Arithmetic trace formulas and Kloostermania

X. Li

Department of Mathematics, Columbia University, 2990 Broadway, NY, 10027;
Email: xqli@math.columbia.edu

Contents

1 Introduction

Automorphic forms live on locally symmetric, homogeneous Riemannian spaces which are natural generalizations of periodic functions on a circle. One of the most important tool in the theory of the Fourier analysis on a circle is the Poisson summation formula which says that for any functions f in the Schwartz class, we have

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n)$$

where

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x)\exp(-2\pi ixy)dx$$

is the Fourier transform of f. The proof is given by the Fourier expansion of the periodic functions $F(x) =: \sum_{n\in\mathbb{Z}} f(n+x)$ and taking $x = 0$.

Although we have the celebrated Selberg trace formula, which are natural generalizations of the Poisson summation formula, we will concentrate on the more arithmetic generalizations -the Petterson formula and the Kuznetsov formula which connect Fourier coefficients of automorphic forms with Kloosterman sums. Kloosterman sums are some exponential sums on algebraic curves or varieties, sharp bounds of which are given using algebraic geometry. This allows one to study spectral problems and Fourier coefficients of automorphic forms borrowing the rich fruits from algebraic geometry and the deep machinery from the classical theory of exponential sums.

The Peterson-Kuznetsov formula approach has, mainly through the work of Iwaniec and his Collaborators Bombieri, Deshoulliers, Duke, Friedlander, ... became a fundamental tool in modern analytic number theory. The theory they have developed, which some have termed $'$Klootermania$'$, has striking applications including mean value theorems of the Riemann zeta function over short intervals, extending Bombieri-Vinogradov's mean value theorem beyond the capability of the Riemann hypothesis, achieving subconvexity bounds of L-functions, etc. Another important application of Kuznetsov-Bruggman's formula made by Motohashi is the asymptotic formula of the 4th moment of the Riemann zeta function.

In these lecture notes, our main goal is to introduce these valuable tools: Peterson-Kuznetsov-Kloosterman approach to the beginners in the theory of automorphic forms. We develop the whole theory from scratch. We give detailed proofs of Peterson's formula and Kuznetsov's formula. In proving Kuznetsov's formula, we take the approach of Zagier with the intention to generalize it to higher rank groups. As for applications, we include the recent work of the author with Iwaniec [IL] on orthogonality of Fourier coefficients of automorphic forms and the work with Sarnak [LS] on number variance for the modular group $SL(2,\mathbb{Z})$, in the end, we sketch the proof of Iwaniec's mean value theorem of the 4th moment of the Riemann zeta function over a short interval. In the appendix we sketch the proof of Selberg's trace formula for the modular group $SL(\mathbb{Z})$. Although many complicated and technical proofs, the basic theory is developed in details, so essentially it is self-contained.

One should be aware that many other great progress in the analytic theory of automorphic forms were made through other approaches. One of the examples is the recent remarkable progress on the generalized Ramanujan conjecture made by Luo-Rudnick-Sarnak [LRS], Kim-Sarnak [KS] using the approach of L-functions and converse theorems, see the nice expository paper [Sa4] and the references therein.

2 Holomorphic automorphic forms

2.1 The Poincare upper half plane and congruence subgroups

Let $\mathbb{H} = \{z = x + iy | y > 0\}$ be the upper half plane, it becomes a Riemannian manifold of curvature -1 with the metric derived from the Poincare differential form

$$ds^2 = y^{-2}(dx^2 + dy^2),$$

we henceforth call $\mathbb{H}$ the Poincare upper half plane. The distance function on $\mathbb{H}$ is given by

$$d(z_1, z_2) = \log \frac{|z_1 - \bar{z}_2| + |z_1 - z_2|}{|z_1 - \bar{z}_2| - |z_1 - z_2|}.$$

Another simpler formula is

$$\cosh d(z_1, z_2) = 1 + 1/2u(z_1, z_2)$$

with

$$u(z_1, z_2) = \frac{|z_1 - z_2|^2}{\Im z_1 \Im z_2}.$$

We are interested in the quotient space of subgroups of $SL(2, \mathbb{R})$ acting on $\mathbb{H}$ by Mobius transformations

$$\gamma z = \frac{az + b}{cz + d}$$

for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $SL(2, \mathbb{R})$. Although the mordern theory of automorphic forms was driven by consideration of more general groups, we are mainly concerned with congruence subgroups of $SL(2, \mathbb{Z})$ for many applications in arithmetic. For $q \geqslant 1$, the principal congruence subgroup of level q is

$$\Gamma(q) := \{\gamma \in SL(2, \mathbb{Z}) | \gamma \equiv I (\text{mod } q)\}$$

where I is the identity matrix. Any subgroup of the modular group of $SL(2,\mathbb{Z})$ which contains $\Gamma(q)$ is called a congruence subgroup of level q. The most important ones are

$$\Gamma_0(q) := \{\gamma \in SL(2,\mathbb{Z}) | \gamma \equiv \begin{pmatrix} * & * \\ & * \end{pmatrix} (\mathrm{mod}\ q)\}$$

and

$$\Gamma_1(q) := \{\gamma \in SL(2,\mathbb{Z}) | \gamma \equiv \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix} (\mathrm{mod}\ q), \}$$

obviously $\Gamma_0(1) = SL(2,\mathbb{Z})$.
The quotient space $\Gamma_0(q) \setminus \mathbb{H}$ is a finite area, noncompact, hyperbolic surface. Specifically, one can show

$$\mathrm{Vol}(SL(2,\mathbb{Z}) \setminus \mathbb{H}) = \frac{\pi}{3}.$$

We can choose a fundamental domain for the action of $\Gamma_0(q)$ on $\mathbb{H}$ formed by a closed polygon whose sides meet at vertices called cusps. Cusps are also described as fixed points of the parabolic elements in $\Gamma_0(q)$ whose traces are 2 or -2. Clearly for

$$\Gamma_\infty := \{\pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} | n \in \mathbb{Z}\},$$

clearly $\Gamma_\infty \infty = \infty$, so ∞ is a cusp for $\Gamma_0(q)$. All the other cusps of $\Gamma_0(q)$ are equivalent to ∞ under the action of the modular group $SL(2,\mathbb{Z})$. The fundamental domain is noncompact because of the existence of cusps. The existence of cusps makes the analysis much more interesting. For more details of this section, one is suggested to read [Iw2].

2.2 Hilbert space of the holomorphic automorphic forms

Let $\chi(\mathrm{mod}\ q)$ be a Dirichlet character and $k \geqslant 0$ be an integer with

$$\chi(-1) = (-1)^k,$$

for a function f on $\mathbb{H}$ and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{R})$, we set

$$(f|_\gamma)(z) = (cz+d)^{-k} f\Big(\frac{az+b}{cz+d}\Big)$$

as the so-called slash operator, then it is easy to check that $(f|_{\gamma_1})|_{\gamma_2} = f|_{(\gamma_1\gamma_2)}$ for any $\gamma_1, \gamma_2 \in SL(2,\mathbb{R})$.
A holomorphic automorphic form of weight k attached to $\chi(\mathrm{mod} q)$ is a holomorphic function on $\mathbb{H}$ satisfying the automorphy conditions

$$f|_\gamma(z) = \chi(d) f(z)$$

for any $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q)$. Moreover f is required to be holomorphic at every cusp Ω. This means the following: Choose $\phi \in SL(2, \mathbb{R})$ such that $\phi(\Omega) = \infty$, then $f|_{\phi^{-1}}|_{\phi\Gamma_\infty\phi^{-1}} = f|_{\Gamma_\infty\phi^{-1}} = f|_{\phi^{-1}}$. Since $\phi\Gamma_\infty\phi^{-1}$ contains a translation $z| \to z + t$ for some real $t > 0$, it follows that $f|_{\phi^{-1}}$ has a Fourier expansion $\sum_{n\in\mathbb{Z}} a_n e(zt^{-1})$, we say f is holomorphic at Ω if $a_n = 0$ for $n < 0$. If f has no zero term in the Fourier expansion at each cusp, we say f is a cusp form.

All the cusp forms form a finite dimensional Hilbert space $S_k(\Gamma_0(q), \chi)$ equipped with the Petersson's inner product

$$< f, g >= \int_{\Gamma_0(q)\backslash\mathbb{H}} f(z)\bar{g}(z)y^k d\mu z$$

where $\Gamma_0(q) \setminus \mathbb{H}$ is a fundamental domain and $d\mu z = y^{-2}dxdy$ is the $SL(2, \mathbb{R})$ invariant measure. For $k \geqslant 2$,

$$\dim S_k(\Gamma_0(q), \chi) \asymp kq \prod_{k|q} \left(1 + \frac{1}{p}\right)$$

which can be proved by the Riemann-Roch theorem (see [Sh], pp. 34) or Petersson's formula we are going to develop. It is important to note that for weight 1 forms, all the above mentioned technology fail, the first nontrivial bound was given by the ingenious work of Duke [Du] and was improved later on by Michel and Venketash [MV].

Exercise: Prove that

$$S_k(\Gamma_1(q)) = \bigoplus_{\chi (\bmod q)} S_k(\Gamma_0(q), \chi)$$

where the sum is over all Dirichlet characters modulo q.

2.3 Hecke operators and primitive forms

It was known from Ramanujan that Fourier coefficients of automorphic forms have multiplicativity properties, it will be nice to choose an orthonormal basis whose Fourier coefficients have multiplicativity properties, this is accomplished by the introduction of Hecke operators. Let $n \geqslant 1, k \geqslant 0$ and χ a Dirichelt character modulo q, the Hecke operator T_n is defined on functions $f : \mathbb{H} \to \mathbb{C}$ which are periodic of period one by

$$(T_n f)(z) = \frac{1}{n} \sum_{ad=n} \chi(a)a^k \sum_{0\leqslant b<d} f\left(\frac{az+b}{d}\right).$$

One can show that these operators are multiplicative, which means the following:

$$T_m T_n = \sum_{d|(m,n)} \chi(d)d^{k-1}T_{mn}d^{-2},$$

hence the Hecke operators commute with each other. T_n maps cusp forms to cusp forms and for $(n,q)=1$ it is a normal operator, that means

$$< T_n f, g >=< f, \bar{\chi}(n) T_n g >,$$

for $f,g \in S_k(\Gamma_0(q),\chi)$. Therefore in the space of $S_k(\Gamma_0(q),\chi)$, we can choose an orthonormal basis which consists of eigenfunctions of all the Hecke operators T_n with $(n,q)=1$. One might hope to find an orthonormal basis consisting of eigenfunctions of all the Hecke operators T_n, but in general this is not true, here comes Atkin-Lehner, W. Li's theory on primitive forms (i.e, the so-called new forms).

Let $q'd|q, q' < q$ and $\chi'(\text{mod } q')$ be the Dirichlet character introduced by $\chi(\text{mod } q)$, then f in $S_k(\Gamma_0(q),\chi)$ is called an old form if $f(z)=g(dz)$ with $g(z) \in S_k(\Gamma_0(q'),\chi'$. The orthogonal complement of the linear subspace of old forms in $S_k(\Gamma_0(q),\chi)$ is the linear space of new forms $S'_k(\Gamma_0(q),\chi)$. It turns out by the multiplicity one theorem, see Chapter VII of [La], in $S'_k(\Gamma_0(q),\chi)$, one can indeed choose an orthonormal basis consisting of eigenfunctions of all Hecke operators T_n which are henceforth call them Hecke eigencusp forms. It follows that if χ is a primitive character modulo q, all the forms in $S_k(\Gamma_0(q),\chi)$ are primitive, we then can find a basis formed by Hecke eigencusp forms.

One immediate important consequence for a new form f being an Hecke eigencusp form is the multiplicativity of its Fourier coefficients, in more concrete words, assume $T_n f = \lambda(n) f$ for $n \neq 1$ and $f(z)$ has the following Fourier series expansion at the cusp ∞ :

$$f(z) = \sum_{n \geqslant 1} \hat{f}(n) e(nz),$$

then $\lambda(n)\hat{f}(1) = \hat{f}(n)$ for all $n \geqslant 1$. One can check easily that $\hat{f}(1) \neq 0$ and the following multiplicativity holds:

$$\lambda(m)\lambda(n) = \sum_{d|(m,n)} \chi(d) d^{k-1} \lambda(mnd^{-2})$$

for any $m,n \geqslant 1$.

Due to the multiplicativity of the Hecke eigenvalues of a new form f, the corresponding Dirichlet series - the so-called L-function

$$L_f(s) = \sum_{n=1}^{\infty} \lambda(n) n^{-s}$$

for $\Re s > 1$ has an Euler product of degree 2

$$L_f(s) = \prod_p (1 - \lambda(p) p^{-s} + \chi(p) p^{k-1-2s})^{-1}$$

which is one of the most important properties of L functions.

2.4 Poincare series and Petersson's formula

One of the most important techniques to construct automorphic forms is by averaging. The Poincare series are constructed in this way: for $m \geqslant 0$ let

$$P_m(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(q)} \bar{\chi}(\gamma) j(\gamma, z)^{-k} e(m\gamma z)$$

with $j(\gamma, z) = cz + d$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}), k \geqslant 2$. These are the Poincare series at the cusp ∞. If $m = 0$,

$$P_0(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(q)} \bar{\chi}(\gamma) j(\gamma, z)^{-k}$$

is the classical Eisenstein series of weight k, usually denoted as $E(z)$. Indeed, the Poincare series are automorphic forms of weight k over the group $\Gamma_0(q)$ since they are averages of $e(mz)$ over the group $\Gamma_0(q)$. By scaling, one can construct other Poincare series at any other cusps of $\Gamma_0(q)$.

Next we are going to develop the Fourier expansion of the Poincare series. We are interested in the Fourier expansion of $P_m(z)$ mainly because it gives rise to the classical :

$$S_\chi(m, n; c) = \sum_{a\bar{a} \equiv 1 (\mathrm{mod}\ c)} \bar{\chi}(a) e\Big(\frac{am + \bar{a}n}{c}\Big). \tag{2.1}$$

The Fourier series expansion of $P_m(z)$ is a direct consequence of the Bruhat decomposition of $\Gamma_0(q)$ which states

Lemma 2.1.

$$\Gamma_0(q) = \Gamma_\infty \bigcup \Big(\bigcup_{\substack{c>0 \\ q|c}} \bigcup_{\substack{d(mod\ c) \\ (d,c)=1}} \Gamma_\infty \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Gamma_\infty \Big)$$

where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q)$.

The proof is obvious since multiplying a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\Gamma_0(q)$ by a matrix in Γ_∞ on the left or on the right simply shifts a and d. So the double coset decomposition is determined uniquely by the pairs $\{(c, d) | q|c, d(\mathrm{mod} c), (d, c) = 1\}$.

For $m \geqslant 0$, assume $P_m(z)$ has the following Fourier expansion:

$$P_m(z) = \sum_{n \in \mathbb{Z}} a_m(n) e(nz),$$

then clearly

$$a_m(n) = \int_0^1 P_m(z) e(-nz) dx$$

due to the simple fact that

$$\int_0^1 e((n-n')x)dx = \begin{cases} 1 & \text{if} \quad n = n' \\ 0 & \text{otherwise.} \end{cases}$$

Applying the Bruhat decomposition to the definition of $P_m(z)$, we have

$$P_m(z) = e(mz) + \sum_{\substack{c>0 \\ c\equiv 0 (\mathrm{mod}\, q)}} \sum_{\substack{d (\mathrm{mod}\, c) \\ (d,c)=1}} \bar{\chi}(d) I(c,d;z) \tag{2.2}$$

where

$$I(c,d;z) = \sum_{l\in\mathbb{Z}} \Big(c(z+l)+d\Big)^{-k} e\Big(m\Big(\frac{a}{c} - \frac{1}{c(c(z+l)+d)}\Big)\Big),$$

this can be easily checked because

$$e\Big(m\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & l \\ & 1 \end{pmatrix} z\Big) = e\Big(m\Big(\frac{a}{c} - \frac{1}{c(c(z+l)+d)}\Big)\Big). \tag{2.3}$$

In view of the above formula, to compute $a_m(n)$, we are led to do the integral

$$J(c,d,m,l) = \int_0^1 (c(z+l)+d)^{-k} e\Big(\frac{-m}{c(c(z+l)+d)} - nz\Big) dx. \tag{2.4}$$

By a change of variable, the above integral is

$$\frac{1}{c}\int_{cl+d}^{c(l+1)+d} (x+iy)^{-k} e\Big(\frac{-m}{c(x+iy)} - \frac{n(x+iy)}{c} + \frac{nd}{c}\Big) dx$$

which we denote it as $e\Big(\frac{nd}{c}\Big) J^*(c,d,m,l)$, the factor $e\Big(\frac{nd}{c}\Big)$ in (2.3) and $\bar{\chi}(d)$ in (2.2) give rise to the Kloosterman sum $S_\chi(m,n;c)$. Since

$$\sum_{l\in\mathbb{Z}} J^*(c,d,m,l) = \frac{1}{c}\int_{\Im z = cy} z^{-k} e\Big(\frac{-m}{cz} - \frac{nz}{c}\Big) dz$$

$$= \begin{cases} 0 & \text{if } n \leqslant 0 \\ \Big(\frac{2\pi}{ic}\Big)^k \frac{n^{k-1}}{\Gamma(k)} & \text{if } n > 0, m = 0, \\ \frac{2\pi}{i^k c}\Big(\frac{m}{n}\Big)^{\frac{k-1}{2}} J_{k-1}\Big(\frac{4\pi\sqrt{mn}}{c}\Big) & \text{if } n > 0, m > 0, \end{cases} \tag{2.5}$$

where $J_\mu(z)$ is the $J-$ Bessel function of order μ, the first formula is by contour integration, the second formula is by ([GR], 8.315.1) and the last one is by ([GR], 8.412.2.) We conclude that

Theorem 2.1.

$$a_m(n) = \int_0^1 P_m(z)e(-nz)dx$$

$$(2.6)\quad = \begin{cases} 1 & \textit{if } n = m = 0 \\ \left(\frac{2\pi}{i}\right)^k \frac{n^{k-1}}{\Gamma(k)} \sum\limits_{\substack{c>0 \\ c\equiv 0(mod\ q)}} c^{-k} S(0,n;c) & \textit{if } n > 0, m = 0, \\ \left(\frac{m}{n}\right)^{\frac{k-1}{2}} \left(\delta_{mn} + 2\pi i^{-k} \sum\limits_{\substack{c>0 \\ q|c}} c^{-1} S_\chi(m,n;c) J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right) & \textit{if } n > 0, m > 0, \\ 0 & \textit{otherwise} \end{cases}$$

where δ_{mn} *is the Kronecker symbol.*

The following corollary follows from the above theorem:

Corollary 2.1. *For* $m \geqslant 1, P_m(z) \in S_k(\Gamma_0(q), \chi)$.

Another important property of Poincare series is that they pick up Fourier coefficients of automorphic forms in the following sense: Let

$$f(z) = \sum_{n\geqslant 1}^{\infty} \hat{f}(n)e(nz) \tag{2.7}$$

be the Fourier expansion of a cusp form f in $S_k(\Gamma_0(q), \chi)$, we have

Proposition 2.1. $< f, P_m >= \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \hat{f}(m)$ *where* $m \geqslant 1$.

Proof: By the definition of the Petersson inner product and the definition of the Poincare series, we have

$$\begin{aligned} &< f, P_m > \\ &= \int_{\Gamma_0(q)\backslash\mathbb{H}} y^k f(z) \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(q)} \chi(\gamma)\overline{j(\gamma,z)}^{-k} e(-m\overline{\gamma z})d\mu z \\ &= \int_{\Gamma_\infty\backslash\mathbb{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(q)} \Im(\gamma z)^k f(\gamma z)e(-m\overline{\gamma z})d\mu z \end{aligned}$$

where we used the automorphy condition of $f(z)$. By unfolding the inetgral, we have

$$\begin{aligned} < f, P_m >&= \int_{\Gamma_\infty\backslash\mathbb{H}} y^k f(z)e(-\overline{mz})d\mu z \\ &= \int_0^\infty \int_0^1 f(z)e(-m\bar{z})\frac{dxdy}{y^2} = \int_0^\infty \hat{f}(m)exp(-4\pi my)\frac{dy}{y^{2-k}} \\ &= \hat{f}(m)\frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \end{aligned}$$

as asserted. □
Hence $P_m(z)$ with $m \geqslant 1$ expands the whole space of $S_k(\Gamma_0(q),\chi)$. Otherwise, there is a cusp form in $S_k(\Gamma_0(q),\chi)$ all of whose Fourier coefficients are 0, obviously, that is impossible.
Let $\mathcal{F}$ be an orthonormal basis of $S_k(\Gamma_0(q),\chi)$, for each f in $\mathcal{F}$, we have the Fourier expansion (2.7), by the Parseval formula and the above proposition, we have

$$\begin{aligned} < P_m, P_n > &= \sum_{f\in\mathcal{F}} < P_m, f >< f, P_n > \\ &= \frac{\Gamma(k-1)^2}{(4\pi)^{2k-2}(mn)^{k-1}} \sum_{f\in\mathcal{F}} \bar{\hat{f}}(m)f(n). \end{aligned} \tag{2.8}$$

On the other hand, we can use the Fourier expansion of poincare series and the above proposition to compute the inner product:

$$\begin{aligned} < P_m, P_n > &= \frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \hat{P}_m(n) \\ &= \frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \Big(\frac{m}{n}\Big)^{\frac{k-1}{2}} \Big(\delta_{mn} + \sigma_\chi(m,n)\Big) \end{aligned} \tag{2.9}$$

where

$$\sigma_\chi(m,n) = 2\pi i^{-k} \sum_{\substack{c>0\\ q|c}} c^{-1} S_\chi(m,n;c) J_{k-1}\Big(\frac{4\pi\sqrt{mn}}{c}\Big). \tag{2.10}$$

Comparing (2.7) with (2.9), we end up with the Petersson formula

Theorem 2.2. *For $m,n \geqslant 1, k \geqslant 2$, we have*

$$\frac{\Gamma(k-1)}{4\pi\sqrt{mn}^{k-1}} \sum_{f\in\mathcal{F}} \bar{\hat{f}}(m)f(n) = \delta_{mn} + \sigma_\chi(m,n)$$

where $\sigma_\chi(m,n)$ as defined above.

