TRACE FORMULAE

DORIAN GOLDFELD

NOTES TAKEN BY PAK-HIN LEE

Abstract. Here are the notes I took for Dorian Goldfeld's course on trace formulae offered at Columbia University in Fall 2013 (MATH G6674: Topics in Number Theory). The course was focussed on trace formulae and covered:

- Petersson Trace Formula
- Kuznetsov Trace Formula
- Theta Functions
- Degenerate Kuznetsov Trace Formula
- Jacquet's Relative Trace Formula
- Selberg Trace Formula
- Arthur–Selberg Trace Formula
- Jacquet–Langlands Correspondence
- Beyond Endoscopy

with applications to classical problems in analytic number theory.

Due to my own lack of understanding of the materials, I have inevitably introduced both mathematical and typographical errors in these notes. Please send corrections and comments to phlee@math.columbia.edu.

Contents

Last updated: September 6, 2014.

1. Lecture 1 (September 10, 2013)

1.1. **Introduction.** The notion of a trace formula arises from matrices. The trace of a matrix is the sum of its diagonal entries, which is also the sum of its eigenvalues. If we are in an infinite-dimensional space, there could be infinitely many eigenvalues so we have to introduce a convergence factor

$$\mathrm{Trace}_f(T) = \sum f(\lambda).$$

This works well if we have countably many eigenvalues. If there are uncountably many, we have to use an integral

$$\mathrm{Trace}_f(T) = \int f(\lambda)d\lambda.$$

The Poisson Summation Formula says

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

which can be thought of as a trace formula.

We will talk about the Petersson Trace Formula. Petersson is responsible for three major things — the trace formula, the Weil–Petersson metric, and the Petersson inner product. I will introduce the trace formula by first talking about automorphic forms in general.

1.2. **Automorphic Forms in General.** Let X be a topological space, and G be a topological group acting on X. For $g \in G$, $x \in X$, we often write gx instead of $g \circ x \in X$. We are interested in discrete actions of G on X, i.e. the intersection $B \cap (gB) \neq \emptyset$ for only finitely many $g \in G$. We will study actions of arithmetic groups, which are discrete. From now on, assume G acts discretely on X.

$F : X \to \mathbb{C}$ is an automorphic function if

$$F(gx) = \psi(g, x)F(x)$$

for all $g \in G$, $x \in X$. (We will focus on complex-valued automorphic functions only.) ψ is called the factor of automorphy. Usually we put further conditions on F, e.g. we may ask that F is holomorphic or smooth, or that it satisfies certain differential equations. What kind of factor of automorphy can we get?

Let $g, g' \in G$. Then there are two ways to do $F(g \cdot g'x)$:

$$F((g \cdot g') \circ x) = F(g \circ (g' \circ x)).$$

This implies certain conditions on ψ:

$$\psi(gg', x)F(x) = \psi(g, g'x)F(g'x) = \psi(g, g'x)\psi(g', x)F(x).$$

If $F(x) \neq 0$, this means that

$$\psi(gg', x) = \psi(g, g'x)\psi(g', x)$$

This is a 1-cocycle relation in the cohomology of groups.

Let us give an example of such an automorphic function. Let $X = \mathfrak{h} := \{x + iy : x \in \mathbb{R}, y > 0\}$ be the upper half plane, and $G = \mathrm{SL}(2, \mathbb{Z})$. If $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $z \in \mathfrak{h}$, then $gx = \frac{az+b}{cx+d}$ is an action.

Definition 1.1. $j(g,z) := cz+d$.

j satisfies the equation

$$j(gg',z) = j(g,g'z)\cdot j(g',z).$$

We can now construct an example of an automorphic function.

$$E_k(z) = \sum_{\gamma\in\Gamma_\infty\backslash \mathrm{SL}(2,\mathbb{Z})} \frac{1}{j(\gamma,z)^k}$$

where $\Gamma_\infty = \left\{\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix} : n\in\mathbb{Z}\right\}$. We have to mod out by Γ_∞ because $j\left(\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix}, z\right) = 1$ and there would be infinitely many terms repeating. How do we know this sum is non-zero? In fact,

$$E_k(z) = \sum_{\substack{c,d\\(c,d)=1}} \frac{1}{(cz+d)^k}.$$

This is zero if k is odd, so we want k even, and $k>2$ (for convergence). To see this is non-zero, let z tend to $i\infty$. Let us prove that it is automorphic. For $\alpha\in\mathrm{SL}(2,\mathbb{Z})$, we have

$$j(\gamma,\alpha z) = \frac{j(\gamma\alpha,z)}{j(\alpha,z)}$$

by the cocycle condition, and so

$$E_k(\alpha z) = \sum_\gamma \frac{1}{j(\gamma,\alpha z)^k} = j(\alpha,z)^k\sum_\gamma\frac{1}{j(\gamma\alpha,z)^k} = j(\alpha,z)^kE_k(z).$$

1.3. **Petersson Trace Formula.** We will now write down the Petersson Trace Formula explicitly. Let

$$\Gamma_0(N) = \left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in\mathrm{SL}(2,\mathbb{Z}) : c\equiv 0 \pmod N\right\}.$$

If $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \in \mathbb{C}^\times$ is a Dirichlet character mod N, define $\widetilde{\chi} : \Gamma_0(N)\to\mathbb{C}^\times$ by

$$\widetilde{\chi}\left(\begin{pmatrix}a & b\\ c & d\end{pmatrix}\right) := \chi(d).$$

Let $S_k(\Gamma_0(N),\chi)$ be the set of holomorphic functions $f:\mathfrak{h}\to\mathbb{C}$ satisfying

$$f(\gamma z) = \widetilde{\chi}(\gamma)j(\gamma,z)^kf(z)$$

for all $\gamma\in\Gamma_0(N)$ and $z\in\mathfrak{h}$, and that f is cuspidal, i.e.

$$\lim_{y\to\infty} f(u+iy) = 0$$

for all $u\in\mathbb{Q}$.

Theorem 1.2 (Petersson–Hecke)**.** $\dim S_k(\Gamma_0(N),\chi)<\infty$.

The Petersson Trace Formula says

$$\frac{\Gamma(k-1)}{(4\pi\sqrt{mn})^{k-1}} \sum_{\substack{f\in S_k\\ \text{orthonormal basis}}} \overline{A_f(m)}A_f(n) = \delta_{m,n} + \frac{1}{(2\pi i)^k}\sum_{c\equiv 0(N)} \frac{S_\chi(m,n,c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

for $m,n\geq 1$. Let us now define the functions involved:

- If $f \in S_k$, then $f\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} z\right) = f(z)$ so there is an expansion

$$f(z) = \sum_{m=1}^{\infty} A_f(m) e^{2\pi i m z}$$

for some functions $A_f(m)$.
- S_χ is the Kloosterman sum

$$S_\chi(m,n,c) = \sum_{\substack{a=1 \\ (a,c)=1 \\ a\cdot\bar{a}\equiv 1(c)}}^{c} \overline{\chi(a)} e^{2\pi i \frac{am+\bar{a}n}{c}}.$$

- $\delta_{m,n}$ is the delta function

$$\delta_{m,n} = \begin{cases} 1 & \text{if } m = n, \\ 0 & \text{if } m \neq n. \end{cases}$$

- Finally, J_k is the Bessel function

$$J_k(z) = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!\Gamma(l+k+1)} \left(\frac{z}{2}\right)^{2l+k}.$$

This is the simplest form of trace formula, and was generalized by Selberg, later by Jacquet and Arthur. The Kloosterman sum $S_\chi(m,n,c)$ is associated with algebraic geometry. There is a famous paper by Weil which counts the number of points on a curve mod p in terms of Kloosterman sums. Weil's estimate

$$S_\chi(m,n,c) \ll c^{\frac{1}{2}+\epsilon}$$

is equivalent to the Riemann hypothesis for curves over finite fields. The Bessel functions are certain matrix coefficients of automorphic representations. Selberg was the first to prove that trace formulae are analogs on GL(2) of the Poisson summation formula.

We will now prove the Petersson Trace Formula.

Proof. The idea is to construct an automorphic function (Poincaré series) and then compute it in two ways:

(1) using Fourier expansions;
(2) using spectral expansions.

The Poincaré series is

$$P_m(z,k,\chi) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \overline{\widetilde{\chi}(\gamma)} j(\gamma,z)^{-k} e^{2\pi i m \gamma z}.$$

Since j is a cocycle, this is indeed an automorphic function:

$$P_m(\alpha z,k,\chi) = \widetilde{\chi}(\alpha) j(\alpha,z)^k P_m(z,k,\chi)$$

for all $\alpha \in \Gamma_0(N)$ and $z \in \mathfrak{h}$. It is not at all obvious that this function is not identically zero. If you can prove this for a fixed m, I will get your paper published on the Annals. What is known is that P_m is non-zero for many values of m.

(1) Fourier expansions

By periodicity, we have

$$P_m(z,k,\chi) = \sum_{n=1}^{\infty} \hat{P_m}(n) e^{2\pi i n z}$$

where $\hat{P_m}(n)$ is the n-th Fourier coefficient. The Fourier theorem implies

$$\hat{P_m}(n) = \lim_{y \to 0} \int_0^1 P_m(x+iy,k,\chi) e^{-2\pi i n x} dx.$$

The integral is equal to

$$\int_0^1 \sum_{\substack{c,d \\ c \equiv 0 (N) \\ (c,d)=1}} \frac{\overline{\chi(d)}}{(cz+d)^k} e^{2\pi i m \frac{az+b}{cz+d}} e^{-2\pi i n x} dx$$

(where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$ — given $(c,d)=1$, we can pick such a and b).

The first step is to write $d = lc + r$ where $1 \le r < c$ and $(r,c) = 1$. We know that $ad - bc = 1$, so $b = \frac{ad-1}{c} = \frac{a(lc+r)-1}{c}$. This implies that $az + b = az + al + \frac{ar-1}{c}$, so the integral is equal to

$$\begin{aligned}
&\int_0^1 \sum_{c \equiv 0 (N)} \sum_{l \in \mathbb{Z}} \sum_{\substack{r=1 \\ (r,c)=1}}^{c} \frac{\overline{\chi}(d)}{(cz+lc+r)^k} e^{2\pi i m \frac{az+al+\frac{ar-1}{c}}{cz+lc+r}} e^{-2\pi i n x} dx \\
&= \sum_{c \equiv 0 (N)} \sum_{l \in \mathbb{Z}} \sum_{\substack{r=1 \\ (r,c)=1}}^{c} \overline{\chi}(r) \int_{l+\frac{r}{c}}^{1+l+\frac{r}{c}} \frac{1}{(cz)^k} e^{2\pi i m \frac{az - \frac{1}{c}}{cz}} e^{-2\pi i n (x - l - \frac{r}{c})} dx \\
&= \sum_{c \equiv 0 (N)} \sum_{l \in \mathbb{Z}} \sum_{\substack{r=1 \\ (r,c)=1}}^{c} \frac{\overline{\chi}(r) e^{\frac{2\pi i r n}{c}}}{|c|^k} \int_{l+\frac{r}{c}}^{1+l+\frac{r}{c}} \frac{e^{2\pi i m (\frac{a}{c} - \frac{1}{c^2 z})} e^{-2\pi i n x}}{z^k} dx \\
&= \sum_{c \equiv 0 (N)} \sum_{\substack{r=1 \\ (r,c)=1}}^{c} \frac{\overline{\chi}(r) e^{\frac{2\pi i r n}{c}}}{|c|^k} \int_{-\infty}^{\infty} \frac{e^{2\pi i m (\frac{a}{c} - \frac{1}{c^2 z})} e^{-2\pi i n x}}{z^k} dx \\
&= \sum_{c \equiv 0 (N)} \frac{1}{|c|^k} \sum_{\substack{r=1 \\ (r,c)=1}}^{c} \overline{\chi}(r) e^{2\pi i \frac{rn + \bar{r} m}{c}} \int_{-\infty}^{\infty} \frac{e^{-\frac{2\pi i m}{|c|^2 z}} e^{-2\pi i n x}}{z^k} dx.
\end{aligned}$$

This is the first half of the calculation. We will do the spectral expansion next time.

Let us explain how we get the terms $A_f(m)$. Pick $m, n \ge 1$, $N \ge 1$, $\chi \pmod{N}$ and $k > 2$. The Petersson inner product is given as follows: for $F, G \in S_k$,

$$\langle F, G \rangle = \int_{\Gamma_0(N) \backslash \mathfrak{h}} F(z) \cdot \overline{G(z)} (\operatorname{Im} z)^k \frac{dx dy}{y^2}.$$

Next time we will use the Fourier coefficients that we computed. □

2. Lecture 2 (September 12, 2013)

2.1. **Petersson Trace Formula.** Last time we did the Petersson trace formula, and we will finish the proof today. We will actually do it in a more general setting than last time. Recall that we had level N, a character $\chi \pmod N$ and weight k. By level N we mean the subgroup $\Gamma_0(N)$. By a character we mean $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$, which lifts to $\widetilde{\chi} : \Gamma_0(N) \to \mathbb{C}^\times$ via $\widetilde{\chi}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) := \chi(d)$. By weight k we mean the cocycle $j(\gamma, z)^k$, where $j(\gamma, z) = cz + d$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$.

There is a one-to-one correspondence between automorphic representations and cusp forms on GL(2). There are two types of cusp forms:

- Holomorphic in z, which were completely developed by Hecke in the 1930's. Recall that they are holomorphic functions on $\mathfrak{h}$ satisfying
$$f(\gamma z) = \widetilde{\chi}(\gamma) j(\gamma, z)^k f(z)$$
for all $\gamma \in \Gamma_0(N)$ and
$$\lim_{y \to \infty} f(u + iy) = 0$$
for all $u \in \mathbb{Q}$.
- Non-holomorphic in z, which were introduced by Maass in the 1940's. Note that holomorphic functions satisfy the differential equation
$$\frac{\partial}{\partial \overline{z}} F(z) = 0$$
where
$$\frac{\partial}{\partial \overline{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$
For example, $\frac{\partial}{\partial \overline{z}} z^m = 0$. For the non-holomorphic cusp forms, we want them to satisfy the differential equation
$$\Delta f = \lambda f$$
where
$$\Delta = y^2 \left(\frac{\partial}{\partial z} \cdot \frac{\partial}{\partial \overline{z}}\right).$$

On the other side of the correspondence, we have adelic irreducible automorphic cuspidal representations. The modern reference is Jacquet and Langlands' book.

Holomorphic cusp forms correspond to discrete series representations, and non-holomorphic cusp forms correspond to non-discrete representations. Automorphic representations are easier to generalize to higher rank groups. On the other hand, applications to number theory all use the classical language.

What is the most general cocycle we can look at? We must have $j(\gamma, z)$, but instead of raising it to an integer power we can consider $j(\gamma, z)^c$ where $c \in \mathbb{C}$. Most generally, consider $\psi_c(\gamma) j(\gamma, z)^c$. In Maass' book, he considers $c \in \mathbb{Q}$ and lists all possibilities for $\psi_c(\gamma)$. There is an adelic theory for $c = \frac{k}{2}$, $k \in \mathbb{Z}$. This is the theory of metaplectic forms using a double cover of GL(2). We will talk about the metaplectic groups when we do theta functions.

Let us consider cocycles for half-integral weights. The level is required to be $4N$. Consider $\psi(\gamma)j(\gamma,z)^{\frac{k}{2}}$, where ψ is the Shimura symbol

$$\psi\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) := \left(\frac{c}{d}\right) \cdot \epsilon_d^{-1}$$

where $\epsilon_d = \begin{cases} 1 & \text{if } d \equiv 1 \pmod 4, \\ i & \text{if } d \equiv -1 \pmod 4 \end{cases}$, and $\left(\frac{*}{d}\right)$ is the quadratic Dirichlet character mod d.

Let us return to the Petersson Trace Formula (for level N, character χ and weight k) — we allow half-integral weights, but then N has to be divisible by 4. Let $S_k(\Gamma_0(N),\chi)$ be the space of cusp forms. Then

$$\frac{\Gamma(k-1)}{(4\pi\sqrt{mn})^{k-1}} \sum_{\substack{f\in S_k(\Gamma_0(N),\chi) \\ \text{orthonormal basis}}} \overline{A_f(m)}A_f(n) = \delta_{m,n} + \frac{1}{(2\pi i)^k} \sum_{c\equiv 0(N)} \frac{S_\chi(m,n,c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

Recall the Bessel function

$$J_k(z) = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!\Gamma(l+k+1)} \left(\frac{z}{2}\right)^{2l+k}.$$

Last time we introduced the Poincaré series and computed its Fourier coefficients.

$$P_m(z,k,\chi) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \overline{\widetilde{\chi}(\gamma)} j(\gamma,z)^{-k} e^{2\pi i m\gamma z} = \sum_{n=1}^{\infty} \hat{P_m}(n) e^{2\pi i n z}$$

(If we want to include half-integral weights, then we insert the Shimura symbol $\overline{\psi(\gamma)}$ in the sum.) We will not finish the computations from last time, but it turns out that

$$\hat{P_m}(n) = \delta_{m,n} + \frac{1}{(2\pi i)^k} \sum_{c\equiv 0(N)} \frac{S_\chi(m,n,c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

(The term $\delta_{m,n}$ comes from the special case when $c=0$. Last time we only computed the sum over $c \neq 0$.)

Now we will talk about how to get the left side. Let $f_1, f_2, \cdots, f_r$ be an orthonormal basis of $S_k(\Gamma_0(N),\chi)$. Recall the Petersson Inner Product: for $F, G \in S_k$,

$$\langle F, G\rangle = \int_{\Gamma_0(N)\backslash\mathfrak{h}} F(z)\cdot\overline{G(z)} y^k \frac{dxdy}{y^2}$$

where $z = x+iy$. The differential $\dfrac{dxdy}{y^2} = \dfrac{i}{2}\dfrac{dz\wedge d\overline{z}}{(\operatorname{Im} z)^2}$ is invariant under $z \mapsto \dfrac{az+b}{cz+d}$ for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$, and so defines an invariant Haar measure on $\Gamma_0(N)\backslash\mathfrak{h}$. Thus

$$\langle f_i, f_j\rangle = \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i\neq j. \end{cases}$$

This implies that for all $F \in S_k$, we have

$$F(z) = \sum_{j=1}^{r} \langle F, f_j \rangle f_j(z).$$

Applying this to the Poincaré series,

$$P_m(z, \chi) = \sum_{j=1}^{r} \langle P_m(*, \chi), f_j \rangle f_j(z).$$

This is the spectral computation. We have

$$\hat{P_m}(n) = \sum_{j=1}^{r} \langle P_m(*, \chi), f_j \rangle A_{f_j}(n).$$

It remains to compute

$$\langle P_m(*, \chi), f_j \rangle = \int_{\Gamma_0(N)\backslash\mathfrak{h}} P_m(z, \chi) \overline{f(z)} y^k \frac{dxdy}{y^2}.$$

Since f is a cusp form, it has a Fourier expansion $f(z) = \sum_{l=1}^{\infty} A_f(l) e^{2\pi i l z}$. The space $\Gamma_0(N)\backslash\mathfrak{h}$ is quite complicated — its fundamental domain can be obtained by choosing coset representatives for $\Gamma_0(N)\backslash\Gamma_0(1)$. Let us continue with the computation

$$\begin{aligned}
&= \int_{\Gamma_0(N)\backslash\mathfrak{h}} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \overline{\widetilde{\chi}(\gamma)} j(\gamma, z)^{-k} e^{2\pi i m \gamma z} \overline{f(z)} (\operatorname{Im} z)^k \frac{dxdy}{y^2} \\
&= \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \int_{\gamma \cdot (\Gamma_0(N)\backslash\mathfrak{h})} \overline{\widetilde{\chi}(\gamma)} j(\gamma, \gamma^{-1} z)^{-k} e^{2\pi i m z} \overline{f(\gamma^{-1} z)} (\operatorname{Im} \gamma^{-1} z)^k \frac{dxdy}{y^2}.
\end{aligned}$$

The cocycle condition implies

$$1 = j(\gamma\gamma^{-1}, z) = j(\gamma, \gamma^{-1} z) j(\gamma^{-1}, z)$$

and so

$$j(\gamma, \gamma^{-1} z) = \frac{1}{j(\gamma^{-1}, z)}.$$

Since f is modular, $f(\gamma^{-1} z) = \widetilde{\chi}(\gamma^{-1}) j(\gamma^{-1}, z)^k f(z)$ and the above is equal to

$$\begin{aligned}
&= \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \int_{\gamma \cdot (\Gamma_0(N)\backslash\mathfrak{h})} j(\gamma^{-1}, z)^k e^{2\pi i m z} \overline{j(\gamma^{-1}, z)}^k \overline{f(z)} (\operatorname{Im} \gamma^{-1} z)^k \frac{dxdy}{y^2} \\
&= \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \int_{\gamma \cdot (\Gamma_0(N)\backslash\mathfrak{h})} e^{2\pi i m z} \overline{f(z)} (\operatorname{Im} z)^k \frac{dxdy}{y^2}
\end{aligned}$$

since $\operatorname{Im} \frac{az+b}{cz+d} = \frac{\operatorname{Im} z}{|cz+d|^2}$, which can be shown by a brute-force computation. The above becomes

$$\begin{aligned}
&= \int_{\Gamma_\infty \backslash \mathfrak{h}} e^{2\pi i m x} e^{-2\pi m y} \overline{f(z)} y^k \frac{dxdy}{y^2} \\
&= \int_{x=0}^{1} \int_{y=0}^{\infty} e^{-2\pi m y} y^k e^{2\pi i m x} \sum_{l=1}^{\infty} A_f(l) e^{-2\pi i l x} e^{-2\pi l y} \frac{dxdy}{y^2}
\end{aligned}$$

$$
\begin{aligned}
&= A_f(m) \int_0^\infty e^{-4\pi m y} y^{k-1} \frac{dy}{y} \\
&= \frac{A_f(m)}{(4\pi m)^{k-1}} \Gamma(k-1).
\end{aligned}
$$

The factor $\sqrt{mn}$ comes from renormalization, but we will not try to fix that now.

