

PETERSSON AND KUZNETSOV TRACE FORMULAS

JIANYA LIU AND YANGBO YE

In memory of Professor Armand Borel

With whom we worked together

At the University of Hong Kong over a three-year period

This article is an introduction to the Petersson trace formula and Kuznetsov trace formula, both of which are now important, standard techniques in analytic number theory. To illustrate their applications to modular forms, we will explain their role in a proof of subconvexity bounds for Rankin-Selberg L -functions $L(s, f \otimes g)$ on the critical line $\sigma = 1/2$, where here and throughout, we write $s = \sigma + it$ for $s \in \mathbb{C}$. Here f is a cusp form whose weight (if f is holomorphic) or Laplace eigenvalue (if f is nonholomorphic) tends to ∞ , while g is a fixed cusp form. There are nine sections:

1. Introduction
2. Poincaré series
3. The Petersson trace formula
4. Maass forms
5. The Kuznetsov trace formula
6. Automorphic L -functions
7. Number theoretical background
8. Subconvexity bounds for Rankin-Selberg L -functions
9. Kuznetsov formula in the proof of subconvexity bounds

1. Introduction. Let f be a holomorphic cusp form of weight k for a Hecke congruence subgroup, and $a_f(n)$ its n th Fourier coefficient. Petersson [14] in 1932 expressed the sum of $a_f(n)\bar{a}_f(m)$ over f in terms of the Bessel function J_{k-1} and Kloosterman sum. This Petersson formula actually gives the arithmetic information contained in the Kloosterman sum a spectral expansion over holomorphic cusp forms. In some sense, this is the first relative trace formula, about 24 years earlier than the trace formula which was proved by Selberg [21].

In 1980 Kuznetsov [10] established a similar spectral expansion over Maass cusp forms for the arithmetic information contained in Kloosterman sums. The spectral side also contains a continuous spectrum over Eisenstein series. This Kuznetsov formula is a relative trace formula of a new type.

The Petersson and Kuznetsov trace formulas have important applications in number theory. They are now standard techniques in the theory of modular

Partially supported by China NNSF Grant Number 10125101.

forms. In this article, we will give one example to illustrate their applications. We will outline their applications in a proof of subconvexity bounds for Rankin-Selberg L -functions. From this proof, one can see that in order to use the Petersson or Kuznetsov trace formula, we have to do an averaging argument which results in huge waste in estimation. Then we will see that the Petersson and Kuznetsov formulas are so powerful that even after this waste, we can still obtain nontrivial bounds.

This article has therefore two purposes in mind. The first is to give the reader an introduction to the Petersson (§3) and Kuznetsov (§5) trace formulas. We will prove the Petersson formula but only state the Kuznetsov formula without proof.

The second purpose is to illustrate the role played by the Kuznetsov trace formula in the proof of a subconvexity bound for Rankin-Selberg L -functions for cusp forms. We will take this opportunity to point out that using a recent result of Blomer [2], a subconvexity bound for Rankin-Selberg L -functions $L(s, f \otimes g)$ for cusp forms f and g can be improved to the same bound as in the case when f is Maass and g is holomorphic. We also want to point out that the subconvexity bound we proved in [11] for Maass cusp form f with Laplace eigenvalue $1/4 + k^2$ and a fixed holomorphic cusp form g is actually

$$(1.1) \quad L\left(\frac{1}{2} + it, f \otimes g\right) \ll k^{(15+2\theta)/16+\varepsilon},$$

where $\theta = 7/64$ and $\varepsilon > 0$ is arbitrarily small. See §9 for detailed computation.

The parameter θ in (1.1) is indeed the exponent in a bound toward the Generalized Ramanujan Conjecture (GRC) (see (8.2) below). The trivial bound has $\theta = 1/2$. The trivial, i.e., so-called convexity bound for $L(1/2 + it, f \otimes g)$ is $O(k)$. Therefore, any nontrivial $\theta < 1/2$ toward Ramanujan in (8.2) will give us a nontrivial, subconvexity bound for the Rankin-Selberg L -functions in (1.1). In this sense, the Ramanujan conjecture is irrelevant to subconvexity bounds for our Rankin-Selberg L -functions, according to our Theorems 8.1 and 8.4.

This is an important observation. Indeed, it is also believed (Sarnak [20]) that the Ramanujan conjecture is irrelevant to the Lindelöf Hypothesis which predicts

$$(1.2) \quad L\left(\frac{1}{2} + it, f \otimes g\right) \ll k^\varepsilon.$$

Recall that for an L -function $L(s, \pi)$ beyond $GL_2(\mathbb{Q})$, the only known nontrivial θ toward Ramanujan is $\theta = 1/2 - 1/(m^2 + 1)$, when π is an automorphic irreducible cuspidal representation of GL_m (Rudnick and Sarnak [16] and Luo, Rudnick, and Sarnak [12]). Any subconvexity bound for L -functions beyond $GL_2(\mathbb{Q})$ might therefore require an estimate like ours. This is indeed the meaning of our theorems, while a mere improvement of a subconvexity bound is not deemed important.

In this article we will not give definitions and basic properties of holomorphic forms. These will be covered by a separate article in this volume. We will not give a complete proof of Kuznetsov trace formula either. For a detailed proof see e.g. Iwaniec [4] and Motohashi [13].

2. Poincaré series. Fix a Hecke congruence subgroup

$$\Gamma = \Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}); c \equiv 0 \pmod{N} \right\}.$$

When $N = 1$, $\Gamma_0(1) = SL_2(\mathbb{Z})$ is the modular group. One can prove that

$$[\Gamma_0(1) : \Gamma_0(N)] = N \prod_{p|N} \left(1 + \frac{1}{p}\right).$$

We will consider holomorphic cusp forms of weight k for the Hecke congruence subgroup $\Gamma_0(N)$; denote their space by $S_k(\Gamma_0(N))$. There is a dimension formula for this space. It is known that

$$\dim S_k(\Gamma_0(N)) \asymp kN \prod_{p|N} \left(1 + \frac{1}{p}\right).$$

In particular when $N = 1$,

$$\dim S_k(\Gamma_0(1)) = \frac{k}{12} + O(1).$$

The actual dimension formula depends on different cases of the weight k : (i) k even, (ii) k odd, and (iii) $2k$ odd. The last case is the so-called half-integral weight, which can only happen when $4|N$. For details, see Sarnak [17] and Iwaniec [5].

We will in this section introduce a special kind of holomorphic cusp forms which generate $S_k(\Gamma_0(N))$: Poincaré series. First we need an extension of the Legendre symbol $\left(\frac{a}{d}\right)$ (Shimura [23]): For b odd, we set

- (i) $\left(\frac{a}{b}\right) = 0$ if $(a, b) \neq 1$;
- (ii) $\left(\frac{a}{b}\right)$ is the Legendre symbol if b is an odd prime;
- (iii) for fixed $b > 0$, $\left(\frac{a}{b}\right)$ as a function of a is a character mod b ;
- (iv) for fixed $a \neq 0$, $\left(\frac{a}{b}\right)$ as a function of b is a character mod $4a$;
- (v) $\left(\frac{a}{-1}\right) = 1$ if $a > 0$, and $= -1$ if $a < 0$;
- (vi) $\left(\frac{0}{\pm 1}\right) = 1$.

Define

$$j(\gamma, z) = \left(\frac{c}{d}\right) \varepsilon_d^{-1} (cz + d)^{1/2}$$

for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, where for odd d , $\varepsilon_d = 1$ if $d \equiv 1 \pmod{4}$, and $= i$ if $d \equiv 3 \pmod{4}$.

Denote $\Gamma_\infty = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \right\} \subset \Gamma_0(N)$. The Poincaré series is defined by

$$(2.2) \quad P_m(z, k) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma, z)^{-2k} e(m\gamma z),$$

where $e(z) = e^{2\pi iz}$.

We note that for $m = 0$, $P_0(z, k)$ is a holomorphic Eisenstein series. Consequently for general m , the Poincaré series $P_m(z, k)$ is dominated by the corresponding non-holomorphic Eisenstein series and hence converges absolutely.

Note that the cusps for $\Gamma_0(1)$ are ∞ and all the rational numbers, and they are all equivalent to ∞ . For $N > 1$, the cusps are still ∞ and all rational numbers, but not all of them are now equivalent to ∞ under $\Gamma_0(N)$. Indeed, the cusps equivalent to ∞ are $-d/c$ with $(c, d) = 1$ and $c \equiv 0 \pmod{N}$.

One can prove (cf. [17]) that the Poincaré series $P_m(z, k)$ vanishes at every cusp not equivalent to ∞ . When $m > 0$, $P_m(z, k)$ also vanishes at every cusp equivalent to ∞ , and hence is a cusp form in $S_k(\Gamma_0(N))$. The exceptional case is $P_0(z, k)$: It is not a cusp form but an Eisenstein series as we pointed out previously.