Exercise: Use Petersson's formula, for q a prime, $k \geqslant 2$, χ a Dirichlet character modulo q, prove that $\dim S_k(\Gamma_0(q),\chi) \asymp kq$, try to get a good error term.

2.5 Orthogonalities of Hecke eigenvalues

In classical analytic number theory, the following so-called large sieve inequality for Dirichlet characters are well known: For any complex vectors $\alpha = \{a_n\}$, we have

$$\frac{1}{Q^2} \sum_{q\leqslant Q} \sum_{\chi(\mathrm{mod}\ q)}^{*} \Big| \sum_{N<n\leqslant 2N} a_n\chi(n)\Big|^2 \leqslant (1+\frac{N}{Q^2})\|\alpha\|^2 \tag{2.11}$$

where the superscript * restricts the summation χ primitive and $\|\alpha\|$ is the l_2- norm. There are many proofs, see ([Mon1], for example), we are going to give a proof due to Bombieri which uses the powerful tool-the duality principle:

Lemma 2.2. *([Mon1])*

Let C_{nr}, $1 \leqslant n \leqslant N$, $1 \leqslant r \leqslant R$ be any complex numbers and $D > 0$, the following are equivalent:

1) For any complex vectors $x = \{x_n\}, n \leqslant N$,

$$\sum_{r\leqslant R} |\sum_{n\leqslant N} C_{nr}x_n|^2 \leq D\|x\|^2. \tag{2.12}$$

2) For any complex vectors $y = \{y_r\}, r \leqslant R$,

$$\sum_{n\leqslant N} |\sum_{r\leqslant R} C_{nr}y_r|^2 \leq D\|y\|^2. \tag{2.13}$$

Proof. Suppose 1) holds, let L be the left side of (2.13), expand the square in (2.13), the left side of (2.13) equals

$$\sum_n\sum_r\sum_{r'} C_{nr}\bar{C}_{nr'}y_r\bar{y}_{r'}$$

which is less than or equal to

$$\|y\|D^{\frac{1}{2}}L^{\frac{1}{2}}$$

by the Cauchy inequality and 1). It follows that

$$L \leqslant D\|y\|^2.$$

2) is proved.

Similarly, we can derive 1) assuming 2) holds. This completes the proof of the lemma. □

Now we are ready to give Bombieri's unpublished proof of a slightly weaker version of (2.11):

Proposition 2.2. *For any complex vectors $\alpha = \{a_n\}$, we have*

$$\frac{1}{Q^2}\sum_{q\leqslant Q}\ \sideset{}{^*}\sum_{\chi (mod\ q)} |\sum_{N<n\leqslant 2N} a_n\chi(n)|^2 \ll_\varepsilon Q^\varepsilon\Big(1+\frac{N}{Q^2}\Big)\|\alpha\|^2. \tag{2.14}$$

Proof. The proof has two steps:

• The first step:

By the duality principle, it suffices to prove the following dual inequality: For any complex numbers $c_{q,\chi}$, we have

$$\sum_{N<n\leqslant 2N} |\sum_{q\leqslant Q}\ \sideset{}{^*}\sum_{\chi(\text{mod }q)} c_{q,\chi}\chi(n)|^2 g\Big(\frac{n}{N}\Big) \ll_\varepsilon Q^\varepsilon(Q^2+N)\|c\|^2 \tag{2.15}$$

where we attached a smooth factor $g(x)$ which is 1 on $(1,2)$and 0 on $(-\infty,0)\cup(3,\infty)$.
Expanding out the square on the left, the left of (2.15) is equal to

$$\sum_{q_1}\sum_{q_2}\sum_{\chi_1(\mathrm{mod} q_1)}{}^{*}\sum_{\chi_2(\mathrm{mod} q_2)}{}^{*} c_{q_1,\chi_1}\overline{c_{q_2,\chi_2}}\sum_n \chi_1(n)\overline{\chi_2}(n). \tag{2.16}$$

Now we recall that primitive characters can be expanded as sums of exponential sums, i.e., if χ is primitive, then

$$\chi(n)=\frac{1}{\tau(\bar\chi)}\sum_{-\frac{q}{2}<a\leqslant\frac{q}{2}}\bar\chi(a)e\Big(\frac{an}{q}\Big),$$

where $\tau(\chi)$ is the Gauss sum. It then follows that

$$\begin{aligned}&\sum_n \chi_1(n)\overline{\chi_2}(n)\\ &\quad=\frac{1}{\tau(\bar\chi_1)\tau(\overline{\chi}_2)}\sum_a\sum_b\bar\chi_1(a)\chi(b)\sum_n e\Big(\Big(\frac{a}{q_1}-\frac{b}{q_2}\Big)n\Big)g\Big(\frac{n}{N}\Big).\end{aligned}\tag{2.17}$$

Now assuming

$$N\geqslant Q^{2+\varepsilon}, \tag{2.18}$$

the above is equal to

$$\frac{1}{\tau(\bar\chi_1)\tau(\bar\chi_2)}\sum_a\sum_b\bar\chi_1(a)\chi_2(b)\int_{-\infty}^{\infty}e\Big(\Big(\frac{a}{q_1}-\frac{b}{q_2}\Big)x\Big)g\Big(\frac{x}{N}\Big)dx+O\Big(\frac{1}{N^A}\Big), \tag{2.19}$$

the integral is negligable, i.e, $O(N^{-A})$ with A any large number unless $q_1=q_2$ and $a=b$. In this case, the main contribution comes from the diagonal term, the left of (2.15) is $\ll N\|c\|^2$, where we have used $|\tau(\chi)|=\sqrt{q}$ for primitive character $\chi(\mathrm{mod}\ q)$. By the duality principle, we proved (2.15) under the condition(2.18).

• The second step:
If (2.18) doesn't hold, we are going to choose a number $P\gg\log Q$, then

$$\begin{aligned}&\sum_{q\leqslant Q}\sum_{\chi(\mathrm{mod}\ q)}{}^{*}|\sum_{N<n\leqslant 2N}a_n\chi(n)|^2\\ &\qquad\ll\frac{1}{P}\sum_q\sum_{\substack{P<p\leqslant 2P\\ p\ \mathrm{prime},(p,q)=1}}\log p\sum_{\chi(\mathrm{mod}\ q)}{}^{*}|\sum_{Np<m\leqslant 2Np}a'(m)\chi(m)|^2\end{aligned}$$

where

$$a'(m)=\begin{cases}a_{\frac{m}{p}} & \text{if } p|m\\ 0 & \text{otherwise.}\end{cases}$$

and we used an important property of χ, that is,

$$|\chi(np)| = |\chi(n)|,$$

if $(p,q)=1$. Fix $P = \left[\frac{Q^{2+\varepsilon}}{N}\right]+1$, then $NP \asymp Q^{2+\varepsilon}$ which satisfies the condition (2.18), so by the result what we have proved in the last step, we have the above is

$$\ll \frac{1}{P}\sum_{P<p\leqslant 2P} \log p(Q^2+Np)\|a'\|^2$$

which is bouned by $Q^{2+\varepsilon}\|a\|^2$ because $\|a'\| = \|a\|$. This finishes the second step. Considering the above two steps, Proposition 2.2 is proved in the full generality. □

Proposition 2.2 was applied by Bombieri to prove the famous mean value theorem on primes. In histroy, such inequalities were called large sieve inequalities (just a name). Let's give more explanations:

Let $\mathcal{F}$ be a family of arithmetic objects. To each $f \in \mathcal{F}$, we assume that a sequence $\{f(n)\}$ with n positive integers is associated and $f(n)$ has the size of $O_f(n^\varepsilon)$ for all $n \geqslant 1$, the large sieve inequality for $\mathcal{F}$ means the following:

For any complex vectors $\alpha = \{a_n\}$ with $1 \leqslant n \leqslant N$ there exists a constant $C = C(\mathcal{F}, N)$ which depends only on $\mathcal{F}$ and N such that

$$\frac{1}{|\mathcal{F}|}\sum_{f\in\mathcal{F}} |\sum_{n\leqslant N} a_n f(n)|^2 \leqslant C(\mathcal{F},N)\sum_{n\leqslant N}|a_n|^2. \tag{2.20}$$

It is called an asymptotic large sieve if the inequality is replaced by an asymptotic formula as $|\mathcal{F}|$ and N go to infinity.

If we expand the square on the left of (2.20), we get a bilinear form in a_m and a_n, the diagonal term $m=n$ gives us the contribution $(1+O(N|\mathcal{F}|^{-\infty}))\|\dashv\|^{\in}$, so the best estimate one can hope for $C(\mathcal{F},N)$ is roughly,

$$C(\mathcal{F},N) \asymp 1 + O\Big(\frac{N}{|\mathcal{F}|}\Big). \tag{2.21}$$

What's the significance of the bound (2.21)?

1. With the bound (2.21), (2.20) says that the projections of any complex vectors $\{a_n\}$ on $\{f(n)\}$ are small quite often comparing to the norm of the vectors $\{a_n\}$, so we say $\{f(n)\}$ are orthogonal in some sense as f varies over $\mathcal{F}$.
2. If we take $a_n = 1$ for all $n \geqslant 1$, it follows that on the average the linear form

$$\mathcal{L}_f(\alpha_0) := \sum_{n\leqslant N} f(n) \leqslant N^{\frac{1}{2}+\varepsilon},$$

if $|\mathcal{F}| \leqslant \mathcal{N}$. For each individual form, this bound is the consequence of the big open Lindelof hypothesis for the L function $L(f,s) = \sum_{n\geqslant 1} f(n)n^{-s}$ which says that $L(f,s) \ll_{f,\varepsilon} q(f,s)^\varepsilon$ on the line $\Re s = \frac{1}{2}$, where $q(f,s)$ is some constant

depending on f and s.
3. Take $a_n = \mu(n)$ the Mobius function, we have, on the average the linear form:

$$\mathcal{L}_f(\alpha_1) := \sum_{n\leqslant N} \mu(n)f(n) \leqslant N^{\frac{1}{2}+\varepsilon},$$

if $|\mathcal{F}| \leqslant N$. For each individual form in $\mathcal{F}$, this estimate is equivalent to the Riemann hypothesis for the L function $L(f,s)$. This also explains why (2.20) often serves a substitute of the Riemann hypothesis for Dirichlet L functions.
Hecke eigenvalues, or more generally, Fourier coefficients of cusp forms are natural generalizations of Dirichlet characters, so one might hope such large sieve inequalities (2.20) with the bound (2.21) still holds. Indeed, in some cases, it is true as the following theorem says. For each f in an orthonormal basis $\mathcal{F}_\chi$ of $S_k(\Gamma_0(q),\chi)$, with $k\geqslant 2$, $f(n)$ be its Fourier coefficients, set

$$\psi_f(n) = \Big(\frac{\Gamma(k-1)}{(4\pi n)^{k-1}}\Big)^{\frac{1}{2}}\hat{f}(n)$$

be the nomalized Fourier coefficients, for $N\leqslant q$, we have

Theorem 2.3. *([Iw])*

$$\sum_{f\in\mathcal{F}_\chi} |\sum_{n\leqslant N} a_n\psi_f(n)|^2 = \Big(1+O_k\Big(\frac{N}{q}\Big)\Big)\sum_{n\leqslant N}|a_n|^2 \tag{2.22}$$

for any complex numbers a_n.

Proof. According to Petersson's formula, the left side of (2.22) is equal to

$$\|a\|^2 + \sum_{m\leqslant N}\sum_{n\leqslant N}\bar{a}_m a_n \sigma_\chi(m,n)$$

with $\sigma_\chi(m,n)$ is defined in (2.10).
Expanding the Kloosterman sum and by the Cauchy ineqality, it is easy to see that the bilinear form

$$\begin{aligned}&\sum_m\sum_n a_m\bar{a}_n S_\chi(m,n;c)\\ &\leqslant \sum_{d(\mathrm{mod}\ c)} |\sum_{n\leqslant N} a_n e\Big(\frac{dn}{c}\Big)|^2 \leqslant (c+N)\|a\|^2.\end{aligned} \tag{2.23}$$

To utilize the above argument, we have to separate the variables m and n in the argument of $J-$ Bessel function. For that purpose, we use the Mellin integral representation

$$J_{k-1}(x) = \frac{1}{2\pi i}\int_{(\sigma)}\gamma(r)x^r dr,\ \ 0<\sigma<k-1$$

with

$$\gamma(r) = 2^{r-1}\Gamma\Big(\frac{k-r-1}{2}\Big)\Gamma^{-1}\Big(\frac{k+r+1}{2}\Big).$$

This represents the Bessel function $J_{k-1}(x)$ by the power function x^r which is useful to separate variables, particularly when x is a product of many independent parameters. However we lose control of the size of the function. We recover this control by moving the line of integration $\Re r = \sigma$ to the contour $\mathcal{L}_X$ with $\Re r = \sigma = -1$ if $|\Im r| \leqslant X$ and $\Re r = \sigma = 1$ if $|\Im r| > X$. Then write

$$J_{k-1}(x) = \frac{1}{2\pi i}\int_{\mathcal{L}_X} \gamma_X(r)\Big(\frac{x}{X}\Big)^r dr$$

with $\gamma_X(r) = \gamma(r)X^r$. For $r \in \mathcal{L}_X$ we have $\gamma_X(r) \ll X(X+|r|)^{-2}$. Hence

$$\int_{\mathcal{L}_X} |\gamma_X(r)||dr| \ll 1.$$

Separating variables m and n as above and by the arguments (2.23), we get what we claimed. $\square$

Remarks. The above bound is consistent with the classical large sieve: (2.20) and (2.21), so it serves the role as the average form of Riemann hypothesis. It also shows $\psi_f(n)$ has the average size of $q^{-\frac{1}{2}}$ assuming that $q \gg n$.

The larger family is put into the play, the stronger orthogonality should come out, which means that we are supposed to get a better error term than the one in the above theorem. Driven by this principle, in [IL], we performed extra averaging over characters $\chi(\mathrm{mod} q)$. we did get some extra saving, but surprisingly not the maximal one as expected by the bound of the classical large sieve (2.20). More precisely, using Petersson's formula, Plancherel's formula and other harmonic analysis, we proved the following: Let

$$\mathcal{L}_f(\alpha) = \sum_{N<m\leqslant 2N} a_m\psi_f(m), \tag{2.24}$$

and

$$\mathcal{P}_{ht}(\alpha) = \sum_{N<n\leqslant 2N} a_n S(h\bar{q}, n; t) J_{k-1}\Big(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\Big). \tag{2.25}$$

Theorem 2.4. *[IL] Let q be prime and $\alpha = (a_n)$ be any sequence of complex numbers for $N < n \leqslant 2N$. Then $k \geqslant 2$ and $1 \leqslant H \leqslant T$, then*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi(mod\ q)\\ \chi(-1)=(-1)^k}} \sum_{f\in\mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2$$
$$= \frac{1}{q}\sum_{\substack{t\leqslant T\\ (t,q)=1}} \Big(\frac{2\pi}{t}\Big)^2 \sum_{h\leqslant H} |\mathcal{P}_{ht}(\alpha)|^2 + O\Big(\Big(1+\frac{N}{q^2}+\sqrt{\frac{N}{qH}}\Big)N^\varepsilon\|\alpha\|^2\Big)$$

with any $\varepsilon > 0$, the implied constant depending only on k and ε.

To estimate how large the main term is, we are led to the following large sieve inequality for Kloosterman sums and $J-$ Bessel functions, set

$$\mathcal{A}(H,T,N)=\sum_{H\leqslant h<2H}\ \sum_{\substack{T\leqslant t<2T\\(t,q)=1}}|\mathcal{P}_{ht}(\alpha)|^2 \tag{2.26}$$

for any $H,T,N\geqslant 1$, where $\mathcal{P}_{ht}(\alpha)$ is the linear form (2.25), we have the following

Proposition 2.3. *Let* $1\leqslant T\leqslant \frac{N}{q},\ \ 1\leqslant H\leqslant \left(\frac{N}{q}\right)^{1+\varepsilon}, \alpha=(a_n)$ *be a sequence of complex numers, then*

$$\mathcal{A}(H,T,N)\ll qT^2\Big(1+\frac{N}{q^2}+\sqrt{\frac{N}{qH}}\Big)N^{\varepsilon}\|\alpha\|^2 \tag{2.27}$$

where the implied constant depends only on k and ε.

This can be proved directly, but we would like to take the dual approach in order to illustrate the techniques of the duality principle.

Proof. Because a_n are arbitrary numbers it suffices to estimate the dual form

$$\mathcal{A}^*(H,T,N)=\sum_n\Big|\sum_h\sum_t\gamma(h,t)S(h\bar{q},n;t)J_{k-1}\Big(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\Big)\Big|^2 \tag{2.28}$$

where $\gamma(h,t)$ are arbitrary complex numbers for $H\leqslant h<2H,\ T\leqslant t<2T$ with $(t,q)=1$. Our estimate must depend on the $\gamma(h,t)$ only by way of being proportional to the l_2-norm

$$\|\gamma\|^2=\sum_h\sum_t|\gamma(h,t)|^2, \tag{2.29}$$

otherwise the result would be useless for applications to the original form $\mathcal{A}(H,T,N)$.