The Ramanujan–Petersson Conjecture says that if $f \in S_k(\Gamma_0(N), \chi)$, then

$$
|A_f(n)| \ll n^{\frac{k-1}{2}+\epsilon}
$$

This was proved by Deligne for f holomorphic. It is enough to prove

$$
|A_f(p)| \leq 2p^{\frac{k-1}{2}}
$$

for p prime. This can be thought of as the error term for counting points on an algebraic variety mod p.

Using the trace formula, we get that

$$
|A_f(n)|^2 \ll \left| \sum_{c \equiv 0(N)} \frac{S_\chi(n,n,c)}{c} J_k\left(\frac{4\pi n}{c}\right) \right|
$$

but this gives a weaker bound than the Ramanujan–Petersson Conjecture. In the case of half-integral weights, for which Deligne's proof does not work, there is an exact formula by Salié. The Salié sum looks like

$$
\sum_{\substack{a=1 \\ (a,c)=1}}^{c} \left(\frac{a}{c}\right) e^{\frac{2\pi i (an + \bar{a}n)}{c}}.
$$

Kloosterman sums are recently generalized as Kloosterman sheaves by Ngô. There should be a theory of Salié sums for metaplectic groups, but no one has done that yet.

This ends the discussion of the Petersson Trace Formula.

2.2. **Kuznetsov Trace Formula.** The Kuznetsov trace formula came out in around 1978. I am skipping the Selberg trace formula (1950) for the moment. It was only after Selberg that people realized what a trace formula was. Jacquet asked me lots of questions in 1979 about the Kuznetsov Trace Formula, and vastly generalized it in around 1979 or 1980 to the Relative Trace Formula in the language of representation theory.

Again, we will be looking at level N, character χ mod N, and weight k. I will not put in the Shimura symbol explicitly, but it also works for the half-integral weight case. The idea is to develop the Petersson trace formula for non-holomorphic forms.

I want to first talk about cusps, but it is time.

3. Lecture 3 (September 17, 2013)

I've been doing work of Hecke and Petersson from the 1930's. Now I want to move to Maass (1940's). Hecke and Petersson studied holomorphic modular forms. The Petersson trace formula is one of the deepest formulas for the finite-dimensional space of modular forms. Maass introduced non-holomorphic modular forms. We will briefly review Maass' theory and then apply it.

3.1. **Maass Forms.** Let $\Gamma \subset \mathrm{SL}(2,\mathbb{R})$ be a discrete subgroup. The example I'll be looking at is the congruence subgroup $\Gamma = \Gamma_0(N)$. On GL(2) (which has rank 1 as a Lie group), there exist non-congruence discrete subgroups. Selberg conjectured that for higher rank groups, the only discrete subgroups of finite index are congruence subgroups. This was proved by Margulis.

A Maass form ϕ is

(1) a complex-valued function $\phi : \mathfrak{h} \to \mathbb{C}$;
(2) $\phi(\gamma z) = \phi(z)$ for all $\gamma \in \Gamma$;
(3) $\Delta\phi = \lambda_\phi \phi$ for some $\lambda_\phi \in \mathbb{R}$, where $\Delta = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ is the Laplacian;
(4) ϕ has moderate growth at the cusps of Γ.

Now we have to talk about cusps. A cusp is $\kappa \in \mathbb{R} \cup \{\infty\}$ such that[1]

$$\kappa = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \infty = \lim_{y\to\infty} \frac{aiy+b}{ciy+d} = \frac{a}{c}$$

for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$. A cusp for a congruence subgroup must be a rational number, but for a non-congruence subgroup, a cusp could be irrational. There are finitely many inequivalent cusps.

$\Gamma\backslash\mathfrak{h}$ is not compact, so we compactify by adding in the cusps: $\Gamma\backslash\mathfrak{h}^*$ where $\mathfrak{h}^* = \mathfrak{h} \cup \{i\infty, \kappa_1, \kappa_2, \cdots\}$. It requires some work to show this is a complex manifold. I think Shimura was the first to write down the complete proof.

As long as we're away from the cusps, ϕ is bounded by some constant. Let κ be a cusp. Then there exists $\sigma \in \mathrm{SL}(2,\mathbb{R})$ such that $\sigma\infty = \kappa$. Let $\Gamma_\kappa = \{\gamma \in \Gamma : \gamma\kappa = \kappa\}$ be the stabilizer of the cusp κ. It is always possible to choose σ such that $\sigma^{-1}\Gamma_\kappa\sigma = \Gamma_\infty$, which is equivalent to $\Gamma_\kappa = \sigma\Gamma_\infty\sigma^{-1}$.

Let's look the function ϕ at a cusp. $\phi(\sigma z)$ is periodic in z, i.e.

$$\phi(\sigma(z+1)) = \phi(\sigma z).$$

This is because

$$\phi\left(\sigma\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} z\right) = \phi\left(\sigma\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\sigma^{-1}\sigma z\right) = \phi(\sigma z)$$

since $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_\infty$. Thus there is a Fourier expansion

$$\phi(\sigma z) = \sum_{m\in\mathbb{Z}} a_\phi(m,y)e^{2\pi i m x}.$$

We say ϕ has moderate growth at a cusp k if there exists a constant $B > 0$ such that $|\phi(\sigma z)| \ll y^B$ as $y \to \infty$ (and x fixed). This implies that $|a_\phi(m,y)| \ll y^B e^{-cmy}$ as $y \to \infty$ and $c > 0$. Let's explain how we get a bound like this. If the function is not an eigenfunction for the Laplacian (i.e. we have an automorphic function, not automorphic form), this estimate does not hold.

Since Δ is invariant under $\mathrm{SL}(2,\mathbb{R})$, We have

$$\Delta\phi(\sigma z) = \lambda_\phi \cdot \phi(\sigma z),$$

[1]The correct definition is that κ be fixed by some parabolic element of Γ.

i.e.

$$\begin{aligned}\Delta\phi(\sigma z) &= -y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\sum_m a_\phi(m,y)e^{2\pi imx}\\ &= \sum_m -y^2(a''_\phi(m,y)e^{2\pi imx}-4\pi^2m^2a_\phi(m,y)e^{2\pi imx})\\ &= \lambda_\phi\cdot\phi(\sigma z),\end{aligned}$$

and hence

$$-y^2(a''_\phi(m,y)-4\pi^2m^2a_\phi(m,y))=\lambda_\phi a_\phi(m,y).$$

This is Whittaker's differential equation, and there are two solutions, one with exponential decay in y and one with exponential growth in y. Since ϕ has moderate growth, $a_\phi(m,y)$ must have exponential decay in y.

Now the question is, do Maass forms exist? The answer is yes, thanks to the Eisenstein series. But then we can ask, do Maass cusp forms exist? A Maass cusp form is a Maass form such that $a_\phi(0,y)=0$ for all cusps κ.

The simplest discrete subgroup one can look at is $\mathrm{SL}(2,\mathbb{Z})$. In 1940, Maass actually constructed cusp forms for $\Gamma_0(N)$ for certain N. In my book *Automorphic Forms and L-functions for the Group* $\mathrm{GL}(n,\mathbb{R})$, I gave the construction. But Maass was unable to construct it for $\mathrm{SL}(2,\mathbb{Z})$. in 1950, Selberg proved infinitely many Maass forms for Γ a congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$ (in particular, for $\mathrm{SL}(2,\mathbb{Z})$). Not only did he prove there are infinitely many, but he could count how many there were. This was one of the first applications of his trace formula.

The final question is, can Maass cusp forms exist for non-congruence subgroups? Around 1980 (at Selberg's 60th birthday), Sarnak and Philips developed the theory of spectral deformation, and conjectured that for the general non-congruence subgroup there are no Maass cusp forms. This is now wide open.

3.2. **Eisenstein Series.** Let $\Gamma\subset\mathrm{SL}(2,\mathbb{R})$ be a discrete subgroup. We will denote $z\in\mathfrak{h}$ and $s\in\mathbb{C}$. Define

$$E(z,s)=\sum_{\gamma\in\Gamma_\infty\backslash\Gamma}\mathrm{Im}(\gamma z)^s=\sum_{\left(\begin{smallmatrix}* & *\\ c & d\end{smallmatrix}\right)\in\Gamma}\frac{y^s}{|cz+d|^{2s}}.$$

This converges for $\mathrm{Re}(s)>1$ (using the fact that Γ is discrete). Note that

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

Since $\Delta_z=\Delta_{\sigma z}$ for all $\sigma\in\mathrm{SL}(2,\mathbb{R})$, we have

$$\Delta(\mathrm{Im}\,\gamma z)^s=s(1-s)(\mathrm{Im}\,\gamma z)^s$$

and so

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

By definition it is easy to see that $E(\gamma z,s)=E(z,s)$. The only question is whether $E(z,s)$ has moderate growth. Since it satisfies the Whittaker differential equation, it is of either moderate or exponential growth.

Theorem 3.1. *$E(z,s)$ has moderate growth.*

There are other Eisenstein series. Actually we can construct an Eisenstein series for each cusp $\kappa_1, \kappa_2, \cdots, \kappa_r$ (a set of inequivalent cusps). Let κ be a cusp, and $\sigma\infty = \kappa$, $\sigma^{-1}\Gamma_\kappa\sigma = \Gamma_\infty$. We define

$$E_\kappa(z,s) = \sum_{\gamma \in \Gamma_\kappa \backslash \Gamma} (\mathrm{Im}(\sigma^{-1}\gamma z))^s.$$

Then we claim that $E_\kappa(\alpha z, s) = E_\kappa(z,s)$ for all $\alpha \in \Gamma$. Indeed, we have

$$E_\kappa(\alpha z, s) = \sum_{\gamma \in \Gamma_\kappa \backslash \Gamma} (\mathrm{Im}(\sigma^{-1}\gamma\alpha z))^s = E_\kappa(z,s).$$

Assume Γ has r inequivalent cusps $\kappa_1 = \infty, \kappa_2, \cdots, \kappa_r$. Then there exist r different Eisenstein seires $E_1(z,s), \cdots, E_r(z,s)$. Each $E_i(z,s)$ has a Fourier expansion at the cusp κ_j

$$E_i(\sigma_j z, s) = \sum_m A_{m,i,j}(y,s) e^{2\pi i m x}.$$

Theorem 3.2. *$E_i(z,s)$ has moderate growth for $i = 1, 2, \cdots, r$.*

We don't have enough time to prove this now. Let me talk about the Selberg Spectral Decomposition — every automorphic function of moderate growth on $\Gamma\backslash\mathfrak{h}$ can be written as a linear combination of Maass cusp forms and integrals of Eisenstein series

$$\sum_{i=1}^{r} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} (*) E_i(z,s)\, ds$$

and sums of residues of Eisenstein series. This can be made very explicit. The Eisenstein series at the different cusps all come into the picture. Thus Γ has:

- a discrete spectrum, given by "Maass cusp forms";
- a continuous spectrum, given by $E_i(z,s)$ for $i = 1, 2 \cdots, r$; and
- a residual spectrum, given by $\operatorname*{Res}_{s=\text{pole}} E_i(z,s)$ for $i = 1, 2, \cdots, r$.

4. Lecture 4 (September 19, 2013)

4.1. **Eisenstein Series.** Last time we were doing the Eisenstein series for $\Gamma \subset \mathrm{SL}(2,\mathbb{R})$, where Γ is a discrete subgroup. For simplicity we assume that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Let $c_1, c_2, \cdots, c_r$ be inequivalent cusps for Γ. For each cusp c, the stabilizer is $\Gamma_c = \{\gamma \in \Gamma : \gamma c = c\}$. At each cusp c_i, we can pick $\sigma_i \in \mathrm{SL}(2,\mathbb{R})$ such that $\sigma_i c_i = \infty$ and $\sigma_i^{-1}\Gamma_{c_i}\sigma_i = \Gamma_\infty = \left\{\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}\right\}$. The Eisenstein series is

$$E_i(z,s) = \sum_{\gamma \in \Gamma_c \backslash \Gamma} \mathrm{Im}(\sigma_i^{-1}\gamma z)^s$$

which is convergent for $\mathrm{Re}(s) > 1$. Each $E_i(z,s)$ has a Fourier expansion at a cusp c_j

$$E_i(\sigma_j z, s) = \sum_{m \in \mathbb{Z}} A_{ij}(m, y, s) e^{2\pi i m x}.$$

Very important is the constant term $A_{ij}(0, y, s)$.

For the case of GL(2), we can compute these $A_{i,j}$ very explicitly, and they turn out to be Bessel functions multiplied by certain divisor sums. Thus we can prove things every easily, e.g. the functional equation and moderate growth, by explicit computations. If we go to higher rank groups, then the explicit calculations break down. The only thing people can compute explicitly in general is the constant term. That's the problem Langlands gave to Shahidi. Once we know the constant term and we know the function is an eigenfunction of a differential operator, we can prove all the properties we want. The first proof was found by Selberg. Langlands, in his book on L-functions, took Selberg's proof and generalized it to all reductive groups. Later Selberg found a second proof who only had one step (without first finding the constant term). In the 1970's and 80's, Sarnak and Phillips were going through that proof, but it didn't seem to work for all cases.

Anyway, now we are going to prove everything using the expansion. We have the following theorem, first proved by Kubota in his book *Elementary theory of Eisenstein series.* I think the proof was known to Selberg, but he never wrote it up.

Theorem 4.1 (Kubota)**.**

$$A_{ij}(m,y,s) = \delta_{0,m}y^s + y^{1-s}\sum_c \left(\sum_{\substack{d \pmod c \\ \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma_\infty \backslash \sigma_i^{-1}\Gamma\sigma_j/\Gamma_\infty}} \frac{1}{|c|^{2s}} e^{\frac{2\pi i m d}{c}} \right) \int_{-\infty}^{\infty} \frac{e^{-2\pi i m x y}}{(x^2+1)^s}\,dx.$$

The integral is a Whittaker function

$$\int_{-\infty}^{\infty} \frac{e^{-2\pi i x y}}{(x^2+1)^s}\,dx = \frac{2\pi^s |y|^{s-\frac{1}{2}}}{\Gamma(s)} K_{s-\frac{1}{2}}(2\pi|y|)$$

where K is the Bessel function with exponential decay as $|y| \to \infty$. When $m = 0$, there is a simpler formula

$$\int_{-\infty}^{\infty} \frac{dx}{(x^2+1)^s} = \sqrt{\pi}\frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)}.$$

Theorem 4.2. *For any fixed $s \in \mathbb{C}$, the Eisenstein series is of moderate growth if and only if $|E_i(z,s)| \ll y^B$ (as $y \to \infty$) for some $B = B(s) > 1$.*

Proof. Assume we have proved the Fourier expansion (Theorem 1). Then

$$|A_{ij}(m,y,s)| \ll y^{\mathrm{Re}(s)}$$

for $\mathrm{Re}(s) > 1$, because

$$|A_{ij}(m,y,s)| \ll y^{\mathrm{Re}(s)} + y^{1-\mathrm{Re}(s)} \sum_c \frac{c}{|c|^{2\,\mathrm{Re}(s)}} \cdot y^{\mathrm{Re}(s)-\frac{1}{2}} |K_{s-\frac{1}{2}}(2\pi|m|y)|.$$

□

Next I want to talk about the functional equation of Eisenstein series. The individual Eisenstein series do not have a functional equation. We have to define a vector of Eisenstein series.

Definition 4.3. $\mathcal{E}(z,s) = (E_1(z,s), E_2(z,s), \cdots, E_r(z,s))$.

Theorem 4.4 (Functional Equation (Selberg))**.**

$$\mathcal{E}(z,s) = \Phi(s)\mathcal{E}(z,1-s).$$

Here $\Phi(s) = (\phi_{ij}(s))_{\substack{i=1,\cdots,r\\ j=1,\cdots,r}}$ *is the Scattering Matrix, where*

$$\phi_{ij}(s) := \sqrt{\pi}\frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)}\sum_{c=1}^{\infty}\frac{1}{|c|^{2s}} \sum_{\substack{d \pmod c \\ \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right)\in\sigma_i^{-1}\Gamma\sigma_j}} 1$$

is the Fourier coefficient of y^{1-s}*.*

All the higher terms satisfy the functional equation because $K_s(y) = K_{-s}(y)$, so the only problem is with the constant term. The key step in the proof is that Φ is a symmetric matrix. We are looking at the kind of sums

$$\sum_{\left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right)\in\sigma_i^{-1}\Gamma\sigma_j} 1,$$

i.e. we are counting the number of representations of $\begin{pmatrix} * & * \\ c & d \end{pmatrix}$ in distinct cosets of $\Gamma_\infty\backslash\sigma_i^{-1}\Gamma\sigma_j/\Gamma_\infty$ for a fixed c. Now apply the map $g \mapsto g^{-1}$ where $g \in \Gamma$, and we can see that we get the exact same counting. Thus Φ is symmetric

$$\Phi(s) = {}^t\Phi(s)$$

and Φ is unitary at $s = \frac{1}{2} + it$ for $t \in \mathbb{R}$.

Selberg's second proof is really amazing — it doesn't use this formula, but instead uses Fredholm theory which is pure analysis.

We have established that the Eisenstein series has moderate growth for $\mathrm{Re}(s) > 1$ and satisfies a functional equation. Combining the two gives moderate growth in the region $\mathrm{Re}(s) < 0$, and by the maximum principle we get moderate growth in between.

All that remains to prove is the Fourier expansion. Last time I proved the Fourier expansion of non-holomorphic Poincaré series, which is elementary. Now we are going to give essentially the same proof, in a group-theoretic setting.

We define the Poincaré series

$$P_m(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^s e^{2\pi i m\gamma z}.$$

When $m = 0$, this gives the Eisenstein series. We can do this for an arbitrary cusp, but let's just look at the cusp at ∞ for simplicity. It has a Fourier expansion

$$P_m(z,s) = \sum_{n\in\mathbb{Z}} \hat{P}_m(n,y,s)e^{2\pi i n x}$$

where

$$\begin{aligned}\hat{P}_m(n,y,s) &= \int_0^1 P_m(z,s)e^{-2\pi i n x}dx \\ &= \int_0^1 \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} (\mathrm{Im}\,\gamma z)^s e^{2\pi i m\gamma z}e^{-2\pi i n x}dx.\end{aligned}$$

We need the Bruhat decomposition of Γ (which is only assumed to be a discrete subgroup). There are basically two kinds of decompositions for discrete groups — double cosets or conjugacy classes. The Bruhat decomposition is a double coset decomposition. Very roughly, the Selberg trace formula involves conjugacy classes and the relative trace formula involves double cosets. The identity coset is $\Gamma_\infty \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Gamma_\infty$ and the other coset is $\Gamma_\infty \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Gamma_\infty$ for $c \neq 0$. We have

$$\Gamma = \left(\Gamma_\infty \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Gamma_\infty\right) \cup \bigcup_{c \neq 0} \left(\Gamma_\infty \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Gamma_\infty\right)$$

where $\Gamma_\infty \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Gamma_\infty = \Gamma_\infty \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \Gamma_\infty$ if and only if $c = c'$ and $d \equiv d' \pmod{c}$. The Bruhat decomposition is

$$\Gamma = \left(\Gamma_\infty \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Gamma_\infty\right) \cup \bigcup_{\substack{\left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma \\ c \neq 0 \\ d \pmod{c}}} \left(\Gamma_\infty \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Gamma_\infty\right).$$

Now we can continue with the computation.

$$\begin{aligned}
\hat{P}_m(n, y, s) &= \int_{-\infty}^{\infty} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_\infty} (\operatorname{Im} \gamma z)^s e^{2\pi i m \gamma z} e^{-2\pi i n x} dx \\
&= \delta_{0,m} y^s e^{-2\pi m y} + \int_{-\infty}^{\infty} \sum_{\substack{\left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma \\ c \neq 0 \\ d \pmod{c}}} (\operatorname{Im} \gamma z)^s e^{2\pi i m \gamma z} e^{-2\pi i m x} dx \\
&= \delta_{0,m} y^s e^{-2\pi m y} + \sum_{\substack{\left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma \\ c \neq 0 \\ d \pmod{c}}} \int_{-\infty}^{\infty} \frac{y^s}{((cx+d)^2 + c^2 y^2)^s} e^{2\pi i m \frac{az+b}{cz+d}} e^{-2\pi i n x} dx.
\end{aligned}$$

For the Eisenstein series, $m = 0$. We substitute $x \mapsto x - \frac{d}{c}$ and then $x \mapsto xy$ to get

$$\hat{P}_m(n, y, s) = \delta_{0,m} y^s e^{-2\pi m y} + \sum_{c=1}^{\infty} \frac{y^{1-s}}{c^{2s}} \sum_{\substack{d \pmod{c} \\ \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma}} e^{\frac{2\pi i n d}{c}} \int_{-\infty}^{\infty} \frac{e^{-2\pi i n x y}}{(1+x^2)^s} dx.$$

4.2. **Selberg Spectral Decomposition.** Let $\Gamma \subset \mathrm{SL}(2, \mathbb{R})$ be a discrete subgroup, $c_1, c_2, \cdots, c_r \in \mathbb{R} \cup \{\infty\}$ be inequivalent cusps of Γ, and $E_i(z, s)$ $(i = 1, 2, \cdots, r)$ be the Eisenstein series at each cusp c_i. Let $\eta_j(z)$ $(j = 1, 2, 3, \cdots)$ be a basis of Maass cusp forms for Γ. Let $\eta_0(z)$ be the constant function.