THEOREM 2.1. *Poincaré series $P_m(z, k)$, $m > 0$, span the space $S_k(\Gamma_0(N))$.*

Our main goal in this section is to prove Theorem 2.1. Before we do that, a natural question to ask is how to find a basis of $S_k(\Gamma_0(N))$ consisting of $P_m(z, k)$, $m > 0$. For the case of modular group, it is known that $P_m(z, k)$, $1 \leq m \leq \dim S_k(\Gamma_0(N))$, form a basis. For $N > 1$, this question is still open.

Proof of Theorem 2.1. Let $f, g \in S_k(\Gamma_0(N))$ and $z = x + iy \in \mathbb{H} = \{z \in \mathbb{C} \mid y > 0\}$. Define the Petersson inner product

$$(2.4) \quad \langle f, g \rangle = \int_{\Gamma \backslash \mathbb{H}} f(z) \bar{g}(z) y^k \frac{dx dy}{y^2}.$$

This integral is well-defined because the integrand $f(z) \bar{g}(z) y^k$ and the measure $dx dy / y^2$ are both invariant under the action of Γ . This integral converges absolutely because cusp forms are rapidly decreasing at each cusp.

For $m > 0$ we compute

$$(2.5) \quad \langle P_m, f \rangle = \int_{\Gamma \backslash \mathbb{H}} P_m(z, k) \bar{f}(z) y^k \frac{dx dy}{y^2}.$$

Substituting the expansion of $P_m(z, k)$ in (2.2) into (2.5), we get

$$(2.6) \quad \begin{aligned} \langle P_m, f \rangle &= \int_{\Gamma \backslash \mathbb{H}} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma, z)^{-2k} e(m\gamma z) \bar{f}(z) y^k \frac{dx dy}{y^2} \\ &= \int_{\Gamma_\infty \backslash \mathbb{H}} e(mz) \bar{f}(z) y^k \frac{dx dy}{y^2}. \end{aligned}$$

Now we use the Fourier expansion of f :

$$(2.7) \quad f(z) = \sum_{n \geq 1} a_f(n) e(nz).$$

Then (2.6) becomes

$$\langle P_m, f \rangle = \int_0^\infty \int_0^1 e(mz) \sum_{n \geq 1} \bar{a}_f(n) \bar{e}(nz) y^k \frac{dx dy}{y^2}.$$

The inner integral on $[0, 1]$ is nonzero only if $n = m$. Consequently

$$(2.8) \quad \langle P_m, f \rangle = \bar{a}_f(m) \int_0^\infty e^{-4\pi m y} y^k \frac{dx dy}{y^2} = \frac{\bar{a}_f(m)}{(4\pi m)^{k-1}} \Gamma(k-1).$$

From (2.8) we know that if a cusp form f is orthogonal to $P_m(z, k)$ for a given $m > 0$, then $a_f(m) = 0$. If f is orthogonal to every $P_m(z, k)$ for $m > 0$, then $a_f(m) = 0$ for every $m > 0$, and hence $f = 0$. This proves that $P_m(z, k)$, $m > 0$, span $S_k(\Gamma_0(N))$. \square

3. The Petersson trace formula. Now we compute Fourier coefficients of Poincaré series $P_m(z, k)$:

$$\begin{aligned} \widehat{P}_m(n) &= \lim_{y \rightarrow 0^+} \int_0^1 P_m(z, k) e(-nx) dx \\ &= \lim_{y \rightarrow 0^+} \int_0^1 \sum_{\gamma \in \Gamma_\infty \setminus \Gamma} j(\gamma, z)^{-2k} e(m\gamma z) e(-nx) dx. \end{aligned}$$

We note that $\gamma \in \Gamma_\infty \setminus \Gamma$ implies $\gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $c \neq 0$. If $\gamma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then $c = 0$ and $d = 1$; hence $\left(\frac{c}{d}\right) = 1$, $\varepsilon_d = 1$, and $j\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, z\right) = 1$. Consequently

$$(3.1) \quad \widehat{P}_m(n) = \lim_{y \rightarrow 0^+} \left(\int_0^1 e(mz) e(-nx) dx \right.$$

$$(3.2) \quad \left. + \sum_{\gamma} \int_0^1 j(\gamma, z)^{-2k} e(m\gamma z) e(-nx) dx \right),$$

where the sum is taken over $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\infty \setminus \Gamma$ with $c \neq 0$ and $c \equiv 0 \pmod{N}$.

The integral in (3.1) equals

$$\int_0^1 e((m-n)x)e^{-2\pi my} dx = \delta_{m,n}e^{-2\pi my}.$$

For the expression in (3.2), we observe that

$$\begin{aligned} & \sum_{r \in \mathbb{Z}} \int_0^1 j\left(\gamma \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}, z\right)^{-2k} e\left(m\gamma \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} z - nx\right) dx \\ &= \int_{-\infty}^{\infty} j(\gamma, z)^{-2k} e(m\gamma z - nx) dx. \end{aligned}$$

Consequently (3.2) becomes

$$\sum_{\substack{d \pmod{c} \\ (c,d)=1}} \left(\frac{c}{d}\right)^{-2k} \varepsilon_d^{2k} \int_{-\infty}^{\infty} (cz+d)^{-k} e\left(m\frac{az+b}{cz+d} - nx\right) dx,$$

where the outer sum is taken over $c > 0$, $c \equiv 0 \pmod{N}$, and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_{\infty} \backslash \Gamma / \Gamma_{\infty}$.

From

$$\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+rc & b+rd \\ c & d \end{pmatrix}$$

we know that a and b are uniquely determined by c and d after taking the quotient on the left by Γ_{∞} . From

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} * & * \\ c & d+rc \end{pmatrix}$$

we know that d is determined uniquely modulo c after taking a quotient on the right by Γ_{∞} . By

$$\frac{az+b}{cz+d} = \frac{a}{c} - \frac{1}{c(cz+d)},$$

(3.2) equals

$$\begin{aligned} & \sum_{\substack{c>0 \\ N|c}} \sum_{\substack{d \pmod{c} \\ (c,d)=1}} \left(\frac{c}{d}\right)^{-2k} \varepsilon_d^{2k} \int_{-\infty}^{\infty} (cz+d)^{-k} e\left(\frac{ma}{c} - \frac{m}{c(cz+d)} - nx\right) dx \\ &= \sum_{\substack{c>0 \\ N|c}} c^{-k} \sum_{\substack{d \pmod{c} \\ (c,d)=1}} e\left(\frac{ma}{c}\right) \left(\frac{c}{d}\right)^{-2k} \varepsilon_d^{2k} \int_{-\infty}^{\infty} \left(z + \frac{d}{c}\right)^{-k} e\left(-\frac{m}{c^2(z+d/c)} - nx\right) dx. \end{aligned}$$

Changing variables from x to $x + d/c$, we get

$$(3.3) \quad \sum_{\substack{c>0 \\ N|c}} c^{-k} \sum_{\substack{d \pmod{c} \\ (c,d)=1}} e\left(\frac{ma+nd}{c}\right) \left(\frac{c}{d}\right)^{-2k} \varepsilon_d^{2k} \int_{-\infty}^{\infty} z^{-k} e\left(-\frac{m}{c^2 z} - nx\right) dx.$$

Taking $y \rightarrow 0$ and writing $\bar{d} = a$ with $d\bar{d} \equiv 1 \pmod{c}$, we can rewrite (3.3) and get

$$\begin{aligned} \widehat{P}_m(n) &= \delta_{m,n} + \sum_{\substack{c>0 \\ N|c}} c^{-k} \sum_{\substack{d \pmod{c} \\ (c,d)=1}} \left(\frac{c}{d}\right)^{-2k} \varepsilon_d^{2k} e\left(\frac{m\bar{d}+nd}{c}\right) \\ &\quad \times 2\pi \left(\frac{nc^2}{m}\right)^{(k-1)/2} i^{-k} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right). \end{aligned}$$

Define a generalized Kloosterman sum

$$K(m, n; c) = \sum_{\substack{d \pmod{c} \\ (c,d)=1}} \left(\frac{c}{d}\right)^{-2k} \varepsilon_d^{2k} e\left(\frac{m\bar{d}+nd}{c}\right).$$

Note that if the weight k is even, then

$$K(m, n; c) = \sum_{\substack{d \pmod{c} \\ (c,d)=1}} e\left(\frac{m\bar{d}+nd}{c}\right) = S(m, n; c),$$

the classical Kloosterman sum. Consequently for $y = 0$ and $z = x$, the n th Fourier coefficient of the m th Poincaré series equals

$$(3.4) \quad \widehat{P}_m(n) = \left(\frac{n}{m}\right)^{(k-1)/2} \left(\delta_{m,n} + 2\pi i^{-k} \sum_{\substack{c>0 \\ N|c}} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right) \frac{K(m, n; c)}{c} \right).$$

We can compute $\widehat{P}_m(n)$ in another way. Let \mathcal{F} be an orthonormal basis of $S_k(\Gamma_0(N))$ with respect the Petersson inner product defined in (2.4). Take any $f \in \mathcal{F}$ with Fourier expansion (2.7). Then from (2.8)

$$\langle P_m, f \rangle = \frac{\bar{a}_f(m)}{(4\pi m)^{k-1}} \Gamma(k-1)$$

we get a spectral expansion of the Poincaré series

$$(3.5) \quad P_m(z, k) = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \sum_{f \in \mathcal{F}} \bar{a}_f(m) f(z).$$

Computing the n th Fourier coefficient of the two sides of (3.5) we get

$$(3.6) \quad \widehat{P}_m(n) = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \sum_{f \in \mathcal{F}} \bar{a}_f(m) a_f(n).$$

Comparing the two expressions of $\widehat{P}_m(n)$ in (3.4) and (3.6), we complete the proof of the Petersson trace formula.