Before opening the square we enlarge and smooth the outer summation

$$\mathcal{A}^*(H,T,N)\leqslant\sum_n g\Big(\frac{n}{N}\Big)|\sum_h\sum_t\gamma(h,t)S(h\bar{q},n;t)J_{k-1}\Big(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\Big)|^2.$$

Here $g(x)$ is a smooth function supported on $[\frac{1}{2},\frac{5}{2}]$. Now this equals

$$\sum_{h_1}\ \sum_{t_1}\sum_{h_2}\sum_{t_2}\gamma(h_1,t_1)\bar{\gamma}(h_2,t_2)\sum_{d_1(t_1)}{}^{*}\sum_{d_2(t_2)}{}^{*}e\Big(\frac{h_1\overline{qd_1}}{t_1}-\frac{h_2\overline{qd_2}}{t_2}\Big)$$
$$\cdot\sum_n g\Big(\frac{n}{N}\Big)e\Big(n\Big(\frac{d_1}{t_1}-\frac{d_2}{t_2}\Big)\Big)J_{k-1}\Big(\frac{4\pi}{t_1}\sqrt{\frac{h_1n}{q}}\Big)J_{k-1}\Big(\frac{4\pi}{t_2}\sqrt{\frac{h_2n}{q}}\Big).$$

The terms with $t_1=t_2, d_1=d_2$ contribute

$$\mathcal{A}_0^*(H,T,N)=\sum_n g\Big(\frac{n}{N}\Big)\sum_t\sum_{d(t)}{}^{*}\Big|\sum_h\gamma(h,t)e\Big(\frac{dn}{t}\Big)J_{k-1}\Big(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\Big)\Big|^2.$$

Before squaring out we expand the summation over $d(\mathrm{mod}\ \ t),(d,t)=1$ to all residue classes modulo t getting

$$\mathcal{A}_0^*(H,T,N)\leqslant\sum_t t\sum_{h_1}\sum_{h_2\equiv h_1(t)}|\gamma_1(h_1,t)\gamma(h_2,t)|$$
$$\cdot\Big|\sum_n g\Big(\frac{n}{N}\Big)J_{k-1}\Big(\frac{4\pi}{t}\sqrt{\frac{h_1n}{q}}\Big)J_{k-1}\Big(\frac{4\pi}{t}\sqrt{\frac{h_2n}{q}}\Big)\Big|.$$

Assuming $N\geqslant 128HT^{-2}q^{-1}$ (this is very weak condition) we find that the sum over n is equal to the corresponding integral

$$\int_0^\infty g\Big(\frac{x}{N}\Big)J\Big(\frac{4\pi}{t}\sqrt{\frac{h_1x}{q}}\Big)J\Big(\frac{4\pi}{t}\sqrt{\frac{h_2x}{q}}\Big)dx \tag{2.30}$$

up to an error term $O(N^{-A})$ which is negligible . Moreover the above integral has no stationary phase, so

$$(2.30)\quad\ll N\Big(1+\frac{|\sqrt{h_1}-\sqrt{h_2}|}{T}\sqrt{\frac{N}{q}}\Big)^{-2}\min(1,T\sqrt{\frac{q}{HN}})$$
$$\asymp T\sqrt{\frac{qN}{H}}\Big(1+\Big(\frac{h_1-h_2}{T}\Big)^2\frac{N}{qH}\Big)^{-1}\Big(1+\frac{qT^2}{TN}\Big)^{-\frac{1}{2}}.$$

Then, given h_1 the resulting sum over $h_2\equiv h_1(\mathrm{mod}\ t)$ yields

$$\sum_{h_2}\Big(1+\Big(\frac{h_1-h_2}{T}\Big)^2\frac{N}{qH}\Big)^{-1}\ll 1+\sqrt{\frac{qH}{N}}.$$

Hence we conclude that the terms $t_1=t_2, d_1=d_2$ contribute at most

$$\mathcal{A}_0^*(H,T,N)\ll qT^2\Big(1+\frac{N}{qH}\Big)^{\frac{1}{2}}\Big(1+\frac{qT^2}{HN}\Big)^{-\frac{1}{2}}\|\gamma\|^2. \tag{2.31}$$

We are left with the terms not satisfying

$$\frac{d_1}{t_1}\equiv\frac{d_2}{t_2}(\mathrm{mod}\ 1) \tag{2.32}$$

As above the sum over n equals to the corresponding integral

$$\int_0^\infty g\Big(\frac{x}{N}\Big)e\Big(x\|\frac{d_1}{t_1}-\frac{d_2}{t_2}\|\Big)J\Big(\frac{4\pi}{t_1}\sqrt{\frac{h_1x}{q}}\Big)J\Big(\frac{4\pi}{t_2}\sqrt{\frac{h_2x}{q}}\Big)dx. \tag{2.33}$$

up to an error term $O(N^{-A})$ which is negligible. Now, however, one may encounter a stationary phase, unless we assume that the linear part in the exponential dominates over the square root parts in the exponentials coming from the Bessel functions . Note that

$$\|\frac{d_1}{t_1}-\frac{d_2}{t_2}\|\geqslant\frac{1}{t_1t_2}\geqslant\frac{1}{4T^2}.$$

Therefore, in order to miss the stationary point, we assume that N is sufficiently large, say

$$N \geqslant \Big(R + 128\frac{H}{q}\Big)T^2 \tag{2.34}$$

where $R \geqslant 1$ is at our disposal. Then the integral (E1) is $\ll NR^{-A}$ for any $A \geqslant 0$ by partial integration A times. Hence the terms (2.32) contributes to $\mathcal{A}^*(H,T,N)$ at most

$$\sum_{h_1}\sum_{t_1}\sum_{h_2}\sum_{t_2} |\gamma(h_1,t_1)\gamma(h_2,t_2)| t_1 t_2 N R^{-A} \ll HT^3NR^{-A}\|\gamma\|^2.$$

Adding this bound to (2.33) we obtain a bound for $\mathcal{A}_0^*(H,T,N)$. Finally applying the duality principle we deduce the following bound for the original sum $\mathcal{A}(H,T,N)$.

$$\mathcal{A}(H,T,N) \ll \{qT^2\Big(1+\frac{N}{qH}\Big)^{\frac12}\Big(1+\frac{qT^2}{HN}\Big)^{-\frac12} + HT^3NR^{-A}\}\|\alpha\|^2. \tag{2.35}$$

where A is any positive number, the implied constant depends only on k, A.

The condition (2.34) can be relaxed by the following device which is reminiscent to the one used in [FV].

$$\sum_{\substack{P<p\leqslant 2P\\ p\nmid t}} \log p \gg \log P \quad \text{ if } \ P \gg \log t.$$

Hence

$$\mathcal{A}(H,T,N) \ll \frac{1}{P}\sum_{P<p\leqslant 2P} (\log p) \sum_h \sum_{(t,p)=1} |a_n S(h\bar{q},n;t) J\Big(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\Big)|^2.$$

Given p we can write $S(h\bar{q},n;t) = S(h\overline{qp},np;t)$ and $J\Big(\frac{4\pi}{t}\sqrt{\frac{hn}{q}}\Big) = J\Big(\frac{4\pi}{t}\sqrt{\frac{hnp}{qp}}\Big)$. Hence the resulting sum is of the same type as $\mathcal{A}(H,T,N)$, but with q replaced by qp, N replaced by Np and the sequence $\alpha = (a_n)$ replaced by the lacunary sequence (a_{nq}) of the same l_2-norm. The above arguments are applicable in the modified situation provided

$$NP \geqslant \Big(R + 128\frac{H}{qP}\Big)T^2 \tag{2.36}$$

getting

$$\mathcal{A}(H,T,N) \ll \Big\{qPT^2\Big(1+\frac{N}{qH}\Big)^{\frac12}\Big(1+\frac{qT^2}{HN}\Big)^{-\frac12} + HT^3NPR^{-A}\Big\}\|\alpha\|^2. \tag{2.37}$$

We take

$$P = \frac{RT^2}{N} + 12T\sqrt{\frac{H}{qN}} + \log T \tag{2.38}$$

which satisfies (2.36). Then we take $R = N^\varepsilon$, so

$$\begin{aligned} P &\ll \Big(1+\frac{T}{q}\Big)N^\varepsilon \\ \Big(1+\frac{T}{q}\Big)\Big(1+\frac{qT^2}{HN}\Big)^{-\frac{1}{2}} &\ll \Big(1+\frac{HN}{q^3}\Big)^{\frac{1}{2}} \\ \Big(1+\frac{N}{qH}\Big)^{\frac{1}{2}}\Big(1+\frac{HN}{q^2}\Big)^{\frac{1}{2}} &\ll \Big(1+\frac{N}{q^2}+\sqrt{\frac{N}{qH}}\Big)N^\varepsilon. \end{aligned}$$

Then we are done. □

We have the following corollary:

Corollary 2.2. *Let q be prime and $\alpha = (a_n)$ be any sequence of complex numbers for $N < n \leqslant 2N$. Then $k \geqslant 2$ and $1 \leqslant H \leqslant T$, then*

$$\frac{2}{\phi(q)} \sum_{\substack{\chi (mod\ q) \\ \chi(-1)=(-1)^k}} \sum_{f\in\mathcal{F}_\chi} |\mathcal{L}_f(\alpha)|^2| \ll_{\varepsilon,k} \Big(\frac{N}{q^2} + \sqrt{\frac{N}{q}} + 1\Big)N^\varepsilon \|\alpha\|^2. \tag{2.39}$$

Remark. From the above corollary, in the most interesting range $N \asymp q^2$, the main term is bounded by $\sqrt{q}\|\alpha\|^2$ which is much larger than the bound $\|\alpha\|^2$ as expected from the classical large sieve (2.20). The large sieve in this case fails its power as a substitute as the Riemann hypothsis! The following proposition shows the upper bound we gave is optimal: Taking the vector $\alpha_0 = (a_0(n)), N < n \leqslant 2N$ with

$$a_0(n) = S(h_0\bar{q}, n; t_0)J\Big(\frac{4\pi}{t_0}\sqrt{\frac{h_0 n}{q}}\Big) \tag{2.40}$$

for $N < n \leqslant 2N$,

$$qt_0^2N^{-1+\varepsilon} \leqslant h_0 \leqslant \min\Big(\frac{N}{q}, \frac{q^3}{N}\Big)N^{-\varepsilon}, \tag{2.41}$$

one can show

Proposition 2.4.

$$\frac{2}{\phi(q)} \sum_{\substack{\chi (mod\ q) \\ \chi(-1)=(-1)^k}} \sum_{f\in\mathcal{F}_\chi} |\mathcal{L}_f(\alpha_0)|^2 \geqslant \frac{\phi(t_0)}{t_0}\sqrt{\frac{N}{qh_0}}(1+O(N^{-\varepsilon}))\|\alpha_0\|^2 \tag{2.42}$$

where the implied constant depends only on ε.

Remarks. This interesting phenomena shows the Fourier coefficients of f point towards the direction of families of vectors formed by products of Kloosterman sums and Bessel functions.

3 Nonholomorphic automorphic forms

3.1 Invariant differential operators

Let's back to the Poincare half plane $\mathbb{H}$. The hyperbolic differential operators derived from the differential $ds^2 = y^{-2}(dx^2 + dy^2)$ is given by

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

It is easy to check that Δ is $SL(2, \mathbb{R})$ invariant which means that for any functions f on $\mathbb{H}$,

$$\Delta(f(\gamma z)) = (\Delta f)(\gamma z)$$

for all $\gamma \in SL(2, \mathbb{R})$, and the algebra of all the differential operaotrs is generated by Δ.

Definition 3.1. *A function $f : \mathbb{H} \to \mathbb{C}$ with continuous partial derivatives of order 2 is an eigenfunction of Δ of type ν if*

$$\Delta f = \nu(1 - \nu)f.$$

Since Δ is an elliptic operator, all its eigenfunctions are real analytic. One can easily check that $I_\nu(z) := y^\nu$ is an eigenfunction of Δ of type ν.

3.2 Fourier expansions of nonholomorphic automorphic forms

From now on, let's fix $\Gamma = SL(2, \mathbb{Z})$ for simplicity although there is no essential difficulty to generalize the whole theory to any congruence subgroups.

Definition 3.2. *An automorphic form f of type ν acting on $\mathbb{H}$ is an eigenfunction of Δ of type ν satisfying the automorphy condition*

$$f(\gamma z) = f(z)$$

for any $\gamma \in \Gamma$.

$L^2(\Gamma \setminus \mathbb{H})$ equipped with the Petersson inner product is an Hilbert space. The main goal of the Harmonic analysis on $\Gamma \setminus \mathbb{H}$ is to decompose $L^2(\Gamma \setminus \mathbb{H})$ into a direct sum of eigenspaces of Δ, i.e., the so-called spectral decomposition.
An automorphic form of type ν is a periodic function, so it must have such a Fourier expansion

$$f(z) = \sum_{m \in \mathbb{Z}} a_m(y)e(mx). \tag{3.1}$$

Set $W_\nu^m(z) = a_m(y)e(mx)$, then it satisfies the following conditions:

$$\Delta W_\nu^m(z) = \nu(1 - \nu)W_\nu^m(z), \tag{3.2}$$

$$W_\nu^m(z + u) = W_\nu^m(z)e(mu), \tag{3.3}$$

for any $u \in \mathbb{R}$. From the classical theory of differential equations, we have

Theorem 3.1. *(Multiplicity one theorem) The function f satisfying (3.2) and (3.3) which grows not too fast, namely $f(z) = o(e^{2\pi y})$ as $y \to \infty$ is unique up to a multiple of a constant. More precisely, $f(z)$ must be a constant multiple of*

$$W_\nu^m(z) = \sqrt{2}(\pi|m|)^{\nu-\frac{1}{2}}\Gamma(\nu)^{-1}\sqrt{2\pi y}K_{\nu-\frac{1}{2}}(2\pi|m|y)e(mx) \tag{3.4}$$

with $K_\nu(y)$ the classical $K-$ Bessel function.

Remark. $W_\nu^m(z)$ is usually called the Whittaker function of type ν associated to the character $e(mx)$.
From the above observation, we have the following

Corollary 3.1. *Any automorphic form of type ν with the growth condition $o(e^{2\pi z})$, as $y \to \infty$ has the Fourier-Whittaker expansion*

$$f(z) = f(y) + \sum_{n\neq 0} a_n\sqrt{y}K_{\nu-\frac{1}{2}(2\pi|n|y)}e(nx). \tag{3.5}$$

Definition 3.3. *A Maass cusp form f is a bounded automorphic form whose zero term in the Fourier expansion is 0, i.e, $f(y) = 0$.*

For the modular surface $\Gamma = SL(2,\mathbb{Z})$, the first eigenvalue corresponding to Maass cusp forms is $\geqslant 0.91$ as numerical analysis shows. While for the congruence subgroups Selberg's conjecture says that the lowest laplacian eigenvalue is $\geqslant \frac{1}{4}$, the best upper bound up to now is due to Kim and Sarnak [KS].

3.3 Poincare series and Eisenstein series

By analogy with the classical Poincare series, following Selberg, we construct real analytic Poincare series using the following datas:
• The power function y^s;
• The so-called $E-$ function which is bounded on $\mathbb{H}$ and satisfies $E_m(z+u) = e(mu)E_m(z)$ for all $u \in \mathbb{R}$ and $z \in \mathbb{H}$.

Definition 3.4. *The real analytic Poincare series is defined as*

$$P_m(z,s,E) =: \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \Im(\gamma z)^s E_m(\gamma z). \tag{3.6}$$

It is easy to see that $P_m(z,s,E)$ is absolutely convergent in the half plane $\Re s > 1$.

Definition 3.5. *The Eisenstein series is*

$$E(z,s) =: \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (\Im\gamma z)^s \tag{3.7}$$

which is absolutely convergent in the half plane $\Re s > 1$.

Obviously the Eisenstein series is a Poincare series with $m = 0$.
We start to derive the Fourier expansion of the Poincare series:
Set

$$a_n(y) = \int_0^1 P_m(z,s,E)e(-nx)dx, \tag{3.8}$$

by the Bruhat decomposition,

$$P_m(z,s,E) = y^s E_m(iy)e(mx) + \sum_{c>0} \sum_{\substack{d(\bmod c) \\ (d,c)=1}} I_m(c,d,z,s,E) \tag{3.9}$$

with

$$\begin{aligned} I_m(c,d,z,s,E) &= y^s \sum_{l\in\mathbb{Z}} \frac{1}{|c(z+l)+d|^{2s}} E_m\Big(\frac{a}{c} - \frac{1}{c(c(z+l)+d)}\Big) \\ &= y^s e\Big(\frac{ma}{c}\Big) \sum_{l\in\mathbb{Z}} \frac{1}{(c(x+l)+d)^2 + c^2y^2)^s} E_m\Big(- \frac{1}{c(c(z+l)+d)}\Big) \end{aligned}$$

by the property of the E function. To compute $a_n(y)$, it leads to compute
(3.10)

$$J_m(c,d,y,s,E) =: \int_0^1 \frac{1}{(c(x+l)+d)^2 + c^2y^2)^s} E_m\Big(-\frac{1}{c(c(z+l)+d)}\Big)e(-nx)dx,$$

by a change of variable, the above equals

$$\frac{1}{c}e\Big(\frac{nd}{c}\Big)\int_{cl+d}^{c(l+1)+d} (x^2+c^2y^2)^{-s} E_m\Big(\frac{-1}{c(x+icy)}\Big)e\Big(-\frac{nx}{c}\Big)dx = \frac{1}{c}e\Big(\frac{nd}{c}\Big)J_m^*(c,d,y,s,E)$$

the factor $e\Big(\frac{nd}{c}\Big)$ and $e\Big(\frac{md}{c}\Big)$ toghether gives us the Kloosterman sum $S(m,n;c)$ as d runs through the residue class modulo c such that $(d,c) = 1$, since

$$\begin{aligned} \sum_{l\in\mathbb{Z}} J_m^*(c,d,y,s,E) &= \int_{-\infty}^{\infty} (x^2+c^2y^2)^{-s} E_m\Big(\frac{-1}{c(x+icy)}\Big)e\Big(\frac{-nx}{c}\Big)dx \\ &= (cy)^{1-2s}\int_{-\infty}^{\infty} (x^2+1)^{-s} E_m\Big(\frac{-1}{c^2y(x+i)}\Big)e(-nxy)dx \\ &=: (cy)^{1-2s} Q_m(c,y,s,E), \end{aligned}$$

we have the following theorem:

Theorem 3.2. *For* $\Re s \geqslant 1$, $n \in \mathbb{Z}$,

$$\begin{aligned} a_n(y) &= \int_0^1 P_m(z,s,E)e(-nx)dx \\ &= y^s E(imy)\delta_{nm} + y^{1-s}\sum_{c>0} \frac{1}{c^{2s}} S(m,n;c) Q_m(c,y,s,E). \end{aligned}$$

In case $m = 0$, it is easy to show

$$Q_0(c, y, s, E) = \int_{-\infty}^{\infty} (x^2 + 1)^{-s} e(-nxy) dx$$

$$= \begin{cases} 2\pi^s |ny|^{s-\frac{1}{2}} \Gamma(s)^{-1} K_{s-\frac{1}{2}}(2\pi|ny|), & \text{if } n \neq 0 \\ \sqrt{\pi}\Gamma(s - \frac{1}{2})\Gamma(s)^{-1} & \text{otherwise.} \end{cases} \tag{3.11}$$

If $m = 0$, the Kloosterman sum degenerates to the Ramanujan sum

$$S(0, n; c) = \sum_{a (\bmod c)}^{*} e\Big(\frac{dn}{c}\Big) = \sum_{d|(c,n)} \mu\Big(\frac{c}{d}\Big) d,$$

and $S(0, 0; c) = \phi(c)$ the Euler function, it follows that

$$\sum_{c=1}^{\infty} c^{-2s} S(0, n; c) = \zeta(2s)^{-1} \sum_{d|n} d^{1-2s}$$

and

$$\sum_{c=1}^{\infty} c^{-2s} S(0, 0; c) = \zeta(2s)^{-1} \zeta(2s - 1),$$

where $\zeta(s)$ is the Riemann zeta function, now we end up with the following Fourier expansion of the Eisenstein series:

Corollary 3.2. *For $\Re s > 1$, we have*

$$E(z, s) = y^s + \rho(s) y^{1-s} + \sum_{n \neq 0} \rho(n, s) \sqrt{2\pi y} K_{s-\frac{1}{2}}(2\pi|n|y) e(nx) \tag{3.12}$$

where

$$\rho(s) = \pi^{\frac{1}{2}} \Gamma(s - \frac{1}{2}) \Gamma(s)^{-1} \zeta(2s - 1) \zeta(2s)^{-1}, \tag{3.13}$$

$$\rho(n, s) = \sqrt{2} |n|^s \Gamma(s)^{-1} \pi^{s-\frac{1}{2}} \zeta(2s)^{-1} \sum_{d|n} d^{1-2s}. \tag{3.14}$$

From the above Fourier expansion of the Eisenstein series, we can derive its meromorphic continuation and functional equation:

Theorem 3.3. *The Eisenstein series can continued meromorphically to the whole complex plane. The modified Eisenstein series $E^*(z, s) = \pi^{-s}\Gamma(s)\zeta(2s)E(z, s)$ is regular except for simple poles at $s = 0$ and 1 and satisfies the functional equation:$E^*(z, s) = E^*(z, 1 - s)$ for all $z \in \mathbb{H}$. Specifically $E(z, s)$ is holomorphic as a function of s on $\Re s = \frac{1}{2}$.*

As the classical Poincare series, another important property of real analytic Poincare series is that it picks up Fourier coefficients of automorphic forms. Particularly, for $E_m(z) = e(mz)$,, i.e,

$$P_m(z, s, E) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\Im(\gamma z))^s e(m\gamma z), \tag{3.15}$$

for f a Maass cusp form of type ν which has a Fourier expansion (3.1) with $a_0(y) = 0$ we have

Proposition 3.1.

$$< P_m(z, s, E), f >= (2\pi)^{\frac{3}{2}} \sqrt{m} (4\pi m)^{-s} \bar{a}_m \frac{\Gamma(s - \bar{\nu})\Gamma(s - \nu)}{\Gamma(s)}$$

and

$$\frac{1}{\pi} < P_m(z, s, E), E(z, \frac{1}{2} + ir) >= 2^{2-2s} (n\pi)^{\frac{1}{2}-s-ir} \sigma_{2ir}(n) \frac{\Gamma(s - \frac{1}{2} + ir)\Gamma(s - \frac{1}{2} - ir)}{\Gamma(s)\Gamma(\frac{1}{2} - ir)\zeta(1 - 2ir)}$$

with $\sigma_s(n) = \sum_{0<d|n} d^s$.

With the help of the above formulas and the spectral theory decomposition in the next section one can obtain the meromorphic continuation of $P_m(z, s, E)$ as Selberg did [Se].