As I mentioned last time, it is very hard to prove that Maass cusp forms exist. We only expect them for congruence subgroups.

We have the space

$$\mathcal{L}^2(\Gamma \backslash \mathfrak{h}) = \left\{ F : \Gamma \backslash \mathfrak{h} \to \mathbb{C} : \int_{\Gamma \backslash \mathfrak{h}} |F(z)|^2 \frac{dx dy}{y^2} < \infty \right\}.$$

The spectral decomposition says that if $F \in \mathcal{L}^2(\Gamma\backslash\mathfrak{h})$, then

$$F(z) = \sum_{j=1}^{\infty} \langle F, \eta_j \rangle \frac{\eta_j(z)}{\langle \eta_j, \eta_j \rangle} + \sum_{i=1}^{r} \frac{1}{4\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \langle F, E_i(*, s)\rangle E_i(z, s) ds + \sum_{k=1}^{l} \langle F, R_k \rangle \frac{R_k(z)}{\langle R_k, R_k \rangle}$$

where $R_k(z)$ is the residue at a pole of some E_i

$$R_k(z) = \operatorname*{Res}_{s=s_0} E_i(z, s).$$

This was first prove by Selberg, but I will not prove this. There is a nice exposition by Müller.

For congruence subgroups, the only residual term comes from the constant function, so we can rewrite this sum as

$$\sum_{j=0}^{\infty} \langle F, \eta_j \rangle \frac{\eta_j(z)}{\langle \eta_j, \eta_j \rangle} + \sum_{i=1}^{r} \frac{1}{4\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \langle F, E_i(*, s)\rangle E_i(z, s) ds.$$

In the case of GL(2) where we can compute Fourier expansions explicitly, the proof is very easy. In my book I gave a proof for GL(3). The most general proof can be found in Arthur's article in the Corvallis proceedings.

5. Lecture 5 (September 19, 2013)

5.1. **Kuznetsov Trace Formula.** The Kuznetsov trace formula is a generalization of the Petersson trace formula, except that it uses non-holomorphic functions. Take $\Gamma = \Gamma_0(N) \in \mathrm{SL}(2, \mathbb{Z})$. We can work more generally but let's stick to this case. We have the standard Poincaré series

$$P_m(z, s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\operatorname{Im} \gamma z)^s e^{2\pi i m \gamma z}$$

and also the more general Poincaré series

$$\mathcal{P}_m(z, p) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} p(2\pi m \operatorname{Im}(\gamma z)) e^{2\pi i m \gamma z},$$

where $p : \mathbb{R} \to \mathbb{C}$ satisfies $|p(y)| \ll y^{1+\epsilon}$ as $y \to \infty$. This condition ensures that the series converges.

The KTF is obtained by computing

$$\langle \mathcal{P}_m(*, p), \mathcal{Q}_n(*, q)\rangle = \int_{\Gamma\backslash\mathfrak{h}} \mathcal{P}_m(z, p) \overline{\mathcal{Q}_n(z, q)} \frac{dxdy}{y^2}$$

in two different ways, where $\mathcal{Q}_n(z, q) = \sum q(2\pi n \operatorname{Im}(\gamma z)) e^{2\pi i n \gamma z}$ is another Poincaré series,

Let's first do the spectral computation. Let $\eta_j(z)$ $(j = 1, 2, \cdots)$ be Maass forms for $\Gamma_0(N)$. There is an expansion

$$\eta_j(z) = \sum_{l \neq 0} A_j(l) \sqrt{y} K_{it_j}(2\pi |l| y) e^{2\pi i l x}$$

where A_j are the arithmetic Fourier coefficients and K_{it_j} is the Whittaker function.

We need to know that these Poincaré series are in L^2. Let's just assume p is chosen such that $\mathcal{P}_m \in L^2(\Gamma\backslash\mathfrak{h})$. In fact $|p(y)| \ll y^{1+\epsilon}$ is enough, because there is a Fourier expansion

$$\mathcal{P}_m(z, p) = y^s e^{-2\pi m y} + \text{higher terms}$$

where the higher terms are very small, and the first term has decay in $y \to \infty$. It is easy to see that at the cusp at infinity, the function is L^2. We also have to check the other cusps, but the Fourier expansions are similar, possibly without the first term.

Since $\mathcal{P}_m \in L^2$, it has a spectral expansion

$$\mathcal{P}_m(z,p) = \sum_{j=1}^{\infty} \langle \mathcal{P}_m, \eta_j \rangle \frac{\eta_j(z)}{\langle \eta_j, \eta_j \rangle} + \sum_{k=1}^{r} \frac{1}{4\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \langle \mathcal{P}_m, E_k(*,s) \rangle E_k(z,s) ds$$

(the only residue for $\Gamma_0(N)$ is the constant function, but that is orthogonal to the Eisenstein series anyway). We have to compute these inner products explicitly, but we can already say

$$\langle \mathcal{P}_m, \mathcal{Q}_n \rangle = \sum_{j=1}^{\infty} \frac{\langle \mathcal{P}_m, \eta_j \rangle \overline{\langle \mathcal{Q}_n, \eta_j \rangle}}{\langle \eta_j, \eta_j \rangle} + \sum_{k=1}^{r} \frac{1}{4\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \langle \mathcal{P}_m, E_k(*,s) \rangle \overline{\langle \mathcal{Q}_n, E_k(*,s) \rangle} ds.$$

This is the key formula.

We need to compute the inner products

$$\begin{aligned}
\langle \mathcal{P}_m, \eta_j \rangle &= \int_{\Gamma \backslash \mathfrak{h}} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} p(2\pi m \operatorname{Im} \gamma z) e^{2\pi i m \gamma z} \cdot \overline{\eta_j(z)} \frac{dxdy}{y^2} \\
&= \int_{\Gamma_\infty \backslash \mathfrak{h}} p(2\pi m y) e^{2\pi i m z} \overline{\eta_j(z)} \frac{dxdy}{y^2} \\
&= \int_{x=0}^{1} \int_{y=0}^{\infty} p(2\pi m y) e^{2\pi i m x} e^{-2\pi m y} \sum_{l \neq 0} \overline{A_j(l)} \sqrt{y} K_{it_j}(2\pi |l| y) e^{-2\pi i l x} \frac{dxdy}{y^2}
\end{aligned}$$

(We know that the Bessel function satisfies $K_{it}(y) = K_{-it}(y) = \overline{K_{it}(y)}$, which is the functional equation of the GL(2) Whittaker function. In fact $A_j(l)$ are real too because the Maass forms are self-dual, but we don't need to assume that.)

$$\begin{aligned}
&= \int_0^{\infty} p(2\pi m y) e^{-2\pi m y} A_j(m) \sqrt{y} K_{it_j}(2\pi m y) \frac{dy}{y^2} \\
&= \frac{A_j(m)}{\sqrt{2\pi m}} \cdot \int_0^{\infty} p(y) y^{-\frac{1}{2}} K_{it_j}(y) \frac{dy}{y}.
\end{aligned}$$

Thus,

$$\langle \mathcal{P}_m, \mathcal{Q}_n \rangle = \sum_{j=1}^{\infty} \frac{A_j(m)\overline{A_j(n)}}{2\pi\sqrt{mn}\langle \eta_j, \eta_j \rangle} p^{\#}(t_j) \cdot \overline{q^{\#}(t_j)} + \sum_{k=1}^{r} \frac{1}{4\pi i} \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} E_k(m,s) \overline{E_k(n,s)} p^{\#}(*) \overline{q^{\#}(*)} ds$$

where

$$p^{\#}(t) = \int_0^{\infty} p(y) y^{-\frac{1}{2}} K_{it}(y) \frac{dy}{y}$$

is the Whittaker transform of p.

Let's work out the last term more precisely. Recall that we have

$$\Delta \eta_j = \left(\frac{1}{4} + t^2 \right) \eta_j$$

and

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

Substituting $s = \frac{1}{2} + it$ will give the second term.

In conclusion, we have obtained the spectral side of the KTF

$$\langle \mathcal{P}_m, \mathcal{Q}_n \rangle = \sum_{j=1}^{\infty} \frac{A_j(m)\overline{A_j(n)}}{2\pi\sqrt{mn}\langle \eta_j, \eta_j \rangle} p^{\#}(t_j)\overline{q^{\#}(t_j)} + \sum_{k=1}^{r} \frac{1}{4\pi} \int_{-i\infty}^{i\infty} E_k\left(m, \frac{1}{2} + it\right) \overline{E_k\left(n, \frac{1}{2} + it\right)} p^{\#}(t)\overline{q^{\#}(t)} dt.$$

Selberg was the first one to prove infinitely many Maass forms, using the Selberg trace formula. At a conference 12 years ago I asked if we can use the Kuznetsov trace formula to prove the same thing. The answer turned out to be yes.

Now we move to the geometric side of KTF. We evaluate $\langle \mathcal{P}_m, \mathcal{Q}_n \rangle$ using the Fourier expansion of $\mathcal{P}_m$

$$\mathcal{P}_m(z, p) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} p(2\pi m \operatorname{Im}(\gamma z)) e^{2\pi i m \gamma z} = \sum_{l \in \mathbb{Z}} B_m(l, p, y) e^{2\pi i l x}$$

where $B_0(l, p, y) = p(2\pi i l y) e^{-2\pi l y} + \cdots$. Basically $B_m(l, p, y)$ are related to Kloosterman sums, which are related to algebraic geometry and explain why this is called the geometric side.

We have

$$\begin{aligned}
\langle \mathcal{P}_m, \mathcal{Q}_n \rangle &= \int_{\Gamma \backslash \mathfrak{h}} \mathcal{P}_m(z, p) \overline{\mathcal{Q}_n(z, q)} \frac{dxdy}{y^2} \\
&= \int_0^1 \int_0^\infty p(2\pi m y) e^{2\pi i m z} \overline{\mathcal{Q}_n(z, q)} \frac{dxdy}{y^2} \\
&= \int_0^1 \int_0^\infty p(2\pi m y) e^{2\pi i m x} e^{-2\pi m y} \sum_l \overline{B_n(l, q, y)} e^{-2\pi i l x} dx \frac{dy}{y^2} \\
&= \int_0^\infty p(2\pi m y) e^{-2\pi m y} \overline{B_n(m, q, y)} \frac{dy}{y^2}.
\end{aligned}$$

Now we can write down the Kuznetsov trace formula

$$\begin{aligned}
\sum_{j=1}^{\infty} \frac{A_j(m)\overline{A_j(n)} p^{\#}(t_j)\overline{q^{\#}(t_j)}}{2\pi\sqrt{mn}\langle \eta_j, \eta_j \rangle} &+ \frac{1}{4\pi} \int_{-\infty}^{\infty} \sum_{k=1}^{r} E_k\left(m, \frac{1}{2} + it\right) \overline{E_k\left(n, \frac{1}{2} + it\right)} p^{\#}(t)\overline{q^{\#}(t)} dt \\
&= \int_0^\infty p(2\pi m y) e^{-2\pi m y} B_n(m, q, y) \frac{dy}{y^2}
\end{aligned}$$

where $B_n(m, q, y)$ is an infinite sum of $S(m, n, c)$.

On GL(2), there are many applications of this formula. The first was by Deshouillers and Iwaniec in 1980's. Last time I mentioned that Sarnak and Phillips conjectured there are no Maass forms for non-congruence subgroups. In order to prove this they needed certain special values of L-functions to be non-zero. Deshouillers–Iwaniec used the KTF to prove these values are non-zero infinitely often.

When Kuznetsov first came up with his trace formula, he thought it was so powerful that he used it to prove four other major conjectures, including the Ramanujan and Lehmer conjectures. These proofs even got published in some obscure Russian journals (Kuznetsov was residing in Siberia), but I don't think they were correct. Of course the trace formula was correct. One application that was correct was concerned with the Linnik–Selberg conjecture.

Conjecture 5.1 (Linnik–Selberg). *For every $\epsilon > 0$, we have*

$$\sum_{c=1}^{x} \frac{S(m,n,c)}{c} \ll x^{\epsilon}$$

as $x \to \infty$.

Kuznetsov proved this for $\epsilon = \frac{1}{6}$, which is the best bound so far.

Another application of the KTF was a result of Sarnak–Luo–Iwaniec.

6. Lecture 6 (September 26, 2013)

6.1. **Takhtajan–Vinogradov Trace Formula.** Today we will talk about an interesting application that is not well-known. The reference is *The Gauss–Hasse hypothesis on real quadratic fields with class number one* by L. Takhtajan and A.I. Vinogradov, published in Crelle's Journal in 1982. This application has to do with the following conjecture of Gauss:

Conjecture 6.1 (Gauss, 1801). *There are infinitely many real quadratic fields with class number equal to* 1*.*

I think this is the only conjecture in the *Disquisitiones* that remains unproved.

We will use the theta function

$$\theta(z) = \sum_{n=-\infty}^{\infty} e^{2\pi i n^2 z}$$

which satisfies

$$\theta\left(\frac{az+b}{cz+d}\right) = \epsilon_d \left(\frac{d}{c}\right) (cz+d)^{\frac{1}{2}} \theta(z)$$

for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4)$, $\left(\frac{d}{c}\right)$ is a real character mod c and $\epsilon_d = \begin{cases} 1 & \text{if } d \equiv 1 \pmod 4, \\ i & \text{if } d \equiv 3 \pmod 4. \end{cases}$

Consider the real quadratic field $\mathbb{Q}(\sqrt{d})$, where $d > 0$. Takhtajan and Vinogradov considered the following inner product

$$\langle P_1(*,s), \overline{\theta(*)}\theta(*d)\,\mathrm{Im}(*)^{\frac{1}{2}} \rangle = \int_{\Gamma_0(4)\backslash\mathfrak{h}} P_1(z,s)\overline{\theta(z)}\theta(dz) y^{\frac{1}{2}} \frac{dxdy}{y^2}. \tag{1}$$

Note that $\theta(z)\overline{\theta(dz)}y^{\frac{1}{2}}$ is invariant under the action of $\Gamma_0(4d^2)$. Here

$$P_1(z,s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(4d^2)} \mathrm{Im}(\gamma z)^s e^{2\pi i \gamma z}.$$

Let's first do the geometric computation to compute (1) using Fourier expansions.

$$\begin{aligned}
\langle \ , \ \rangle &= \int_{\Gamma_0(4d^2)\backslash\mathfrak{h}} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(4d^2)} \mathrm{Im}(\gamma z)^s e^{2\pi i \gamma z} \cdot (\overline{\theta(z)}\theta(dz)y^{\frac{1}{2}}) \frac{dxdy}{y^2} \\
&= \int_{\Gamma_\infty\backslash\mathfrak{h}} y^s e^{2\pi i z} \overline{\theta(z)}\theta(dz) y^{\frac{1}{2}} \frac{dxdy}{y^2} \\
&= \int_{x=0}^{1}\int_{y=0}^{\infty} e^{2\pi i x} e^{-2\pi y} \left(\sum_{m=-\infty}^{\infty} e^{-2\pi i m^2 x} e^{-2\pi m^2 y}\right) \left(\sum_{n=-\infty}^{\infty} e^{2\pi i n^2 dx} e^{-2\pi n^2 dy}\right) y^{s-\frac{1}{2}} \frac{dxdy}{y}.
\end{aligned}$$

Note that the x-integral gives $m^2 - dn^2 = 1$, which is Pell's equation. Thus

$$\langle\ ,\ \rangle = \int_{y=0}^{\infty} \sum_{\substack{m,n \\ m^2-dn^2=1}} e^{-2\pi y(m^2+dn^2+1)} y^{s-\frac{1}{2}} \frac{dy}{y}$$
$$= \sum_{\substack{m,n \\ m^2-dn^2=1}} \frac{1}{(2\pi(m^2+dn^2+1))^{s-\frac{1}{2}}} \Gamma\left(s - \frac{1}{2}\right).$$

So we have proved:

Theorem 6.2.

$$\langle P_1(*,s), \overline{\theta}\theta_d \operatorname{Im}()^{\frac{1}{2}} \rangle = \frac{\Gamma\left(s-\frac{1}{2}\right)}{(2\pi)^{s-\frac{1}{2}}} \sum_{m^2-dn^2=1} \frac{1}{(m^2+dn^2+1)^{s-\frac{1}{2}}}.$$

We need the following old theorem about Pell's equation.

Theorem 6.3 (Pell's Equation)**.** *All solutions to Pell's equation $x^2 - dy^2 = 1$ are of the form $(\pm x, \pm y)$ where*

$$x + y\sqrt{d} = (x_0 + y_0\sqrt{d})^k$$

for some $k \geq 0$ and $x_0 + y_0\sqrt{d}$ is a fundamental unit of $\mathbb{Q}(\sqrt{d})$.

Thus $N(x_0 + y_0\sqrt{d}) = x_0^2 - y_0^2 d = 1$. We usually call $x_0 + y_0\sqrt{d} = \epsilon_d$. We have

$$\epsilon_d^{-1} = \frac{1}{x_0 + y_0\sqrt{d}} = \frac{x_0 - y_0\sqrt{d}}{x_0^2 - y_0^2 d} = x_0 - y_0\sqrt{d}$$

and so

$$\epsilon_d + \epsilon_d^{-1} = 2x_0,$$
$$\epsilon_d^k + \epsilon_d^{-k} = 2x.$$

Now we can express the inner product as

$$\frac{\Gamma(s-\frac{1}{2})}{(2\pi)^{s-\frac{1}{2}}} \sum_{m^2-dn^2=1} \frac{1}{(m^2+dn^2+1)^{s-\frac{1}{2}}} = \frac{\Gamma(s-\frac{1}{2})}{\pi^{s-\frac{1}{2}}} \left(\frac{1}{4} + \sum_{k=1}^{\infty} \frac{1}{(\epsilon_d^k + \epsilon_d^{-k})^{2s-1}}\right).$$

Let's look at

$$\sum_{k=1}^{\infty} \frac{1}{(\epsilon_d^k + \epsilon_d^{-k})^{2s-1}} = \sum_{k=1}^{\infty} \frac{1}{\epsilon_d^{k(2s-1)}} (1 + \epsilon_d^{-2k})^{1-2s} = \sum_{k=1}^{\infty} \left(\epsilon_d^{(1-2s)k} + O\left(\frac{2s-1}{\epsilon_d^{(2s+1)k}}\right)\right),$$

i.e.

$$\sum_{k=1}^{\infty} \frac{1}{(\epsilon_d^k + \epsilon_d^{-k})^{2s-1}} = \sum_{k=1}^{\infty} \frac{1}{\epsilon_d^{k(2s-1)}} + h(s)$$

where $h(s)$ is holomorphic for $\operatorname{Re}(s) > 0$, and

$$\sum_{k=1}^{\infty} \frac{1}{\epsilon_d^{k(2s-1)}} = \frac{1}{1 - \epsilon_d^{2s-1}}$$

has a simple pole at $s = \frac{1}{2}$ with residue equal to $(*) \log \epsilon_d$.

Theorem 6.4. *The inner product $\langle P_1(*,s), \overline{\theta}\theta_d \operatorname{Im}()^{\frac{1}{2}}\rangle$ has a double pole at $s = \frac{1}{2}$ with residue $(*) \log \epsilon_d$ (a constant multiple of the regular of $\mathbb{Q}(\sqrt{d})$).*

This is the geometric side of the trace formula. Now let's look at the spectral side of the Takhtajan-Vinogradov trace formula.

When we look at the spectral side only Maass forms with eigenvalues $\frac{1}{4}$ will produce double poles at $s = \frac{1}{2}$. What are special about these? They are the cohomological forms on GL(2); in the theory of automorphic forms, only cohomological forms can be associated to Galois representations.

Let $\eta_1(z), \eta_2(z), \cdots$ be a basis of Maass forms for $\Gamma_0(4d^2)$. We have the expansions

$$\eta_j(z) = \sum_{l \neq 0} A_j(l)\sqrt{y}K_{it_j}(2\pi|l|y)e^{2\pi i l x}.$$

We will conisder the spectral expansion of $P_1(z,s)$

$$P_1(z,s) = \sum_{j=1}^{\infty} \langle P_1(*,s), \eta_j\rangle \frac{\eta_j(z)}{\langle \eta_j, \eta_j\rangle} + \frac{1}{4\pi i}\sum_{l=1}^{r}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \langle P_1, E_l\rangle E_l(z,s)ds.$$

Then

$$\begin{aligned}\langle P_j(*,s), \eta_j\rangle &= \int_0^1\int_0^{\infty} y^s e^{2\pi i z}\overline{\eta_j(z)}\frac{dxdy}{y^2} \\ &= A_j(1)\int_0^{\infty} y^s e^{-2\pi y}\sqrt{y}K_{it_j}(2\pi y)\frac{dy}{y^2}.\end{aligned}$$

This integral can be found in the Russian book *Table of Integrals, Series and Products*, and is equal to

$$A_j(1)\cdot(*)\cdot\Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right).$$

When does $\Gamma\left(\frac{s-\frac{1}{2}+it_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-it_j}{2}\right)$ have a double pole at $s = \frac{1}{2}$? Only when $t_j = 0$. Remember that

$$\Delta\eta_j = \left(\frac{1}{4} + t_j^2\right)\eta_j,$$

so $t_j = 0$ if and only if η_j has eigenvalue $\frac{1}{4}$.