THEOREM 3.1. *Let \mathcal{F} be an orthonormal basis of $S_k(\Gamma_0(N))$. Then*

$$(3.7) \quad \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \sum_{f \in \mathcal{F}} \bar{a}_f(m) a_f(n) \\ = \left(\frac{n}{m}\right)^{(k-1)/2} \left(\delta_{m,n} + 2\pi i^{-k} \sum_{\substack{c>0 \\ N|c}} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right) \frac{K(m,n;c)}{c} \right).$$

We call the left side of (3.7) the spectral side and the right side the geometric side of the Petersson formula. Note that the spectral side is a finite sum.

4. Maass forms. The standard Laplace operator on the complex plane \mathbb{C} is defined by

$$\Delta^e = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2},$$

but on the upper half plane \mathbb{H} , the Laplace operator is given by

$$\Delta = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right).$$

The operator Δ is invariant under the action of $SL_2(\mathbb{R})$, i.e. for any $z \in \mathbb{H}$ and any element $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$,

$$(4.1) \quad (\Delta f)(gz) = \Delta(f(gz))$$

where $gz = (az + b)/(cz + d)$. It is well known that $SL_2(\mathbb{R})$ is generated by $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$, $b \in \mathbb{R}$. Thus in order to show (4.1), it is sufficient to check it for the generators.

DEFINITION 4.1. *A real-analytic function $f \neq 0$ is called a Maass form for the group $SL_2(\mathbb{Z})$ if it satisfies the following properties:*

(i) *For all $g \in SL_2(\mathbb{Z})$ and all $z \in \mathbb{H}$,*

$$f(gz) = f(z);$$

(ii) *f is an eigenfunction of the non-Euclidean Laplace operator,*

$$\Delta f = \lambda f,$$

where λ is an eigenvalue of Δ ;

(iii) *There exists a positive integer N such that*

$$f(z) \ll y^N, \quad y \rightarrow +\infty;$$

(iv) *If we want to consider square integrable Maass forms in $L^2(SL_2(\mathbb{R}) \backslash \mathbb{H})$, we further require*

$$\int_{\Gamma \backslash \mathbb{H}} |f(z)|^2 dz < +\infty,$$

where $dz = dx dy / y^2$ is the hyperbolic measure on \mathbb{H} .

DEFINITION 4.2. A Maass form f is said to be a cusp form if

$$\int_0^1 f\left(\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} z\right) db = \int_0^1 f(z+b) db = 0$$

holds for all $z \in \mathbb{H}$.

As an example of Maass forms, we consider non-analytic Eisenstein series

$$(4.2) \quad E^*(z, s) = \frac{\pi^{-s}}{2} \Gamma(s) \sum_{\substack{m, n \in \mathbb{Z} \\ (m, n) \neq (0, 0)}} \frac{y^s}{|mz + n|^{2s}}.$$

Let $s = \sigma + it$. Then we have

$$\frac{y^s}{|mz + n|^{2s}} \ll \frac{|y|^\sigma}{(m|y|)^{2\sigma}} \ll |y|^{-\sigma} m^{-2\sigma}.$$

Thus (4.2) absolutely converges in $\sigma > 1$. But it is not analytic in $z \in \mathbb{H}$.

We introduce $E(z, s)$ by the following formula

$$E^*(z, s) = E(z, s)\xi(s),$$

where $\xi(s) = \pi^{-s}\Gamma(s)\zeta(2s)$ and $\zeta(s)$ is the Riemann zeta-function. Then we have

$$E(z, s) = y^s + \frac{1}{2} \sum_{\substack{(c, d)=1 \\ c \neq 0}} \frac{y^s}{|cz + d|^{2s}} = \frac{1}{2} \sum_{\Gamma_\infty \backslash \Gamma} (\text{Im } gz)^s,$$

where Γ_∞ is the subgroup of $SL_2(\mathbb{Z})$ generated by $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

From the definition of $E(z, s)$, it is easy to see that $E(z, s)$ is invariant under $SL_2(\mathbb{Z})$. We can also prove that $E(z, s)$ is an eigenfunction of Δ in $\sigma > 1$. In fact, we have

$$\Delta y^s = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) (y^s) = s(1-s)y^s,$$

so that y^s is an eigenfunction of Δ with eigenvalue $\lambda = s(1-s)$. Since Δ is invariant under $SL_2(\mathbb{Z})$,

$$(\text{Im } gz)^s = \frac{y^s}{|cz + d|^{2s}}$$

is also an eigenfunction of Δ with eigenvalue $\lambda = s(1-s)$. Therefore

$$\Delta E(z, s) = s(1-s)E(z, s),$$

as required.

Now we introduce the Fourier expansion for Maass forms. Let $f(z)$ be a Maass form. Since $f(gz) = f(z)$ for $g = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL_2(\mathbb{Z})$, i.e.

$$f(z+1) = f(x+1+iy) = f(z) = f(x+iy),$$

$f(z)$ has the Fourier expansion

$$f(z) = \sum_{n \in \mathbb{Z}} c_n(y) e^{2\pi i n x} = \sum_{n \in \mathbb{Z}} c_n(y) e(n x).$$

Thus it remains to determine the coefficients $c_n(y)$.

By Definition 4.1, we have $\Delta f = \lambda f$, and hence $c_n(y)$ satisfies

$$-y^2 c_n'' + 4\pi^2 n^2 y^2 c_n = \lambda c_n.$$

If $n \neq 0$, this is a modified Bessel equation, which has the general solution

$$(4.3) \quad c_n(y) = a_n \sqrt{y} K_{i\nu}(2\pi|n|y) + b_n \sqrt{y} I_{i\nu}(2\pi|n|y).$$

Here $\nu = \sqrt{\lambda - 1/4}$, $K_{i\nu}$ is the modified Bessel function of the third kind which is rapidly decreasing when $y \rightarrow \infty$, and $I_{i\nu}(y)$ in (4.3) is the modified Bessel function of the first kind which is rapidly increasing. From (iii) in Definition 4.1, we see that b_n must equal zero.

If $n = 0$, the equation (4.3) has the solution $c_0(y) = y^{1/2+i\nu}$, where $\nu = \pm \sqrt{\lambda - 1/4}$. If $\lambda \geq 1/4$, then ν is real, and consequently,

$$\int_{\Gamma \backslash \mathbb{H}} |C_0(y)|^2 dz \geq \int_{-1/2}^{1/2} dx \int_1^\infty \frac{|y^{1+2i\nu}|}{y^2} dy \geq \int_1^\infty \frac{dy}{y} = \infty.$$

For a square-integrable Maass form $f \in L^2(\Gamma \backslash \mathbb{H})$, we then must have $c_0(y) = 0$. Therefore for $\lambda \geq 1/4$, a square-integrable Maass form must be a cusp form, and its Fourier expansion takes the form

$$f(z) = \sum_{n \neq 0} a_n \sqrt{y} K_{i\nu}(2\pi|n|y) e(n x).$$

It has been shown that the eigenvalue of any Maass cusp form for $SL_2(\mathbb{Z})$ satisfies $\lambda > 3\pi^2/2$, and the first eigenvalue can actually be computed: $\lambda_1 = 91.14\dots$. Selberg conjectured that the eigenvalue of any Maass cusp form for $\Gamma_0(N)$ satisfies $\lambda \geq 1/4$, and he proved $\lambda \geq 3/16$.

5. The Kuznetsov trace formula. We state without proof the following Kuznetsov formula. Let $\{f_j\}$ be an orthonormal basis of the space of Maass forms for the Hecke congruence subgroup $\Gamma_0(N)$. Denote by $1/4 + k_j^2$ the Laplace eigenvalue for f_j . Let

$$f_j(z) = (y \cosh \pi k_j)^{1/2} \sum_{n \neq 0} \lambda_j(n) K_{ik_j}(2\pi|n|y) e(n x)$$

be the Fourier expansion of f_j . Denote $\zeta_N(s) = \prod_{p|N} (1 - p^{-s})^{-1}$.