3.4 Selberg's spectral decomposition

In this section, we are going to prove Selberg's spectral decomposition of $L^2(\Gamma \setminus \mathbb{H})$ which states that

$$L^2(\Gamma \setminus \mathbb{H}) = \mathbb{C} \oplus L^2_{\text{cusp}} \oplus L^2_{\text{cont}}$$

where $\mathbb{C}$ is the space of constant functions, L^2_{cusp} is the space of Maass cusp forms and L^2_{cont} is the space of integrals of the Eisenstein series which are square integrable. the existence of continuous spectrum L^2_{cont} is due to the fact that $\Gamma \setminus \mathbb{H}$ has cusps which make it noncompact. More concretely, it states the following

Theorem 3.4. *Let* $u_0(z) = \sqrt{\frac{3}{\pi}}$ *and* u_j *for* $j \geqslant 1$ *is an orthonormal basis of the space of Maass cusp forms, for any* $f \in L^2(\Gamma \setminus \mathbb{H})$, *we have*

$$f(z) = \sum_{j \geqslant 0} < f, u_j > u_j(z) + \frac{1}{4\pi} \int_{\mathbb{R}} < f, E(z, \frac{1}{2} + ir) > E(z, \frac{1}{2} + ir > dr.$$

To prove the Selberg spectral decomposition theorem, we need the following lemma:

Lemma 3.1. *For any bounded* $f \in L^2(\Gamma \setminus \mathbb{H})$ *and* $< f, 1 >= 0$, *let*

$$\check{f}(z) = f(z) - \frac{1}{4\pi} \int_{\mathbb{R}} < f, E(z, \frac{1}{2} + ir) > E(z, \frac{1}{2} + ir > dr,$$

then $\check{f}(z)$ *is a Maass cusp form.*

The reference of the following proof is [Za] or [Go]:

Proof. Let

$$f(z) = \sum_{n \in \mathbb{Z}} a_n(y) e(nx)$$

be the Fourier expansion of f, for $\Re s > 1$, by unfolding the integral, we have

$$\begin{aligned} < f, E(, \bar{s}) >&= \int_{\Gamma \setminus \mathbb{H}} f(z) \sum_{\gamma \in \Gamma_\infty \setminus \Gamma} (\Im \gamma z)^s d\mu z \\ &= \int_{\Gamma_\infty \setminus \mathbb{H}} f(z) y^s d\mu z = \int_0^\infty \int_0^1 f(z) y^s \frac{dxdy}{y^2} \\ &= \int_0^\infty a_0(y) y^s \frac{dy}{y^2} =: M(s-1). \end{aligned}$$

Since $\mathrm{Res}_{s=1} E(z,s) = \frac{3}{\pi}$ and $< f, \frac{3}{\pi} >= 0$, $M(s-1)$ has no pole at $s=1$. It follows that $M(s-1)$ is holomorphic for $\Re s \geqslant \frac{1}{2}$. Also $M(s-1)$ has the following functional equation inheriting the functional equation of $E(z,s)$:

$$M(s-1) = \phi(s) M(1-s) \tag{3.16}$$

or equivalently

$$M(-s) = \phi(1-s) M(s-1). \tag{3.17}$$

By Mellin inversion, for $c > 1$,

$$a_0(y) = \frac{1}{2\pi i} \int_{\Re s = c} M(s-1) y^{1-s} ds.$$

Since $M(s-1)$ is holomorphic for $\Re s \geqslant \frac{1}{2}$, we can move the line of integration to $\Re s = \frac{1}{2}$. We then have

$$\begin{aligned} a_0(y) &= \frac{1}{2\pi i} \int_{\Re s = \frac{1}{2}} M(s-1) y^{1-s} ds \\ &= \frac{1}{2\pi i} \int_{\Re s = \frac{1}{2}} M(-s) y^s ds \end{aligned}$$

by changing variable $s \to 1 - s$.

Applying the functional equation in Theorem 3.3, we have

$$\begin{aligned} a_0(y) &= \frac{1}{4\pi i} \int_{\Re s = \frac{1}{2}} M(-s)(y^s + \phi(s) y^{1-s}) ds \\ &= \frac{1}{4\pi i} \int_{\Re s = \frac{1}{2}} M(\bar{s} - 1)(y^s + \phi(s) y^{1-s}) ds \end{aligned}$$

Recall that $y^s + \phi(s)y^{1-s}$ is the constant term of $E(z,s)$, the above means that the constant term of $\check{f}(z)$ is 0. So we are done. □
One can also show the resolvent of the Laplacian is a compact operator on $L^2_{\text{cusp}}(\Gamma \setminus \mathbb{H})$, so Δ has pure point spectrum on this space, Selberg's spetral decomposition follows from this observation and the above lemma.

3.5 Kuznetsov's formula

This section is devoted to the derivation of Kuznetsov's formula which is a generalization of Poisson's summation formula to nonholomorphic forms. We will use Zagier's unpublished proof because it is easy to be generalized to $GL(n)$.
We need to introduce some invariant integral operators. for each k on $\mathbb{H} \times \mathbb{H}$ which is point pair invariant, i.e, $k(\gamma z, \gamma z') = k(z, z')$ for any $\gamma \in SL(2,\mathbb{R})$, obviously, k only depends on the hyperbolic distance between two points, we hence denote it as $k(u(z,z'))$. Define an integral operator L_k which acts on $f \in L^2(\Gamma \setminus \mathbb{H})$ by

$$(L_k f)(z) := \int_{\mathbb{H}} f(w)k(z,w)d\mu w \tag{3.18}$$
$$= \sum_{\gamma\in\Gamma} \int_{\Gamma\setminus\mathbb{H}} f(\gamma w)k(z,\gamma w)d\mu w$$
$$= \int_{\Gamma\setminus\mathbb{H}} f(w)K(z,w)d\mu w$$

where

$$K(z,w) := \sum_{\gamma\in\Gamma} k(z,\gamma w) \tag{3.19}$$

is called an automorphic kernel.
It is easy to check that these integral operators L_k are invariant operators, i.e,

$$(L_k f)(\gamma z) = L_k(f(\gamma z)) \tag{3.20}$$

for all $\gamma \in SL(2,\mathbb{R})$ and they commute with the Laplacian. Furthermore, we have

Proposition 3.2. *Let ϕ be a Maass cusp form of type ν, then*

$$(L_k\phi)(z) = \hat{k}(\nu)\phi(z)$$

where $\hat{k}(\nu)$ depends only on k and ν. More precisely,

$$\hat{k}(\nu) = (L_k I_\nu)(i) = \int_{\mathbb{H}} k(w,i)(\Im w)^\nu dw.$$

$\hat{k}(\nu)$ is usually called the Selberg (or Harish-Chandra) transform of k. It has an inversion formula given by

$$k(u) = \frac{1}{4\pi}\int_{-\infty}^{\infty} F_\nu(u)\hat{k}(\nu)\tanh(\pi t)tdt \tag{3.21}$$

with $\nu = \frac{1}{2} + it$, for any $k(u) \in L^2(\mathbb{R})$, where $F_\nu(u) = P_{-\nu}(2u+1)$ being the spherical functions given explicitly by

$$F_\nu(u) = \frac{1}{\pi}\int_0^\pi (2u+1+2\sqrt{u(u+1)}\cos\theta)^{-\nu}d\theta, \tag{3.22}$$

and $\tanh(\pi t)tdt$ is the spectral measure.
The following propositon states the properties of the Selberg transform:

Proposition 3.3. *Set* $Q(w) = \int_w^\infty \frac{k(t)}{\sqrt{t-w}}dt, g(u) = Q(e^u + e^{-u} - 2)$, *then* $h(r) = \int_{-\infty}^\infty g(u)e\left(\frac{ru}{2\pi}\right)du$ *is the Selberg transform of* k. *Conversely,* $k(t) = -\frac{1}{\pi}\int_t^\infty \frac{dQ(w)}{\sqrt{w-t}}$.

Given $k(u) \in C_c^\infty(\mathbb{R})$, we will compute the $(m,n)-$ th Fourier coefficients of $K(z,z')$ with $m,n \geqslant 1$ in two ways. One way is to use the Bruhat decomposition and the other is to use the spectral decomposition of $L^2(\Gamma\setminus\mathbb{H})$. After these two computations are made, we end up with an identity which is termed as the pre-Kuznetsov formula.
To start, define

$$P_m(y,z') := \int_0^1 K(z,z')e(mx)dx, \tag{3.23}$$

we have

Lemma 3.2. *The function* $P_m(y,z')$ *is a Poincare series in* z'.

Proof. We rewrite $P_m(y,z')$ as

$$\begin{aligned} P_m(y,z') &= \sum_{\gamma\in\Gamma}\int_0^1 k(z,\gamma z')e(mx)dx \\ &= \sum_{\gamma\in\Gamma_\infty\setminus\Gamma}\int_{-\infty}^\infty k(z,\gamma z')e(mx)dx. \end{aligned} \tag{3.24}$$

Define

$$E_m(y,z') = \int_{-\infty}^\infty k(z,z')e(mx)dx,$$

then for $u \in \mathbb{R}$

$$\begin{aligned} E_m(y,z'+u) &= \int_{-\infty}^\infty k(z,z'+u)e(mx)dx \\ &= \int_{-\infty}^\infty k(z-u,z')e(mx)dx \\ &= e(mu)E_m(y,z') \end{aligned} \tag{3.25}$$

which shows that $E_m(y,z')$ is an $E-$ function in z'. Now the lemma follows from the definition of Poincare series.

It follows from Theorem 3.2 and the above lemma that

$$\int_0^1 P_m(y,z')e(-nx')dx' = \delta_{nm}\int_{-\infty}^{\infty} k(z,iy')e(mx)dx$$

$$+\sum_{c>0} S(m,n;c)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k\Big(z,\frac{-1}{c^2(x'+iy')}\Big)e(mx-nx')dxdx'. \tag{3.26}$$

On the other hand, since $K(z,z')\in L^2(\Gamma\setminus\mathbb{H})$ in z or z', by Selberg's spectral decomposition,

$$K(z,z') = \sum_{j\geqslant 0}\hat{k}(\nu_j)\phi_j(z')\bar{\phi}_j(z) + \frac{1}{4}\int_{\Re s=\frac{1}{2}}\hat{k}(s)E(z',s)\overline{E(z,s)}ds, \tag{3.27}$$

we can decompute the Fourier coefficients of $P_m(y,z')$ by the above spectral expansion of K :

$$\int_0^1 P_m(y,z')e(-nx')dx'$$

$$= \int_0^1\int_0^1 K(z,z'e(mx-nx')dxdx'$$

$$= 2\pi\sqrt{yy'}[\sum_{j\geqslant 1}\hat{k}(\nu_j)\rho_j(n)\bar{\rho}_j(m)K_{\nu_j-\frac{1}{2}}(2\pi|n|)y')\bar{K}_{\nu_j-\frac{1}{2}}(2\pi|m|y)$$

$$+\int_{\Re s=\frac{1}{2}}\hat{k}(s)\rho(n,s)\bar{\rho}(m,s)K_{s-\frac{1}{2}}(2\pi|n|y')\bar{K}_{s-\frac{1}{2}}(2\pi|m|y)ds] \tag{3.28}$$

where

$$\rho_j(l)\sqrt{y^*}K_{\nu-\frac{1}{2}}(2\pi|l|y^*) = \int_0^1 \phi_j(z^*)e(-lx)dx, \tag{3.29}$$

with $y^*=y$ or y' and $z^*=z$ or z' and $\rho(l,s)$ is defined by (3.14).

Remark. On the right of (3.28), the constant function $\phi_0(z)$ and $\phi_0(z')$ don't contribute because their $(m,n)-th$ Fourier coefficients are 0, for $m,n\geqslant 1$.

Comparing (3.26) with (3.28), we arrive at the following pre-Kuznetsov formula:

Proposition 3.4. *(Pre-Kuznetsov formula) For $k\in C_c^{\infty}(\mathbb{R})$, $\rho_j(l)$ and $\rho(l,s)$*

are defined by (3.29) and (3.14) , $m, n > 1$, we have

$$
\begin{aligned}
&2\pi\sqrt{yy'}\Big(\sum_{j\geqslant 1}\hat{k}(\nu_j)\rho_j(n)\bar{\rho}_j(m)K_{\nu_j-\frac{1}{2}}(2\pi|n|)y')\bar{K}_{\nu_j-\frac{1}{2}}(2\pi|m|y)\\
&\quad+\int_{\Re s=\frac{1}{2}}\hat{k}(s)\rho(n,s)\bar{\rho}(m,s)K_{s-\frac{1}{2}}(2\pi|n|y')\bar{K}_{s-\frac{1}{2}}(2\pi|m|y)ds\Big)\\
&=\delta_{m,n}\int_{-\infty}^{\infty}k(x+iy,iy')e(mx)+\sum_{c>0}S(m,n;c)H(y,y',c,n,m) \qquad (3.30)
\end{aligned}
$$

where

$$
H(y,y',c,n,m)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}k\Big(z,\frac{-1}{c^2(x'+iy')}\Big)e(mx-nx')dxdx'. \qquad (3.31)
$$

To obtain a handy formula, we intend to remove those $K-$ Bessel functions in the pre-Kuznetsov formula. For that purpose, we set

$$
ny'=my:=v \qquad (3.32)
$$

multiplying both sides of (3.30) by v^{-1} and integrate both sides with respect to v from 0 to ∞, due to

$$
\int_0^{\infty}|K_{iu}(x)|^2dx=\frac{\pi^2}{4}(\cosh\pi u)^{-1},
$$

for $u\in\mathbb{R}$, we have

$$
\begin{aligned}
&\frac{\pi^2}{4}\sqrt{\frac{1}{nm}}\Big(\sum_{j\geqslant 1}\hat{k}(\nu_j)\rho_j(n)\bar{\rho}_j(m)(\cosh\pi t_j)^{-1}\\
&\quad+\int_{\Re s=\frac{1}{2}}\hat{k}(s)\rho(n,s)\bar{\rho}(m,s)(\cosh\pi t)^{-1}dt\Big)\\
&=\frac{1}{2\pi}\Big(\delta_{m,n}\int_0^{\infty}\int_{-\infty}^{\infty}v^{-1}k(x+im^{-1}v,in^{-1}v)e(mx)dxdv\\
&\quad+\sum_{c>0}S(m,n;c)\int_0^{\infty}v^{-1}H(m^{-1}v,n^{-1}v,c)dv\Big). \qquad (3.33)
\end{aligned}
$$

After the following long computations, it turns out all of the above integrals can be explicitly given in terms of $\hat{k}$.

Lemma 3.3. *[Za]*

$$
\int_0^{\infty}\int_{-\infty}^{\infty}v^{-1}k(x+im^{-1}v,im^{-1}v)e(mx)dxdv=\frac{1}{8m\pi}\int_{-\infty}^{\infty}h(t)t\tanh\pi tdt.
$$

Proof. By (3.57), for $0 < \Re s < 1$,

$$\begin{aligned}\psi(s) &:= \int_0^\infty \int_{-\infty}^\infty v^{s-1} k(\alpha x^2) e(xv) dx dv \\ &= 2\int_0^\infty (2\pi|x|)^{-s}\Gamma(s)\cos\frac{\pi s}{2} k(\alpha x^2) dx \\ &= \frac{\Gamma(s)\cos\frac{\pi s}{2}}{(2\pi)^{s+1}} \int_0^\infty x^{-s} \int_{-\infty}^\infty P_{-\frac{1}{2}-it}(2\alpha x^2+1) h(t)\tanh(\pi t) t dt dx \\ &= (4\alpha)^{\frac{s}{2}-\frac{1}{2}} \frac{\Gamma(s)\cos\frac{\pi s}{2}}{2(2\pi)^{s+1}} \int_0^\infty u^{-\frac{s+1}{2}} \int_{-\infty}^\infty P_{-\frac{1}{2}+it}(\frac{u}{2}+1) h(t)\tanh(\pi t) t dt du,\end{aligned}$$

where we used the following formula:

$$\int_0^\infty v^{s-1} e(xv) dv = (2\pi|x|^{-s})\Gamma(s)\cos\frac{\pi s}{2}.$$

Interchanging the order of integration again and using the formula ([GR 7.134]):

$$\int_0^\infty u^{-\frac{s+1}{2}} P_{-\frac{1}{2}+it}\Big(\frac{u}{2}+1\Big) du = 2^{1-s} \frac{\Gamma\Big(\frac{1-s}{2}\Big)\Gamma\Big(\frac{s}{2}+ir\Big)\Gamma\Big(\frac{s}{2}-ir\Big)}{\Gamma\Big(\frac{1+s}{2}\Big)\Gamma\Big(\frac{1}{2}+ir\Big)\Gamma\Big(\frac{1}{2}-ir\Big)}$$

for $0 < \Re s < 1$, we have

$$\psi(s) = (4\alpha)^{\frac{s}{2}-\frac{1}{2}} \frac{2^{-s-2}\Gamma\Big(\frac{s}{2}\Big)}{\pi^{s+\frac{3}{2}}\Gamma\Big(\frac{1+s}{2}\Big)} \int_{-\infty}^\infty \Gamma\Big(\frac{s}{2}-it\Big)\Gamma\Big(\frac{s}{2}+it\Big) h(t) t \sinh \pi t dt. \tag{3.34}$$

Since both sides are holomorphic for $\Re s > 0$, the above formula holds whenever $\Re s > 0$. Particularly, if we take $s = 1$, we have

$$\psi(1) = \frac{1}{8\pi}\int_{-\infty}^\infty h(t) t \tanh \pi t dt, \tag{3.35}$$

it follows that

$$\begin{aligned}&\int_0^\infty \int_{-\infty}^\infty v^{-1} k\Big(\frac{m^2x^2}{v^2}\Big) e(mx) dx dv \\ &\quad = \frac{1}{m}\int_0^\infty \int_{-\infty}^\infty k(x^2) e(xv) dx dv \\ &\quad = \frac{1}{8\pi m}\int_{-\infty}^\infty h(t) t \tanh \pi t dt.\end{aligned} \tag{3.36}$$

□

Lemma 3.4.

$$\int_0^\infty v^{-1} H(m^{-1}v, n^{-1}v, c) dv = \frac{1}{c}\sqrt{\frac{1}{nm}} \int_{-\infty}^\infty J_{2it}\Big(\frac{4\pi\sqrt{mn}}{c}\Big)\frac{h(t)t}{\cosh t} dt$$

Proof.

$$\int_0^\infty v^{-1}H(m^{-1}v,n^{-1}v,c)dv$$

$$(3.37) = \int_0^\infty v^{-1}\int_{-\infty}^\infty\int_{-\infty}^\infty k\Big(x+i\frac{v}{m},\frac{-1}{c^2(x'(x'+ivn^{-1})}\Big)e(mx-nx')dxdx'dv.$$

Making the substitution,

$$x=\frac{1}{c}\sqrt{\frac{n}{m}}(u+t),\ x'=\frac{1}{c}\sqrt{\frac{m}{n}}u,\ v=\frac{\sqrt{mn}}{c}v',$$

then the above integral becomes

$$\frac{1}{c^2}\int_{-\infty}^\infty e\Big(\frac{\sqrt{mn}}{c}t\Big)V(t)dt$$

where

$$\begin{aligned} V(t) &= \int_0^\infty\int_{-\infty}^\infty k\Big(\frac{1}{c}\sqrt{\frac{n}{m}}(u+t+iv'),\frac{-1}{c\sqrt{\frac{m}{n}}(u+iv')}\Big)\frac{dudv'}{v'} \\ &= c\sqrt{\frac{m}{n}}\int_{\mathbb{H}}k\Big(z+\frac{1}{c}\sqrt{\frac{n}{m}}t,\frac{-n}{mc^2z}\Big)\frac{dxdy}{y} \qquad (3.38)\end{aligned}$$

with $z=x+iy$.
The next two paragraphs are devoted to computing the integral $V(t)$:
Let $\alpha=\frac{1}{c}\sqrt{\frac{n}{m}},\beta=\frac{n}{mc^2}$ and

$$V^*(\alpha,\beta)=\int_{\mathbb{H}}k\Big(z+\alpha,\frac{-\beta}{z}\Big)y\frac{dxdy}{y^2}.$$

• If $|\alpha|\leqslant 2\beta$, making the substitutaion $z\to\frac{z+\frac{\alpha}{2}-i\sqrt{\beta-\frac{\alpha^2}{4}}}{z+\frac{\alpha}{2}+i\sqrt{\beta-\frac{\alpha^2}{4}}}=re^{i\theta}$ which maps $\mathbb{H}$ to the unit disc, we have

$$\begin{aligned} V^*(\alpha,\beta) &= 4\sqrt{\beta-\frac{\alpha^2}{4}}\int_0^{2\pi}\int_0^1 k\Big(\frac{16(\beta-\frac{\alpha^2}{4})r^2}{\beta(r^2-1)^2}\Big)\frac{1}{1-2r\cos\theta+r^2}\frac{rdrd\theta}{1-r^2} \\ &= 8\pi\sqrt{\beta-\frac{\alpha^2}{4}}\int_0^1 k\Big(\frac{16(\beta-\frac{\alpha^2}{4})r^2}{\beta(r^2-1)^2}\Big)\frac{rdr}{(1-r^2)^2}\end{aligned}$$