On the other hand, the double pole contribution $s = \frac{1}{2}$ from the Eisenstein series is 0.

We take residue of the double pole at $s = \frac{1}{2}$ on both sides of the Takhtajan-Vinogradov trace formula leading to identities of the following type. Let $f_1(z), f_2(z), \cdots, f_n(z)$ be cohomological Maass forms with eigenvalue $\frac{1}{4}$ for $\Gamma_0(4d^2)$ (by cohomological I mean they are cusp forms which are eigenforms for all the Hecke operators, and have eigenvalue $\frac{1}{4}$). Consider the Dirichlet L-functions

$$L(s,\chi_j) = \sum_{m=1}^{\infty}\frac{\chi_j(m)}{m^s}$$

associated with characters of the ideal class group of $\mathbb{Q}(\sqrt{d})$. Let h be the class number of $\mathbb{Q}(\sqrt{d})$. Then

$$\sum_{j=1}^{h-1} L(1,\chi_j) = \frac{\sqrt{d}L(1,\chi_d)}{2}\sum_{j=1}^{n} A_{f_j}(1)(*)$$

where the left side is summed over the non-trivial characters, χ_d is the Dirichlet character (so that $\zeta_{\mathbb{Q}(\sqrt{d})}(s) = \zeta(s)L(s,\chi_d)$), and $(*)$ is some simple factor (but hard to define).

n should be equal to h, but Takhtajan and Vinogradov could only prove $n \geq h$.

In all trace formulae I have seen, both sides involve infinite sums. I asked if there exists a trace formula with a finite sum on the geometric side, and found one which I called the degenerate trace formula. We will talk about this next time.

7. Lecture 7 (October 1, 2013)

7.1. **Theta Functions.** Today we will study theta functions as examples of automorphic forms. In his paper *Indefinite quadratische formen and funktionentheorie* in Math. Ann., Siegel (1951, 1952) constructed theta functions, generalizing Maass' earlier construction. He proved things like

$$\theta\left(\frac{az+b}{cz+d}\right) = \epsilon\psi\begin{pmatrix} a & b \\ c & d\end{pmatrix}(cz+d)^{\frac{1}{2}}\theta(z)$$

where ϵ is some 8-th root of unity. The problem of solving for ϵ explicitly was solved by Stark (1982) in *On the transformation formula for the symplectic theta functions and applications* in J. Fac. Sci. Tokyo, but this was only for the symplectic case. If you look at the binary quadratic form $ax^2+bxy+cy^2$ (whose discriminant is $D = b^2-4ac$), $D<0$ is the positive definite case, with $ax^2+bxy+cy^2$ always positive. We can ask how often $ax^2+bxy+cy^2$ equals a fixed integer, which happens finitely many times. But this may be infinitely often in the indefinite case $D>0$.

For the simple theta function $\sum e^{2\pi i n^2 z}$, Shimura was able to compute the 8-th root of unity. Stark asked his student Friedberg to compute this for theta functions for indefinite forms, and the result was published in *On theta functions associated to indefinite quadratic forms* in J. Number Theory. I'm not going to give proofs but I will just give a review of these papers.

Let's begin with the symplectic theta functions.

7.2. **Symplectic Theta Functions.** Let $\mathrm{Sp}_n(\mathbb{R})$ be the symplectic group consisting of $2n\times 2n$ matrices $M = \begin{pmatrix} A & B \\ C & D\end{pmatrix}$ satisfying the identity

$$^{t}MJM = J = \begin{pmatrix} 0 & -I_n \\ I_n & 0\end{pmatrix}$$

where I_n is the $n\times n$ identity matrix. Hecke and Maass developed the theory of automorphic forms on the upper half plane. Siegel generalized this to higher dimensions.

The Siegel upper half plane $\mathfrak{h}_n$ is the set of $n\times n$ symmetric matrices Z with $\mathrm{Im}(Z)>0$.

The discrete subgroup $\Gamma^{(n)}$ is defined to be $\mathrm{Sp}_n(\mathbb{Z})$, which acts on the Siegel upper half plane: for $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma^{(n)}$ and $Z \in \mathfrak{h}_n$, the action is given by

$$MZ = (AZ+B)(CZ+D)^{-1}.$$

Now we will construct the theta function. We need a special subgroup (even in GL(2), we need to get to level 4 to get theta functions). We define the theta subgroup $\Gamma_\theta^{(n)} \subset \Gamma^{(n)}$ to be the set of all $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma^{(n)}(\mathbb{Z})$ such that A^tB and C^tD have even diagonal entries.

Definition 7.1 (Siegel Theta)**.** Let u, v be column vectors in $\mathbb{C}^n$. Let $Z \in \mathfrak{h}_n$. We define

$$\theta\left(Z, \begin{pmatrix} u \\ v \end{pmatrix}\right) := \sum_{m \in \mathbb{Z}^n} e^{2\pi i[{}^t(m+v)Z(m+v) - 2{}^tmu - {}^tvu]}.$$

We have the following

Theorem 7.2 (Eichler)**.** *For $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_\theta^{(n)}$,*

$$\theta\left(MZ, M\begin{pmatrix} u \\ v \end{pmatrix}\right) = \chi(M)(\det(CZ+D))^{\frac{1}{2}}\theta\left(Z, \begin{pmatrix} u \\ v \end{pmatrix}\right)$$

for some $\chi(M)$ an 8-th root of unity.

χ was determined by Stark. The formula for χ is very complicated. It was quite a technical feat to get that 8-th root of unity.

The fact that θ is automorphic basically comes from the Poisson summation formula.

7.3. **Theta Functions associated to Indefinite Quadratic Forms.** The main problem with the indefinite case is that we cannot write down a series like above, since the expression does not converge. Siegel showed that indefinite forms are connected to certain definite forms, via "majorants".

Let K be a totally real algebraic number field of degree r_1, $\mathcal{O}_K$ be the ring of integers of K, δ_K be the different, and D_K be the discriminant of K. For $\alpha \in K$, let $\alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(r_1)}$ be the conjugates. We put $\Gamma = \mathrm{SL}_2(\mathcal{O}_K)$. For an integral ideal $\mathfrak{m}$, let

$$\Gamma_0(\mathfrak{m}) = \left\{ M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} : M \in \Gamma, \gamma \in \mathfrak{m} \right\}.$$

The upper half plane is $\mathfrak{h}^{r_1}$, which is just r_1 copies of the standard $\mathfrak{h} = \{x+iy : x \in \mathbb{R}, y > 0\}$. Let $z = (z_1, z_2, \cdots, z_{r_1}) \in \mathfrak{h}^{r_1}$. The action of $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ on $\mathfrak{h}^{r_1}$ is given by

$$M \circ z := (M^{(1)} \circ z_1, M^{(2)} \circ z_2, \cdots, M^{(r_1)} \circ z_{r_1})$$

where $M^{(j)} = \begin{pmatrix} \alpha^{(j)} & \beta^{(j)} \\ \gamma^{(j)} & \delta^{(j)} \end{pmatrix}$ and

$$M^{(j)} \circ z_j = (\alpha^{(j)}z_j + \beta^{(j)})(\gamma^{(j)}z_j + \delta^{(j)})^{-1}$$

is the usual GL(2)-action.

Let Q be a symmetric $n \times n$ matrix defining the quadratic form

$$Q[x] := {}^t x Q x$$

for $x \in \mathbb{R}^n$. If Q has entries in $\mathcal{O}_K$ and diagonal entries divisible by 2, we say Q is of level $N \in \mathcal{O}_K$ if

- NQ^{-1} has entries in $\mathcal{O}_K$, and 2 divides the diagonal entries of NQ^{-1};
- for $M \in \mathcal{O}_K$, $N \mid M$ whenever MQ^{-1} has entries in $\mathcal{O}_K$ and 2 divides the diagonal entries in MQ^{-1}.

Definition 7.3 (Signature)**.** $Q^{(j)}$ has signature (p,q) for $j = 1, 2, \cdots, r$ if there exists $L_j \in \mathrm{GL}(n, \mathbb{R})$ such that

$$Q^{(j)} = {}^t L_j \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} L_j.$$

We have now set up generalizations of congruence subgroups of level N. The problem is that this is an indefinite quadratic form, so we cannot construct a theta function in the usual way. We will first construct a symplectic theta function.

Definition 7.4 (Siegel's Majorant)**.** Let Q be a symmetric $n \times n$ matrix of signature (p,q). Let $R_j = {}^t L_j \cdot L_j$. Then R_j is a majorant for $Q^{(j)}$ if ${}^t R_j = R_j > 0$ and

$$R_j {Q^{(j)}}^{-1} R_j = Q^{(j)}.$$

Example 7.5. Let $Q = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix}$. Then $R = I_{p+q}$ is a majorant.

Majorants always exist but may not be unique, which gives lots of trouble.

Now I can define Siegel's theta function (indefinite case).

Let Q be an $n \times n$ symmetric matrix with entries in $\mathcal{O}_K$ such that 2 divides all diagonal entries and Q is of level $N \in \mathcal{O}_K$. Assume Q has signature (p,q). Let $(u_1, \cdots, u_{r_1})$ and $(v_1, \cdots, v_{r_1})$ be in $\mathbb{C}^{r_1}$. Let $\mathfrak{J} \in \mathcal{O}_K$ be an integral ideal. Let $z = (z_1, \cdots, z_{r_1}) \in \mathfrak{h}^{r_1}$, where $z_j = x_j + iy_j$. Then we define the theta function

$$\theta_Q\left(z, \begin{pmatrix} u \\ v \end{pmatrix}\right) = \left(\prod_{j=1}^{r_1} y_j\right)^{\frac{q}{2}} \sum_{\lambda \in \mathfrak{J}^n} e^{\pi i [\sum_{j=1}^{r_1} Q^{(j)}(\lambda^{(j)}+v_j)x_j + iR_j(\lambda^{(j)}+v_j)y_j - 2^t\lambda^{(j)}Q^{(j)}u - {}^t v_j Q^{(j)} u_j]}.$$

Remarkably it can be shown that $\theta_Q\left(z, \begin{pmatrix} u \\ v \end{pmatrix}\right)$ is a symplectic theta function. Friedberg determined the 8-th root of unity for this theta function in his thesis.

Theta functions have been generalized to the adelic setting.

Theta functions and modular forms of half integral weight are not covered by the Langlands program. It may be possible to use ideas from Mumford's book *Tata Lectures on Theta* to give a geometric interpretation of majorants.

7.4. **The simplest Theta Functions.** These will be weight $\frac{1}{2}$ modular forms for $\Gamma_0(4N)$. We are looking for holomorphic functions

$$f\left(\frac{az+b}{cz+d}\right) = \chi\begin{pmatrix} a & b \\ c & d \end{pmatrix}(cz+d)^{\frac{1}{2}} f(z)$$

for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$ and $\chi \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ some 8-th root of unity.

We can ask:

- Can we find a basis for the space? What is the dimension?
- Can theta functions be cusp forms?

The answers to all these questions were solved by Serre-Stark (1976). It is a surprising fact that theta functions can be cusp forms! The Ramanujan conjecture says that if $f(z) = \sum a_n e^{2\pi i n z}$ is a holomorphic cusp form of integral weight, then $|a_n| \ll n^{\frac{k-1}{2}+\epsilon}$. Here, the Ramanujan conjecture fails for θ but it holds "on average" by Piatetski-Shapiro. One of the goals of the Langlands conjectures is to prove the Ramanjan conjecture, which is another reason for why there isn't a Langlands conjecture for half-integral weights.

Next time I will talk about the Serre–Stark paper.

8. Lecture 8 (October 3, 2013)

8.1. **Modular forms of weight $\frac{1}{2}$.** Today I will discuss the Serre-Stark paper "*Modular forms of weight $\frac{1}{2}$*" in Lecture Notes in Mathematics 627, *Modular forms of one variable VI* (1976). I will not prove everything, but only the more important results.

Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4N)$ and $j(\gamma, z) = \epsilon_d \left(\frac{d}{c}\right)(cz+d)^{\frac{1}{2}}$ be the weight $\frac{1}{2}$ cocycle, where $\left(\frac{d}{c}\right)$ is the quadratic symbol. We define the action

$$(f|\gamma)(z) := j(\gamma, z)^{-1} f(\gamma z).$$

Let $\chi : (\mathbb{Z}/4N\mathbb{Z})^\times \to \mathbb{C}^\times$ be a Dirichlet character. We let $M_0(4N, \frac{1}{2}, \chi)$ be the space of weight $\frac{1}{2}$ modular forms for $\Gamma_0(4N)$ with character χ, i.e. functions satisfying

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

Let $M_1(4N, \frac{1}{2}) = \bigoplus_{\chi \pmod{4N}} M_0(4N, \frac{1}{2}, \chi)$. If $f \in M_1$, then $(f|\gamma)(z) = f(z)$ for all $\gamma \in \Gamma_1(4N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{4N} \right\}$. (It is because of this relation that we just need to study Γ_0 instead of Γ_1.)

The first theorem of Serre–Stark is as follows:

Theorem 8.1. *A basis for $M_1(4N, \frac{1}{2})$ is given by the theta functions*

$$\theta(\psi, tz) = \sum_{n \in \mathbb{Z}} \psi(n) e^{2\pi i n^2 t z}$$

where

1. $\psi : (\mathbb{Z}/4N\mathbb{Z})^\times \to \mathbb{C}^\times$ *is an even character of conductor* r_ψ*;*
2. $r_\psi^2 t \mid N$*;*
3. *if* $\theta \in M_0(4N, \frac{1}{2}, \chi)$*, then* $\chi(n) = \psi(n)\left(\frac{t}{n}\right)$ *for all* $(n, 4N) = 1$*.*

This gives a very explicit basis. The simplest proof to this theorem is by Hecke operators, which can be found in the appendix to the Serre–Stark paper by Deligne.

We are led to the interesting question: can any of the theta functions in Theorem 8.1 be cusp forms? The answer is yes.

Definition 8.2. We say an even character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ is *not totally even* if in the decomposition $N = \prod_{i=1}^r p_i^{e_i}$, $\chi = \prod_{i=1}^r \chi_{p_i^{e_i}}$, one of $\chi_{p_i^{e_i}}$ is not even, i.e. $\chi_{p_i^{e_i}}(-1) = -1$.

Theorem 8.3. *Let $\chi : (\mathbb{Z}/4N\mathbb{Z})^\times \to \mathbb{C}^\times$ be even. Then a basis for the cusp forms $S(\Gamma_0(4N), \frac{1}{2}, \chi)$ is the set of theta functions*

$$\theta(\psi, tz) = \sum_n \psi(n) e^{2\pi i n^2 tz}$$

where

1. *ψ is an even character of conductor r_ψ and ψ is not totally even;*
2. *$r_\psi^2 t \mid N$;*
3. *$\chi(n) = \psi(n)\left(\frac{t}{n}\right)$ for all $(n, 4N) = 1$.*

We can ask the question: what is the first (i.e. lowest level $4N$) theta function that is a cusp form? It can be written down very explicitly. Let $4N = 576 = 24^2$ and $\psi_{12} : (\mathbb{Z}/12\mathbb{Z})^\times \to \mathbb{C}^\times$ be given by $\psi_{12}(1) = \psi_{12}(11) = 1$ and $\psi_{12}(5) = \psi_{12}(7) = -1$. Then the theta function is given by

$$\theta(\psi, z) = \sum_{n \neq 0} \psi_{12}(n) e^{2\pi i n^2 z} = e^{2\pi i z} - e^{2\pi i \cdot 25 z} - e^{2\pi i \cdot 49 z} + e^{2\pi i \cdot 121 z} + \cdots .$$

Theorem 8.1 implies we can twist by a character $\chi \pmod q$ and get a cusp form

$$\theta(\psi_{12} \cdot \chi, tz) = \sum_{n \neq 0} \psi_{12}(n) \chi^2(n) e^{2\pi i n^2 tz}$$

of level $576tq^2$ with character χ. This will satisfy

$$\theta\left(\psi_{12} \cdot \chi^2, t \cdot \frac{az+b}{cz+d}\right) = \chi(d)\epsilon_d \left(\frac{d}{c}\right)(cz+d)^{\frac{1}{2}} \theta(\psi_{12}\chi^2, tz)$$

for all $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4Ntq^2)$. These are all the cusp forms of weight $\frac{1}{2}$, so we can think of θ as the analog of the Ramanujan cusp form of weight 12.

I will give a complete proof of Theorem 8.3.

8.2. **Proof of Theorem 8.3.** Let $f(z) \in M_1(N, \frac{1}{2})$ (we will assume $4 \mid N$ throughout) have Fourier expansion

$$f(z) = \sum_{l=0}^{\infty} a(l) e^{2\pi i l z}.$$

For $M \geq 1$, let $\epsilon_M : \mathbb{Z}/M\mathbb{Z} \to \mathbb{C}^\times$ be a complex-valued periodic function with period M, i.e. $\epsilon_M(z + M) = \epsilon_M(z)$.

Definition 8.4. The twist of f by ϵ_M is

$$(f * \epsilon_M)(z) := \sum_{l=0}^{\infty} a(l) \epsilon_M(l) e^{2\pi i l z}.$$

We remark that θ is not a twist of the classical theta function $\sum e^{2\pi i n^2 z}$ by ψ_{12} — if it were, the coefficient of $e^{2\pi i n^2 z}$ would have to be $\psi_{12}(n^2)$ instead of $\psi_{12}(n)$. This makes an important difference!

The finite Fourier transform is given by

$$\hat{\epsilon}_M(l) = \frac{1}{M} \sum_{m \in \mathbb{Z}/M\mathbb{Z}} \epsilon_M(m) e^{-\frac{2\pi i m l}{M}}$$

with inverse transform

$$\epsilon_M(l) = \sum_{m \in \mathbb{Z}/M\mathbb{Z}} \hat{\epsilon}_M(m) e^{\frac{2\pi i m l}{M}}.$$

I claim that we have

$$(f * \epsilon_M)(z) = \sum_{m \in \mathbb{Z}/M\mathbb{Z}} \hat{\epsilon}_M(m) f\left(z + \frac{m}{M}\right).$$

This is very easy to prove; you should be able to do it mentally.

Proposition 8.5. *Fix $M \geq 1$, $M \mid N$. The following are equivalent:*

(1) *f vanishes at all cusps $\frac{m}{M}$, where $1 \leq m \leq M$ and $(m, M) = 1$.*
(2) *For every $\epsilon_M : \mathbb{Z}/M\mathbb{Z} \to \mathbb{C}^\times$, the function*

$$\phi_{f * \epsilon_M}(s) := \sum_{l=1}^{\infty} a(l) \epsilon_M(l) l^{-s}$$

is holomorphic at $s = \frac{1}{2}$.

Let us recall what cusps are. Every cusp of $\Gamma_0(N)$ is of the form $\frac{m}{M}$ with $M \mid N$, $(m, M) = 1$. Two cusps $\frac{m}{M}, \frac{m'}{M'}$ are equivalent if and only if $M = M'$ and $m \equiv m' \pmod{\left(M, \frac{N}{M}\right)}$. For example, 0 and 1 are equivalent under $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Proof. We first prove this for $M = 1$. (1) is equivalent to the statement that f vanishes at the cusp 0. (2) is equivalent to the statement that $\phi_f(s)$ vanishes at the cusp 0, since ϕ_f is just the standard Mellin transform of f.

Let $w_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$. Then

$$g(z) := (f|w_N)(z) = N^{-\frac{1}{4}} (-iz)^{-\frac{1}{2}} f\left(-\frac{1}{Nz}\right)$$

has a Fourier expansion

$$g(z) = \sum_{n=0}^{\infty} b(n) e^{2\pi i n z}.$$

The cusp 0 is equivalent to ∞ under w_N, so (1) is equivalent to the statement that $g(z)$ vanishes at ∞, i.e. $b(0) = 0$.

On the other hand, the functional equation relating $\phi_f(s)$ and $\phi_g(\frac{1}{2} - s)$ implies that (2) is equivalent to saying $(2\pi)^{-s}\Gamma(s)\phi_g(s)$ is holomorphic at $s = 0$, i.e. $\phi_g(0) = 0$.

So we have proved the Proposition when $M = 1$.

Next assume $M > 1$. Apply the above ideas to $f * \epsilon_M$ with N replaced by $N \cdot M^2$ (it is not hard to show $f * \epsilon_M$ is automorphic of level NM^2). Then conditions (1) and (2) are equivalent to:

(3) For all $\epsilon_M : \mathbb{Z}/M\mathbb{Z} \to \mathbb{C}^\times$, the modular form $f * \epsilon_M$ vanishes at the cusp 0.
(4) For all $m \in \mathbb{Z}/M\mathbb{Z}$, the modular form $f(z + \frac{m}{M})$ vanishes at the cusp 0.