THEOREM 5.1 (KUZNETSOV [10]). *Let $h(r)$ be an even function of complex variable, which is analytic in $-\Delta \leq \text{Im } r \leq \Delta$ for some $\Delta \geq 1/4$. Assume in this region that $h(r) \ll r^{-2-\delta}$ for some $\delta > 0$ as $r \rightarrow \infty$. Then for any $n, m \geq 1$,*

$$(5.1) \quad \sum_{f_j} h(k_j) \lambda_j(m) \overline{\lambda_j(n)} + \frac{1}{\pi} \int_{\mathbb{R}} d_{ir}(n) d_{ir}(m) h(r) \frac{|\zeta_N(1+2ir)|^2}{|\zeta(1+2ir)|^2} dr$$

$$= \frac{\delta_{n,m}}{\pi^2} \int_{\mathbb{R}} \tanh(\pi r) h(r) dr$$

$$+ \frac{2i}{\pi} \sum_{c \geq 1} \frac{S(n, m; c)}{c} \int_{\mathbb{R}} J_{2ir} \left(\frac{4\pi\sqrt{mn}}{c} \right) \frac{h(r)r}{\cosh(\pi r)} dr,$$

where $d_v(n) = \sum_{ab=|n|} (a/b)^v$, and $S(m, n; c)$ is the classical Kloosterman sum.

Again the left side of (5.1) is the spectral side, and the right the geometric side of the trace formula. The integral on the spectral side represents the continuous spectrum of the Laplace operator. Indeed, the divisor function $d_{ir}(n)$ is the Fourier coefficients of Eisenstein series.

The integral on the left side converges absolutely. This is because first the zeta-function $\zeta(s)$ does not vanish on the line $\sigma = 1$, and secondly $|\zeta(1+2ir)| \geq c \log^{-2/3}(2+|r|)$ for some $c > 0$. The absolute convergence is then from the fact that $d_{ir}(n) \ll_n 1$ and $h(r) \ll r^{-2-\delta}$. The same bound for $h(r)$ also give us convergence of the first integral on the right side of (5.1).

Different from the Petersson formula, the sum of Fourier coefficients on the left side of (5.1) is an infinite sum. For a proof of the Kuznetsov trace formula, as well as a discussion of convergence, see Kuznetsov [10], Iwaniec [4], or Motohashi [13], Chapter 2.

The normalization for f_j is crucial here. Without suitable normalization of the Maass forms f_j , there would be no reason to have this sum convergent. In fact, the Kuznetsov trace formula (5.1) is only valid for Maass forms f_j which form an orthonormal basis of the space of the Maass cusp forms.

6. Automorphic L -functions. First let us recall the Riemann zeta-function

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

for $\sigma > 1$. The zeta-function has a functional equation and analytic continuation to the whole complex plane \mathbb{C} , with only a simple pole at $s = 1$. It has simple zeros at negative even integers, which are called trivial zeros. Other zeros of $\zeta(s)$ lie in the critical strip $0 \leq \sigma \leq 1$, and are called nontrivial zeros. The Riemann Hypothesis predicts that all nontrivial zeros of $\zeta(s)$ are on the line $\sigma = 1/2$.

Note that

$$\zeta(s)\left(1 - \frac{2}{2^s}\right) = \sum_{n \geq 1} \frac{1}{n^s} - 2 \sum_{n \geq 1} \frac{1}{(2n)^s} = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n^s}.$$

Since the series on the right side converges for real $s > 0$, we get an expression of $\zeta(s)$ for real $s > 0$

$$(6.1) \quad \zeta(s) = \left(1 - \frac{2}{2^s}\right)^{-1} \sum_{n \geq 1} \frac{(-1)^{n-1}}{n^s}.$$

By a general argument on the convergence of Dirichlet series, the right side of (6.1) also converges for all values of s with $\sigma > 0$ (cf. Titchmarsh [24], pp.16-17).

Next let us look at Dirichlet L -functions. Let N be a positive integer, and χ a character modulo N . Define $\chi(p) = 0$ if $p|N$. The Dirichlet L -function in this case is

$$L(s, \chi) = \sum_{n \geq 1} \frac{\chi(n)}{n^s} = \prod_p (1 - \chi(p)p^{-s})^{-1} = \prod_{p \nmid N} (1 - \chi(p)p^{-s})^{-1}$$

for $\sigma > 1$. $L(s, \chi)$ also has a functional equation and analytic continuation to \mathbb{C} . $L(s, \chi)$ is entire when χ is not a principal character. When χ is the principal character mod N , $L(s, \chi)$ is essentially the Riemann zeta function. Again, $L(s, \chi)$ has trivial zeros and nontrivial zeros. The Generalized Riemann Hypothesis claims that all nontrivial zeros have real part equal to $1/2$.

Now we turn to L -functions attached to cusp forms for the Hecke congruence subgroup $\Gamma_0(N)$. Let $f \in S_k(\Gamma_0(N))$ be a holomorphic cusp form with Fourier expansion

$$(6.2) \quad f(z) = \sum_{n \geq 1} n^{(k-1)/2} \lambda_f(n) e(nz).$$

Or let f be a Maass cusp form with Laplace eigenvalue $1/4 + \nu^2$ and Fourier expansion

$$(6.3) \quad f(z) = \sum_{n \neq 0} \sqrt{|y|} \lambda_f(n) K_{i\nu}(2\pi|n|y) e(nx).$$

in (6.2) and (6.3) we normalized f so that $\lambda_f(1) = 1$. Then the L -function attached to f is defined as

$$(6.4) \quad L(s, f) = \sum_{n \geq 1} \frac{\lambda_f(n)}{n^s}$$

$$(6.5) \quad = \prod_p (1 - \lambda_f(p)p^{-s} + \chi_0(p)p^{-2s})^{-1},$$

where χ_0 is the principal character modulo N . Note that we have normalized f by $\lambda_f(1) = 1$ so that its L -function has the first term equal to 1.

By a trivial bound for $\lambda_f(p)$, the sum in (6.4) and product in (6.5) converge absolutely for σ sufficiently large. Using the Rankin-Selberg method, one can further prove that they indeed converge absolutely for $\sigma > 1$.

The product in (6.5) equals

$$L(s, f) = \prod_{p \nmid N} (1 - \lambda_f(p)p^{-s} + \chi_0(p)p^{-2s})^{-1} \prod_{p|N} (1 - \lambda_f(p)p^{-s})^{-1}.$$

For $p \nmid N$, we define $\alpha_f(p)$ and $\beta_f(p)$ by

$$\alpha_f(p) + \beta_f(p) = \lambda_f(p), \quad \alpha_f(p)\beta_f(p) = 1.$$

For $p|N$, we set $\alpha_f(p) = \lambda_f(p)$ and $\beta_f(p) = 0$. Then the L -function can be written as

$$L(s, f) = \prod_p (1 - \alpha_f(p)p^{-s})^{-1} (1 - \beta_f(p)p^{-s})^{-1}.$$

The Ramanujan conjecture predicts that for unramified p , i.e., for $p \nmid N$, $|\alpha(p)| = |\beta(p)| = 1$. This conjecture was proved by Deligne for holomorphic cusp forms, but is still open for Maass cusp forms.

The function $L(s, f)$ has a functional equation and analytic continuation. It is indeed an entire function. We also know that it is nonzero on $\sigma = 1$. The Generalized Riemann Hypothesis predicts that all nontrivial zeros of $L(s, f)$ are on the line $\sigma = 1/2$.

The next L -function we will consider is the Rankin-Selberg L -function. Let f and g be two cusp forms, either holomorphic or Maass. Their Rankin-Selberg L -function is defined by

$$(6.6) \quad L(s, f \otimes g) = \prod_p (1 - \alpha_f(p)\alpha_g(p)p^{-s})^{-1} (1 - \alpha_f(p)\beta_g(p)p^{-s})^{-1} \\ \times (1 - \beta_f(p)\alpha_g(p)p^{-s})^{-1} (1 - \beta_f(p)\beta_g(p)p^{-s})^{-1}$$

$$(6.7) \quad = \zeta(2s) \sum_{n \geq 1} \frac{\lambda_f(n)\lambda_g(n)}{n^s}.$$

The product in (6.6) and the series in (6.7) are absolutely convergent for $\sigma > 1$. Because of our normalization of f and g , the first term in $L(s, f \otimes g)$ is again equal to 1.

The function $L(s, f \otimes g)$ has a functional equation and analytic continuation. When $f = g$, it has a simple pole at $s = 1$.