By changing variable $x=2A\frac{1+r^2}{1-r^2}$, with $A=\sqrt{\frac{\beta-\frac{\alpha^2}{4}}{\beta}}$, we have

$$V^*(\alpha,\beta)=\pi\sqrt{\beta}\int_A^\infty k(x^2-4A^2)dx. \qquad (3.39)$$

• If $|\alpha| \geqslant 2\beta$, by changing variable, $z \to \frac{z+\frac{\alpha}{2}-\sqrt{\frac{\alpha^2}{4}-\beta}}{z+\frac{\alpha}{2}+\sqrt{\frac{\alpha^2}{4}-\beta}} = \zeta + i\eta$ mapping the upper half plane to itself, setting

$$B = \sqrt{\frac{\frac{\alpha^2}{4} - \beta}{\beta}}$$

we have

$$\begin{aligned} & V^*(\alpha,\beta)2\sqrt{\beta - \frac{\alpha^2}{4}} \int_{\mathbb{H}} k\Big(B^2\frac{\zeta^2+\eta^2}{\eta^2}\Big)\frac{1}{(1-\zeta)^2+\eta^2}\frac{d\zeta d\eta}{\eta} \\ (3.40) \qquad & = \frac{\pi}{2}\sqrt{\beta}\int_{-\infty}^{\infty} k(x^2+4\beta^2)dx. \end{aligned}$$

From (3.39) and (3.40), we conclude that

$$(3.41) \qquad V(t) = \begin{cases} \frac{\pi}{2}\int_{x^2+t^2\geqslant 4} k(x^2+t^2-4)dx & \text{if } \ |t| \leqslant 2 \\ \frac{\pi}{2}\int_{-\infty}^{\infty} k(x^2+t^2-4)dx & \text{if } \ |t| \geqslant 2. \end{cases}$$

In the second case, $|t| \geqslant 2$,, assuming $t = 2\cosh\frac{\alpha}{2}$, then

$$\begin{aligned} V(t) &= \frac{\pi}{2}\int_{-\infty}^{\infty} k(x^2+4\sinh^2\frac{\alpha}{2})dx \\ &= \frac{\pi}{2}Q(4\sinh^2\frac{\alpha}{2}) = \frac{\pi}{2}g(\alpha) \\ (3.42) \qquad &= \frac{1}{4}\int_{-\infty}^{\infty} h(r)\cos r\alpha dr \end{aligned}$$

In the first case, $|t| \leqslant 2$, by changing variable, we have

$$V(t) = \frac{\pi}{4}\int_0^{\infty}\int_0^{\infty} k(u)\frac{du}{\sqrt{u+c(t)}}$$

where $c(t) = 4 - t^2$. According to Proposition 3.3, the above is

$$\frac{-1}{4}\int_0^{\infty}\frac{1}{\sqrt{t+c(t)}}\int_t^{\infty}\frac{dQ(w)}{\sqrt{w-t}}dt = \frac{-1}{4}\int_0^{\infty}\int_0^{w}\frac{dtdQ(w)}{\sqrt{(t+c(t))(w-t)}}$$

by partial differential, it is equal to

$$\begin{aligned} & \frac{1}{4}\int_0^{\infty} Q(w)d\Big(\int_0^{w}\frac{dt}{\sqrt{(t+c(t))(w-t)}}\Big) \\ & = \frac{1}{4}\int_0^{\infty} Q(w)d\Big(\frac{\pi}{2} + \arcsin\frac{w-c(t)}{w+c(t)}\Big) \\ & = \frac{1}{4}\int_0^{\infty} Q(w)\frac{\sqrt{c(t)}}{\sqrt{w}(w+c(t))}dw. \end{aligned}$$

In terms of $g(u)$ and $h(r)$, the above equals

$$\begin{aligned}&\frac{1}{4}\sqrt{c(t)}\int_0^\infty g(u)\frac{e^{\frac{u}{2}}+e^{-\frac{u}{2}}}{(e^{\frac{u}{2}}-e^{-\frac{u}{2}})^2+c(t)}du\\ &\quad=\frac{\sqrt{c(t)}}{16\pi}\int_{-\infty}^\infty h(r)\int_{-\infty}^\infty \frac{e^{\frac{u}{2}}+e^{\frac{-u}{2}}}{(e^{\frac{u}{2}}-e^{-\frac{u}{2}})^2+c(t)}e\Big(\frac{ru}{2\pi}\Big)dudr.\end{aligned}$$

The inner integral above can be computed by the residue theorem, we obtain

$$V(t)=\frac{1}{4}\int_{-\infty}^\infty \frac{e^{-2l\pi r}+e^{-(2\pi-2l\pi)r}}{1+e^{-2\pi r}}h(r)dr$$

with $-c(t)=(e^{il\pi}-e^{-il\pi})^2$. Writing $t=2\cos\frac{\alpha}{2}$, i.e, $t^2-4=(e^{\frac{\alpha}{2}i}-e^{-\frac{\alpha}{2}i})^2$, we have

$$V(t)=\frac{1}{4}\int_{-\infty}^\infty h(r)\frac{\cosh(\pi-\alpha)r}{\cosh\pi r}dr. \tag{3.43}$$

From the above we have

$$\begin{aligned}&\int_{-\infty}^\infty V(t)e^{ixt}dt\\ &\quad=\frac{1}{4}\int_{-\infty}^\infty h(r)\cosh^{-1}\pi r\Big(-\int_{-\pi}^{\pi}\cosh(\pi-\alpha)re^{2ix\cos\frac{\alpha}{2}}\sin\frac{\alpha}{2}d\alpha\\ &\quad+\int_{-\infty}^\infty \cosh\pi r e^{ir\alpha+2ix\cosh\frac{\alpha}{2}}\sinh\frac{\alpha}{2}d\alpha\Big)dr\end{aligned}\tag{3.44}$$

by (4.4), the following integral representation of $J-$ Bessel function,

$$J_\nu(z)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(z\sin\theta-\nu\theta)d\theta-\frac{\sin\nu\pi}{\pi}\int_0^\infty e^{-z\sinh u-\nu u}du$$

and the recurrence formula

$$J_{\nu-1}(z)+J_{\nu+1}(z)=2\nu z^{-1}J_\nu(z),$$

the lemma follows easily. □

We arrive at the Kuznetsov formula which is stated in terms of the Selberg transform $h(t)$ of $k(u)$:

Theorem 3.5. *(Kuznetsov's formula) Let $h(r)$ be an even function of a complex variable r which is holomorphic in the strip $|\Im r|\leqslant\frac{1}{2}+\delta$ for some $\delta>0$ which is holomorphic in the strip $|\Im r|\leqslant\frac{1}{2}+\delta$ for some $\delta>0$ and $h(r)\ll_\varepsilon(|r|+1)^{-2-\varepsilon}$, in the above strip, then for any positive integers m,n, we have*

$$\begin{aligned}&\sum_{j\geqslant 1}\frac{\rho_j(n)\bar\rho_j(m)}{\cosh\pi r}h(r_j)+\frac{1}{\pi}\int_{-\infty}^\infty\Big(\frac{m}{n}\Big)^{ir}\sigma_{2ir}(n)\sigma_{-2ir}(m)\frac{h(r)dr}{|\zeta(1+2ir)|^2}\\ &\quad=\frac{\delta_{n,m}}{\pi^2}\int_{-\infty}^\infty r\tanh(\pi r)h(r)dr+\sum_{c\geqslant 1}\frac{1}{c}S(n,m;c)h^+\Big(\frac{4\pi\sqrt{mn}}{c}\Big)\end{aligned}\tag{3.45}$$

where

$$h^+(x) = \frac{2i}{\pi}\int_{-\infty}^{\infty} J_{2ir}(x)\frac{r}{\cosh \pi r}h(r)dr.$$

Remarks: 1. Actually we derived the above formula assuming $k(u) \in C_c^\infty(\mathbb{R})$, but one can easily justify the proof for $h(t)$ satisfying the conditions in the theorem.
2. The constant function in the spectrum doesn't show up in Kuznetsov's formula because the $m, n-th$ Fourier coefficient with $m, n > 1$ is 0.
If we take a special function

$$h(r) := H(r,t) = \frac{\cosh \pi r}{\cosh \pi(t-r)\cosh \pi(t+r)} \tag{3.46}$$

where t is a free parameter, one can show

$$\int_{-\infty}^{\infty} r\tanh(\pi r)H(r,t)dr = \frac{t}{\sinh \pi t} \tag{3.47}$$

and

$$h^+(x) = \frac{2i}{\pi}\int_{-\infty}^{\infty} J_{2ir}(x)\frac{r}{\cosh \pi r}H(r,t)dr = \frac{x}{\pi}D_{2it}(x) \tag{3.48}$$

where

$$D_{2it}(x) = -\frac{2it}{\sinh \pi t}\int_{-i}^{i} K_{2it}(xv)\frac{dv}{v}, \tag{3.49}$$

then we have

Corollary 3.3. *For $m, n \geqslant 1$, we have*

$$\begin{aligned} &\sum_{j\geqslant 1}\frac{\rho_j(n)\bar{\rho}_j(m)}{\cosh \pi r}H(\kappa_j,t) + \frac{1}{\pi}\int_{-\infty}^{\infty}\left(\frac{m}{n}\right)^{ir}\sigma_{2ir}(n)\sigma_{-2ir}(m)\frac{H(r,t)dr}{|\zeta(1+2ir)|^2} \\ &= \frac{\delta_{n,m}}{\pi^2}\frac{t}{\sinh \pi t} + \sum_{c\geqslant 1}\frac{4\pi\sqrt{mn}}{c^2}S(n,m;c)D_{2it}\Big(\frac{4\pi\sqrt{mn}}{c}\Big). \end{aligned} \tag{3.50}$$

In applications, we multiply $H(r,t)$ by a function $f(t)$, then integrating with respect to the free variable t gives us our favorable test functions on the spectral side.

3.6 Kuznetsov's inversion formula

Kloosterman sums as special kind of exponential sums have important applications in number theory. To study sums of Kloosterman sums using Kuznetsov type formula, we need flexibility on the choice of the test functions on the Kloosterman sum side. For that purpose, Kuznetsov derived an inversion formula. It is interesting that Petersson's formula also comes into the picture. Let ϕ_{jk} with

$j \geqslant 1$ be an orthonormal basis of the space of holomorphic forms of weight k and

$$\psi_{jk}(l) = \Big(\frac{\pi^{-k}\Gamma(k)}{(4l)^{k-1}}\Big)^{\frac{1}{2}} \hat{\phi}_{jk}(l) \tag{3.51}$$

be the normalized Fourier coefficients, recall that $\rho_j(l)$ and $\rho(l,s)$ are defined in (3.29) and (3.14), we have the following

Theorem 3.6. *(Kuznetsov's inversion formula) Let $f(x) \in C^2[0,\infty)$ which satisfies*

$$f(0) = f'(0) = 0,\ f^{(j)} \ll_\delta (x+1)^{-2-\delta},\ 0 \leqslant j \leqslant 2,\ \delta > 0,$$

then for any positive integers m, n, we have

$$\begin{aligned}\sum_{c\geqslant 1} &\frac{4\pi\sqrt{mn}}{c^2} S(m,n;c) f\Big(\frac{4\pi\sqrt{mn}}{c}\Big)\\ &= \sum_{j\geqslant 1} \frac{\rho_j(n)\bar{\rho}_j(m)}{\cosh \pi r} T_f(t_j) + \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{\bar{\rho}(m,t)\rho(n,t)}{\cosh \pi t} T_f(t)dt\\ &+ \sum_{0<k\equiv 0(mod\ 2)} \sum_{j\geqslant 1} i^k \bar{\psi}_{jk}(m)\psi_{jk}(n) N_f(k-1),\end{aligned} \tag{3.52}$$

where

$$T_f(t) = \int_0^\infty f(x) B_{2it}(x) x^{-1} dx \tag{3.53}$$

with $B_{2it}(x) = \pi(J_{-2it}(x) - J_{2it}(x))(2i\sinh \pi t)^{-1}$; we also have

$$\begin{aligned}\sum_{c\geqslant 1} &\frac{4\pi\sqrt{mn}}{c^2} S(m,-n;c) f\Big(\frac{4\pi\sqrt{mn}}{c}\Big)\\ &= \sum_{j\geqslant 1} \frac{\rho_j(n)\bar{\rho}_j(m)}{\cosh \pi r} K_f(t_j) + \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{\bar{\rho}(m,t)\rho(n,t)}{\cosh \pi t} K_f(t)dt\end{aligned} \tag{3.54}$$

where

$$K_f(t) = \frac{4}{\pi}\cosh \pi t \int_0^\infty K_{2it}(x) f(x) x^{-1}. \tag{3.55}$$

The main lemma we are going to use is the following Bessel expansion of a function f :

Lemma 3.5. *Suppose that $f(x) \in C^2[0,\infty), f(0) = 0, f^{(j)}(x) \ll_B (|x|+1)^{-B}$ for $0 \leqslant j \leqslant 2$, we have*

$$\begin{aligned} f(x) &= \sum_{k\geqslant 0} 2(2k+1) N_f(2k+1) J_{2k+1}(x)\\ &+ \int_0^\infty T_f(t) B_{2it}(x) \tanh \pi t t dt.\end{aligned} \tag{3.56}$$

Sketch of the proof: Let

$$H_f(y) = \int_0^\infty f(x) J_0(xy) dx$$

be the Hankel transform of f, then it has an inversion formula

$$\begin{aligned} f(x) &= \int_0^\infty H_f(y) J_0(xy) xy dy \\ &= \int_0^1 H_f(y) J_0(xy) xy dy + \int_1^\infty H_f(y) J_0(xy) xy dy. \end{aligned} \tag{3.57}$$

The fact

$$\int_0^1 H_f(y) J_0(xy) dy = \sum_{k \geqslant 0} 2(2k+1) N_f(2k+1) J_{2k+1}(x)$$

and

$$\int_0^1 H_f(y) J_0(xy) dy = \int_0^\infty T_f(t) B_{2it}(x) \tanh \pi t \, t dt$$

follow from the recurrence formulas of the $J-$ Bessel function:

$$(z^\nu J_\nu(z))' = z^\nu J_{\nu-1}(z),$$

$$J_{\nu-1}(z) + J_{\nu+1}(z) = 2\nu z^{-1} J_\nu(z),$$

and

$$(z^{-\nu} J_\nu(z))' = -z^{-\nu} J_{\nu+1}(z).$$

We are done. □

Be the above lemma and discontinuous integrals of Weber-schafheitbin ([Er, vol. 2], 7.74), one can verify that

$$\frac{2}{\pi} \int_0^\infty T_f(t) \tanh \pi t \, t dt = \frac{1}{\pi} \int_0^\infty f(x) J_0(x) dx \tag{3.58}$$

which we call f^∞.

On the other hand, by the recurrence relation for $J-$ Bessel function we have

$$\begin{aligned} &\sum_{1 \leqslant l \equiv 1 (\mathrm{mod}\ 2)} (2\pi i^k)^{-1} 2l N_f(l) \\ &= \frac{1}{\pi} \int_0^\infty (\sum_{r \geqslant 1} (-1)^r (2r-1) J_{2r-1}(x)) f(x) x^{-1} dx \\ &= -\frac{1}{2\pi} \int_0^\infty J_0(x) f(x) dx = -f^\infty \end{aligned} \tag{3.59}$$

Taking $h(t) = T_f(t)$ in the Kuznetsov formula, applying Petersson's formula for each space $S_{2k+1}(\Gamma)$ and then adding up all the k, the δ symbol parts are cancelled out because of (3.58) and (3.59), the other terms contribute exactly as those in the Kuznetsov inversion formula. This finishes the proof of the Kuznetsov inversion formula.

Remarks 1. By choosing a suitable test function on the Kloosterman sum side in the Kuznetsov formula and estimating the other side by mean value theorems of Fourier coefficients of Maass cusp forms and holomorphic forms using Kuznetsov's formula, Petersson's formula and Weil's bound for Kloosterman sums:

$$S(m,n;c) \ll c^{\frac{1}{2}}(m,n,c)^{\frac{1}{2}}\tau(c), \tag{3.60}$$

Kuznetsov [Ku] proved the following

$$\sum_{c \leqslant C} \frac{S(m,n;c)}{c^{\frac{1}{2}}} \ll_{\varepsilon,m,n} c^{\frac{2}{3}+\varepsilon} \tag{3.61}$$

for any $\varepsilon > 0$.

2. By a clever control of the growth rate of the Kloosterman zeta function

$$Z_{m,n}(s) = \sum_{c>0} \frac{S(m,n;c)}{c^{2s}}$$

on the vertical lines, Goldfeld and Sarnak [GS] was also able to prove (3.61) without using Kuznetsov's formula and the inversion formula.

3. Recently Foury and Michel [FM] showed

$$c_0(k)\frac{x}{\log x}(\log\log x)^k \leqslant \sum_{v \leqslant x} \frac{|S(1,1;v)|}{v^{\frac{1}{2}}},$$

so indeed (3.61) shows that cancellations in the sums of Kloosterman sums due to the variation of the signs of Kloosterman sums not simply due to the variation of the absoute values of Kloosterman sums exist.

3. Similar cancellations in the sums of Kloosterman sums with arithmetic progressions was recently proved by Luo-Rudnick-Sarnak [LRS] as a result of their bounds towards the Selberg eigenvalue conjecture.

4 Applications of Kuznetsov's formula

4.1 The multiplicity of the spectrum

It is believed from numerical experiments that the cuspidal spectrum of the modular surface is simple. This is an important conjecture because of the Pillips-Sarnak spectral deformation theory (see [PS]). However the progress on this problem is elusive and the avalible tools are very restrictive. In this section using Kuznetsov's formula we are going to obtain a bound (best up to now) on the multiplicity of the cuspidal spectrum. This section is taken from [Sa3] and

[LS].
Fix h a Schwartz class function, even, nonegative on $\mathbb{R}$ with $\int_{-\infty}^{\infty} h(x)dx = 1$ and the support of its Fourier transform $\hat{h}(\zeta) = \int_{-\infty}^{\infty} h(x)e(-x\zeta)dx$ in $(-1,1)$, we use this h to define a smoothed count of the Laplacian eigenvalues in the cuspidal spectrum. of the modular surface. For $1 \leqslant L \leqslant t^{1-\varepsilon}$, set

$$N_h(t,L) = \sum_{j\geqslant 1} h(L(t-t_j)),$$

where the summing runs through the cuspidal spectrum. Thus $N_h(t,L)$ counts the number of the Laplacian eigenvalues within L^{-1} of t and the larger we can take L the more information about the local distribution of the spectrum can be determined. A simple application of the Selberg trace formula shows that for $1 \leqslant L \leqslant \frac{\log t}{\pi}$,

$$N_h(t,L) \sim \frac{t}{6L}, \tag{4.1}$$

for $1 \leqslant L \leqslant \frac{\log t}{\pi}$. While by using the Kuznetsov formula we can extend the range suitably:

Theorem 4.1. *For $1 \leqslant L \leqslant \frac{2\log t}{\pi}$, we have*

$$N_h(t,L) \sim \frac{t}{6L} \tag{4.2}$$

as $t \to \infty$.

Corollary 4.1. *Let $m(t)$ be the multiplicity of the eigenvalue $\frac{1}{4} + t^2$, then*

$$\limsup_{t\to\infty} \frac{m(t)\log t}{t} \leqslant \frac{\pi}{12}. \tag{4.3}$$

While from (4.1), one can only obtain

$$\limsup_{t\to\infty} \frac{m(t)\log t}{t} \leqslant \frac{\pi}{6}.$$

The bound (4.1) is embarrasingly far from the believed $m(t) \leqslant 1$, but it is the best bound that we know.
The Kuznetsov formula involves sums over the spectrum weighted by Fourier coefficients of Maass cusp forms. These weights need to be removed (which is nontrivial) since the count $N_h(t,L)$ involves no weights. The gain in using Kuznetsov's formula comes from Weil's bound of the Kloosterman sums. The idea of introducing these weights and then removing them was first used by Iwaniec [Iw] in connection with improving the error term in counting closed geodesics on the modular surface.