This finishes the proof. □

By this proposition, we have

Corollary 8.6. *The following are equivalent:*

(1) *f is a cusp form.*
(2) *For all $\epsilon_M : \mathbb{Z}/M\mathbb{Z} \to \mathbb{C}^\times$, the function $\phi_{f*\epsilon_M}(s) = \sum_{l=1}^{\infty} a(l)\epsilon_M(l)l^{-s}$ is holomorphic at $s = \frac{1}{2}$.*

There is one last step in the proof — we need to use the condition of an even character being not totally even. Finally, we have the following

Proposition 8.7. *Let ψ be an even character which is not totally even. Then $\theta(\psi, z)$ is a cusp form.*

Proof. Let $\epsilon_M : \mathbb{Z}/M\mathbb{Z} \to \mathbb{C}^\times$. We must show that

$$H(s) = \sum_{n=1}^{\infty} \epsilon_M(n^2)\psi(n)n^{-2s}$$

is holomorphic at $s = \frac{1}{2}$. But this function is equal to

$$H(s) = \sum_{m \in \mathbb{Z}/M\mathbb{Z}} \epsilon_M(m^2)\psi(m) \sum_{\substack{n \equiv m \pmod M \\ n \geq 1}} n^{-2s}.$$

Note that the function $\sum_{\substack{n \equiv m \pmod M \\ n \geq 1}} n^{-2s}$ has a simple pole at $s = \frac{1}{2}$ with residue $\frac{1}{M}$. Thus $H(s)$ has a simple pole at $s = \frac{1}{2}$ with residue

$$\frac{1}{M} \sum_{m \in \mathbb{Z}/M\mathbb{Z}} \epsilon_M(m^2)\psi(m).$$

We need to show that if ψ is even but not totally even, then the residue is equal to 0.

We know that since ψ is not totally even, there exists a prime $p \mid r(\psi)$ such that the p-th component of ψ is an odd character. Write $M = p^a \cdot M'$ where $(M', p) = 1$. Then

$$(\mathbb{Z}/M\mathbb{Z})^\times = (\mathbb{Z}/p^a\mathbb{Z})^\times \times (\mathbb{Z}/M'\mathbb{Z})^\times. \tag{2}$$

Let $x_p \in (\mathbb{Z}/M\mathbb{Z})^\times$ whose first component in (2) is -1 and second component is $+1$. Since x_p is invertible, the sum $\sum_m$ doesn't change if m is replaced by $x_p m$, so the sum in the residue is equal to

$$\sum_{m \in \mathbb{Z}/M\mathbb{Z}} \epsilon_M((x_p m)^2) \cdot \psi(x_p m) = \sum \epsilon_M(m^2)\psi(x_p m) = -\sum \epsilon_M(m^2)\psi(m)$$

which implies it is 0. □

9. Lecture 9 (October 8, 2013)

9.1. **Adelic Poincaré series on** $\mathrm{GL}(n)$**.** I will talk about some recent joint work with Michael Woodbury, where we look at Poincaré series from the adelic point of view. I will also move from $\mathrm{GL}(2)$ to $\mathrm{GL}(n)$. Basically this is the relative trace formula for the unipotent subgroup

$$\begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & * \\ & & & 1 \end{pmatrix}.$$

A lot of people have been studying the relative trace formula for the unitary group, but this is really different. On the unitary group, there are clever tricks to get rid of the continuous spectrum, but the continuous spectrum plays a crucial role for the unipotent group and cannot be avoided.

To keep things simple we will work over the adele ring $\mathbb{A} = \mathbb{A}_{\mathbb{Q}}$ over $\mathbb{Q}$, but there is no problem working over any number field. v will denote either ∞ or a prime p. $\mathbb{Q}_v$ is the completion of $\mathbb{Q}$ at v, i.e.

$$\mathbb{Q}_v = \begin{cases} \mathbb{Q}_p & \text{if } v = p = \text{prime}, \\ \mathbb{R} & \text{if } v = \infty. \end{cases}$$

We have the unipotent and diagonal subgroups

$$U = \begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & * \\ & & & 1 \end{pmatrix}, T = \begin{pmatrix} * & & & \\ & * & & \\ & & * & \\ & & & * \end{pmatrix}.$$

Let us introduce the maximal compact subgroups

$$K_v = \begin{cases} \mathrm{GL}(n, \mathbb{Z}_p) & \text{if } v = p, \\ \mathrm{O}_n(\mathbb{R}) & \text{if } v = \infty. \end{cases}$$

and put

$$K = \prod_v K_v.$$

Let Z be the center of $G = \mathrm{GL}(n)$.

The Iwasawa decomposition is $G = UTK = TUK$. For a proof, you can refer to my book.

For $H \subset G$, we want to construct functions $f : Z(H)(\mathbb{A})H(\mathbb{Q})\backslash H(\mathbb{A}) \to V$, where $Z(H)(\mathbb{A})$ is the center of H with elements in $\mathbb{A}$, and $V \cong \mathbb{C}^k$ is some finite-dimensional vector space over $\mathbb{C}$. If $k = 1$, we are basically looking at weight 0; I will explain more on that.

Every function of this type is invariant under the group $Z(H)(\mathbb{A})H(\mathbb{Q})$, i.e.

$$f(\gamma g) = f(g)$$

for all $\gamma \in H(\mathbb{Q})$, and

$$f(zg) = f(g)$$

for all $z \in Z(H)(\mathbb{A})$. If we just consider the case $k = 1$, then we have the Poincaré series, but we want to construct something more general. We could also have a non-trivial central character.

We will construct a Poincaré series as follows. Let $V = \prod_v V_v$, and $\sigma_v : K_v \to \mathrm{GL}(V_v)$ be a finite-dimensional representation of K_v. We assume σ_v is trivial except for finitely many v. Put $(\sigma, V) = \bigotimes_v (\sigma_v, V_v)$. We will have a central character

$$\chi = \bigotimes_v \chi_v : G(\mathbb{Q})\backslash G(\mathbb{A}) \to \mathbb{C}^\times,$$

i.e. χ is the character which extends the central character of the representation σ.

We shall assume there exists an inner product $\langle \ , \ \rangle_{V_v}$ on each V_v, and set

$$\langle \ , \ \rangle_V = \prod_v \langle \ , \ \rangle_{V_v}.$$

Further we assume (σ, V) is a unitary representation.

Now we need a character on the unipotent group U. Take $a = (a_1, a_2, \cdots, a_{n-1}) \in \mathbb{Q}^{n-1}$. We want to define a character $\psi_a : U \to \mathbb{C}$. Each $u \in U(\mathbb{A})$ locally looks like

$$u_v = \begin{pmatrix} 1 & u_{1,2} & & & \\ & 1 & u_{2,3} & & \\ & & \ddots & \ddots & \\ & & & 1 & u_{n-1,n} \\ & & & & 1 \end{pmatrix}$$

and we set

$$\psi_v(u_v) = \begin{cases} \exp\left(-2\pi i(a_1 u_{1,2} + a_2 u_{2,3} + \cdots + a_{n-1} u_{n-1,n})\right) & \text{if } v = p, \\ \exp\left(2\pi i(a_1 u_{1,2} + a_2 u_{2,3} + \cdots + a_{n-1} u_{n-1,n})\right) & \text{if } v = \infty. \end{cases}$$

It is easy to see that

$$\psi_v(u_v \cdot u_v') = \psi_v(u_v) \cdot \psi_v(u_v')$$

and so ψ_v gives a character. We define

$$\psi = \prod_v \psi_v.$$

By the Iwasawa decomposition $G = UTK$, we define ψ to satisfy

$$\psi_v(u_v \tau_v \kappa_v) = \psi_v(u_v)$$

for all $\tau_v \in T_v$ and $\kappa_v \in K_v$.

Next we need a toric function on

$$T = \left\{ \begin{pmatrix} t_1 & & & 0 \\ & t_2 & & \\ & & \ddots & \\ 0 & & & t_n \end{pmatrix} \right\}.$$

Definition 9.1 (Toric norm)**.** For $t_v = \begin{pmatrix} t_{v,1} & & & 0 \\ & t_{v,2} & & \\ & & \ddots & \\ 0 & & & t_{v,n} \end{pmatrix}$, we define

$$||t_v||_{\text{tor},v} = \prod_{i=1}^{n} \left| \frac{t_{v,i}}{(\det t_v)^{\frac{1}{n}}} \right|_v .$$

Note that $||zt_v||_{\text{tor},v} = ||t_v||_{\text{tor},v}$ for all z in the center. Then we extend $||\ ||_{\text{tor},v}$ with the Iwasawa decomposition

$$||ut\kappa||_{\text{tor},v} := ||t||_{\text{tor},v}$$

for all $u \in U_v$ and $\kappa \in K_v$. Finally we define the full toric norm

$$||\ ||_{\text{tor}} := \prod_v ||\ ||_{\text{tor},v}.$$

We can now construct an Eisenstein series

$$E(g, s, \chi) = \sum_{\gamma \in Z(\mathbb{Q})U(\mathbb{Q})\backslash G(\mathbb{Q})} \chi(\gamma g)||\gamma g||^s_{\text{tor}},$$

where $s = (s_1, s_2, \cdots, s_n) \in \mathbb{C}^n$ with $\text{Re}(s_i) \gg 1$, and

$$||g||^s_{\text{tor}} = \prod_v \prod_{i=1}^{n} \left| \frac{t_{v,i}}{\det(t_v)^{\frac{1}{n}}} \right|_v^{s_i}$$

for the Iwasawa decomposition $g = tu\kappa$. We will not study the Eisenstein series; we just want to construct the Poincaré series.

To construct a Poincaré series, we introduce one more function $H : \text{GL}(n, \mathbb{A}_\mathbb{Q}) \to V \cong \mathbb{C}^k$. We require

$$H(zut\kappa) = \sigma^{-1}(\kappa)H(t)$$

where z is central, u is unipotent, t is toric and κ is in the maximal compact. Thus H is really a function on T. Moreover, we require that H is factorizable, i.e.

$$H = \prod_v H_v.$$

We can think of κ as the analog of the weight of a modular form.

We now have the ingredients to define the Poincaré series. They were first defined in a paper written by Bump, Friedberg and me in the early 1980's, but we did it for totally ramified extensions only. Glenn Stevens immediately generalized it.

Definition 9.2 (Poincaré Series (Stevens, 1980's))**.**

$$P_{H,\psi_a,\sigma}(g, s) := \sum_{\gamma \in Z(\mathbb{Q})U(\mathbb{Q})\backslash G(\mathbb{Q})} \chi(\gamma g)\psi_a(\gamma g)H(\gamma g)||\gamma g||^s_{\text{tor}},$$

where $s = (s_1, s_2, \cdots, s_n) \in \mathbb{C}^n$ with $\text{Re}(s_i) \gg 1$, and H is chosen such that this converges absolutely (e.g. if H is compactly supported or absolutely bounded).

The Poincaré series satisfies the following properties:

- $P(\gamma g, s) = P(g, s)$ for all $g \in \text{GL}(n, \mathbb{A}_\mathbb{Q})$ and $\gamma \in G(\mathbb{Q})$;

- $P(t\kappa, s) = \sigma^{-1}(\kappa)P(t,s)$ (analog of weight k modular forms on GL(2));
- More generally, $P(gzk, s) = \chi(z)\sigma^{-1}(\kappa)P(g,s)$;
- If we choose H carefully (e.g. H compactly supported), then $P(g,s) \in \mathcal{L}^2$.

We want to obtain the trace formula by taking the inner product of two of these Poincaré series, using the spectral decomposition and Fourier expansions. The Fourier expansion will in particular show $P(g,s) \in \mathcal{L}^2$.

Let's look at the Hilbert space $\mathcal{L}^2(\mathrm{GL}(n,\mathbb{Q})\backslash \mathrm{GL}(n,\mathbb{A}_\mathbb{Q}), \sigma, \chi)$ consisting of all square-integrable functions $F : \mathrm{GL}(n,\mathbb{Q})\backslash \mathrm{GL}(n,\mathbb{A}_\mathbb{Q}) \to V \cong \mathbb{C}^k$ satisfying

$$F(gz\kappa) = \chi(z)\sigma^{-1}(\kappa)F(g),$$

with inner product given by

$$\langle F_1, F_2\rangle = \int_{\mathrm{GL}(n,\mathbb{Q})\backslash \mathrm{GL}(n,\mathbb{A}_\mathbb{Q})} \prod_v \langle F_1(g), F_2(g)\rangle_{V_v} dg.$$

Remark. For all but finitely many v, we have $k = 1$ and we can choose

$$\langle F_1(g), F_2(g)\rangle_{V_v} = F_1(g) \cdot \overline{F_2(g)}.$$

Theorem 9.3. *Let $\phi \in \mathcal{L}^2(G(\mathbb{Q})\backslash G(\mathbb{A}), \sigma, \chi)$ be an automorphic form whose Whittaker function, defined to be*

$$\int_{U(\mathbb{Q})\backslash U(\mathbb{A})} \phi(ug)\overline{\psi_a(u)}du,$$

is factorizable as $\prod_v W_v$, then

$$\langle P_{H,\psi_a,\sigma}(*, s), \phi\rangle = \prod_v I_v$$

is a product of toric integrals, where for $F_v : \mathbb{Q}_v \to V_v$,

$$I(F_v, W_v, v) = \int_{Z(\mathbb{Q}_v)\backslash T(\mathbb{Q}_v)} \langle F_v(t_v), W_v(t_v)\rangle_{V_v} ||t_v||^s_{tor,v} dt_v$$

We will prove this next time and compute the Fourier expansion.

10. Lecture 10 (October 10, 2013)

Unfortunately I was unable to attend the lecture.

11. Lecture 11 (October 17, 2013)

11.1. **Adelic Poincaré series on GL(n).** Recall that we are looking at $\mathrm{GL}(n,\mathbb{A}_\mathbb{Q})$. Each $a = (a_1, \cdots, a_{n-1}) \in \mathbb{A}^{n-1}$ defines a character on the unipotent group $\psi_a : U \to \mathbb{C}$ as follows. For

$$u = \begin{pmatrix} 1 & u_{1,2} & & & \\ & 1 & u_{2,3} & & \\ & & \ddots & \ddots & \\ & & & 1 & u_{n-1,n} \\ & & & & 1 \end{pmatrix},$$

we define $\psi_{a,v}(u) = \exp(\pm 2\pi i(a_{1,v}u_{1,2} + \cdots + a_{n-1,v}u_{n-1,n}))$ and $\psi_a = \prod \psi_{a,v}$.

The Poincaré series is

$$P_a(g,s) = \sum_{\gamma \in Z(\mathbb{Q})U_n(\mathbb{Q})\backslash \mathrm{GL}(n,\mathbb{Q})} \chi(\gamma g)\psi_a(\gamma g)H(\gamma g)||\gamma g||_{\mathrm{tor}}^s$$

where $s = (s_1, \cdots, s_n) \in \mathbb{C}^n$. For $t = (t_1, \cdots, t_n) \in T$, we set $||t|| = \prod_{i=1}^{n} \left(\frac{t_i}{*}\right)^{s_i}$.

Last time we computed the Fourier expansion of the Poincaré series.

Theorem 11.1.

$$\int_{U(\mathbb{Q})\backslash U(\mathbb{A})} P_a(gu,s)\overline{\psi_b}(u)du = \sum_{w\in W} \sum_{\tau \in Z(\mathbb{Q})\backslash T(\mathbb{Q})} \prod_v \delta_v(a,b,w\tau)K_v(g_v,s,a,b,w,\tau)$$

where $\delta_v = \begin{cases} 1 & \text{if } \psi_{a,v}(w\tau u\tau^{-1}w^{-1}) = \psi_{b,v}(u), \\ 0 & \text{otherwise,} \end{cases}$ K_v *is the Kloosterman integral*

$$K_v(g_v,s,a,b,w,\tau) = \int_{\overline{U}_w(\mathbb{Q})} \chi_v(w\tau u g_v)\psi_{a,v}(w\tau u g_v)H_v(w\tau u g_v)||w\tau u g_v||^s du,$$

and $U = U_w \cdot \overline{U}_w$*, with* $U_w = (w^{-1}Uw) \cap U$*,* $\overline{U}_w = (w^{-1t}Uw) \cap U$*.*

In the 3 by 3 case, we have

$$U_w = \begin{pmatrix} 1 & 0 & * \\ & 1 & * \\ & & 1 \end{pmatrix}, \overline{U}_w = \begin{pmatrix} 1 & * & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}.$$

Today we will study the Kloosterman integrals at $v = p$.

Theorem 11.2.

$$K_p(t,s,a,b,w,\tau) = \sum_{t_1 \in T(\mathbb{Z}_p)\backslash T(\mathbb{Q}_p)} \sum_{\substack{u \in U(\mathbb{Z}_p)\backslash U(\mathbb{Q}_p) \\ u' \in \overline{U}_w(\mathbb{Z}_p)\backslash \overline{U}_w(\mathbb{Q}_p) \\ ut_1^{-1}w\tau t u' \in \mathrm{GL}(n,\mathbb{Z}_p)}} \overline{\psi_{a,p}(t_1 u t_1^{-1})}\psi_{b,p}(tu't^{-1})\sigma_p^{-1}(ut_1 w\tau t u'\kappa)H_p(t_1)||t_1||_{tor,p}^s$$

Let me first assume this and show that it reduces to the usual Kloosterman sum in the case of GL(2).

Let $u(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ and $t(y,y') = \begin{pmatrix} y & 0 \\ 0 & y' \end{pmatrix} \in T$. We have the classical Kloosterman sum

$$S(a,b,n) = \sum_{x \in (\mathbb{Z}/n\mathbb{Z})^\times} e^{2\pi i\left(\frac{ax+bx^{-1}}{n}\right)}.$$

It satisfies the following well-known multiplicative property, which we will not prove.

Proposition 11.3.

$$S(a,b,mn) = S(a\overline{m}, b\overline{m}, n) \cdot S(a\overline{n}, b\overline{n}, m)$$

where $(m,n) = 1$*,* $n \cdot \overline{n} \equiv 1 \pmod{m}$ *and* $m \cdot \overline{m} \equiv 1 \pmod{n}$*.*

The Weyl group is $W = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$. Denoting $e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, we have

$$U_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} = U, \quad \overline{U}_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} = \{e\},$$
$$U_{\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)} = \{e\}, \quad \overline{U}_{\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)} = U.$$

The key proposition is

Proposition 11.4. *Suppose* $\tau = t(n,m) = \begin{pmatrix} n & 0 \\ 0 & m \end{pmatrix}$ *and* $H_p(t) = \begin{cases} 1 & \text{if } t \in \mathbb{Z}_p, \\ 0 & \text{if } t \notin \mathbb{Z}_p. \end{cases}$ *Then*

$$K_p \left(e, \tau, a, b, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, s \right) = 0$$

unless $v_p(n) = -v_p(m) \geq 0$, *in which case*

$$K_p \left(e, \tau, a, b, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, s \right) = S(-a(n^*)^{-1}, bm^*, p^{v_p(n)})$$

where $n = n^* p^{v_p(n)}$ *and* $m = m^* p^{v_p(m)}$.

Since H_p is supported on $\mathbb{Z}_p$ and σ is trivial, the theorem gives

$$K_p = \sum_{\substack{x,x' \in \mathbb{Z}_p \backslash \mathbb{Q}_p \\ u(x) w\tau u(x') \in \mathrm{GL}(2,\mathbb{Z}_p)}} \overline{\psi_{a,p}(u(x))} \psi_{b,p}(u(x')).$$

If $u(x)w\tau u(x') \in \mathrm{GL}(2,\mathbb{Z}_p)$, then $v_p(n) = l$ and $-v_p(m) \geq 0$, so $x = \alpha p^l$ and $x' = d - m\alpha - 1$.

Again we will not go through the details, but it is possible to prove the following

Proposition 11.5. *Let* $\tau = \begin{pmatrix} n & \\ & n^{-1} \end{pmatrix}$ *where* $n \in \mathbb{Z}$, σ *be trivial, and* $H_p(t) = \begin{cases} 1 & \text{if } t \in \mathbb{Z}_p, \\ 0 & \text{otherwise.} \end{cases}$ *Then*

$$\prod_p K_p = S(a,b,n).$$

The proof uses the multiplicative property of the Kloosterman sums.