When $f \in S_k(\Gamma_0(N))$ and $g \in S_l(\Gamma_0(N))$ are both holomorphic cusp forms, the functional equation is

$$(6.8) \quad L(s, f \otimes g) = \gamma(s)L(1-s, f \otimes g),$$

where

$$(6.9) \quad \gamma(s) = (2\pi)^{4s-2} \frac{\Gamma(-s + (k+l)/2)\Gamma(1-s + (k-l)/2)}{\Gamma(s-1 + (k+l)/2)\Gamma(s + (k-l)/2)}.$$

We will also need an approximate functional equation when l is fixed but $k \rightarrow \infty$. Applying Stirling's formula

$$\Gamma(z) = e^{-z} e^{(z-1/2)\log z} (2\pi)^{1/2} \left(1 + \frac{z^{-1}}{12} + \frac{z^{-2}}{288} + O(z^{-3})\right)$$

to (6.9), we get

$$\gamma(s) = \left(\frac{16\pi^2}{(k+l-2)(k-l+2)}\right)^{2s-1} (1 + \eta_k(s))$$

where $\eta_k(s) \rightarrow 0$ when $k \rightarrow \infty$. In fact, one can prove that $\eta_k(s) \ll (1+|s|)^3/k^2$. Then the approximate functional equation is

$$L(s, f \otimes g) \sim \left(\frac{16\pi^2}{(k+l-2)(k-l+2)}\right)^{2s-1} L(1-s, f \otimes g).$$

Similar functional equations and approximate functional equations exist for f and g being both Maass, or one Maass and one holomorphic. Like in (6.1), we can use these approximate functional equation to get an approximation to the central value of $L(s, f \otimes g)$ at $s = 1/2$:

$$(6.10) \quad L\left(\frac{1}{2}, f \otimes g\right) = 2 \sum_{b \geq 1} \frac{1}{b} \sum_{a \geq 1} \frac{\lambda_f(a)\lambda_g(a)}{\sqrt{a}} V\left(\frac{ab^2}{X}\right) + O(k^\varepsilon)$$

where $X = (k+l-2)(k-l+2)/(16\pi^2)$ and

$$(6.11) \quad V(y) = \frac{1}{2\pi i} \int_{\sigma=1} G(s) y^{-s} \frac{ds}{s}$$

for a good function $G(s)$.

7. Number theoretical background. Gauss and Legendre conjectured that

$$(7.1) \quad \sum_{p \leq x} 1 \sim \frac{x}{\log x}, \quad x \rightarrow \infty,$$

which is called the prime number theorem. Riemann established a deep connection between the distribution of primes and distribution of zeros of the zeta-function. The following are some basic properties of the zeta-function.

The zeta-function satisfies the functional equation

$$\zeta(s) = \pi^{s-1/2} \frac{\Gamma\left(\frac{1-s}{2}\right)}{\Gamma\left(\frac{s}{2}\right)} \zeta(1-s).$$

This gives the analytic continuation of $\zeta(s)$ to the whole plane. $\{-2n\}_{n=1}^{\infty}$ comprises all the trivial zeros of $\zeta(s)$. There are infinitely many non-trivial zeros, all of which lie in the critical strip $0 \leq \sigma \leq 1$.

Let $\Lambda(n)$ be the von Mangoldt function, i.e. $\Lambda(n) = \log p$ for $n = p^a$ with $a \geq 1$, and zero otherwise. Then the prime number theorem is equivalent to

$$(7.2) \quad \sum_{n \leq x} \Lambda(n) \sim x.$$

The connection between the Riemann zeta-function and primes is given by the so-called explicit formula, which states that, for $2 \leq T \leq x$,

$$(7.3) \quad \sum_{n \leq x} \Lambda(n) = x - \sum_{|\gamma| \leq T} \frac{x^\rho}{\rho} + O\left(\frac{x \log^2 x}{T}\right),$$

where the sum is taken over nontrivial zeros $\rho = \beta + i\gamma$ of $\zeta(s)$ with $|\gamma| \leq T$.

If there is a constant $B < 1$ such that all the non-trivial zeros satisfy $\beta \leq B$, then the sum over ρ in (7.3) is

$$\ll \sum_{|\gamma| \leq T} \frac{x^\beta}{|\rho|} \leq x^B \sum_{|\gamma| \leq T} \frac{1}{|\rho|} \ll x^B \log^2 x,$$

and this gives the prime number theorem in the form

$$(7.4) \quad \sum_{n \leq x} \Lambda(n) = x + O(x^B \log^2 x).$$

However, we cannot establish such zero-free region as above for $\beta \leq B < 1$. In 1896, Hadamard and de la Vallée Poussin proved that $\zeta(1+it) \neq 0$. This weak information on the zeros of $\zeta(s)$ is sufficient to prove the prime number theorem (7.2). De la Vallée Poussin proved that $\zeta(s)$ is zero-free in the region

$$\sigma \geq 1 - \frac{c}{\log(|t| + 2)}.$$

The Riemann Hypothesis says that all non-trivial zeros of $\zeta(s)$ lie on the vertical line $\sigma = 1/2$. This is equivalent to

$$\sum_{n \leq x} \Lambda(n) = x + O(x^{1/2} \log^2 x).$$

By the Phragmén-Lindelöf method, we have

$$\zeta\left(\frac{1}{2} + it\right) \ll |t|^{1/4+\varepsilon}.$$

Consequently,

$$\zeta\left(\frac{1}{2} + it\right) \ll |t|^{1/4}$$

is called the convexity bound. In 1921, Weyl improved the bound to $|t|^{1/6+\varepsilon}$. Over the years, there have been many improvements. Under the Riemann Hypothesis, one can get

$$\zeta\left(\frac{1}{2} + it\right) \ll |t|^\varepsilon,$$

which is known as the Lindelöf Hypothesis.

8. Subconvexity bounds for Rankin-Selberg L -functions. Let f and g be two cusp forms, either holomorphic or Maass. Recall their Fourier expansions (6.2) and (6.3), and Rankin-Selberg L -function (6.6) and (6.7).

In the weight aspect of the holomorphic cusp form f , the convexity bound is, for fixed g and $t \in \mathbb{R}$,

$$L\left(\frac{1}{2} + it, f \otimes g\right) \ll k,$$

as the weight k of f going to infinity. In [19], Sarnak established the subconvexity bound

$$(8.1) \quad L\left(\frac{1}{2} + it, f \otimes g\right) \ll k^{576/601+\varepsilon}.$$

Kowalski, Michel, and Vanderkam [9] proved a subconvexity bound for a Rankin-Selberg L -function $L(1/2 + it, f \otimes g)$ as f varies over holomorphic new forms for $\Gamma_0(N)$ as the level N tends to ∞ , with fixed t, g , and the weight of f .

The following Theorem 8.1 was proved by the authors in [11], with the correct subconvexity bound obtainable using the techniques there. First we need a bound toward the Generalized Ramanujan Conjecture (GRC). In terms of representation theory, let π be an automorphic cuspidal representation of $GL_2(\mathbb{Q}_{\mathbb{A}})$ with unitary central character and local Hecke eigenvalues $\alpha_\pi^{(j)}(p)$ for $p < \infty$ and $\mu_\pi^{(j)}(\infty)$ for $p = \infty, j = 1, 2$. Then bounds toward the GRC are

$$(8.2) \quad \begin{aligned} |\alpha_\pi^{(j)}(p)| &\leq p^\theta \quad \text{for } p \text{ at which } \pi \text{ is unramified,} \\ |\operatorname{Re}(\mu_\pi^{(j)}(\infty))| &\leq \theta \quad \text{if } \pi \text{ is unramified at } \infty. \end{aligned}$$

These bounds were proved for $\theta = 1/4$ by Selberg and Kuznetsov [10], for $\theta = 1/5$ by Shahidi [22] and Luo, Rudnick, and Sarnak [12], for $\theta = 1/9$ by Kim and Shahidi [8], and most recently for $\theta = 7/64$ by Kim and Sarnak [7]. The GRC predicts that $\theta = 0$.

THEOREM 8.1 (LIU AND YE [11]). *Let g be a fixed holomorphic cusp form for $SL_2(\mathbb{Z})$. Let f be a Maass cusp form for $SL_2(\mathbb{Z})$ with Laplace eigenvalue $1/4 + k^2$. Then we have for any $\varepsilon > 0$ and $t \in \mathbb{R}$ that*

$$L\left(\frac{1}{2} + it, f \otimes g\right) \ll_{\varepsilon, t, g} k^{(15+2\theta)/16+\varepsilon},$$

where θ is given by (8.2) and we can take $\theta = 7/64$.