We need to introduce the Rankin-Selberg $L-$ functions. Since the Hecke operators similarly defined as in the theory of holomorphic forms commute with the Laplacian, we can choose a common basis on the cuspidal spectrum for both. Assuming $\{u_j\}$ is such a basis with normalized Fourier coefficients $\nu_j(n)$ such that

$$\nu_j(n) = \Big(\frac{\pi}{\cosh \pi t_j}\Big)^{\frac{1}{2}} \rho_j(n).$$

Then for $n \geqslant 1$,

$$\nu_j(n) = \nu_j(1)\lambda_j(n)$$

where $\lambda_j(n)$ is u_j's eigenvalue for T_n. In particular, they satisfy

$$\lambda_j(n)\lambda_j(m) = \sum_{d|(m,n)} \lambda_j\Big(\frac{mn}{d^2}\Big).$$

For $\Re s > 1$, we define the Rankin-Selberg $L-$ functions $R_j(s)$ by

$$R_j(s) = \sum_{n \leqslant 1} |\nu_j(n)|^2 n^{-s}.$$

One can obatin the meromorphic continuation and functional equation of $R_j(s)$ by the corresponding properties of Eisensten series $E(z,s)$ due to the following

$$\begin{aligned} &\pi^{-1-s}\Gamma(s)^{-1}\Gamma\Big(\frac{s}{2}\Big)^2\Gamma\Big(\frac{s}{2}+it_j\Big)\Gamma\Big(\frac{s}{2}-it_j\Big)\cosh \pi t_j R_j(s) \\ &= 8\int_{\Gamma\backslash\mathbb{H}} |u_j(z)|^2 E(z,s) d\mu z. \end{aligned} \tag{4.4}$$

The above integral representation of the rankin-Selberg $L-$ function is proved by the following standard unfolding technique: The right side of (4.4) is equal to

$$\begin{aligned} &\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\Gamma\backslash\mathbb{H}} (u_j(z)|^2 (\Im\gamma z)^s d\mu z \\ &= \int_0^\infty \int_0^1 |u_j(z)|^2 y^s d\mu z \\ &= \sum_{n\geqslant 1} |\rho_j(n)|^2 \int_0^\infty |y^{\frac{1}{2}} K_{it_j}(2\pi n y)|^2 y^{s-2} dy, \end{aligned}$$

(4.4) follows from the above calculations and the following formula

$$\int_0^\infty K_\mu(x)K_\nu(x)x^{s-1}dx = 2^{s-3}\Gamma(s)^{-1}\prod \Gamma\Big(\frac{s\pm\mu\pm\nu}{2}\Big)$$

for $\Re s > |\Re\mu| + |\Re\nu|$. By Theorem 3.3, it is easy to prove the following: $L_j(s) =: \zeta(2s)R_j(s)$ has meromorphic continuation to the whole complex plane with a simple pole at $s = 1$ and residue 2. Set

$$\Lambda_j(s) = L_j(\infty, s)L_j(s)$$

with

$$L_j(\infty, s) = \pi^{-2s}\Gamma\Big(\frac{s}{2}\Big)^2\Gamma\Big(\frac{s}{2}+it_j\Big)\Gamma\Big(\frac{s}{2}-it_j\Big),$$

we have the functional equation

$$\Lambda_j(s) = \Lambda_j(1-s).$$

In the critical strip $0 < \Re s < 1$, we have the convexity bound:

Proposition 4.1. *For* $0 < \beta < 1$ *and* $\Re s = \beta$, *we have*

$$|R_j(s)| \ll_\varepsilon |t_j|^{1-\beta+\varepsilon}|s|^{2(1-\beta+\varepsilon)}.$$

Proof. By a bound due to Iwaniec ([Iw4], pp. 130),

$$\sum_{0<n<x} |\lambda_j(n)|^2 \ll x(\lambda_j N)^\varepsilon. \tag{4.5}$$

and the bound $|\nu_j(1)| \ll \lambda_j^\varepsilon$ of Hoffstein and Lockhart we have $R_j(s) \ll_\varepsilon \lambda_j^\varepsilon$ on the line $\Re s = 1+\varepsilon$. Hence on $\Re s = 1+\varepsilon$, we have

$$|L_j(s)| \ll |R_j(s)| \ll_\varepsilon |t_j|^\varepsilon.$$

Now applying the functional equation (..) and Stirling's formula we get

$$L_j(s)| \ll_\varepsilon |t_j|^{1+3\varepsilon}|s|^{2+4\varepsilon}$$

on $\Re s = -\varepsilon$. The conclusion therefore follows from the Phragmen-Lindelof principle. □

We will also make use of Luo's zero density theorem for the family $L_j(s)$ as a replacement of Lindelof hypothesis:

Proposition 4.2. *([Lu]) Let* $\eta > 0$ *be a sufficiently small constant, then at most* $t^{\frac{1}{5}}$ *of the* $R_j^{'}(s)$ *with* $|t_j| \leqslant t$ *have a zero in the rectangle*

$$1-\eta \leqslant \Re s \leqslant 1, \ \ |\Im s| \leqslant \log^3 t. \tag{4.6}$$

Furthermore, all but at most $t^{\frac{1}{5}}$ *of these* $R_j^{'}s$ *with* $|t_j| \leqslant t$ *satisfy*

$$|(s-1)R_j(s)| \ll_\varepsilon (|st|)^\varepsilon$$

for

$$1-\frac{\eta}{2} \leqslant \Re s \leqslant 1, \ \ |\Im s| \leqslant \log^2 t.$$

The proof needs Hadamard three circle theorem and classical techniques in the large sieve theory.

Now . the proofs of Theorem 4.1 and corollary 4.1. Let $h(x)$ satisfy the conditions stated before (4.1)and let

$$h_{t,L} = h(L(t-x)) + h(L(t+x)) \tag{4.7}$$

where t is large and $L \ll t^{\varepsilon}$. It is easy to see that $h_{t,L}$ can serve as a test function in the Kuznetsov formula. Note that

$$N_h(t,L) = \sum_{j\geqslant 1} h_{t,L}(t_j) + O(1). \tag{4.8}$$

• Removing the weights:
In using Kuznetsov's formula for the test function $h_{t,L}$, we introduce the weights $\nu_j(n)|^2$ and then remove them by averaging over $n \leqslant N$ (with N to be chosen.) According to Luo's zero density theorem with $2t$ instead of t, we split the set of $|t_j| \leqslant 2t$ into two sets G_1 and G_2. G_1 contains those t_j for which $R_j(s)$ ha sno zeroes in the rectangle (4.6) and G_2 contains the rest, by his result, $|G_2| \leqslant (2t)^{\frac{1}{5}}$. Now consider

$$\Omega_j(N) := \sum_{n=1}^{\infty} |\nu_j(n)|^2 e^{-\frac{n}{N}} = \frac{1}{2\pi i}\int_{\sigma=2} \Gamma(s)R_j(s)N^s ds.$$

Shifting the contour to $\Re s = \beta = 1-\delta$ with $\delta < \frac{\eta}{2}$ and η as in Luo's theorem and picking up a pole at $s=1$, we have

$$\Omega_j(N) = \frac{12N}{\pi} + I_j(N)$$

with

$$I_j(N) = \frac{1}{2\pi i}\int_{\Re s=\beta} \Gamma(s)R_j(s)N^s ds.$$

For $t_j \in G_1$ we apply Luo's theorem which gives

$$I_j(N) \ll_{\varepsilon} N^{\beta}t^{\varepsilon}.$$

For $t_j \in G_2$, we simply apply the convexity bound and find thta

$$I_j(N) \ll_{\varepsilon} N^{\beta}|t_j|^{1-\beta+\varepsilon}.$$

Hence, combing all of the above, we have

$$\begin{aligned} M_h(t,L) &:= \frac{1}{N}\sum_j h_{t,L}(t_j)\Omega_j(N) \\ &= \frac{12}{\pi}N_h(t,L) + \frac{1}{N}\sum_j h_{t,L}(t_j)I_j(N) + O(1) \\ &= \frac{12}{\pi}N_h(t,L) + \frac{1}{N}\sum_{j\in G_1} + \frac{1}{N}\sum_{j\in G_2} + O(1) \\ &= \frac{1}{N}N_h(t,L) + O_{\varepsilon}(N^{-\delta}t^{1+\varepsilon}+1). \end{aligned} \tag{4.9}$$

• Estimation of the continuous spectrum:
In using Kuznetsov's formula, the contribution of the continuous spectrum after

summing over n is

$$\frac{1}{4\pi N}\sum_{n\geqslant 1} e^{-\frac{n}{N}}\int_{-\infty}^{\infty} h_{t,L}(x)|\eta(n,x)|^2 dx$$
$$N^{\varepsilon}\int_{-\infty}^{\infty}|h_{t,L}(x)|\log^2(1+2|x|)dx$$
$$\ll N^{\varepsilon}L^{-1}\log(1+2|t|)$$

by the bounds $\zeta(1+2ix)\gg\log(1+2|x|)^{-1}$ and $\tau(n)\ll n^{\varepsilon}$, where $\tau(n)$ is the divisor function.
Now we are going to estimate the contribution of each term on the spectrum side of the Kuznetsov formula still for the test function $h_{t,L}$:
• Contribution of the diagonal term:

$$\frac{1}{\pi}\int_{-\infty}^{\infty} x\tanh\pi x h_{t,L}(x)dx$$
$$\frac{1}{\pi}\int_{-\infty}^{\infty}(h(L(t-x))+h(L(t+x)))dx\sim\frac{2t}{L}\hat{h}(0)$$

with a negligable error term.
Recall that to remove the arithmetic weights $\nu_j(n)$ in the Kuznetsov formula, we sum over n smoothly. It follows from the above, the contribution from the diagonal term after summing over n is:

$$\sim\frac{1}{N}\sum_{n\geqslant 1} e^{-\frac{n}{N}}\frac{2t}{\pi L}\hat{h}(0)\sim\frac{2t}{\pi L}\hat{h}(0).$$

• Contribution from sums of Kloosterman sums:
First, we need the behavior of $J_{2iy}(2x)$ for $y\geqslant x$. Let $z=\sqrt{x^2+y^2}$, then for $y\geqslant x$, we have the following asymptotic expansion ([..]):

$$J_{2iy}(2x)=(2\pi^{\frac{1}{2}})^{-1}z^{-\frac{1}{2}}e^{-\frac{\pi}{4}i}exp(\pi y)e(z\pi^{-1}-y\pi^{-1}\log((z-y)x^{-1}))$$
$$\{1+y^{-1}p_1(yz^{-1})+...+y^{-A}p_A(yz^{-1})+O_A(y^{-A-1})\}$$

where $p_i(x)$ are polynomials of degree i.
Appealing the above expansion of the Bessel function and setting $h_*(y)=h(L(t-y))$ we have

Lemma 4.1. *For $x\leqslant 2t$,*

$$h_*^+(x)\sim i\pi^{-\frac{3}{2}}e^{\frac{i\pi}{4}}t^{\frac{1}{2}}L^{-1}\Big(\frac{ex}{4t}\Big)^{2it}\hat{h}\Big(\frac{\log\frac{4t}{ex}}{\pi L}\Big)$$
$$+L^{-1}t^{\frac{1}{2}}\Big(\frac{ex}{4t}\Big)^{2it}\sum_{k\geqslant 1}(xt^{-1})^k\sum_{m\geqslant 0}\alpha_m(L^{-1},t^{-1})\hat{h}^{(m)}\Big(\frac{\log\frac{4t}{ex}}{\pi L}\Big)$$

where $\alpha_m(L^{-1},t^{-1})$ are polynomials in L^{-1} and t^{-1} and the asymptotic expansion when terminated at say $k\leqslant B$, leaves a remainder of $O(x^Bt^{-B})$.

From now on, in the case of x very small comparing to t, we will only examine the leading term in the above series, the higher order terms can be handled similarly. we always terminate at some fixed order B which is large enough so that the remainder is negligible for our purpose.
Suppose our original test function h satifies the support of $\hat{h} \subset [-b, b] \subset (-1, 1)$. Let δ_1 be small with $0 < \delta_1 < 1 - b$ and let $N = t^{\delta_1}$ and $1 \leqslant L \leqslant \frac{2\log t}{\pi}$, then

$$\frac{1}{N}\sum_{n\geqslant 1} e^{-\frac{n}{N}} \sum_{c\geqslant 1} \frac{S(n,n;c)}{c} h_{t,L}^{+}\Big(\frac{4\pi n}{n}\Big)$$

$$\ll \frac{1}{N}\sum_{n\leqslant N^{1+\varepsilon}} \sum_{c\geqslant 1} \frac{t^{\frac{1}{2}}}{L}\frac{(n,c)^{\frac{1}{2}}}{c^{\frac{1}{2}}}\tau(c)\Big|\hat{h}_{t,L}\Big(\frac{\log\frac{ct}{emn}}{\pi L}\Big)\Big|$$

$$\ll_{\varepsilon} L^{-1}N^{\varepsilon}t^{b+\frac{\delta_1}{2}} = o\Big(\frac{t}{L}\Big). \tag{4.10}$$

where we have invoked Weil's bound for the Kloosterman sum

$$S(n,n;c) \ll c^{\frac{1}{2}}(n,c)^{\frac{1}{2}}\tau(c).$$

Combining all of the above, we end up with

$$M_h(t,L) = \frac{2t}{\pi L} + o\Big(\frac{t}{L}\Big).$$

This toghether with (4.9) (of N) yields that for $1 \leqslant L \leqslant (2\log t)\pi^{-1}$,

$$N_h(t,L) = \frac{t}{6L} + o\Big(\frac{t}{L}\Big).$$

This completes the proof of Theorem 4.1.□
To deduce Corollary 4.1 from this

$$h(0)m(t) \leqslant \sum_{j\geqslant 1} h((t_j - t)L) \sim \frac{t}{6L}\hat{h}(0).$$

taking $L = \frac{2\log t}{\pi}$ (that is as large as allowed):

$$\limsup_{t\to\infty} \frac{m(t)\log t}{t} \leqslant \frac{\pi}{12}\frac{\hat{h}(0)}{h(0)}.$$

As shown in[ILS, pp. 115],

$$\min_{\substack{h\geqslant 0\\ \mathrm{supp}\hat{h}\subset[-1,1]}} \frac{\hat{h}(0)}{h(0)} = 1.$$

So

$$\limsup_{t\to\infty} \frac{m(t)\log t}{t} \leqslant \frac{\pi}{12}.$$

Corollary 4.1 is proved.□

4.2 The Poisson distribution of the Laplacian eigenvalues

This section is closely related to the last section which is also taken from [Sa3] and [LS]. The techniques is too complicated to include here, so we simply state the results, the interested reader is suggested to read the paper [LS]. We continue to use the notations in the last section.
There have been many conjectural and related numerical developments concerning the spectrum of the modualr surface (see [Sa1], [St]). Except the one that the spectrum is simple, there is another important one that the local scaled spacing distributions are "Poissonian"rather than the Gaussian orthogonal distribution which is expected for the generic hyperbolic surface.
In thi section, we study the smoothed number variance. As in the last section we take test functions h satisfying the same conditions , define the smoothed counting function $N_h(t,L)$ for $1 \leqslant L \leqslant t^{1-\varepsilon}$ as we showed in the last section, for $1 \leqslant L \leqslant \frac{2\log t}{\pi}$,

$$N_h(t,L) \sim \frac{t}{6L}.$$

We set

$$V_h(T,L) := \frac{1}{T}\int_0^\infty \psi\Big(\frac{t}{T}\Big)\Big(N_h(t,L) - \frac{t}{6L}\Big)^2 dt$$

to be the smoothed number variance with $\psi \geqslant 0$ be a fixed smooth function with support in $(1,2)$ such that $\int_0^\infty \psi(x)dx = 1$, our main result is the determination of $V_h(T,L)$ in a small window $L \in \Big[\frac{(1+\delta)}{\pi}\log T, \frac{(1+\frac{1}{121})}{\pi}\log T\Big]$. The result indicates a Poissonian number variance which emerges from a detailed analysis of the off-diagonal terms whose contribution turns out to be significant. The following is our main theorem:

Theorem 4.2. *For ψ stated above, $\frac{(1+\delta)}{\pi}\log T \leqslant L \leqslant \frac{(1+\frac{1}{121})}{\pi}\log T$, we have*

$$V_h(T,L) = \frac{\int_0^\infty \zeta^5\psi(\zeta)d\zeta}{\pi^6}\frac{T}{L^2}\sum_{v\geqslant 1}\sum_{(u,v)=1}\prod_{p|v}(1-p^{-2})^{-2}\frac{S(u,u;v)^2}{u^2v^2}\Big|\hat{h}\Big(\frac{\log\frac{Tv}{u}}{\pi L}\Big)\Big|^2 + O_\varepsilon\Big(\frac{T}{L^{2-\varepsilon}}\Big), \tag{4.11}$$

where $S(u,u;v)$ is the classical Kloosterman sum.

Note that the series on the right side above consists of positive terms. Thus its asymptotic behavior depends on the average sizes of Kloosterman sums. This is a quite subtle issue and it has been addressed recently by Fouvry and Michel [FM]. They show that

$$\sum_{v\leqslant x}\frac{S(1,1;v)^2}{v^2} \gg exp\Big(\log\log x)^{\frac{5}{17}}\Big)$$

and

$$\sum_{v \leqslant x} \frac{S(u,u;v)^2}{v^2} \ll u^{\varepsilon} (\log\log x)^3 \log x.$$

From this, we deduce that if the support of $\hat{h}$ is close enough to ± 1 then the first term on the right in Theorem .... satisfies

$$\frac{T}{L^2} exp\Big((\log\log T)^{\frac{5}{17}} \Big) \ll R \ll \frac{T}{L} (\log L)^3.$$

In particular, it is the main term! As a consequence, we have for such h

Corollary 4.2. *Fix $\delta > 0$, then for*

$$\frac{1+\delta}{\pi} \log T \leqslant L \leqslant \frac{1+\frac{1}{121}}{\pi} \log T$$

we have

$$\frac{T}{L^2} exp\Big((\log\log T)^{\frac{5}{17}} \Big) \ll V_h(T,L) \ll \frac{T}{L} (\log L)^3.$$

It seems reasonable to conjecture that as $x \to \infty$,

$$\text{Conjecture A:} \quad \sum_{v \leqslant x} \frac{S(u,u;v)^2}{v^2} \sim A \log x$$

for a nonzero constant A. This combines with Theorem 4.2 would lead to $V_h(T,L) \sim c\frac{T}{L}$, for a nonzero constant c (and L restricted as Theorem 4.2). That is to say that at least for L in this window the number variance is Poissionian. The extension of the range of L to the window specified in Theorem... is the analogue of extending the range in Montgomery's pair correlation conjecture for the zeroes of the Riemann zeta function [Mo2] to the region $\alpha > 1$ (see [Pe2] for the analogue of Montgomery's analysis in the context of the eigenvalues of a hyperbolic surface). In the case of the zeroes of zeta such an extension would follow from a quantitative version of the Hardy-Littlkewood prime 2-tuple conjecture. In our case of the eigenvalues of the modular surface we handle similar off-diagonal shifted sums. We briefly say some words about the proof of the theorem. As the last section, instead of using Selberg's trace formula, we use the Kuznetsov formula. As we noted earlier, Theorem... involves extending L to be large enough to see the Poissonian number variance. Not surprisingly this analysis requires understanding the contributions from the off-diagonal terms. These do in fact contribute to the main term and certain cancellations among these are crucia. we handle these off-diagonal terms emerging from the Kuznetsov formula using the circle method and in particular the smooth $'\delta'$ method developed by Duke, Friedlander and Iwaniec [DFI].

There is another related important problem: let $\lambda_j = \frac{1}{4} + t_j^2, 0 < \lambda_1 \leqslant \lambda_2 ...$ be the eigenvalues of the Laplacian on the cuspidal subspace of $L^2(\Gamma \setminus \mathbb{H})$. Using his trace formula Selberg [He] showed that thses obey a Weyl law:

$$N(T) = \sum_{0 \leqslant t_j \leqslant T} 1 \sim \frac{T^2}{12}$$

as $T \to \infty$. In fact he showed more. Define the remainder term $S(T)$ by

$$N(T) := \frac{T^2}{12} + k_1 T \log T + k_2 T + k_3 - \omega(T) + S(T) \tag{4.12}$$

where k_1, k_2 and k_3 are the constants given in [He, pp. 466],

$$\omega(T) = \frac{1}{4\pi} \int_{-T}^{T} \frac{\phi'}{\phi} \left(\frac{1}{2} + it \right) dt \tag{4.13}$$

and

$$\phi(s) = \frac{\zeta(2s-1)}{\zeta(2s)}, \zeta(s) = \pi^{-\frac{s}{2}} \Gamma \left(\frac{s}{2} \right) \zeta(s) \tag{4.14}$$

A simple application of the Selberg trace formula (to the modualr surface X) shows that

$$S(T) = O \left(\frac{T}{\log T} \right). \tag{4.15}$$

On the other hand Selberg established the lower bound (see Hejhal [He])

$$\frac{1}{T} \int_{T}^{2T} |S(T)|^2 dt \gg \frac{T}{(\log T)^2} \tag{4.16}$$

in particular it follows that

$$S(T) = \Omega \left(\frac{T^{\frac{1}{2}}}{\log T} \right). \tag{4.17}$$

(4.16) and (4.17) are all that were known concerning $S(T)$.

One of the consequences of our analysis of the number variance below is the following modest improvement of

Theorem 4.3.

$$\frac{1}{T} \int_{T}^{2T} |S(T)|^2 dt \gg \frac{T}{(\log T)^2} exp \left((\log \log T)^{\frac{5}{17}} \right) \tag{4.18}$$

and correspondingly

$$S(T) = \Omega \left(\frac{T^{\frac{1}{2}}}{\log T} exp \left(\frac{1}{2} (\log \log T)^{\frac{5}{17}} \right) \right). \tag{4.19}$$

If we assume the conjecture A we will have

$$\frac{1}{T} \int_{T}^{2T} S(t)^2 dt \gg \frac{T}{\log T}$$

which could well be the true order of the magnitude for the mean square of $S(t)$. This is a difficult problem, any improvements of (4.16) and (4.17) are impressive.