Proof of Theorem 11.2. Let us prove the formula for K_p in general. Last time we showed that

$$K_p(g,s,a,b,w,\tau) = \int_{\overline{U}_w(\mathbb{Q}_p)} \chi_p(u\tau u' g) \psi_{a,p}(w\tau u' g) H_p(w\tau u' g) ||w\tau u' g||^s_{\mathrm{tor}} \overline{\psi_{b,p}(u')} du'.$$

Let $g = zutk$. Then it is easy to see that

$$K_p(zutk, \cdots) = \chi_p(z) \psi_{b,p}(u) \sigma^{-1}(k) K_p(t, \cdots).$$

Since the normalization is such that $\int_{U(\mathbb{Z}_p)} du = 1$, it follows that

$$K_p(t,s,a,b,w,\tau) = \int_{U(\mathbb{Z}_p)\backslash \overline{U}_w(\mathbb{Q}_p)} \chi_p(w\tau u' t) \psi_{a,p}(w\tau u' t) \sigma^{-1}(k) H_p(w\tau u' t) ||w\tau u' t||^s_{\mathrm{tor},p} \overline{\psi_{b,p}(u')} du'.$$

If we make the change of variables $u' \mapsto tu't^{-1}$, then this becomes

$$\int_{U(\mathbb{Z}_p)\backslash\overline{U}_w(\mathbb{Q}_p)} \chi_p(w\tau tu')\psi_{a,p}(w\tau tu')\sigma^{-1}(k)H_p(w\tau tu')||w\tau tu'||^s_{\mathrm{tor},p}\overline{\psi_{b,p}(tu't^{-1})}du'$$
$$= \sum_{u'\in U(\mathbb{Z}_p)\backslash\overline{U}_w(\mathbb{Q}_p)} \chi_p(w\tau tu')\psi_{a,p}(w\tau tu')\sigma^{-1}(k)H_p(w\tau tu')||w\tau tu'||^s_{\mathrm{tor},p}\overline{\psi_{b,p}(tu't^{-1})}.$$

Next we use the Iwasawa decomposition. If $u' \in \overline{U}_s(\mathbb{Q}_p)$ and $\tau, t \in \mathbb{Z}_p\backslash T(\mathbb{Q}_p)$, then there exists $t_1 \in \mathbb{Z}_p\backslash T(\mathbb{Q}_p)$ such that

$$ut_1^{-1}w\tau tu' \in \mathrm{GL}(n, \mathbb{Z}_p).$$

Thus the sum is equal to

$$\sum_{t_1\in T(\mathbb{Z}_p)\backslash T(\mathbb{Q}_p)} \sum_{\substack{u\in U(\mathbb{Z}_p)\backslash U(\mathbb{Q}_p)\\ u'\in\overline{U}_w(\mathbb{Z}_p)\backslash\overline{U}_w(\mathbb{Q}_p)\\ ut_1^{-1}w\tau tu'\in\mathrm{GL}(n,\mathbb{Z}_p)}} \chi_p(w\tau tu')\psi_{a,p}(w\tau tu')\sigma_p^{-1}(k)H_p(w\tau tu')||w\tau tu'||^s_{\mathrm{tor},p}\overline{\psi_{b,p}(tu't^{-1})}$$

$$= \sum_{t_1\in T(\mathbb{Z}_p)\backslash T(\mathbb{Q}_p)} \sum_{\substack{u\in U(\mathbb{Z}_p)\backslash U(\mathbb{Q}_p)\\ u'\in\overline{U}_w(\mathbb{Z}_p)\backslash\overline{U}_w(\mathbb{Q}_p)\\ ut_1^{-1}w\tau tu'\in\mathrm{GL}(n,\mathbb{Z}_p)}} \begin{array}{l}\chi_p(t_1ut_1^{-1}w\tau tu')\overline{\psi_{a,p}(t_1ut_1^{-1})}\psi_{a,p}(t_1ut_1^{-1}w\tau tu')\sigma_p^{-1}(k)\\ \cdot H_p(t_1ut_1^{-1}w\tau tu')||t_1ut_1^{-1}w\tau tu'||^s_{\mathrm{tor},p}\overline{\psi_{b,p}(tu't^{-1})}\end{array}$$

$$= \sum_{t_1\in T(\mathbb{Z}_p)\backslash T(\mathbb{Q}_p)} \sum_{\substack{u\in U(\mathbb{Z}_p)\backslash U(\mathbb{Q}_p)\\ u'\in\overline{U}_w(\mathbb{Z}_p)\backslash\overline{U}_w(\mathbb{Q}_p)\\ ut_1^{-1}w\tau tu'\in\mathrm{GL}(n,\mathbb{Z}_p)}} \overline{\psi_{a,p}(t_1ut_1^{-1})}\cdot\overline{\psi_{b,p}(tut^{-1})}\sigma_p^{-1}(ut^{-1}w\tau tu'k)H_p(t_1)||t_1||^s_{\mathrm{tor},p}.$$

□

There is an analogous theorem at the archimedean places:

$$K_\infty(zutk, s, a, b, w, \tau) = \chi_\infty(z)\psi_{b,\infty}(u)\int_{\overline{U}_w(\mathbb{R})} \chi_\infty(t_1)\psi_{a,\infty}(u^{-1})\overline{\psi_{b,\infty}(tu't^{-1})}\sigma(k^{-1})H_\infty(t_1)||t_1||^s_{\mathrm{tor},\infty}du'.$$

On $\mathrm{GL}(n)$ where $n \geq 3$, the K_p are hyper-Kloosterman sums of the type studied by Deligne. Using the Riemann hypothesis for algebraic varieties over finite fields (proved by Deligne), it is possible to get very sharp bounds for K_p. The first person to get these bounds was M. Larsen, and they were included as the appendix of a paper I wrote with Bump and Friedberg in which we worked out the archimedean case (Acta Arith., 1988).

Next time I will do the relative trace formula. On one side, we have these Kloosterman integrals. On the other side, there is some spectral information.

12. Lecture 12 (October 22, 2013)

Unfortunately I was unable to attend the lecture.

13. Lecture 13 (October 24, 2013)

13.1. **Selberg–Arthur Trace Formula.** Today I will talk about the Selberg–Arthur trace formula. We will first do it for GL(2) by classical methods (using the upper half plane model), and then I will move on to adelic representations.

Let me first give an overview. Let G be a reductive group acting on a topological space X. We assume that X has further properties so that we can do integration on this space. We are looking at

$$\mathcal{L}^2(G\backslash X) = \left\{ f : X \to \mathbb{C} : \int_{G\backslash X} |f(x)|^2 dx < \infty \right\}$$

which we assume to be a Hilbert space, with inner product given by

$$\langle F_1, F_2 \rangle = \int_{G\backslash X} F_1(x)\overline{F_2(x)}dx$$

for $F_1, F_2 \in \mathcal{L}^2(G\backslash X)$.

We want to use spectral theory, i.e. the study of eigenfunctions of certain operators. The natural choice is to use differential operators, but they are usually unbounded. Selberg's idea was to consider integral operators. Let $k : X \times X \to \mathbb{C}$ be a kernel function, satisfying:

- $k(gx, gy) = k(x, y)$ for all $g \in G$, $x, y \in X$;
- $\int_X \int_X |k(x, y)|^2 dxdy < \infty$ (Hilbert–Schmidt property).

Let us assume there is such a function. Then we get an integral operator.

Definition 13.1 (Integral operator with kernel $k(x, y)$)**.** Let $f \in \mathcal{L}^2(G\backslash X)$. We define the integral operator

$$Kf(y) := \int_{G\backslash X} k(x, y)f(x)dx.$$

We just need to prove one thing:

Proposition 13.2. *$Kf(gy) = Kf(y)$ for all $g \in G$.*

Proof. This is very easy. Since $k(x, gy) = k(g^{-1}x, g^{-1}gy) = k(g^{-1}x, y)$, we have

$$Kf(gy) = \int_{G\backslash X} k(x, gy)f(x)dx = \int_{G\backslash X} k(g^{-1}x, y)f(x)dx = \int_{G\backslash X} k(x, y)f(gx)dx = kf(y)$$

where we made the change of variables $x \mapsto gx$. Here dx is assumed to be an invariant measure. □

Because of the Hilbert–Schmidt property, we can deduce that $K : \mathcal{L}^2(G\backslash X) \to \mathcal{L}^2(G\backslash X)$ is a bounded integral operator. So by the spectral theorem, there exists an orthonormal basis $f_1, f_2, \cdots$ of $\mathcal{L}^2(G\backslash X)$ where $Kf_i = \lambda_i f_i$ for the eigenvalues $\lambda_i \in \mathbb{C}$.

If $k(x, y) = \overline{k(y, x)}$ for all $x, y \in X$, then K will be a self-adjoint operator and the eigenvalues λ_i are real. This is because

$$\begin{aligned}
\langle Kf_i, f_i \rangle &= \int_{G\backslash X} Kf_i(x)\overline{f_i(x)}dx \\
&= \int_{G\backslash X} \left(\int_{G\backslash X} k(y, x)f_i(y)dy \right) \overline{f_i(x)}dx \\
&= \int_{G\backslash X} \int_{G\backslash X} \overline{k(x, y)}f_i(y)\overline{f_i(x)}dxdy \\
&= \int_{G\backslash X} f_i(y)\overline{\int_{G\backslash X} k(x, y)f_i(x)dxdy}
\end{aligned}$$

$$= \int_{G\backslash X} f_i(y)\overline{Kf_i(y)}dy$$
$$= \langle f_i, Kf_i \rangle.$$

Thus $\lambda_i \langle f_i, f_i \rangle = \overline{\lambda_i} \langle f_i, f_i \rangle$ and so $\lambda_i \in \mathbb{R}$.

The trace formula is obtained by integrating $k(x,y)$ on the diagonal

$$\int_{G\backslash X} k(x,x)dx$$

and computing this in two ways: (1) spectral theory; (2) using the geometry of $G\backslash X$. We cannot do (2) yet because this depends on the space X, but the spectral theory computation is general.

We have the basis $f_1, f_2, \cdots \in \mathcal{L}^2(G\backslash X)$ with $Kf_i = \lambda_i f_i$. Consider $k(x,y)$ with $y = y_0$ fixed. As a function of x, $k(x, y_0) \in \mathcal{L}^2(G\backslash X)$. By the spectral theorem, $k(x, y_0)$ has a spectral expansion in x:

$$k(x, y_0) = \sum_{i=1}^{\infty} \langle k(*, y_0), f_i \rangle f_i(x).$$

There is no continuous spectrum because the operator is bounded. Since

$$\langle k(*, y_0), f_i \rangle = \int_{G\backslash X} k(x, y_0)\overline{f_i(x)}dx = \overline{\lambda_i} \cdot \overline{f_i(y_0)},$$

we get the identity

$$k(x,y) = \sum_{i=1}^{\infty} \overline{\lambda_i} f_i(x)\overline{f_i(y)}.$$

Thus

$$\int_{G\backslash X} k(x,x)dx = \sum_{i=1}^{\infty} \overline{\lambda_i} \int_{G\backslash X} f_i(x)\overline{f_i(x)}dx = \sum_{i=1}^{\infty} \overline{\lambda_i}$$

which is the trace.

There is a beautiful book *Bounded Integral Operators on $\mathcal{L}^2$ Spaces* by Halmos and Sunder which discusses these in general. Selberg's idea was to apply this to the group GL(2). Now we will do the simplest case — the Selberg trace formula for $\Gamma = \mathrm{SL}(2, \mathbb{Z})$. In this case we don't have to worry about cusps. It is important to understand what is going on in this simplest case first before jumping to the adelic generalizations.

Let $\phi : \mathbb{R} \to \mathbb{R}^+$ satisfy $\phi(t) \leq \frac{c_0}{(2+|t|)^{1+\epsilon}}$ for some fixed constant $c_0 > 0$.

Definition 13.3 (Selberg's kernel function)**.** Let $z = x + iy, z' = x' + iy' \in \mathfrak{h}$ with $x, x' \in \mathbb{R}$ and $y, y' > 0$. Then

$$k_\phi(z, z') := \phi\left(\frac{|z - z'|^2}{yy'}\right).$$

Remark.

(1) $\frac{|z-z'|^2}{yy'}$ is almost the hyperbolic distance between z, z'.

(2) $k_\phi(z, z')$ is symmetric: $k_\phi(z, z') = k_\phi(z', z) = \overline{k_\phi(z', z)}$.

Proposition 13.4. *For all $g \in \mathrm{SL}(2,\mathbb{R})$, we have*

$$k_\phi(gz, gz') = k_\phi(z, z').$$

Proof. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $k_\phi(gz, gz')$ is equal to

$$\phi\left(\frac{\left|\frac{az+b}{cz+d} - \frac{az'+b}{cz'+d}\right|^2}{\left(\frac{y}{|cz+d|^2}\right)\left(\frac{y'}{|cz'+d|^2}\right)}\right) = \phi\left(\frac{|(az+b)(cz'+d)-(az'+b)(cz+d)|^2}{yy'}\right) = \phi\left(\frac{|z-z'|^2}{yy'}\right).$$

Note we need the fact that $ad - bc = 1$ here. ∎

Definition 13.5. $K_\phi(z, z') = \sum_{\gamma\in\Gamma} k_\phi(\gamma z, z') = \sum_{\gamma\in\Gamma} k_\phi(z, \gamma z')$.

Then K_ϕ determines an integral operator on $\mathcal{L}^2(\Gamma\backslash\mathfrak{h})$. For $f \in \mathcal{L}^2(\Gamma\backslash\mathfrak{h})$, we define

$$K_\phi f(z) := \int_{\Gamma\backslash\mathfrak{h}} K_\phi(z, z') f(z') dz'.$$

Remark. Without that property $\phi(t) \le \frac{c_0}{(2+|t|)^{1+\epsilon}}$, the series defining K_ϕ will not converge absolutely.

We want to compute the trace of K_ϕ. Formally,

$$\operatorname{Trace} K_\phi = \int_{\Gamma\backslash\mathfrak{h}} K_\phi(z, z) dz.$$

But there is a big problem — K_ϕ is not Hilbert–Schmidt and we get ∞. Selberg knows how to fix it in this case and we will follow his notes on the trace formula. This problem is even more serious for $\mathrm{GL}(n)$. Arthur found a fix and it is extremely complicated.

Formally,

$$\begin{aligned} K_\phi f(z) &= \int_{\Gamma\backslash\mathfrak{h}} K_\phi(z, z') f(z) dz' \\ &= \int_{\Gamma\backslash\mathfrak{h}} \sum_{\gamma\in\Gamma} k_\phi(z, \gamma z') f(z') dz' \\ &= \sum_{\gamma\in\Gamma} \int_{\gamma(\Gamma\backslash\mathfrak{h})} k_\phi(z, z') f(z') dz' \\ &= \int_{\mathfrak{h}} k_\phi(z, z') f(z') dz'. \end{aligned}$$

Now we will give Selberg's solution to the convergence problems. He defines two modifications to K_ϕ.

(1) $K_\phi^{\#}(z, z') := K_\phi(z, z') - \sum_{m\in\mathbb{Z}} k_\phi(z, z'+m)$;

(2) $\widetilde{K_\phi}(z, z') := K_\phi(z, z') - \int_0^1 K_\phi(z, z'+t) dt$.

Selberg proves that $K_\phi^\#$ is Hilbert–Schmidt. He also proves $|K_\phi^\# - \widetilde{K_\phi}|$ is small, which implies $\widetilde{K_\phi}$ is Hilbert–Schmidt. Then he works with $\widetilde{K_\phi}$.

Let $\mathcal{L}_0^2(\Gamma\backslash\mathfrak{h})$ be the space of cusp forms for Γ. Then $f \in \mathcal{L}_0^2(\Gamma\backslash\mathfrak{h})$ if and only if $f \in \mathcal{L}^2(\Gamma\backslash\mathfrak{h})$ and the constant term in the Fourier expansion is $\int_0^1 f(z+t)dt = 0$.

Proposition 13.6. *If $f \in \mathcal{L}_0^2(\Gamma\backslash\mathfrak{h})$, then $\widetilde{K_\phi} f \in \mathcal{L}_0^2(\Gamma\backslash\mathfrak{h})$.*

Proof.

$$\widetilde{K_\phi} f(z) = \int_{\Gamma\backslash\mathfrak{h}} \left(K_\phi(z,z') - \int_0^1 K_\phi(z, z'+t)dt \right) f(z')dz'.$$

We can check that

$$\int_0^1 \widetilde{K_\phi} f(z+u)du = \int_0^1 \int_{\Gamma\backslash\mathfrak{h}} \left(K_\phi(z+u,z') - \int_0^1 K_\phi(z+u, z'+t)dt \right) f(z')dz'du = 0.$$

□

We can now compute the trace

$$\operatorname{Trace} \widetilde{K_\phi} = \int_{\Gamma\backslash\mathfrak{h}} \widetilde{K_\phi}(z,z)dz.$$

Let $f \in \mathcal{L}_0^2(\Gamma\backslash\mathfrak{h})$ be an eigenfunction of $\Delta = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ with $\Delta f = \lambda f$. Selberg proves that f must also be an eigenfunction of $\widetilde{K_\phi}$:

$$\widetilde{K_\phi} f = h(\lambda) f.$$

h is called the Selberg transform. Then we have

$$\operatorname{Trace} \widetilde{K_\phi} = \sum_\lambda h(\lambda).$$

Next time I will prove that $\widetilde{K_\phi}$ is Hilbert–Schmidt and show that h is a combination of the Abel transform and the Fourier transform. There is an explicit formula of h by Selberg, which makes his trace formula very powerful.

14. Lecture 14 (October 29, 2013)

Unfortunately I was unable to attend the lecture.

15. Lecture 15 (October 31, 2013)

15.1. **Selberg Trace Formula.** We started out with a test function of the type $\phi(t) \ll \frac{1}{|2+t|^{1+\epsilon}}$ for $t \geq 0$, and considered

$$k_\phi(z,z') = \phi\left(\frac{|z-z'|^2}{yy'}\right)$$

which satisfies $k_\phi(\gamma z, \gamma z') = k_\phi(z,z')$ for all $\gamma \in \mathrm{SL}(2,\mathbb{R})$. We set

$$K_\phi(z,z') = \sum_{\gamma\in\Gamma=\mathrm{SL}(2,\mathbb{Z})} k_\phi(\gamma z, z').$$

The Selberg transform, which was also discovered by Harish-Chandra in a different form, is

$$K_\phi f(z) := \int_{\Gamma\backslash\mathfrak{h}} K_\phi(z,z')f(z')dz'.$$

We have the following

Theorem 15.1. *If $\Delta f = \lambda f$, then $K_\phi f = h_\phi(\lambda) f$ where*

$$h_\phi(\lambda) = \frac{1}{\sqrt{2}}\int_0^\infty t^{ir}\Phi(t-2+t^{-1})\frac{dt}{t}$$

and

$$\Phi(x) = \sqrt{2}\int_0^\infty \phi(y)\frac{dy}{\sqrt{y-x}}.$$

Proof. Let $w = re^{i\theta}$ be in polar coordinates. Assume $f(w)$ is radially symmetric, i.e. $f(re^{i\theta}) = f(r)$ for all $r \in [0,2\pi)$, so f is a function of r only. If $\Delta f = \lambda f$ is regular, i.e. has a power series in r, then up to a constant, f is unique. This follows from the theory of differential equations. Since Δ is a second order differential operator, there are only two solutions. If we write

$$f(w) = r^c(1 + a_1 r + a_2 r^2 + \cdots),$$

then

$$\Delta f = -\frac{1}{4}(1-r^2)\left(\frac{d^2 f}{dr^2} + \frac{1}{r}\frac{df}{dr}\right) = \lambda f.$$

The other solution is $(\log r)f(r)$ which is not regular.

Consider

$$K_\phi f(z) = \iint_{\mathfrak{h}} \phi\left(\frac{|z-z'|^2}{yy_1'}\right) f(z_1)\frac{dx_1 dy_1}{y_1^2}.$$

We want to show that this equals $h_\phi(\lambda)f(z)$ for some unique function h_ϕ. Consider the map $\mathfrak{h} \to U = \{w = re^{i\theta} : 0 \le r \le 1, 0 \le \theta \le 2\pi\}$ given by a linear fractional transformation $z_1 \mapsto \frac{az_1+b}{cz_1+d}$ which sends $z \mapsto 0$. In other words, we want

$$\iint_U \phi\left(\frac{4|w|^2}{1-|w|^2}\right) f^*(w)d^*w \stackrel{?}{=} h_\phi(\lambda)f^*(0)$$

where $f^*(w) = f(z_1)$. Note that $\phi\left(\frac{4|w|^2}{1-|w|^2}\right)$ is radially symmetric, so the above is equivalent to

$$\int_{r=0}^1 \phi\left(\frac{4|w|^2}{1-|w|^2}\right)\int_{\theta=0}^{2\pi} f^*(w)d\theta r dr \stackrel{?}{=} h_\phi(\lambda)f^*(0)$$

where $f^\#(w) := \int_{\theta=0}^{2\pi} f^*(w)d\theta$ is radially symmetric. Thus we want

$$\iint_U \phi\left(\frac{4|w|^2}{1-|w|^2}\right) f^\#(w)dw \stackrel{?}{=} h_\phi(\lambda)f^\#(0),$$

so $h_\phi(\lambda)$ is uniquely determined.

To explicitly compute $K_\phi f$, we choose $f(z) = y^s$ with $\lambda = s(1-s)$. □

Now we want to compute the geometric side of the trace formula. Formally we have

$$\operatorname{Tr} K_\phi = \iint_{\mathcal{D}} K_\phi(z,z)dz = \iint_{\mathcal{D}} \sum_{\gamma\in\Gamma} k_\phi(z,\gamma z)dz$$

where $\mathcal{D} = \Gamma\backslash\mathfrak{h}$. The idea is to break the sum into conjugacy classes. For $\tau \in \Gamma$, define the conjugacy class $[\tau] := \{\alpha\tau\alpha^{-1} : \alpha \in \Gamma\}$ and the centralizer $\Gamma_\tau := \{\sigma : \sigma\tau = \tau\sigma\}$. Then

$$\begin{aligned}
\operatorname{Tr} K_\phi &= \iint_{\mathcal{D}} \sum_{[\tau]} \sum_{\gamma\in\Gamma_\tau\backslash\ \Gamma} k_\phi(z,\gamma\tau\gamma^{-1}z)dz \\
&= \sum_{[\tau]} \sum_{\gamma\in\Gamma_\tau\backslash\Gamma} \iint_{\mathcal{D}} k_\phi(\gamma z,\tau\gamma z)dz \\
&= \sum_{[\tau]} \sum_{\gamma\in\Gamma_\tau\backslash\Gamma} \iint_{\gamma\mathcal{D}} k_\phi(z,\tau z)dz \\
&= \sum_{[\tau]} \iint_{\Gamma_\tau\backslash\mathfrak{h}} k_\phi(z,\tau z)dz.
\end{aligned}$$

Here $O(\tau) := \iint_{\Gamma_\tau\backslash\mathfrak{h}} k_\phi(z,\tau z)dz$ is called the orbital integral.