Here we replaced the bound in [11] by a correct bound here. Note that $L(s, f \otimes g)$ is indeed an L -function attached to a group representation of GL_4 over \mathbb{Q} . A proof of Theorem 8.1 needs GL_4 techniques. Sarnak [19] and the authors [11] accomplished this by using (i) the Petersson or Kuznetsov trace formula once, which is a GL_2 technique, and (ii) an estimation of a shifted sum $\sum_n \lambda_g(n) \bar{\lambda}_g(n+h)$ of Fourier coefficients of cusp form g for $h \neq 0$. This estimation is again a GL_2 technique, proved by Sarnak [19], Appendix.

More precisely, for $\nu_1, \nu_2 > 0$ let us define

$$(8.3) \quad \begin{aligned} & D_g(s, \nu_1, \nu_2, h) \\ &= \sum_{\substack{m, n > 0, \\ \nu_1 m - \nu_2 n = h}} \lambda_g(n) \bar{\lambda}_g(m) \left(\frac{\sqrt{\nu_1 \nu_2 m n}}{\nu_1 m + \nu_2 n} \right)^{l-1} (\nu_1 m + \nu_2 n)^{-s}, \end{aligned}$$

if g is a holomorphic cusp form of weight l , and

$$(8.4) \quad \begin{aligned} & D_g(s, \nu_1, \nu_2, h) \\ &= \sum_{\substack{m, n \neq 0, \\ \nu_1 m - \nu_2 n = h}} \lambda_g(n) \bar{\lambda}_g(m) \left(\frac{\sqrt{\nu_1 \nu_2 |m n|}}{\nu_1 |m| + \nu_2 |n|} \right)^{2il} (\nu_1 |m| + \nu_2 |n|)^{-s}, \end{aligned}$$

if g is a Maass cusp form with Laplace eigenvalue $1/4 + l^2$.

THEOREM 8.2 (SARNAK [19]). *Let g be a holomorphic cusp form of even weight l . For $\sigma > 1$, $\nu_1, \nu_2 > 0$, and $h \in \mathbb{Z}$, define $D_g(s, \nu_1, \nu_2, h)$ as above. Then assuming (8.2) for θ we have that $D_g(s)$ extends to a holomorphic function for $\sigma \geq 1/2 + \theta + \varepsilon$, for any $\varepsilon > 0$. Moreover, in this region it satisfies*

$$(8.5) \quad D_g(s, \nu_1, \nu_2, h) \ll_{g, \varepsilon} (\nu_1 \nu_2)^{-1/2+\varepsilon} |h|^{1/2+\theta+\varepsilon-\sigma} (1 + |t|)^3.$$

THEOREM 8.3 (SARNAK [19]). *Let g be a Maass cusp form of Laplace eigenvalue $1/4 + l^2$. For $\sigma > 1$, $\nu_1, \nu_2 > 0$, and $h \in \mathbb{Z}$ define $D_g(s, \nu_1, \nu_2, h)$ as above. Then assuming (8.2) for θ we have that $D_g(s)$ extends to a holomorphic function for $\sigma \geq 1/2 + \theta + \varepsilon$, for any $\varepsilon > 0$. Moreover, in this region it satisfies*

$$(8.6) \quad D_g(s, \nu_1, \nu_2, h) \ll_{g, \varepsilon} (\nu_1 \nu_2)^{-1/2+\varepsilon} |h|^{1/2+\theta+\varepsilon-\sigma} (1 + |t|)^3 + |h|^{1-\sigma}.$$

Here we can take $\theta = 7/64$ of Kim and Sarnak [7], and set $\sigma = 1/2 + \theta + \varepsilon$. Note that there is an extra term on the right side of (8.6). Recently Blomer [2] removed this extra term so that Theorem 8.3 now has the same bound as in (8.5). Because of this progress, our Theorem 8.1 now holds for the case of g being a Maass cusp form too.

THEOREM 8.4. *Let g be a fixed Maass cusp form for $\Gamma_0(N)$. Let f be a Maass cusp form for $\Gamma_0(N)$ with Laplace eigenvalue $1/4 + k^2$. Then we have for any $\varepsilon > 0$ and $t \in \mathbb{R}$ that*

$$L\left(\frac{1}{2} + it, f \otimes g\right) \ll_{\varepsilon, t, g, N} k^{(15+2\theta)/16+\varepsilon},$$

where θ is given by (8.2) and we can take $\theta = 7/64$.

The proof of Theorem 8.4 as well as Theorem 8.1 follows from the following theorem.

THEOREM 8.5. *Let g be a fixed holomorphic or Maass cusp form for $\Gamma_0(N)$. Let f be a Maass cusp form for $\Gamma_0(N)$ with Laplace eigenvalue $1/4 + k^2$. Then we have for any $\varepsilon > 0$ and $t \in \mathbb{R}$ that*

$$(8.7) \quad \sum_{K-L \leq k \leq K+L} \left| L\left(\frac{1}{2} + it, f \otimes g\right) \right|^2 \ll_{\varepsilon, t, g, N} (KL)^{1+\varepsilon},$$

for $K^{(7+2\theta)/8+\varepsilon} \leq L \leq K^{1-\varepsilon}$. Here θ is given by (8.2) and we can take $\theta = 7/64$.

By Weyl's law, which states that $\#\{k : 1/4 + k^2 \leq T\} \sim cT$, we have

$$\begin{aligned} & \#\{k : K - L \leq k \leq K + L\} \\ &= \#\{k : 1/4 + (K - L)^2 \leq 1/4 + k^2 \leq 1/4 + (K + L)^2\} \\ &\sim c((1/4 + (K + L)^2) - (1/4 + (K - L)^2)) \\ &= 4cKL. \end{aligned}$$

Thus from Theorem 8.5, the generalized Lindelöf hypothesis $L(1/2 + it, f \otimes g) \ll k^\varepsilon$ is true on average for $K - L \leq k \leq K + L$ with L in the range $K^{(7+2\theta)/8+\varepsilon} \leq L \leq K^{1-\varepsilon}$.

Now we derive Theorem 8.4 from Theorem 8.5. We take only one term from the left side of (8.7) and get

$$\left| L\left(\frac{1}{2} + it, f \otimes g\right) \right|^2 \ll (KL)^{1+\varepsilon}$$

for $K - L \leq k \leq K + L$ and $K^{(7+2\theta)/8+\varepsilon} \leq L \leq K^{1-\varepsilon}$. Taking $K = k$ and $L = K^{(7+2\theta)/8+\varepsilon}$, we prove Theorem 8.4.

Now let us turn to an application of our subconvexity bounds in Theorem 8.1 or 8.4. Let f be a Maass Hecke eigenform for $\Gamma_0(N)$ with Laplace eigenvalue λ . Normalize f so that $\mu_f = |f(z)|^2 dx dy / y^2$ is a probability measure on $\Gamma_0(N) \backslash \mathbb{H}$, i.e., $\mu_f(\Gamma_0(N) \backslash \mathbb{H}) = 1$. The equidistribution conjecture (Rudnick and Sarnak [15]) predicts that

$$\mu_f \longrightarrow \text{vol}(\Gamma_0(N) \backslash \mathbb{H})^{-1} \frac{dx dy}{y^2}$$

as λ tends to infinity. According to Sarnak [18] and Watson [25], this conjecture would follow from a subconvexity bound for $L(1/2, f \otimes f \otimes g)$, where f is as above and g is a fixed Maass Hecke eigenform. If f is a CM form corresponding to a representation of the Weil group $W_{\mathbb{Q}}$, i.e., if $L(s, f) = L(s, \eta)$ for some grossencharacter η on a quadratic number field, then the triple Rankin-Selberg L -function can be factored as $L(s, F \otimes g) L(s, g \otimes \chi) L(s, g)$ for a fixed quadratic character χ of conductor N . Here F is a Maass cusp form with Laplace eigenvalue $1/4 + (2k)^2$, if $\lambda = 1/4 + k^2$. This way the equidistribution conjecture for CM Maass forms is reduced to a subconvexity estimate of $L(s, F \otimes g)$. Our Theorem 8.1 or 8.4 therefore implies the following theorem.