4.3 Orthognality of Fourier coefficients of Maass cusp forms

This section is based on Deshouillers and Iwaniec [DI]. In this section we use Kuznetsov's formula to derive some large sieve inequalities of Fourier coefficients of Maass cusp forms which is a preparation for the nxet section. First we need some bounds on bilinear forms of Kloosterman sums:

Lemma 4.2. *For b_m any complex numbers, $\theta > 0$, c a positive integer, let*

$$B(\theta, c, N) = \sum_{N<m,n\leqslant 2N} b_m \bar{b}_n S(m,n;c) e\Big(\frac{2\sqrt{mn}}{c}\theta\Big),$$

then we have

$$1) B(\theta, c, N) \ll c^{\frac{1}{2}} \tau(c)^2 N \|b\|^2$$
$$2) B(\theta, c, N) \ll_\varepsilon (c + N + \sqrt{Nc\theta}) \|b\|^2 \quad \textit{if} \quad N^{1-\varepsilon} \leqslant c$$
$$3) B(\theta, c, N) \ll_\varepsilon \theta^{-1} c^{\frac{1}{2}} N^{\frac{1}{2}+\varepsilon} \|b\|^2 \quad \textit{if} \quad 2 \leqslant c \leqslant N^{1-\varepsilon}, \theta < 2.$$

Proof. 1) Obviously,

$$B(\theta, c, N) \ll \sum_{N<m\leqslant 2N} |b_m|^2 \sum_{N<n\leqslant 2N} |S(m,n;c)|.$$

By Weil's bound for the Kloosterman sum $S(m,n;c) \ll c^{\frac{1}{2}} (m,n;c)^{\frac{1}{2}} \tau(c)$, the above is

$$\ll c^{\frac{1}{2}} \tau(c) \|b\|^2 \sum_{d|c} d^{\frac{1}{2}} \Big(\Big[\frac{N}{c}\Big] + 1\Big)$$

which is bounded by $Nc^{\frac{1}{2}} \tau(c)^2 \|b\|^2$.

2) It is harmless to attach $B(\theta, c, N)$ a smooth function $\eta\Big(\frac{\sqrt{mn}}{N}\Big)$ such that $\eta(x) = 1$ for $1 < x \leqslant 2$ and 0 for $-3 \leqslant x \leqslant 3$, while keeping the condition $N < m, n \leqslant 2N$. Now we use the Mellin inversion to separate variables m and n. Assume

$$\eta(x) e\Big(\frac{2Nx\theta}{c}\Big) = \frac{1}{2\pi i} \int_{\Re s = 1} M(s) x^{-s} ds$$

to be the Mellin inversion of $M(s)$, then

$$M(s) = \int_0^\infty \eta(x) e\Big(\frac{2Nx\theta}{c}\Big) x^{s-1} dx = \int_0^\infty \eta(x) e(f(x)) dx$$

where

$$f(x) = \frac{4\pi N x\theta}{c} + t \log x.$$

If $|t| > \frac{12\pi N\theta}{c}$, then there is no stationary phase in the integral $M(s)$, by partial integration twice, $M(s) \ll (1+|t|)^{-2}$. If $|t| \leqslant \frac{12\pi N\theta}{c} + 1 =: T$, by stationary phase integral,

$$M(s) \ll \frac{\sqrt{t}c}{N\theta} \ll \sqrt{\frac{c}{N\theta}}.$$

Exanding out the Kloosterman sum and using the hybrid large sieve inequlity for Farey points,

$$\int_{-T}^{T} \sideset{}{^*}\sum_{d(\mathrm{mod}\ c)} \Big| \sum_{N<n\leqslant 2N} b_n e\Big(\frac{dn}{c}\Big) n^{it} \Big|^2 dt \ll (cT+N)\|b\|^2,$$

we have

$$\begin{aligned} &B(\theta, c, N) \\ &\qquad \ll \int_{\Re s=1} |M(s)| |\sum_m \sum_n b_m \bar{b}_n \Big(\frac{\sqrt{mn}}{N}\Big)^{-s} S(m,n;c) ds| \\ &\qquad \ll (\sqrt{Nc\theta} + N + c)\|b\|^2. \end{aligned} \tag{4.20}$$

3) Since N is large than c, we must employ the summation of N nontrivially. By Cauchy's inequality,

$$\begin{aligned} B(c,\theta,c,N)^2 &\ll \|b\|^2 \sum_m |\sum_n b_n S(m,n;c) e\Big(\frac{2\sqrt{mn}}{c}\theta\Big)|^2 g\Big(\frac{m}{N}\Big) \\ &= \|b\|^2 \sum_{n_1}\sum_{n_2} b_{n_1}\bar{b}_{n_2} \sum_m S(m,n_1;c)S(m,n_2;c) e\Big(\frac{2\sqrt{m}(\sqrt{n_1}-\sqrt{n_2}}{c}\theta\Big) g\Big(\frac{m}{N}\Big) \\ &\ll \|b\|^4 \max_{n_1} \sum_{n_2} |F(n_1,n_2)| \end{aligned}$$

where $g(x) = 1$ if $1 \leqslant x \leqslant 2$ and 0 if $-3 \leqslant x \leqslant 3$. Set

$$F(n_1,n_2) = \sum_{m\in\mathbb{Z}} S(m,n_1;c)S(m,n_2;c) e\Big(\frac{2\sqrt{m}(\sqrt{n_1}-\sqrt{n_2})}{c}\theta\Big) g\Big(\frac{m}{N}\Big).$$

Expanding out the Kloosterman sum,

$$\begin{aligned} &F(n_1,n_2) \\ &\qquad = \sideset{}{^*}\sum_{-\frac{c}{2}<d_1\leqslant\frac{c}{2}} \sideset{}{^*}\sum_{-\frac{c}{2}<d_1\leqslant\frac{c}{2}} e\Big(\frac{n_1\bar{d}_1 - n_2\bar{d}_2}{c}\Big) \sum_m e(f(m)) g\Big(\frac{m}{N}\Big) \end{aligned}$$

with

$$f(x) = \frac{x(d_1-d_2) + 2\sqrt{x}(\sqrt{n_1}-\sqrt{n_2})\theta}{c}.$$

Since $|f'(x)| < 1$, by the Poisson summation formula

$$\sum_m e(f(m)) g\Big(\frac{m}{N}\Big) = \int_{\mathbb{R}} e(f(x)) g\Big(\frac{x}{N}\Big) dx + O(N^{-1}).$$

If $d_1 \neq d_2$, by partial integration $\left[\frac{A}{\varepsilon}\right]+1$ times, we have

$$\int_{\mathbb{R}} e(f(x))g\Big(\frac{x}{N}\Big)dx \ll N^{-A+1}$$

for any large number A due to the fact $c \leqslant N^{1-\epsilon}$.
If $d_1 = d_2, n_1 \neq n_2$, by partial integration once, we have

$$\int_{\mathbb{R}} e(f(x))g\Big(\frac{x}{N}\Big)dx \ll \frac{c\sqrt{N}}{\theta|\sqrt{n_1}-\sqrt{n_2}|} \ll \frac{c}{\theta|n_1-n_2|}$$

hence

$$\begin{aligned} &B(\theta,c,N)\\ &\quad\ll \|b\|^4(cN+c^2)+\|b\|^4c^2\theta^{-1}\max_{n_1}\sum_{\substack{n_1\neq n_2\\ N<n_2\leqslant 2N}}\frac{1}{|n_1-n_2|}\\ &\quad\ll \|b\|^4cN^{1+\varepsilon}\theta^{-1}. \end{aligned}$$

This finishes the proof. □
Let $\nu_j(n)$ be the normalized Fourier coefficients of Maass cusp forms of type $\frac{1}{2}+i\kappa_j^2$, we have the following large sieve inequality for Maass cusp forms:

Theorem 4.4. *For any complex numbers a_n, we have*

$$\sum_{0\leqslant\kappa_j\leqslant T}|\sum_{n\leqslant N}a_n\nu_j(n)|^2 \ll_\varepsilon (K^2+N)N^\varepsilon\|a\|^2.$$

Remark. By Weyl 's law,

$$\sum_{0\leqslant\kappa\leqslant T}1\sim\frac{T^2}{12},$$

the bound in the above theorem is sharp in the sense of classical large sieve inequalities, recall section 1.5.
Proof of the theorem. We make use of corollary 3.3. We multiply both sides of (3.50) by $t\sinh\pi t e^{-(\frac{t}{k})^2}a_m\bar{a}_n$ and sum over $m,n\in(N,2N]$ and integrate over t on $\mathbb{R}$. By positivity, we can remove the continuous spectrum part, we get

$$\begin{aligned} &\sum_{0<\kappa_j\leqslant K}\kappa_j|\sum_n a_n\nu_j(n)|^2\\ &\quad\ll\sum_{j\geqslant 1}\int_{\mathbb{R}}t\sinh\pi t\; e^{-t^2K^{-2}}H(\kappa_j,t)dt|\sum_n a_n\nu_j(n)|^2\\ &\quad\ll\frac{1}{\pi^2}\int_{\mathbb{R}}t^2e^{-t^2k^{-2}}dt\|a\|^2+\sum_{c\geqslant 1}c^{-1}B(c) \end{aligned}$$

where

$$B(c)=4\pi c^{-1}\sum_m\sum_n a_m\bar{a}_n\sqrt{mn}S(m,n;c)\phi\Big(\frac{4\pi\sqrt{mn}}{c}\Big)$$

with

$$\phi(x) = \int_{\mathbb{R}} t^2 e^{-\frac{t^2}{K^2}} \int_{-i}^{i} K_{2it}(xv)\frac{dv}{v}.$$

For $\Re y > 0$, $K_{2it}(y)$ has the integral representation ([EMOT], pp.82)

$$\begin{aligned} K_{2it(y)} &= \int_0^\infty e^{-y\cosh\zeta}\cos(2t\zeta)d\zeta \\ &= \frac{y}{2t}\int_0^\infty e^{-y\cosh\zeta}\sinh\zeta\sin(2t\zeta)d\zeta \end{aligned} \tag{4.21}$$

by partial integration once. Moreover, we have ([GR], pp. 214)

$$\int_{-\infty}^{\infty} te^{-\frac{t^2}{K^2}}\sin 2t\zeta dt = \sqrt{\pi}K^3\zeta e^{-\zeta^2K^2}.$$

Hence

$$\phi(x) = \sqrt{\pi}iK^3\int_0^\infty e^{-\zeta^2K^2}\zeta\tanh\zeta\sin(x\cosh\zeta)d\zeta \ll 1, \tag{4.22}$$

by partial integration,

$$\begin{aligned} \phi(x) &= \sqrt{\pi}ix^{-1}K^3\int_0^\infty e^{-\zeta^2K^2}(1-\zeta\tanh\zeta-2\zeta^2K^2)\cos(x\tanh\zeta)\frac{d\zeta}{\cosh\zeta} \\ &\ll x^{-1}K^2 \end{aligned} \tag{4.23}$$

By 1) and (4.22), $\sum\limits_{c>N^2} c^{-1}B(c) \ll N^{1+\varepsilon}\|a\|^2$;
By 2) and (4.22), $\sum\limits_{N<c\leqslant N^2} c^{-1}B(c) \ll N^{1+\varepsilon}\|a\|^2$;
By 3) and (4.22), $\sum\limits_{K^{-2}<c\leqslant N} c^{-1}B(c) \ll KN\|a\|^2$;
By 3) and (4.23), $\sum\limits_{c\leqslant NK^{-2}} c^{-1}B(c) \ll KN^{1+\varepsilon}\|a\|^2$.
By partial summation to remove r_j, the conclusion follows.□

4.4 Mean value theorem of the Riemann zeta function

The famous Riemann zeta function $\zeta(s)$ is defined as $\zeta(s) = \sum\limits_{n\geqslant 1} n^{-s}$ for $\Re s > 1$. It has meromorphic continuation to the whole complex plane and has a pole at $s=1$ with residue 1. It satisfies the following functional equation:

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

On the criticle line $\Re s = \frac{1}{2}$, the Lindelof hypothesis predicts that $\zeta(1/2+it) \ll_\varepsilon |t|^\varepsilon$. By playing with functional equation, one can obtain the convexity bound

$\zeta(1/2+it) \ll t^{\frac{1}{4}}$. In this section we are going to see how Iwaniec achieved the subconvexity bound $\zeta(1/2+it) \ll t^{\frac{1}{6}+\varepsilon}$ using Kuznetsov's formulas as a result of the 4th moment of the zeta function over a short interval. Precisely, we are going to sketch the proof of the following theorem of Iwaniec:

Theorem 4.5.

$$\int_T^{T+T^{\frac{2}{3}}} |\zeta(1/2+it)|^4 dt \ll_\varepsilon T^{\frac{2}{3}+\varepsilon}.$$

A direct consequence of the above theorem is the subconvexity bound $\zeta(1/2+it) \ll t^{\frac{1}{6}+\varepsilon}$ for $t \in \mathbb{R}$ due to the following lemma:

Lemma 4.3. *([Iv, pp.9]) For any positive integers $k, 0<\delta<\frac{1}{2}$, we have*

$$|\zeta(1/2+it)|^k \ll_\delta \log T \int_{-\delta}^{\delta} |\zeta(1/2+it+iu)|^k du + (|t|+1)^{-A}$$

wher A is any large number.

Proof. Let r be a parameter which will be chosen later and $X = exp(u_1+u_2+ \ldots + u_r)$. By the residue theorem, we have

$$2\pi i B^r \zeta(s)^k = \int_0^B \ldots \int_0^B \int_{|w|=\delta} \zeta(s+w)^k X^w w^{-1} dw du_1 \ldots du_r. \tag{4.24}$$

Let $\Delta = \int_{-\delta}^{\delta} |\zeta(1/2+it+iu)|^k du$, on the half circle $\Re w < 0$,

$$\begin{aligned} &\int_0^B \ldots \int_0^B \int_{\substack{|w|=\delta \\ \Re w<0}} \zeta(s+w)^k X^w w^{-1} dw du_1 \ldots du_r \\ &= \int_{\substack{|w|=\delta \\ \Re w<0}} \zeta(s+w)^k \Big(\frac{e^{Bw}-1}{w}\Big)^r w^{-1} dw \ll t^{\frac{k}{2}} \Big(\frac{2}{\delta}\Big)^r \end{aligned} \tag{4.25}$$

by the convexity bound $\zeta(s+w) \ll (|t|+1)^{\frac{1}{2}}$.
On the other half circle, $\Re w > 0$, due to

$$\frac{x^w}{w} = \frac{x^w - x^{-w}}{w} + \frac{x^{-w}}{w},$$

correspondingly the integral in (4.24) splits into two parts. for the same reason as the above, the part of $x^{-w}w^{-1}$ contribute to the integral (4.24) $O(t^{\frac{k}{2}} 2^r \delta^{-r})$, Since $(x^w - x^{-w})w^{-1}$ is regular for $\Re w \geqslant 0$, we can move the line of the integration to the segment $[-i\delta, i\delta]$, then

$$\begin{aligned} &\int_0^B \ldots \int_0^B \Big| \int_{\substack{|w|=\delta \\ \Re w<0}} \zeta(s+w)^k \frac{X^w - X^{-w}}{w} dw du_1 \ldots du_r \Big| \\ &\ll \Big| \int_0^B \ldots \int_0^B \log X du_1 \ldots du_r \Big| \Delta \leqslant \Delta B^{1+r} r \end{aligned}$$

hence

$$|\zeta(1/2+it)|^k \ll \Delta Br + B^{-r}\delta^{-r}t^{\frac{k}{2}}.$$

Taking $B = e\delta^{-1}$ and $r = A\log(|t|+1)$ with A a large number, we finish the proof. □

Proposition 4.3. *(Approximate functional equation)*
Let $G(u)$ be any even function which is holomorphic and bounded in the strip $-4 < \Re < 4$ normalized such that $G(0) = 1$. Then for $\Re s = \frac{1}{2}$ we have

$$\zeta(s) = \sum_{n\geqslant 1} n^{-s}V_s(n) + \gamma(s)\sum_{n\geqslant 1} n^{-(1-s)}V_{1-s}(n) + R$$

where

$$V_s(y) = \frac{1}{2\pi i}\int_{(3)} y^{-u}G(u)\frac{\pi^{-\frac{s+u}{2}}\Gamma(\frac{s+u}{2})}{\pi^{\frac{s}{2}}\Gamma(\frac{s}{2})}\frac{du}{u}$$

and

$$\gamma(s) = \frac{\pi^{-\frac{1-s}{2}}\Gamma(\frac{1-s}{2})}{\pi^{\frac{s}{2}}\Gamma(\frac{s}{2})}.$$

Proof. Let $\Lambda(s) = \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\zeta(s)$, consider

$$I(s) = \frac{1}{2\pi i}\int_{(3)} \Lambda(s+u)G(u)\frac{du}{u}.$$

Moving the integration to the line $\Re u = -3$, applying the functional equation $\Lambda(s) = \Lambda(1-s)$ picking uo the pole at $u = 0, 1-s$ and $-s$, we get

$$\Lambda(s) = I(s) + I(1-s) + R\pi^{-\frac{s}{2}}\Gamma\Big(\frac{s}{2}\Big)$$

where

$$R = (\mathrm{Res}_{u=1-s} + \mathrm{Res}_{u=-s})\frac{\Lambda(s+u)}{\pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})}\frac{G(u)}{u}.$$

Introducing the Dirichlet series for $\zeta(s)$ and integrating termwise, we have

$$I(s) = \pi^{-\frac{s}{2}}\Gamma(\frac{s}{2})\sum_{n\geqslant 1} n^{-s}V_s(n).$$

Do the same step for $I(1-s)$ and insert them into Λ, the proposition follows at once. □
Now we fix a special test function $G(u)$

$$G(u) = \Big(\cos\frac{\pi u}{4A}\Big)^{-4A},$$

then by Stirling's formula, one can prove that derivatives of $V_s(y)$ on $\Re s = \frac{1}{2}$ satisfy

$$y^a V_s^{(a)}(y) \ll \Big(1 + \frac{y}{\sqrt{|s|+1}}\Big)^{-A}$$

for $a \geqslant 0$. For this choice of $G(u,\ R = O(exp(-\frac{\pi}{4}|s|))$. In this case, the above proposition says that on the criticle line $\zeta(s)$ can be well approximated by two Dirichlet polynormials each of which has length $\sqrt{|s|+1}$.
Applying a smooth partition of unity, we derive by proposition 4.3 that

$$\zeta(s) \ll \sum_N G_N(t)\Big(1 + \frac{N}{\sqrt{|s|+1}}\Big)^{-A}$$

where

$$G_N(t) = \sum_n n^{-\frac{1}{2}-it} g(n)$$

with $g(x)$ a smooth function supported on $[N, 2N]$ for $N = 2^{\frac{\nu}{2}}, \nu \geqslant -1$, such that $g^{(a)}(x) \ll_{a,A} N^{-a}$ for every $a \geqslant 0$. By Cauchy's inequality,

$$\zeta(s)^2 \ll \sum_N |G_N(t)|^2 \Big(1 + \frac{N}{\sqrt{|s|+1}}\Big)^{-2A} \log 3A.$$

We have

$$\begin{aligned} |G_N(t)|^2 &= \sum_{n_1}\sum_{n_2} (n_1 n_2)^{-frac12} \Big(\frac{n_1}{n_2}\Big)^{it} g(n_1) g(n_2) \\ &= \sum_d \frac{1}{d} g(d) G(t,d) \end{aligned}$$

where

$$G(t,d) = \sum_{n_1} \sum_{(n_2,n_1)=1} (n_1 n_2)^{-\frac{1}{2}} \Big(\frac{n_1}{n_2}\Big)^{it} \bar{g}(n_1 d) g(n_2 d).$$

In order to prove theorem 4.5, it is sufficient to prove

$$\int_T^{T+T_0} |\zeta(\frac{1}{2}+it)|^2 |G_N(t)|^2 dt \ll_\varepsilon T_0 T^\varepsilon \tag{4.26}$$

for $N \leqslant T^{\frac{1}{2}+\varepsilon}$.
Furthermore we smooth out the interval by attaching a smooth truncation function $f(x)$ which is supported on $(-2T_0T^{-1}, 2T_0T^{-1})$ and is 1 on $[0, T_0T^{-1}]$, we are led to prove

$$\int_0^\infty e^{-\frac{2\pi t}{T}} |\zeta(\frac{1}{2}+it)|^2 |G_N(t)|^2 f\Big(\frac{t-T}{T}\Big) dt \ll_\varepsilon T_0 T^\varepsilon \tag{4.27}$$

Let $f(\hat{x}) = \int_{\mathbb{R}} f(y) e(yx) dy$, then $\hat{f}(x) \ll \frac{T_0}{T}\Big(\frac{T}{|x|+1)T_0}\Big)^p$, for any $p \geqslant 0$, it follows by Fourier inversion

$$\begin{aligned} &\int_0^\infty e^{-\frac{2\pi t}{T}} |\zeta(\frac{1}{2}+it)|^2 |G_N(t)|^2 f\Big(\frac{t-T}{T}\Big) dt \\ &= \int_{\mathbb{R}} \hat{f}(x) W(x) dx, \end{aligned} \tag{4.28}$$

where

$$W(x) = \int_0^{\infty} |\zeta(\frac{1}{2} + it)|^2 |G_N(t)|^2 e\Big(\frac{x+i}{T} t\Big) dt.$$

By [Ti, pp. 141], we have the following formula for the Laplace transform of $|\zeta\Big(\frac{1}{2} + it\Big)|^2$:

Lemma 4.4. *For $\Re s > 0$, we have*

$$\int_0^{\infty} |\zeta\Big(\frac{1}{2} + it\Big)|^2 e^{-st} dt = 2\pi e^{is} \sum_{l=1}^{\infty} \tau(l) e(l exp(is)) + p(s)$$

where $p(s)$ is regular when $|s|$ is sufficiently small.