For $\tau \in \mathrm{SL}_2(\mathbb{Z})$, there are three classifications:

- If $\operatorname{Trace}(\tau) = 2$ and $\tau \neq \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, then τ is <u>parabolic</u>.
- If $\operatorname{Trace}(\tau) > 2$, then τ is <u>hyperbolic</u>.
- If $\operatorname{Trace}(\tau) < 2$, then τ is <u>elliptic</u>.

The orbital integrals are completely different in each of the above three cases. For the hyperbolic and elliptic cases, the orbital integral converges absolutely. In the parabolic case, the orbital integral blows up and we need to subtract by a multiple of the Eisenstein series.

First we compute the identity orbital integral:

$$O\left(\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right) = \iint_{\Gamma\backslash\mathfrak{h}} \phi\left(\frac{|z-\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)z|^2}{y^2}\right)dz = \phi(0)\cdot \operatorname{Vol}(\Gamma\backslash\mathfrak{h}) = \frac{3}{\pi}\phi(0).$$

Next let us look at the hyperbolic orbital integral. Let P be a hyperbolic element in $\mathrm{SL}(2,\mathbb{Z})$ (we denote by P because hyperbolic elements are analogous to primes!). Then there exists $\gamma \in \mathrm{SL}(2,\mathbb{R})$ such that $\gamma P\gamma^{-1} = \begin{pmatrix}a & 0\\ 0 & a^{-1}\end{pmatrix}$ with $a > 1$, and

$$\operatorname{Tr}(P) = \operatorname{Tr}(\gamma P\gamma^{-1}) = a + a^{-1} > 2.$$

Selberg defines the norm $NP = a^2$, so we have

$$\operatorname{Tr}(P) = NP^{\frac{1}{2}} + NP^{-\frac{1}{2}}.$$

Now P hyperbolic implies P^l hyperbolic for $l = 0,1,2,\cdots$. Let P_0 generate an infinite cyclic group of hyperbolic elements, and $P = P_0^l$ for some $l \geq 1$. The centralizer of P is $\Gamma_P = \{P_0^l : l \in \mathbb{Z}\}$. We need to compute $\Gamma_P\backslash\mathfrak{h}$. Note that

$$\gamma P_0\gamma^{-1}z = \begin{pmatrix}a & 0\\ 0 & a^{-1}\end{pmatrix}z = a^2 z = NPz.$$

The fundamental domain of $\Gamma_P\backslash\mathfrak{h}$ is the horizontal strip $\{z \in \mathbb{C} : \operatorname{Im} z \in [1, NP_0]\}$. The orbital integral is thus

$$\begin{aligned} O(P) &= \iint_{\Gamma_P\backslash\mathfrak{h}} k_\phi(z, Pz)dz \\ &= \int_{y=1}^{NP_0}\int_{x=-\infty}^{\infty} \phi\left(\frac{|z-Pz|^2}{y^2 NP}\right)\frac{dxdy}{y^2} \\ &= \int_1^{NP_0}\int_{-\infty}^{\infty} \phi\left((NP^{\frac{1}{2}} + NP^{-\frac{1}{2}})^2\left(1+\frac{x^2}{y^2}\right)\right)\frac{dxdy}{y^2} \\ &= \log NP_0 \int_{-\infty}^{\infty} \phi((NP^{\frac{1}{2}} + NP^{-\frac{1}{2}})^2(x^2+1))dx. \end{aligned}$$

This is an Abel transform, which turns into the Selberg transform after we apply an inverse Fourier transform. We are not going to do the details, but the final formula is

$$O(P) = \frac{\log NP}{NP^{\frac{1}{2}} - NP^{-\frac{1}{2}}} h_\phi(\log NP).$$

The other two types of orbital integrals are more complicated, and the Selberg trace formula is

$$\operatorname{Tr}\widetilde{K_\phi} = \phi(0)\operatorname{Vol}(\Gamma\backslash\mathfrak{h}) + \sum_{[P]} \frac{\log NP}{NP^{\frac{1}{2}} - NP^{-\frac{1}{2}}} h_\phi(\log NP) + \sum_{\text{elliptic}} (*) + \left(\sum_{\text{parabolic}} (*) - \int \text{Eisenstein series}\right),$$

where $\operatorname{Tr}\widetilde{K_\phi} = \sum h_\phi(\lambda_i)$ is the spectral side, and the right hand side is the geometric side. Note that K_ϕ has to be replaced by $\widetilde{K_\phi}$ here because we are subtracting the Eisenstein series.

Assume $\Gamma \subset \mathrm{SL}(2,\mathbb{R})$ is cocompact (i.e. $\Gamma\backslash\mathfrak{h}$ is compact) with no elliptic elements. Then the trace formula is very simple:

$$\operatorname{Tr} K_\phi = \sum_{i=1}^{\infty} h_\phi(\lambda_i) = \phi(0)\operatorname{Vol}(\Gamma\backslash\mathfrak{h}) + \sum_{[P]} \frac{\log NP}{NP^{\frac{1}{2}} - NP^{-\frac{1}{2}}} h_\phi(\log NP).$$

This is the case Selberg discusses first in his Tata paper. Selberg stated his results and only released his proofs after others gave more complicated proofs.

The presence of log is suggestive of the explicit formula of the prime number theorem, which we recall now. If we integrate both sides of

$$-\frac{\zeta'}{\zeta}(s) = \sum_p \sum_{k=1}^{\infty} \frac{\log p}{p^{ks}}.$$

against some test function $H(s)$, we get

$$\frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} -\frac{\zeta'}{\zeta}\left(\frac{1}{2}+s\right)H(s)ds = \sum_p \sum_{k=1}^{\infty} \frac{\log p}{p^{k/2}} \cdot \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} e^{-(\log p)ks} H(s)ds.$$

The explicit formula is an identity relating this and the zeros of ζ:

$$\sum_p \sum_{k=1}^{\infty} \widetilde{H}(k\log p) = \sum_{\substack{p \\ \zeta(\frac{1}{2}+p)=0}} (\cdots).$$

Selberg asked if it is possible to construct a zeta function $Z(s)$ using hyperbolic elements $P \in \Gamma$ such that when you compute $\int \frac{Z'}{Z}(s)h(s)ds$ you get $\operatorname{Tr}\widetilde{K_\phi}$. The answer is YES! The zeros of $Z(s)$ are $\frac{1}{2}+i\nu$ where $\frac{1}{4}+\nu^2$ is the eigenvalue of Δ, i.e. there exists a Maass form $f : \mathcal{L}^2(\Gamma\backslash\mathfrak{h}) \to \mathbb{C}$ such that $\Delta f = \left(\frac{1}{4}+\nu^2\right) f$. This implies that the Riemann hypothesis holds for $Z(s)$, but this is slightly different from the classical case in that $Z(s)$ is known to have Siegel zeros, i.e. zeros close to 1.

Langlands once said the Selberg zeta function doesn't exist for him. They are not connected to automorphic forms!

16. Lecture 16 (November 7, 2013)

16.1. Selberg Trace Formula. I'm going to review the Selberg transform, using Selberg's original notations. We have a function $f \in \mathcal{L}^2(\Gamma\backslash\mathfrak{h})$ with $\Delta f = \lambda f$ where $\Delta = -y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)$, and $K_\phi f = h(\lambda) f$ for

$$K_\phi f(z) = \int_{\Gamma\backslash\mathfrak{h}} K_\phi(z,z')f(z')dz'$$

which is the Selberg transform. The Abel transform is

$$\Phi(w) = \int_{-\infty}^{\infty} \phi(x^2+w)dx = \int_{-\infty}^{\infty} \frac{\phi(t)}{\sqrt{t-w}}dt.$$

Selberg introduces the functions

$$g(u) := \Phi(e^u+e^{-u}-2)$$

and

$$h(r) := \int_0^\infty g(u)e^{iru}du.$$

Let us state the Selberg trace formula for co-compact groups Γ. Suppose there is an orthonormal basis of Maass forms for Γ satisfying $\Delta\eta_j = \left(\frac{1}{4}+r_j^2\right)\eta_j$ and $\eta_j(\gamma z) = \eta_j(z)$ for all $\gamma \in \Gamma$. There are no Eisenstein series! Last time I computed the Selberg trace formula, but let me make it more explicit using this basis. The trace is

$$\begin{aligned}\sum_{j=1}^{\infty} h(r_j) &= \int_{\Gamma\backslash\mathfrak{h}} K_\phi(z,z)\frac{dxdy}{y^2} \\ &= \mathrm{Vol}(\Gamma\backslash\mathfrak{h})\phi(0) + \sum_{\substack{\text{primitive hyperbolic} \\ \text{classes } [P_0]}} \sum_{l=1}^{\infty} \int_{-\infty}^{\infty}\int_1^{P_0^2} \phi\left(\frac{(P_0^{2l}-1)(x^2+y^2)}{P_0^{2l}y^2}\right)\frac{dxdy}{y^2}\end{aligned}$$

assuming there are no elliptic elements.

The identity term can be written as

$$\phi(0) = -\frac{1}{\pi}\int_0^\infty \frac{\Phi'(w)}{\sqrt{w}}dw = -\frac{1}{\pi}\int_0^\infty \frac{g'(u)}{e^{\frac{u}{2}}-e^{-\frac{u}{2}}}du = \frac{1}{2\pi^2}\int_0^\infty\int_0^\infty rh(r)\frac{\sin(ru)}{e^{\frac{u}{2}}-e^{-\frac{u}{2}}}dudr$$

because h is the Fourier transform of g. Similarly, we can get rid of the ϕ in the sum over hyperbolic terms and express it in terms of g and h. Thus we get the trace formula for co-compact groups

$$\sum_{j=1}^{\infty} h(r_j) = \mathrm{Vol}(\Gamma\backslash\mathfrak{h}) \cdot \frac{1}{4\pi}\int_{-\infty}^{\infty} r\frac{e^{\pi r}-e^{-\pi r}}{e^{\pi r}+e^{-\pi r}}h(r)dr + \sum_{[P_0]}\sum_{l=1}^{\infty}\frac{\log(NP_0)}{(NP_0)^{\frac{l}{2}}-(NP_0)^{-\frac{l}{2}}}g(l\log NP_0).$$

Selberg asked whether we can construct a zeta function for which the trace formula is the analogue of the explicit formula, which we now review. We have

$$-\frac{\zeta'}{\zeta}(s) = \sum_p \sum_{l=1}^{\infty} \frac{\log p}{p^{ls}}.$$

Then

$$\frac{1}{2\pi i}\int_{\frac{1}{2}+\epsilon-i\infty}^{\frac{1}{2}+\epsilon+i\infty} -\frac{\zeta'}{\zeta}\left(\frac{1}{2}+s\right)\widetilde{H}(s)ds = \sum_p\sum_{l=1}^{\infty}\frac{\log p}{p^{\frac{l}{2}}}\cdot\frac{1}{2\pi i}\int_{\left(\frac{1}{2}+\epsilon\right)} p^{-ls}\widetilde{H}(s)ds = \sum_p\sum_{l=1}^{\infty}\frac{\log p}{p^{\frac{l}{2}}}H(p^l),$$

where the Mellin transform is

$$\widetilde{H}(s) = \int_0^{\infty} y^s H(y)\frac{dy}{y}$$

and the inverse Mellin transform is

$$H(y) = \frac{1}{2\pi i}\int y^{-s}\widetilde{H}(s)ds.$$

Let us compare this with the trace formula. Consider the completed Riemann zeta function $\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s)$ under $s \mapsto 1-s$. The functional equation gives

$$-\frac{\pi}{2}+\frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}\right)+\frac{\zeta'}{\zeta}(s) = \frac{\pi}{2}-\frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{1-s}{2}\right)-\frac{\zeta'}{\zeta}(1-s),$$

i.e.

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

Shifting the line of integration and picking up the residues, we get

$$\frac{1}{2\pi i}\int_{\left(\frac{1}{2}+\epsilon\right)} -\frac{\zeta'}{\zeta}\left(\frac{1}{2}+s\right)\widetilde{H}(s)ds = \widetilde{H}\left(\frac{1}{2}\right) - \sum_{\substack{\alpha \\ \zeta(\frac{1}{2}+\alpha)=0}}\widetilde{H}(\alpha)+\frac{1}{2\pi i}\int_{\left(-\frac{1}{2}-\epsilon\right)} -\frac{\zeta'}{\zeta}\left(\frac{1}{2}+s\right)\widetilde{H}(s)ds.$$

Substituting the above into this and simplifying, the final identity is

$$2\sum_p\sum_{l=1}^{\infty}\frac{\log p}{p^{\frac{l}{2}}}H(p^l) = \widetilde{H}\left(\frac{1}{2}\right) - \sum_{\substack{\alpha \\ \zeta(\frac{1}{2}+\alpha)=0}}\widetilde{H}(\alpha)+\frac{1}{2\pi i}\int\frac{\Gamma'}{\Gamma}\cdots$$

i.e.

$$\sum \widetilde{H}(\alpha) - \widetilde{H}\left(\frac{1}{2}\right) = -2\sum_p\sum_l \frac{\log p}{p^{\frac{l}{2}}}H(p^l)+\frac{1}{2\pi i}\int\frac{\Gamma'}{\Gamma}\cdots.$$

This looks similar to the Selberg trace formula, so the question is whether we can construct an Euler product so that when we take its logarithmic derivative, we get the explicit formula.

Definition 16.1 (Selberg Zeta Function)**.**

$$Z(s) := \prod_{P_0} \prod_{l=0}^{\infty} \left(1 - \frac{1}{NP_0^{s+l}}\right).$$

Recall a hyperbolic element can always be diagonalized

$$\gamma P \gamma^{-1} = \begin{pmatrix} \rho & 0 \\ 0 & \rho^{-1} \end{pmatrix}$$

and we define the norm to be $NP = \rho^2$. If we compute $\int \frac{Z'}{Z}(s)$, we should get the trace formula. To do this we need a lemma.

Lemma 16.2.

$$\frac{Z'}{Z}\left(\frac{1}{2} + s\right) = \sum_{P_0} \sum_{l=1}^{\infty} \frac{\log NP_0}{NP_0^{l(s+\frac{1}{2})} - NP_0^{-l(s+\frac{1}{2})}}.$$

Proof.

$$\begin{aligned}
\frac{Z'}{Z}(s) = \frac{d}{ds}(\log Z(s)) &= \frac{d}{ds} \sum_{P_0} \sum_{l=0}^{\infty} \log(1 - NP_0^{-s-k}) \\
&= \sum_{P_0} \sum_{l=0}^{\infty} \log NP_0 \frac{NP_0^{-s-l}}{1 - NP_0^{-s-l}} \\
&= \sum_{P_0} \sum_{l=0}^{\infty} \log NP_0 \cdot NP_0^{-s-l} \sum_{m \geq 0} NP_0^{-ms-ml} \\
&= \sum_{P_0} \sum_{m \geq 1} \frac{\log NP_0}{NP_0^{ms} - NP_0^{m(s-1)}}.
\end{aligned}$$

□

Now we make a special choice of H so that everything is nice. We will follow Hejhal's book. Let

$$h(r) = \frac{1}{r^2 + \alpha^2} - \frac{1}{r^2 + \beta^2} = \frac{\beta^2 - \alpha^2}{(r^2 + \alpha^2)(r^2 + \beta^2)}$$

where $\alpha, \beta \in \mathbb{C}$ with $\frac{1}{2} < \mathrm{Re}(\alpha) < \mathrm{Re}(\beta)$. Note $h(r) = \Theta(\frac{1}{r^4})$. Then

$$g(u) = \frac{1}{2\alpha} e^{-\alpha(u)} - \frac{1}{2\beta} e^{-\beta(u)}.$$

If we plug these into the trace formula,

$$\begin{aligned}
&\frac{1}{2\alpha} \sum_{P_0} \sum_{l=1}^{\infty} \frac{\log NP_0}{NP_0^{\frac{l}{2}} - NP_0^{-\frac{l}{2}}} \cdot \frac{1}{NP_0^{l\alpha}} - \frac{1}{2\beta} \sum_{P_0} \sum_{l=1}^{\infty} \frac{\log NP_0}{NP_0^{\frac{l}{2}} - NP_0^{-\frac{l}{2}}} \cdot \frac{1}{NP_0^{l\beta}} \\
=& \sum_{j=0}^{\infty} \left(\frac{1}{r_j^2 + \alpha^2} - \frac{1}{r_j^2 + \beta^2}\right) - \frac{\mathrm{Vol}(\Gamma \backslash \gamma)}{4\pi} \int_{-\infty}^{\infty} r \left(\frac{1}{r^2 + \alpha^2} - \frac{1}{r^2 + \beta^2}\right) \tanh(\pi r) dr.
\end{aligned}$$

Next choose $\alpha = s - \frac{1}{2}$ and $\beta \mapsto b - \frac{1}{2}$. By the lemma,

$$\frac{1}{2s-1}\frac{Z'}{Z}\left(s-\frac{1}{2}\right)-\frac{1}{2\beta-1}\frac{Z'}{Z}\left(\beta-\frac{1}{2}\right)=\sum_{j=1}^{\infty}\left(\frac{1}{r_j^2+(s-\frac{1}{2})^2}-\frac{1}{r_j^2+(\beta-\frac{1}{2})^2}\right)-\frac{\mathrm{Vol}}{4\pi}\int\frac{dr}{r^2+(s-\frac{1}{2})^2}.$$

Note that the RHS is invariant under $s \mapsto 1-s$, and it has poles at $s = \frac{1}{2} \pm ir_j$. Thus the logarithmic derivative $\frac{Z'}{Z}$ has simple poles at $s = \frac{1}{2} + ir_j$ and no other poles. If Z has a pole, then $\frac{Z'}{Z}$ has a simple pole there with negative residue, but the residue from the RHS is never negative, so we know Z has no poles.

This is how we prove the analytic continuation, functional equation and the zeros and poles of the Selberg zeta function.

Theorem 16.3. *The Selberg zeta function $Z(s) = \prod_{P_0}\prod_{l=0}^{\infty}(1 - NP_0^{-s-l})$ is an entire function of $s \in \mathbb{C}$ which satisfies*

1. *$Z(s)$ has trivial zeros at $s = -l = -1, -2, -3, \cdots$ with multiplicity $(2g-1)(2l+1)$.*
2. *$s = 0$ is a zero of multiplicity $2g-1$ and $s = 1$ is a zero of multiplicity 1.*
3. *The non-trivial zeros are at $s = \frac{1}{2} \pm ir_j$, where $\frac{1}{4} + r_j^2$ is an eigenvalue of Δ.*
4. *(Functional Equation)* $Z(s) = Z(1-s)\exp\left(\mathrm{Vol}(\Gamma\backslash\mathfrak{h})\int_0^{s-\frac{1}{2}} r\tanh(\pi r)dr\right)$.

The proof can be found in Hejhal's book.

Let me make some remarks:

1. $Z(s)$ satisfies the Riemann hypothesis because Δ is a self-adjoint operator on $\mathcal{L}^2(\Gamma\backslash\mathfrak{h})$, which implies the eigenvalues $\frac{1}{4} + r_j^2$ are real.
2. For $\mathcal{L}^2(\Gamma\backslash\mathfrak{h})$ with Γ a congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$, we have a spectral gap
$$\lambda_j = \frac{1}{4} + r_j^2 > \frac{3}{16}.$$
Selberg conjectured that $\lambda_j \geq \frac{1}{4}$, which implies all trivial zeros lie on the real axis and there are no exceptional zeros.

Every hyperbolic matrix P has two fixed points r and $-r$ on the real axis, and the length of the closed geodesic between r and $-r$ is equal to NP. We have the prime geodesic theorem

$$\sum_{NP\leq x}\log NP \sim x$$

as $x \in \infty$. Since we have the Riemann hypothesis in this case, Selberg showed

$$\sum_{NP\leq x}\log NP \sim x + \Theta(x^{\frac{3}{4}+\epsilon}).$$

We don't get an exponent of $\frac{1}{2}$ because there is one big difference between the Riemann and Selberg zeta functions. We say an entire function $f(s)$ has Hadamard order r if

$$H(s) \ll e^{|s|^r}$$

for all $s \in \mathbb{C}$. The Riemann zeta function $\zeta(s)$ has Hadamard order 1, and the Selberg zeta function $Z(s)$ has Hadamard order 2. For $\zeta(s)$, the Hadamard order comes from $\Gamma(s+1) = s\Gamma(s)$. When this first came out, the question was what this weird integral in the functional equation of $Z(s)$ is. The answer was found by Marie-France Vignéras — it is the Barnes double gamma function $G(s)$ which satisfies

$$G(s+1) = \Gamma(s)G(s).$$

This gives a Hadamard order 2.

Let me mention two generalizations of the Selberg zeta function. The Ihara zeta function is the p-adic version of the Selberg zeta function. The Ruelle zeta function is constructed using dynamical systems.

We only know the Selberg zeta function for GL(2). It is an open problem whether there exists a Selberg zeta function of higher rank.

Next time we will start the adelic version of the Selberg trace formula.