THEOREM 8.6. *The equidistribution conjecture is true for CM Maass forms.*

9. Kuznetsov formula in the proof of subconvexity bounds. Similar to (6.10), for the central value of L -function $L(s, f \otimes g)$ with f being a Maass cusp form, we have

$$\begin{aligned} (9.1) \quad L\left(\frac{1}{2}, f \otimes g\right) &= 2 \sum_{1 \leq b \leq X^{1/2+\varepsilon}} \frac{1}{b} \sum_{a \geq 1} \frac{\lambda_f(a) \lambda_g(a)}{\sqrt{a}} V\left(\frac{ab^2}{X}\right) \\ &\quad + O(k^\varepsilon) \\ &= \frac{2}{\sqrt{X}} \sum_{1 \leq b \leq X^{1/2+\varepsilon}} \sum_{a \geq 1} \lambda_f(a) \lambda_g(a) \\ &\quad \times V\left(\frac{a}{X/b^2}\right) \sqrt{\frac{X/b^2}{a}} + O(k^\varepsilon), \end{aligned}$$

where $X \asymp (k^2 - l^2)/(8\pi^2) \asymp k^2$ and V is the same as in (6.11). Here the outer sum is written as a finite sum because terms with $b > X^{1/2+\varepsilon}$ are negligible. The estimation of (9.1) is thus reduced to that of

$$(9.2) \quad S_Y(f) = \sum_{a \geq 1} \lambda_f(a) \lambda_g(a) H\left(\frac{a}{Y}\right),$$

where $Y = X/b^2$ and

$$H\left(\frac{a}{Y}\right) = V\left(\frac{a}{X/b^2}\right) \sqrt{\frac{X/b^2}{a}}.$$

By an argument of smooth dyadic subdivision we can assume that H is a smooth function of compact support in $[1, 2]$.

No one can get a nontrivial estimate for individual $S_Y(f)$ directly. In order to use the Kuznetsov trace formula, let us consider a smoothly weighted average

$$\begin{aligned} & \sum_{f_j} |S_Y(f_j)|^2 \left(h\left(\frac{K-k_j}{L}\right) + h\left(\frac{K+k_j}{L}\right) \right) \\ &= \sum_{m,n} \lambda_g(n) \bar{\lambda}_g(m) H\left(\frac{n}{Y}\right) \bar{H}\left(\frac{m}{Y}\right) \\ & \quad \times \sum_{f_j} \left(h\left(\frac{K-k_j}{L}\right) + h\left(\frac{K+k_j}{L}\right) \right) \lambda_j(n) \bar{\lambda}_j(m) \end{aligned}$$

Now the Kuznetsov trace formula (5.1) can be applied to the inner sum on the right side above. This will give us

$$(9.3) \quad \begin{aligned} & \sum_{f_j} |S_Y(f_j)|^2 \left(h\left(\frac{K-k_j}{L}\right) + h\left(\frac{K+k_j}{L}\right) \right) \\ &= - \sum_{m,n} \lambda_g(n) \bar{\lambda}_g(m) H\left(\frac{n}{Y}\right) \bar{H}\left(\frac{m}{Y}\right) \end{aligned}$$

$$(9.4) \quad \begin{aligned} & \times \int_{\mathbb{R}} \left(h\left(\frac{K-r}{L}\right) + h\left(\frac{K+r}{L}\right) \right) d_{ir}(n) d_{ir}(m) \frac{|\zeta_N(1+2ir)|^2}{|\zeta(1+2ir)|^2} dr \\ & + \sum_{m,n} \lambda_g(n) \bar{\lambda}_g(m) H\left(\frac{n}{Y}\right) \bar{H}\left(\frac{m}{Y}\right) \end{aligned}$$

$$(9.5) \quad \begin{aligned} & \times \frac{\delta_{n,m}}{\pi} \int_{\mathbb{R}} \tanh(\pi r) \left(h\left(\frac{K-r}{L}\right) + h\left(\frac{K+r}{L}\right) \right) r dr \\ & + 2i \sum_{m,n} \lambda_g(n) \bar{\lambda}_g(m) H\left(\frac{n}{Y}\right) \bar{H}\left(\frac{m}{Y}\right) \sum_{c \geq 1} \frac{S(n, m; c)}{c} \end{aligned}$$

$$(9.6) \quad \times \int_{\mathbb{R}} J_{2ir}\left(\frac{4\pi\sqrt{nm}}{c}\right) \left(h\left(\frac{K-r}{L}\right) + h\left(\frac{K+r}{L}\right) \right) \frac{r dr}{\cosh(\pi r)}.$$

We note that (9.3) is positive, while the expression in (9.4) without the leading negative sign is also positive. In fact, by change of variables, (9.4) equals

$$\begin{aligned} & -\frac{2L}{\pi} \sum_{n,m} \lambda_g(n) \bar{\lambda}_g(m) H\left(\frac{n}{Y}\right) \bar{H}\left(\frac{m}{Y}\right) \\ & \quad \times \int_{\mathbb{R}} d_{i(uL+K)}(n) d_{i(uL+K)}(m) \frac{h(u) du}{|\zeta(1+2i(uL+K))|^2} \\ &= -\frac{2L}{\pi} \int_{\mathbb{R}} \left| \sum_n \lambda_g(n) H\left(\frac{n}{Y}\right) d_{i(uL+K)}(n) \right|^2 \frac{h(u) du}{|\zeta(1+2i(uL+K))|^2}. \end{aligned}$$

Therefore estimation of (9.3) is reduced to estimation of (9.5) and (9.6).

The expression in (9.5) is bounded by

$$O\left(\sum_n |\lambda_g(n)|^2 \left|H\left(\frac{n}{Y}\right)\right|^2 \int_{\mathbb{R}} \left|h\left(\frac{K-r}{L}\right)\right| |r| dr\right).$$

By the Rankin-Selberg method, we can get

$$\sum_n |\lambda_g(n)|^2 \left|H\left(\frac{n}{Y}\right)\right|^2 \ll Y^{1+\varepsilon}.$$

On the other hand,

$$\int_{\mathbb{R}} \left|h\left(\frac{K-r}{L}\right)\right| |r| dr \ll \int_{K-cL}^{K+cL} r dr \ll KL.$$

Therefore (9.5) contributes at most $O(LKY^{1+\varepsilon})$.

Estimation of (9.6) is difficult. In fact, more than half of the pages in Liu and Ye [11], and in Sarnak [19] for the case of holomorphic Rankin-Selberg L -functions, are devoted to (9.6). The techniques include an identity from Bateman [1], vol. 1, p.59,

$$\left(\frac{J_{2ir}(x) - J_{-2ir}(x)}{\sinh(\pi r)}\right)^\wedge(y) = -i \cos(x \cosh(\pi y)),$$

and a Voronoi summation formula (see [19], p.435). The most important technique in estimation of (9.6) is Sarnak's bound for shifted sums of Fourier coefficients of g , as given in Theorems 8.2 and 8.3, and Blomer's improvement [2]. See [19] and [11] for details.

Here we want to point out that all of these techniques reduce the estimation of (9.6) to [11], (4.11):

$$(9.7) \quad \frac{Y^{(\mu-j+1)/2}}{L^{2\nu-1}} \sum_{0 \leq k \leq N} \frac{1}{k!} \left(\frac{L^2}{2\pi\sqrt{Y}}\right)^k \sum_{c \leq Y/(LK^{1-\varepsilon})} c^{j+k-\mu} \sum_{|h| < Y} |P(c, h, Y)|,$$

for $0 \leq 2\mu \leq \nu < N$ and $0 \leq j < 2N$, where according to [11], (4.21),

$$(9.8) \quad P(c, h, Y) \ll KY^{(\mu-j-k-1)/2+\sigma} \left(\left(\frac{|h|}{c}\right)^5 + \dots + 1\right) |h|^{1/2+\theta+\varepsilon-\sigma}.$$

By Blomer's recent results, (9.8) now is valid for holomorphic as well as Maass cusp form g . By [11], §4.12, we can take the innermost sum in (9.7) over $|h| \leq K^{2+\varepsilon} c^2/Y$. By Sarnak's Theorems 8.2 and 8.3, we can take $\sigma = 1/2 + \theta + \varepsilon$ in (9.8).

Consequently the innermost sum in (9.7) is bounded by

$$\begin{aligned}
& \sum_{|h| \leq K^{2+\varepsilon} c^2 / Y} |P(c, h, Y)| \\
& \ll \sum_{|h| \leq K^{2+\varepsilon} c^2 / Y} KY^{(\mu-j-k)/2+\theta+\varepsilon} \left(\left(\frac{|h|}{c} \right)^5 + 1 \right) |h|^\varepsilon \\
& \ll K^{13+\varepsilon} Y^{(\mu-j-k)/2-6+\theta+\varepsilon} c^{7+\varepsilon}.
\end{aligned}$$

Then the sum with respect to c in (9.7) is bounded by

$$\begin{aligned}
& K^{13+\varepsilon} Y^{(\mu-j-k)/2-6+\theta+\varepsilon} \sum_{c \leq Y/(LK^{1-\varepsilon})} c^{j+k-\mu+7+\varepsilon} \\
& \ll K^{13+\varepsilon} Y^{(\mu-j-k)/2-6+\theta+\varepsilon} \left(\frac{Y}{LK^{1-\varepsilon}} \right)^{j+k-\mu+8+\varepsilon} \\
& \ll \frac{K^{\mu-j-k+5+\varepsilon} Y^{(j+k-\mu)/2+2+\theta+\varepsilon}}{L^{j+k-\mu+8+\varepsilon}}.
\end{aligned}$$