Using the above lemma,

$$\begin{aligned} W(x) &= \sum_d d^{-1} g(d) \sum_{n_1} \sum_{(n_2,n_1)=1} n_2^{-1} g(n_1 d) g(n_2 d) 2\pi exp(\pi(x+i)T^{-1}) \\ &\quad V_x(n_1, n_2) + O(N), \end{aligned}$$

where

$$\begin{aligned} V_x(n_1, n_2) &= \sum_{l \geqslant 1} \tau(l) e\Big(l n_1 n_2^{-1} exp(2\pi(x+i)T^{-1})\Big) \\ &= \sum_{l \geqslant 1} \tau(l) e(l n_1 n_2^{-1}) F_l(n_1, n_2, x) \end{aligned}$$

with

$$F_l(n_1, n_2, x) = e(-l n_1 n_2^{-1}(1 - exp(2\pi(x+i)T^{-1}))).$$

In $W(x)$ we split the n_1 sum into residue classes modulo n_2, for each class $n_1 \equiv u$ modulo n_2, we apply Poisson summation, we end up with

$$\begin{aligned} W(x) &= \sum_d d^{-1} g(d) \sum_{n_2} n_2^{-1} g(n_2 d) 2\pi exp(\pi(x+i)T^{-1}) \\ &\quad \sum_{l \geqslant 1} \tau(l) \sideset{}{^*}\sum_{u (\mathrm{mod}\ n_2)} e\Big(\frac{lu}{n_2}\Big) \frac{1}{n_2} \sum_{k \in \mathbb{Z}} e\Big(\frac{ku}{n_2}\Big) V_l(k, n_2, d, x) + O(N) \end{aligned}$$

where

$$V_t(k, n_2, d, x) = \int_{\mathbb{R}} e\Big(-\frac{k}{n_2} y\Big) g(yd) F_l(y, n_2, x) dy.$$

Next we execute the summation over l by means of the following Poisson type formula ([Ju], Theorem 1.7):

Lemma 4.5. *Let $h(t)$ be a smooth, Schwartz function on $\mathbb{R}^+$ and $(u, n_2) = 1$, then we have*

$$\sum_{l \geqslant 1} \tau(l) e\Big(\frac{lu}{n_2}\Big) h(l) = \frac{1}{n_2} \int_{\mathbb{R}^+} (\log t + 2\gamma - 2\log n_2) h(t) dt$$
$$-\frac{2\pi}{n_2} \sum_1^\infty \tau(l) e\Big(\frac{\bar{u}l}{n_2}\Big) \int_0^\infty Y_0\Big(\frac{4\pi}{n_2}\sqrt{lt}\Big) h(t) dt$$
$$+\frac{4}{n_2} \sum_1^\infty \tau(l) e\Big(\frac{\bar{u}l}{n_2}\Big) \int_0^\infty K_0\Big(\frac{4\pi}{n_2}\sqrt{lt}\Big) h(t) dt$$

where $Y_0(x)$ and $K_0(x)$ are Bessel functions of order 0.

Taking $h(t) = V_t(k, n_2, d)$ and inserting the above into $W(x)$ and (4.28) we get from the summation in $u(\mathrm{mod} n_2)$ with $(u, n_2) = 1$ complete Kloosterman sums. we obtain

$$W(x) = 2\pi exp(\pi(x+i)T^{-1}) \sum_d d^{-1} g(d) [H_1(d) + H_2(d) + H_3(d)],$$

where

$$H_1(d) = \sum_{n_2} n_2^{-3} g(n_2 d) \sum_k S(k, 0; n_2) I_1(k, n_2, d)$$

with

$$I_1(k, n_2, d) = \int_0^\infty (\log t + 2\gamma - 2\log n_2) V_t(k, n_2, d) dt;$$

$$H_2(d) = -2\pi \sum_{n_2} n_2^{-3} g(n_2 d) \sum_{k \in \mathbb{Z}} \sum_{l \geqslant 1} \tau(l) S(k, -l; n_2) I_2(k, l, n_2, d)$$

with

$$I_2(k, l, n_2, d) = \int_0^\infty Y_0\Big(\frac{4\pi}{n_2}\sqrt{lt}\Big) V_t(k, n_2, d) dt;$$

$$H_3(d) = 4 \sum_{n_2} n_2^{-3} g(n_2 d) \sum_{k \in \mathbb{Z}} \sum_{l \geqslant 1} \tau(l) S(k, l; n_2) I_3(k, l, n_2, d)$$

with

$$I_3(k, l, n_2, d) = \int_0^\infty K_0\Big(\frac{4\pi}{n_2}\sqrt{lt}\Big) V_t(k, n_2, d) dt.$$

For the Ramanujan sum, we apply the bound $S(k, 0; n_2) \ll (k, n_2)$. For the weighted sums of Kloosterman sums we use Kuznetsov's inversion formula. The last step is to apply the large sieve inequalities of Fourier coefficients of Maass cusp forms in the last section. This finishes the sketch of the proof. □

5 Appendix: Selberg's trace formula

In this appendix, we will state Selberg's trace formula for $\Gamma = SL(2,\mathbb{R})$ and sketch its proof.

The grouip $SL(2,\mathbb{R})$ acts on $\mathbb{H}$ by Mobiüs transformation. Namely, $\gamma z = \frac{az+b}{cz+d}$ for $\gamma \in \Gamma$ and $z \in \mathbb{H}$. We first classify these motions (except the identity motion) into three classes invariant under conjugation. These three conjugacy classes are characterized by their fixed points on $\mathbb{R} \cup \{\infty\} : \hat{\mathbb{R}}$.

1) Parabolic class whose elements have only one fixed points on $\hat{\mathbb{R}}$.

2) Hyperbolic class whose elements have two distinct fixed points on $\hat{\mathbb{R}}$.

3) Elliptic class whose elements have one fixed point in $\mathbb{H}$ and the other one is its conjugate parabolic motions move points by translation. Elliptic motions move points by rotation and hyperbolic motions move points by dilation, the dilation factor p is called the norm of hyperbolic motions which is independent on the choice of representatives.

As the derivatation of Kuzuetsov's formula, we take a kernel function $k(z,w)$ which is point-pair invariant and compactly supported on $\mathbb{R}$ as a function of the distance between k and w. we define the automorphic kernel

$$k(z,w) = \sum_{\gamma\in\Gamma} k(z,\gamma w)$$

The Selberg trace formula is obtained on the one hand, by integrating the spectral decomposition of $k(z,w)$ on the diagonal. i.e.

1)

$$\begin{aligned} TrK_s := \int_{\Gamma\backslash\mathbb{H}} k(z,w)d\mu z &= \int_{\Gamma\backslash\mathbb{H}} \sum_{j\geqslant 0} h(t_j)|h_j(z)|^2 d\mu z \\ &= \sum_{j\geqslant 0} h(t_j); \end{aligned} \tag{5.1}$$

On the other hand,

2)

$$TrK_g := \int_{\Gamma\backslash\mathbb{H}} k(z,z)d\mu z = \sum_{\gamma\in\Gamma} \int_{\Gamma\backslash\mathbb{H}} k(z,\gamma z)d\mu z \tag{5.2}$$

comparing 1) and 2), we can get a trace formula, but it is not quite useful. The integral in 2) is still hard to compute. To go further, we split Γ into conjugacy class l, so

$$TrK_g = \sum_{\mathcal{L}} TrK_{\mathcal{L}}$$

with

$$\begin{aligned} TrK_{\mathcal{L}} &= \sum_{\gamma\in\mathcal{L}} \int_{\Gamma\backslash\mathbb{H}} k(z,\gamma z)d\mu z \\ &= \int_{\mathbb{Z}(\gamma)\backslash\mathbb{H}} k(z,\gamma z)d\mu z \end{aligned}$$

where $Z(\gamma)$ is the contrlizer of $\gamma \in \mathcal{L}$. It is easy to check that the above computations don't depend on the choice of representatives of γ in the class $\mathcal{L}$. It turns out the computation of $TrK_{\mathcal{L}}$ depends on the type of the conjugacy classes:

Furthermore, the set of conjugacy classes which have the same fixed points has a generator in the sense that all the other conjugacy classes are powers of it, we henceforth call it a primitive conjugacy class. The fundamental domains $Z(\gamma) \setminus \mathbb{H}$ for these primitive conjugacy classes are strips or sectors, so all the integrals in 2) then can be computed explicitly in terms of the Selberg transform of k. At this moment, we compare 1) and 2) and end up with the Selberg trace formula with k hiding.

Now in our case, Γ being the modular group, so $\Gamma \setminus \mathbb{H}$ is noncompact. Difficulties occur due to the fact Einsenstein series is not square integrable. Selberg's trick is to use a truncated domain $F(Y)$ which approaches the fundamental domain F of Γ as $Y \to \infty$. We end up with a pre-trace formula with Y on both sides. Comparing the main terms of the identity as $Y \to \infty$, we obtain the Selberg trace formula. For more details, see [Se2], [Iw4] or [He].

Now we simply state the result in the case of $\Gamma = SL(2, \mathbb{R})$.

Let h be an even function of a complex variable r which is holomorphic in the strip $|\Im r| < \frac{1}{2} + \delta$ for some $\delta > 0$ and satifies the following estimate in the above strip:

$$h(r) \ll_\epsilon (|r| + 1)^{-2-\epsilon}$$

$\hat{h}(x) = \int_{\mathbb{R}} h(t)e(-xt)dt$ being its Fourier transform and

$$\phi(s) = \frac{\pi^{-(1-s)}\Gamma(1-s)\zeta(2(1-s))}{\pi^{-s}\Gamma(s)\zeta(2s)}$$

.

Theorem 5.1. *(Selberg's trace formula)*

$$\begin{aligned}
&\sum_{j \geqslant 0} h(t_j) + \frac{1}{4\pi}\frac{-\phi'}{\phi}(\frac{1}{2} + it)h(t)dt \\
&\quad = \frac{1}{12}\int_{\mathbb{R}} h(t)\tanh(\pi t)tdt + \frac{1}{2\pi}\sum_P \frac{\log NP_0}{NP^{1/2} - NP^{-1/2}}\hat{h}(\frac{1}{2\pi}\log NP) \\
&\quad + \sum_R \sum_{0<l<m} \left(2m\frac{\sin \pi l}{m}\right)^{-1} \int_{\mathbb{R}} \frac{e^{-\frac{2\pi t l}{m}}}{1 - e^{-2\pi t}} h(t)dt \\
&\quad - \frac{1}{2\pi}\log 2\hat{h}(0) - \frac{1}{2\pi}\int_{\mathbb{R}} h(t)\frac{\Gamma'(1+it)}{\Gamma(1+it)}dt
\end{aligned}$$

where P runs through the hyperbolic conjugacy classes of Γ and R runs through the elliptic conjugacy classes of Γ.

Remark1. On the left or say the spectral side of the trace formula, the first term comes from the discrete spectrum and the second term is from the contribution

of the continuous spectrum, i.e., the Eisenstein series. One should notice that we have the constant function which corresponds to the term $j = 0$ in the trace formula, while in the Kuzuetsov formula the constant function doesn't contribute which is a nice feature of Kuzuetsov's formula.
2. On the right or say the geometric side of the trace formula, the first term comes from the identity matrices. The second term is from the hyperbolic classes. The third term is from the elliptic classes and the last two terms come from the parabolic classes. One of the immediate results from the trace formula is Weyl's law: Let

$$N(t) = \#\{j : T \geqslant t_j \geqslant 0\}$$

be the counting function of the cuspidal spectrum and

$$M(T) = \frac{1}{4\pi}\int_{-T}^{T} -\frac{-\phi'}{\phi}(1+it)dt$$

be the corresponding analogy of the continuous spectrum. Weyl's law says

Corollary 5.1. *As $T \to \infty$, we have*

$$N(T) + M(T) \sim \frac{T^2}{12}$$

For the modular group, or in general congruence subgroups, one can estimate the contribution from the continuous spectrum $M(T)$ since $\phi(s)$ can be explicitly given and show

$$M(T) \ll T\log T$$

so it follows

Corollary 5.2. *As $T \to \infty$, we have*

$$N(T) \sim \frac{T^2}{12}$$

Proof of Corollary. In the trace formula, take a test function $h(t) = e^{-\frac{t^2}{T^2}}$, then the identity motion contributes

$$\frac{1}{12}\int_0^{\infty} e^{-\frac{t^2}{T^2}} \tanh \pi t t dt = \frac{T^2}{12} + O(1)$$

the hyperbolic and the elliptic motions contribute a bouinded quantity. Using the Stirling formula, one can estimate the contribution from the parabolic motions by $O(T\log T)$.
The corollary follows at once by the classical Tauberian theorem. □
Using the Selberg zeta function, one can get a more precise Weyl's law for the modular surface ([He]).

$$N(T) = \frac{T^2}{12} - \frac{2T\log T}{\pi} + \Big(\frac{2+\log\pi - \log 2}{\pi}\Big)T + O(T/\log T).$$

Note they did it for a general surface $\Gamma \setminus \mathbb{H}$.
Remark 1. Selberg showed the abundance of Maass cusp forms for congruence subgroups simply by the above counting, pradoxically, none of one Maass cusp form form on the modular surface was constructed up to now!

2. It is not true for general Fuschian groups, the cuspidal spectrum is larger than the continuous spectrum. Actually, the famous work of Philippe and Sarnak [PS] shows the abundance of Maass cusp forms is a special phenomena for congruence subgroups.

References

[Br] Bruggeman, R. W. Fourier coefficients of cusp forms. Invent. Math. 45 (1978), no. 1, 1–18.

[CP] Cogdell, James W.; Piatetski-Shapiro, Ilya The arithmetic and spectral analysis of Poincar series. Perspectives in Mathematics, 13. Academic Press, Inc., Boston, MA, 1990.

[DFI] Duke, W, Friedlander, J, Iwaniec, H. A quadratic divisor problem, Invent. math. 115 (1994), 209-217.

[DI] Deshouillers, J.-M.; Iwaniec, H. Kloosterman sums and Fourier coefficients of cusp forms. Invent. Math. 70 (1982/83), no. 2, 219–288.

[DIPS] Deshouillers, J.-M.; Iwaniec, H.; Phillips, R. S.; Sarnak, P. Maass cusp forms. Proc. Nat. Acad. Sci. U.S.A. 82 (1985), no. 11, 3533–3534.

[Du] Duke, W. The dimension of the space of cusp forms of weight one. Internat. Math. Res. Notices 1995, no. 2, 99–109 (electronic).

[EMOT] Erdlyi, Arthur; Magnus, Wilhelm; Oberhettinger, Fritz; Tricomi, Francesco G. Higher transcendental functions. Vols. I, II. Based, in part, on notes left by Harry Bateman. McGraw-Hill Book Company, Inc., New York-Toronto-London, 1953.

[FM] Fouvry, Etienne; Michel, Philippe Sommes de modules de sommes d'exponentielles. (French) [Sums of moduli of exponential sums] Pacific J. Math. 209 (2003), no. 2, 261–288.

[FV] M. Forti, C. Viola, On large sieve type estimates for Dirichlet series operator, Proc. Symp. Pure Math. 24 (1973), 31-49.

[Go] Goldfeld, D. Analytic number theory on $GL(n)$, to appear.

[GS] Goldfeld, D; Sarnak, P. Sums of Kloosterman sums, Invent. Math. 71(1983), 243-250.

[He] Hejhal, Dennis A. The Selberg trace formula for PSL(2, R). Vol. 2. Lecture Notes in Mathematics, 1001. Springer-Verlag, Berlin, 1983.

[ILS] Iwaniec, Henryk; Luo, Wenzhi; Sarnak, Peter Low lying zeros of families of L-functions.

[IK] Iwaniec, Henryk; Kowalski, Emmanuel Analytic number theory. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI, 2004.

[IL] Li, Xiaoqing; Iwaniec, Henryk The orthogonalities of Hecke eigenvalues, submitted.

[Iv] Ivić, A. Lectures on mean values of the Riemann zeta function. Tata Institute of Fundamental Research Lectures on Mathematics and Physics, 82. Published for the Tata Institute of Fundamental Research, Bombay; by Springer-Verlag, Berlin, 1991. Inst. Hautes tudes Sci. Publ. Math. No. 91 (2000), 55–131 (2001).

[Iw1] Iwaniec, Henryk Fourier coefficients of cusp forms and the Riemann zeta-function. Seminar on Number Theory, 1979–1980 (French), Exp. No. 18, 36 pp., Univ. Bordeaux I, Talence, 1980.

[Iw2] Iwaniec, Henryk Spectral theory of automorphic functions and recent developments in analytic number theory. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986), 444–456, Amer. Math. Soc., Providence, RI, 1987.

[Iw3] Iwaniec, Henryk Topics in classical automorphic forms. Graduate Studies in Mathematics, 17. American Mathematical Society, Providence, RI, 1997.

[Iw4] Iwaniec, Henryk Spectral methods of automorphic forms. Second edition. Graduate Studies in Mathematics, 53. American Mathematical Society, Providence, RI; Revista Matemtica Iberoamericana, Madrid, 2002.

[Iw5] Iwaniec, Henryk Automorphic number theory. Current developments in mathematics, 2003, 35–52, Int. Press, Somerville, MA, 2003.

[Ju] Jutila, M. Lectures on a method in the theory of exponential sums, Springer Verlag, 1987.

[KS] Kim, Henry H. Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2. With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak. J. Amer. Math. Soc. 16 (2003), no. 1, 139–183 (electronic).

[Ku] Kuznecov, N. V. The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums. (Russian) Mat. Sb. (N.S.) 111(153) (1980), no. 3, 334–383, 479.

[LRS] Luo, Wenzhi; Rudnick, Zev; Sarnak, Peter On the generalized Ramanujan conjecture for $\mathrm{GL}(n)$. 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.

[LRS] Luo, W.; Rudnick, Z.; Sarnak, P. On Selberg's eigenvalue conjecture. Geom. Funct. Anal. 5 (1995), no. 2, 387–401.

[Lu] Luo, Wenzhi Values of symmetric square L-functions at 1. J. Reine Angew. Math. 506 (1999), 215–235.

[LS] Li, Xiaoqing; Sarnak, Peter On the number variance for $SL(2,\mathbb{Z})\setminus\{\mathbb{H}$, to appear in GAFA.

[Mo1] Montgomery, H. L. Topics in multiplicative number theory, Lecture Notes in Math. 227, Springer Verlag, 1971.

[Mo2] Montgomery, H. L. The pair correlation of zeros of the zeta function. Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), pp. 181–193. Amer. Math. Soc., Providence, R.I., 1973.

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

[MV] Michel, Philippe; Venkatesh, Akshay On the dimension of the space of cusp forms associated to 2-dimensional complex Galois representations. Int. Math. Res. Not. 2002, no. 38, 2021–2027.

[Pm] Peter, Manfred The correlation between multiplicities of closed geodesics on the modular surface. Comm. Math. Phys. 225 (2002), no. 1, 171–189.

[PS] Phillips, R. S.; Sarnak, P. On cusp forms for co-finite subgroups of PSL$(2,R)$. Invent. Math. 80 (1985), no. 2, 339–364.

[Sa1] Sarnak, Peter Spectra of hyperbolic surfaces. Bull. Amer. Math. Soc. (N.S.) 40 (2003), no. 4, 441–478 (electronic).

[Sa2] Sarnak, Peter Kloosterman, quadratic forms and modular forms. Nieuw Arch. Wiskd. (5) 1 (2000), no. 4, 385–389.

[Sa3] Sarnak, Peter Letter to Z. Rudnick, 2002, www.math.princeton.edu/ sarnak

[Sa4] Sarnak, Peter Notes on the generalized Ramanujan conjecture. www.math.princeton.edu/ sarnak

[Se1] Selberg, Atle On the estimation of Fourier coefficients of modular forms. 1965 Proc. Sympos. Pure Math., Vol. VIII pp. 1–15 Amer. Math. Soc., Providence, R.I.

[Se2] Selberg, A. Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. J. Indian Math. Soc. (N.S.) 20 (1956), 47–87.

[St] Stein, G. Diplom. Math. Univ., Hamburg, 1992.

[Ti] Titchmarsh, E. C. The theory of the Riemann zeta function, Clarendon Press, Oxford 1951.

[Za] Zagier, Don Eisenstein series and the Selberg trace formula. I. Automorphic forms, representation theory and arithmetic (Bombay, 1979), pp. 303–355, Tata Inst. Fund. Res. Studies in Math., 10, Tata Inst. Fundamental Res., Bombay, 1981.

[Za2] Zagier, Don Unpublised notes on Kuznetsov's formula.