta Math.*, 49(3-4):407–464, 1927.

[KO99] Winfried Kohnen and Ken Ono. Indivisibility of class numbers of imaginary quadratic fields and orders of Tate-Shafarevich groups of elliptic curves with complex multiplication. *Invent. Math.*, 135(2):387–398, 1999.

[MS09] M. Ram Murty and Kaneenika Sinha. Effective equidistribution of eigenvalues of Hecke operators. *J. Number Theory*, 129(3):681–714, 2009.

[Pet32] Hans Petersson. über die Entwicklungskoeffizienten der automorphen Formen. *Acta Math.*, 58(1):169–215, 1932.

[PT02] Y. Petridis and J. A. Toth. The remainder in Weyl's law for random two-dimensional flat tori. *Geom. Funct. Anal.*, 12(4):756–775, 2002.

[Rud05] Zeév Rudnick. A central limit theorem for the spectrum of the modular domain. *Ann. Henri Poincaré*, 6(5):863–883, 2005.

[Sar90] Peter Sarnak. *Some Applications of Modular Forms*. Cambridge Tracts in Mathematics. Cambridge University Press, 1990.

[Sar95] Peter Sarnak. Arithmetic quantum chaos. In *The Schur lectures (1992) (Tel Aviv)*, volume 8 of *Israel Math. Conf. Proc.*, pages 183–236. Bar-Ilan Univ., Ramat Gan, 1995.

[Sar97] Peter Sarnak. Values at integers of binary quadratic forms. In *Harmonic analysis and number theory (Montreal, PQ, 1996)*, volume 21 of *CMS Conf. Proc.*, pages 181–203. Amer. Math. Soc., Providence, RI, 1997.

[Sar02] Peter Sarnak. *Letter to Z. Rudnick on multiplicities of eigenvalues for the modular surface*, October 2002. https://publications.ias.edu/sarnak/paper/500.

[Sar15] N. T Sardari. Optimal strong approximation for quadratic forms. *ArXiv e-prints*, October 2015.

[Ser97] Jean-Pierre Serre. Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p. *J. Amer. Math. Soc.*, 10(1):75–102, 1997.

[Shi75] Goro Shimura. On the holomorphy of certain Dirichlet series. *Proc. London Math. Soc. (3)*, 31(1):79–98, 1975.

[SST16] Peter Sarnak, Sug Woo Shin, and Nicolas Templier. Families of L-functions and their symmetry. In *Families of automorphic forms and the trace formula*, Simons Symp., pages 531–578. Springer, [Cham], 2016.

[Wey11] H. Weyl. Über die asymptotische Verteilung der Eigenwerte. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.*, 1911:110–117, 1911.