Now (9.7) is bounded by

$$(9.9) \quad \frac{K^{\mu-j+5+\varepsilon} Y^{5/2+\theta+\varepsilon}}{L^{j-\mu+2\nu+7+\varepsilon}} \sum_{0 \leq k \leq N} \frac{1}{k!} \left(\frac{L}{2\pi K} \right)^k.$$

Note that as L being a smaller power of K , we have $L/(2\pi K) < 1$ and the sum in (9.9) is bounded by e . Recall that $0 \leq 2\mu \leq \nu < N$ and $0 \leq j < 2N$. We get $(LK)^\mu / L^{2\nu} \leq 1$ when $L \geq K^{1/3}$. Consequently (9.9) and hence (9.7) and (9.6) are bounded by

$$\frac{K^{5+\varepsilon} Y^{5/2+\theta+\varepsilon}}{L^7} \ll LKY^{1+\varepsilon} \cdot \frac{K^4 Y^{3/2+\theta}}{L^8}.$$

This is the desired $O(LKY^{1+\varepsilon})$, the same bound as for (9.5), if

$$(9.10) \quad K^4 Y^{3/2+\theta} \leq L^8.$$

Note that we need

$$(9.11) \quad Y \ll K^{2+\varepsilon}.$$

Therefore (9.10) is true when $K^{7+2\theta+\varepsilon} \leq L^8$, i.e., when

$$(9.12) \quad K^{(7+2\theta)/8+\varepsilon} \leq L \leq K^{2-\delta}$$

for arbitrarily small $\delta > 0$ and $\varepsilon > 0$. Here θ is given by (8.2), and we may take $\theta = 7/64$.

Consequently,

$$\sum_{f_j} |S_Y(f_j)|^2 \left(h\left(\frac{K-k_j}{L}\right) + h\left(\frac{K+k_j}{L}\right) \right) \ll LKY^{1+\varepsilon},$$

and hence

$$(9.13) \quad \sum_{K-L \leq k_j \leq K+L} |S_Y(f_j)|^2 \ll LKY^{1+\varepsilon}$$

under (9.11) and (9.12). According to the approximation formula of the central value of the L -function in (9.1), we have

$$\begin{aligned} & \sum_{K-L \leq k_j \leq K+L} \left| L\left(\frac{1}{2} + it, f_j \otimes g\right) \right|^2 \\ & \ll \sum_{K-L \leq k_j \leq K+L} \left| \sum_{1 \leq b \leq K^{1+\varepsilon}} \frac{1}{b} \sum_{a \geq 1} \frac{\lambda_j(a) \lambda_g(a)}{\sqrt{a}} V\left(\frac{ab^2}{K^2}\right) \right|^2 \\ & \ll \frac{1}{K^2} \sum_{K-L \leq k_j \leq K+L} \left| \sum_{a \geq 1} \lambda_j(a) \lambda_g(a) \sum_{1 \leq b \leq K^{1+\varepsilon}} \frac{V(ab^2/K^2)}{\sqrt{ab^2/K^2}} \right|^2. \end{aligned}$$

Applying smooth dyadic subdivisions to the function

$$\sum_{1 \leq b \leq K^{1+\varepsilon}} \frac{V(ab^2/K^2)}{\sqrt{ab^2/K^2}},$$

we get

$$\begin{aligned} & \sum_{K-L \leq k_j \leq K+L} \left| L\left(\frac{1}{2} + it, f_j \otimes g\right) \right|^2 \\ & \ll \frac{\log K}{K^2} \sum_{K-L \leq k_j \leq K+L} \max_{1 \leq B \leq K^{2+\varepsilon}} \left| \sum_{a \geq 1} \lambda_j(a) \lambda_g(a) H\left(\frac{a}{K^2/B}\right) \right|^2, \end{aligned}$$

where H is a fixed smooth function of compact support in $[1, 2]$. Using the bound in (9.13) with $Y = K^2/B$, we see that the maximum is from $B = 1$ and hence

$$\begin{aligned} \sum_{K-L \leq k_j \leq K+L} \left| L\left(\frac{1}{2} + it, f_j \otimes g\right) \right|^2 & \ll \frac{\log K}{K^2} \sum_{K-L \leq k_j \leq K+L} |S_{K^2}(f_j)|^2 \\ & \ll \frac{\log K}{K^2} LK(K^2)^{1+\varepsilon} \ll (LK)^{1+\varepsilon} \end{aligned}$$

for L as in (9.12). This proves Theorem 8.5.

Acknowledgments. We wish to express our thanks to the Mathematics Institute at Zhejiang University and the Obermann Center for Advanced Studies at the University of Iowa for warm hospitality and nice working environment.

REFERENCES

- [1] H. Bateman, *Tables of Integral Transforms*, McGraw-Hill, New York, 1954.
- [2] V. Blomer, *Shifted convolution sums and subconvexity bounds for automorphic L -functions*, preprint, 2004.
- [3] J. Hoffstein and P. Lockhart, *Coefficients of Maass forms and the Siegel zero*, Ann. Math., **140** (1994), 161-181.
- [4] H. Iwaniec, *Introduction to the Spectral Theory of Automorphic Forms*, Biblioteca de la Revista Matemática Iberoamericana, Madrid, 1995. Second edition: *Spectral Methods of Automorphic Forms*, Amer. Math. Soc. and Revista Matemática Iberoamericana, Providence, 2002.
- [5] H. Iwaniec, *Topics in Classical Automorphic Forms*, Amer. Math. Soc., Providence, 1997.
- [6] H. Kim, *Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2* , J. Amer. Math. Soc., **16** (2003), 139-183.
- [7] H. Kim and P. Sarnak, *Appendix: Refined estimates towards the Ramanujan and Selberg conjectures*, Appendix to Kim [6].
- [8] H. Kim and F. Shahidi, *Cuspidality of symmetric powers with applications*, Duke Math. J., **112** (2002), 177-197.
- [9] E. Kowalski, P. Michel, and J. Vanderkam, *Rankin-Selberg L -functions in the level aspect*, to appear in Duke Math. J.
- [10] N.V. Kuznetsov, *Petersson's conjecture for cusp forms of weight zero and Linnik's conjecture. Sums of Kloosterman sums*, Mat. Sb. (N.S.), **111 (153)** (1980), 334-383, 479 (in Russian). English translation: Math. USSR Sbornik, **39** (1981), 299-342.
- [11] Jianya Liu and Yangbo Ye, *Subconvexity for Rankin-Selberg L -functions of Maass forms*, Geom. Funct. Anal., **12** (2002), 1296-1323.
- [12] W. Luo, Z. Rudnick, and P. Sarnak, *On the generalized Ramanujan conjecture for $GL(n)$* , in: Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996), 301-310, Proc. Sympos. Pure Math., 66, Part 2, Amer. Math. Soc., Providence, RI, 1999.
- [13] Y. Motohashi, *Spectral theory of the Riemann zeta-function*, Cambridge Tracts in Mathematics, 127, Cambridge University Press, Cambridge, 1997.
- [14] H. Petersson, *Über die Entwicklungskoeffizienten der automorphen Formen*, Acta Math. **58** (1932), 169-215.

- [15] Z. Rudnick and P. Sarnak, *The behavior of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), 195-213.
- [16] Z. Rudnick and P. Sarnak, *Zeros of principal L-functions and random matrix theory*, Duke Math. J. **81** (1996), 269-322.
- [17] P. Sarnak, *Some applications of modular forms*, Cambridge Tracts in Math. **99**, Cambridge Univ. Press, 1990.
- [18] P. Sarnak, *Arithmetic quantum chaos*, the Schur Lectures (Tel Aviv 1992), Israel Math. Conf. vol.8, Bar Ilan, (1995), 183-236.
- [19] P. Sarnak, *Estimates for Rankin-Selberg L-functions and quantum unique ergodicity*, J. Functional Analysis, **184** (2001), 419-453.
- [20] P. Sarnak, *Lecture at Ohio State University*, March, 2003.
- [21] A. Selberg, *Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series*, J. Indian Math. Soc. **20** (1956), 47-87.
- [22] F. Shahidi, *On the Ramanujan conjecture and finiteness of poles for certain L-functions*, Ann. of Math., (2) **127** (1988), no. 3, 547-584.
- [23] G. Shimura, *On Modular forms of half integral weight*. Ann. Math. **97** (1973), 440-481.
- [24] E.C. Titchmarsh, *The theory of the Riemann zeta-function*, Oxford, 1951.
- [25] T. Watson, *Central value of Rankin triple L-function for unramified Maass cusp forms*, thesis, Princeton, 2001.

DEPARTMENT OF MATHEMATICS, SHANDONG UNIVERSITY
 JINAN 250100, CHINA.
 jyliu@sdu.edu.cn

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF IOWA
 IOWA CITY, IOWA 52242-1419, USA.
 yey@math.uiowa.edu