17. Lecture 17 (November 12, 2013)

17.1. **Selberg Trace Formula (General Case).** We have spent a lot of time on Selberg's original version of his trace formula, so now we will jump to the most general case. The actual trace formula will be very simple in the end, except that nothing converges. We will first set up the notations, and then deal with convergence issues.

Let F be an algebraic number field and K/F be a finite extension. Let G be an algebraic group, e.g. GL(2), and G_K be the base change over F. The only difference between G and G_K is that for $G(R)$ we may only consider rings that contain K.

Definition 17.1 (Algebraic torus)**.** A torus T over a field F is an algebraic group over F such that after base change, $T_{\overline{F}} \cong \mathrm{GL}(1)^k$ for some positive integer k.

Here GL(1) is the multiplicative group in algebraic geometry defined by $ab = 1$.

Definition 17.2 (Split torus)**.** A torus T_F over F is split if $T_F \cong \mathrm{GL}(1)^k$.

Definition 17.3 (Anisotropic torus)**.** A torus T is anisotropic if $\mathrm{Hom}(T, \mathrm{GL}(1)) = \{0\}$.

Let us give some examples.

Example 17.4.

(1) GL(1) is a split torus.
(2) Let $\mathrm{SO}(2) = \{g \in \mathrm{SL}(2) : g^{-1} = {}^t g\} \cong \{x^2 + y^2 = 1 : x, y \in F\}$. Then $\mathrm{SO}(2, \mathbb{R})$ is anisotropic and $\mathrm{SO}(2, \mathbb{C}) \cong \mathrm{GL}(1)$ via $a = x + iy$ and $b = x - iy$.

A torus T_F has two pieces of structure:

(1) Base change $T_{\overline{F}}$;
(2) Galois descent needed to recover polynomial equations defining T as an algebraic group.

The next thing we need is a reductive group.

Definition 17.5 (Reductive group)**.** A reductive group G is an algebraic group satisfying $G \subset \mathrm{GL}(n)$ and $G = {}^t G$.

Definition 17.6 (Split reductive group)**.** A reductive group is split if it contains a maximal torus which is split.

I can now list the examples of split reductive groups.

Example 17.7.

- General linear groups: $A_n = \mathrm{SL}(n+1) = \{g \in \mathrm{GL}(n+1) : \det(g) = 1\}$;
- Odd orthogonal groups: $B_n = \mathrm{SO}(2n+1) = \{g \in \mathrm{GL}(2n+1) : {}^t g Q_{2n+1} = Q_{2n+1}, \det(g) = 1\}$, where $Q_{2n+1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & Q_{2n-1} & 0 \\ 1 & 0 & 0 \end{pmatrix}$;
- Symplectic groups: $C_n = \mathrm{Sp}(2n)$;
- Even special orthogonal groups: $D_n = \mathrm{SO}(2n)$.

We will now move to the Selberg trace formula for all these groups. Let G be a reductive group, F be a number field, $\mathbb{A}$ be the adele ring over F, and $\omega : F^\times \backslash \mathbb{A}^\times \to \mathbb{C}^\times$ be a Hecke character. We have the infinite-dimensional vector space:

$$V := \mathcal{L}^2(G(F)\backslash G(\mathbb{A})) := \{f : f(\gamma z g) = \omega^{-1}(z) f(g) \text{ for all } g \in G(\mathbb{A}), z \in Z(\mathbb{A}), \gamma \in G(F)\}$$

where $\mathcal{L}^2$ just means that the integral of the absolute value of f is bounded. We can think of V as the space of automorphic forms.

We have a representation $\pi : G(\mathbb{A}) \to \mathrm{GL}(V)$ defined by right action, i.e. if $f \in \mathcal{L}^2(G(F)\backslash G(\mathbb{A}))$ and $h \in G(\mathbb{A})$, then we define

$$\pi(h) f(g) := f(gh).$$

Using π, we construct an operation $V \to V$ as follows.

Definition 17.8. Let $\phi : G \to \mathbb{C}$ be a function satisfying $\phi(zg) = \omega^{-1}(z)\phi(g)$ for $z \in Z(G)$ and $g \in G$, which is compactly supported modulo $Z(G(\mathbb{A}))$. We define $\pi(\phi) : V \to V$ as follows:

$$\pi(\phi) := \int_{G(\mathbb{A})} \phi(y)\pi(y)dy.$$

As an example, let us compute

$$\pi(\phi)f(g) = \int_{G(\mathbb{A})} \phi(y)\pi(y)f(g)dg = \int_{G(\mathbb{A})} \phi(y) f(gy)dy$$

which implies

$$\pi(\phi)f(zg) = \omega^{-1}(z)\pi(\phi)f(g)$$

and

$$\pi(\phi)f(\gamma g) = \pi(\phi)f(g)$$

for all $\gamma \in G(F)$.

To do the Selberg trace formula, we have to analyze $\pi(\phi)$ further. We compute for $f \in \mathcal{L}^2(G(F)\backslash G(\mathbb{A}))$,

$$\begin{aligned} \pi(\phi)f(x) &= \int_{G(\mathbb{A})} \phi(y) f(xy)dy \\ &= \int_{G(\mathbb{A})} \phi(x^{-1}y) f(y)dy \\ &= \sum_{\gamma \in G(F)} \int_{\gamma \cdot Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} \phi(x^{-1}y) f(y)dy \end{aligned}$$

$$= \int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} \sum_{\gamma \in G(F)} \phi(x^{-1}\gamma y) f(y) dy.$$

Here we see Selberg's kernel function

$$K_\phi(x, y) = \sum_{\gamma \in G(F)} \phi(x^{-1}\gamma y).$$

But there is a big problem — this integral may not converge! We need to modify $K_\phi(x, y)$ into the truncated kernel $K_\phi^T(x, y)$. This is the most difficult part and I will talk about it next time. If K_ϕ is Hilbert–Schmidt, i.e.

$$\int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} \int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} |K_\phi(x, y)|^2 < \infty,$$

then we can do a spectral expansion for K_ϕ. This will not be true, but we will have

$$\int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} \int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} |K_\phi^T(x, y)|^2 < \infty.$$

Now I want to talk about the geometric side. The trace is obtained by

$$\text{Trace} = \int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} K_\phi(x, x) dx.$$

Let's assume for the moment that this converges. Then we can rewrite it as

$$\begin{aligned}
\text{Trace} &= \int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} \sum_{\gamma \in G(F)} \phi(x^{-1}\gamma x) dx \\
&= \int_{Z(\mathbb{A})G(F)\backslash G(\mathbb{A})} \sum_{\substack{\text{conjugacy} \\ \text{classes } [\tau]}} \sum_{\gamma \in G_\tau \backslash [\tau]} \phi(x^{-1}\gamma\tau\gamma^{-1} x) dx \\
&= \sum_{[\tau]} \text{Vol}(Z(\mathbb{A})G_\tau(F)\backslash G_\tau(\mathbb{A})) \int_{G_\tau \backslash G(\mathbb{A})} \phi(x^{-1}\tau x) dx.
\end{aligned}$$

The computation is exactly like before for GL(2). Here $\int_{G_\tau\backslash G(\mathbb{A})} \phi(x^{-1}\tau x) dx$ are called the orbital integrals.

Assuming ϕ is factorizable, i.e. $\phi = \prod_v \phi_v$, we have

$$\int_{G_\tau \backslash G(\mathbb{A})} \phi(x^{-1}\tau x) dx = \prod_v \int_{G_{v,\tau}\backslash G_v(\mathbb{A})} \phi_v(x_v^{-1}\tau_v x_v) dx$$

which is a product of local orbital integrals. The whole problem is to compute these local orbital integrals, which give us the geometric side of the trace formula.

Remark. In their book, Jacquet and Langlands computed the trace formula for two different reductive groups and got matching orbital integrals. The functoriality conjecture, which I won't state precisely, basically says that if you take an automorphic form and some tensor product, you get an automorphic form on a different group. For example, if you have a GL(2)-automorphic form and take the symmetric square lift, Jacquet and Gelbart proved that it gives an automorphic form on GL(3). The problem is stable conjugacy.

Suppose K is an algebraic number field.

Lemma 17.9. *If $A, B \in \mathrm{GL}_n(K)$ are conjugate over $\overline{K}$, i.e.*

$$B = \gamma A \gamma^{-1}$$

for some $\gamma \in \mathrm{GL}_n(\overline{K})$, then there exists $\alpha \in \mathrm{GL}(n, K)$ such that

$$B = \alpha A \alpha^{-1}.$$

But this lemma may break down if G is not GL_n.

Example 17.10. For $G = \mathrm{SL}(2, \mathbb{R})$, $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ are conjugate by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ but not conjugate by an element of $\mathrm{SL}(2, \mathbb{R})$.

Definition 17.11 (Stably conjugate). $A, B \in G(K)$ are stably conjugate if they are conjugate over $G(\overline{K})$.

To match orbital integrals in different groups, it is necessary to write the trace formula in terms of stable conjugacy classes. Langlands and Shelstad conjectured the fundamental lemma which gave explicit comparisons for stable orbital integrals, finally proved by Ngô.

Let me conclude by doing the simplest version of the general case. Let G be a finite group, $V \cong \mathbb{C}^r$ be a finite-dimensional complex vector space and $\pi : G \to \mathrm{GL}(V)$ be a representation. We have the group algebra

$$\mathbb{C}[G] := \{\phi : G \to \mathbb{C}\}$$

which is an algebra with convolution

$$\phi_1 * \phi_2(g) = \sum_{g_1 g_2 = g} \phi_1(g_1)\phi_2(g_2).$$

Example 17.12. We have the trivial representation π_{triv}, and the regular representation π_{reg} given by

$$\pi_{\mathrm{reg}}(h)\phi(g) = \phi(gh).$$

The Selberg trace formula breaks G into conjugacy classes

$$[g] = \{\sigma^{-1} g \sigma : \sigma \in G\}.$$

The class functions are defined to be $\mathrm{Class}[G] := \{\phi : G \to \mathbb{C} : \phi(\sigma g \sigma^{-1}) = \phi(g) \text{ for all } \sigma, g \in G\}$. There is an inner product on class functions

$$\langle \phi_1, \phi_2 \rangle := \frac{1}{|G|} \sum_{g \in G} \phi_1(g)\overline{\phi_2(g)}.$$

Let $\chi_\pi = \mathrm{Tr}(\pi(g))$ be the character associated to π.

Example 17.13. $\chi_{\mathrm{triv}}(g) = 1$ and $\chi_{\mathrm{reg}}(g) = \begin{cases} |G| & \text{if } g = \mathrm{Id}, \\ 0 & \text{otherwise.} \end{cases}$

Definition 17.14.

$$1_{[h]}(g) = \begin{cases} 1 & \text{if } g \in [h], \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\langle 1_{[h_1]}, 1_{[h_2]} \rangle = \begin{cases} \frac{|[h_1]|}{|G|} & \text{if } [h_1] = [h_2], \\ 0 & \text{if } [h_1] \neq [h_2]. \end{cases}$$

The elements $1_{[h]}$ form a basis for class functions.

The trace formula is

$$\sum_i m(\pi_i)\chi_{\pi_i} = \sum_{[g]} \mathrm{Tr}(\pi(g)) \cdot 1_{[g]}$$

where $m(\pi_i)$ is the multiplicity of π_i in π. This is just the spectral theorem. To get the Selberg trace formula, we consider $\Gamma \subset G$. Let $\pi : \Gamma \to \mathrm{GL}(W)$ be a representation, where $W = \mathbb{C}^l$ is a finite-dimensional complex vector space. We induce π to a representation $\pi_{\mathrm{ind}} = \mathrm{Ind}_\Gamma^G(\pi) : G \to \mathrm{GL}(V)$ by

$$V := \{f : g \to W : f(\gamma g) = \pi(\gamma)f(g) \text{ for all } \gamma \in \Gamma, g \in G\}$$

and

$$\pi_{\mathrm{ind}}(h)f(g) := f(gh).$$

Definition 17.15 (Selberg kernel)**.**

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

where $x, y \in \Gamma\backslash G$.

We have

$$\chi_{\mathrm{ind}}(\phi) = \sum_{[\gamma]} \mathrm{Vol}(\Gamma_\gamma\backslash G_\gamma) \sum_{x \in \Gamma_\gamma\backslash G} \phi(x^{-1}\gamma x)$$

where $\sum_{x \in \Gamma_\gamma\backslash G} \phi(x^{-1}\gamma x)$ is the orbital integral.

If we work everything out, we get the Selberg trace formula for finite groups

$$\sum_{\substack{\text{conjugacy} \\ \text{classes } [\gamma]}} \mathrm{Vol}(\Gamma_\gamma\backslash G_\gamma) 1_{[\gamma]}(\phi) = \sum_{\substack{\text{irreducible} \\ \text{representations } \phi_i}} m_i \cdot \chi_i(\phi)$$

where m_i is the multiplicity. This turns out to be equivalent to Frobenius reciprocity.

Next time I will to explain how to do the truncation of the kernel function for GL(2). I will follow the exposition in Gelbart's book.

18. Lecture 18 (November 14, 2013)

Unfortunately I was unable to attend the lecture.

19. Lecture 19 (November 19, 2013)

19.1. **Beyond Endoscopy.** Beyond Endoscopy is the title of a famous paper by Langlands, which is available on the IAS website. Shortly after the paper was published, Sarnak wrote the letter *Comments on Robert Langlands' Lecture "Endoscopy and Beyond"*, which is also available online and has been very influential. Everything I say in the next few lectures will be motivated by Sarnak's letter.

Let me give a rough outline of what Beyond Endoscopy means. The motivation is Langlands' functoriality conjectures. Rather than defining these conjectures in general, which requires a huge amount of notations, let me give an example which is still unsolved. Let π be an automorphic representation on $\mathrm{GL}(n, \mathbb{Q})$. Associated to π we have an L-function $L(s, \pi)$, which will have an Euler product that looks like

$$L(s,\pi) = \prod_p \prod_{i=1}^n \left(1 - \frac{\alpha_{p,i}(\pi)}{p^s}\right)^{-1}$$

where $\alpha_{p,i}(\pi) \in \mathbb{C}$. The key point is that there are n factors at each prime. We have the Ramanujan conjecture: $|\alpha_{p,i}(\pi)| \leq 1$ (usually equal to 1, but 0 when there is ramification). Let's take the tensor product

$$L(s,\pi^{\otimes k}) = \prod_p \prod_{1 \leq i_1, \cdots, i_k \leq n} \left(1 - \frac{\alpha_{p,i_1}(\pi) \cdots \alpha_{p,i_k}(\pi)}{p^s}\right)^{-1}.$$

We could do other kinds of operations than tensor products, e.g. symmetric powers

$$L(s,\mathrm{Sym}^k(\pi)) = \prod_p \prod_{1 \leq i_1 \leq \cdots \leq i_k \leq n} \left(1 - \frac{\alpha_{p,i_1}(\pi) \cdots \alpha_{p,i_k}(\pi)}{p^s}\right)^{-1}.$$

We have the special case of Langlands functoriality

Conjecture 19.1.

(1) *$L(s, \pi^k)$ is automorphic on* $\mathrm{GL}(n^k)$.
(2) *$L(s, \mathrm{Sym}^k(\pi))$ is automorphic on* $\mathrm{GL}\left(\binom{n+k-1}{k}\right)$.[2]

Let me review the known results.

(1) D. Ramakrishnan: if π is on $\mathrm{GL}(2)$, then $\pi \otimes \pi$ is automorphic on $\mathrm{GL}(4)$.
(2) Jacquet–Gelbart: if π is on $\mathrm{GL}(2)$, then $L(s, \mathrm{Sym}^k(\pi))$ is automorphic on $\mathrm{GL}(3)$. Kim–Shahidi: if π is on $\mathrm{GL}(2)$, then $L(s, \mathrm{Sym}^k \pi)$ is automorphic for $k = 2, 3, 4$. They have some partial results for $k = 5$.

There seems to be a real barrier going beyond these results. Langlands proposed a method to try to prove these kinds of conjectures using the trace formula. The basic idea is that we want to compare the trace formulae on two different groups and get some kind of matching. I should say that Ngô's method of proving the fundamental lemma does not help with proving these conjectures. All the known results use the converse theorem. Remarkably, Kim–Shahidi used the exceptional Lie groups to get their results, so it doesn't go any further. Langlands proposed a completely different method. Venkatesh, Herman and Altug have reproved some of the above results using Beyond Endoscopy methods.

The first step is to get analytic continuation of things like $L(s, \pi^k)$ and $L(s, \mathrm{Sym}^k(\pi))$ and location of poles. This seems hopeless at the moment for a single π, but Langlands suggested trying to do it for

$$S(h) = \sum_\pi h(\lambda_\pi) L(s, \mathrm{Sym}^k(\pi))$$

where h is a test function of rapid decay and λ_π is the Laplace eigenvalue of π, called the spectral parameter. The key idea is that we can construct $S(h)$ using the trace formula. In

[2]Here $\binom{n+k-1}{k} = \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq n} 1$.

the trace formula we have the kernel function

$$K(g,g') = \sum_f h(f)\cdot f(g)\overline{f(g')}$$

which gives

$$\mathrm{Tr} = \int K(g,g) = \sum h(f).$$

In order to get $L(s,\mathrm{Sym}^k(\pi))$, we need the Hecke operators T_n. Let

$$T_n f = \alpha(n) f$$

where $\alpha(n)$ is the n-th Fourier coefficient. Let's say $f\in \mathrm{Space}(\pi)$ generates the space. Then

$$L(s,\pi) = \sum_n (T_n f)\cdot n^{-s}.$$

The idea is to apply T_s^k, roughly of the form $\sum \frac{T_n}{n^s}$ (dependent on k), and get

$$\int T_s^k \cdot K(g,g) = \sum h(\pi) L(s,\mathrm{Sym}^k(\pi)).$$

By modifying the trace formula, we can get the sum $S(h)$ on the spectral side. The idea is to compare the trace formula on $\mathrm{GL}(n)$ using $\int T_s^k K(g,g) = \sum h(\pi)L(s,\mathrm{Sym}^k(\pi))$ with the trace formula on $\mathrm{GL}\left(\binom{n+k-1}{k}\right)$, and try to find some matching.

Definition 19.2. A smaller or simpler group G' is endoscopic to G if the representations of G' describe the internal structure of the representations of G.

In this case the trace formula for G and G' can be compared.

Langlands suggested using poles of $L(s,\mathrm{Sym}^k\pi)$ and considering some matching of the form

$$\mathrm{Trace} = \sum_{\pi \text{ on } \mathrm{GL}(n)} h(\pi)\operatorname*{Res}_{s=\alpha} L(s,\mathrm{Sym}^k(\pi)) = \sum_{\pi^* \text{ on } \mathrm{GL}(M)} h(\pi^*)\operatorname*{Res}_{s=\alpha} L(s,\pi^*).$$

In Sarnak's letter to Langlands, he first talks about Andy Booker's paper *A test for identifying Fourier coefficients of automorphic forms and application to Kloosterman sums.* Basically he gives a numerical test for modularity.

Katz made the following conjecture in his book *Exponential Sums and Differential Equations.*

Conjecture 19.3 (Katz)**.** *Let*

$$S(m,n,c) = \sum_{\substack{a=1\\(a,c)=1\\ a\bar{a}\equiv 1 \pmod c}}^{c} e^{2\pi i\cdot\frac{am+\bar{a}n}{c}}$$

be the Kloosterman sum. Then

$$L(s,Kl) = \prod_p \left(1 \pm \frac{S(m,n,p)}{p^{s+\frac12}} + \frac{1}{p^{2s}}\right)^{-1}$$

(which is convergent on $\mathrm{Re}(s)>1$*) is the* L*-function of a Maass form for* $\Gamma_0(N)$ *for some* $N\geq 1$*.*

Remark. In 1999, Chai and Li proved Katz's conjecture for function fields.

Booker numerically tested Katz's conjecture and it came out very negative. Sarnak's idea was to combine Booker's test and Langlands' ideas.

Theorem 19.4 (Booker)**.** *$L(s, Kl)$ cannot be the L-function of a holomorphic modular form on $\Gamma_0(N)$.*

Proof. If it is associated to a holomorphic form, then the Fourier coefficients all lie in a fixed number field $K/\mathbb{Q}$. Suppose that for fixed m, n, all $S(m, n, p) \in K$. Pick $p > 3$ not dividing N. Then

$$S(m, n, p) \in \mathbb{Q}(\zeta_p) \cap K = \mathbb{Q}.$$

On the other hand,

$$\begin{aligned} S(m, n, p) &= \sum_{a=1}^{p-1} e^{2\pi i \cdot \frac{am+\bar{a}n}{p}} \\ &= \sum_{t=1}^{p-1} \left(1 + \left(\frac{t^2 - 4mn}{p}\right)\right) e^{\frac{2\pi i t}{p}} \end{aligned}$$

where $\left(\frac{t^2-4mn}{p}\right)$ is the quadratic symbol. This series has at most $\frac{p-1}{2}$ terms, but this is impossible because the minimal polynomial for ζ_p is $1 + x + \cdots x^{p-1}$. ☐

Let us talk about Booker's numerical test for checking modularity. Let

$$f(x + iy) = \sum_{n \neq 0} a_f(n) \sqrt{y} K_{i\nu}(2\pi |n| y) e^{2\pi i n x}$$

be a Maass form for $\Gamma_0(N)$. The associated L-function is

$$L(s, f) = \sum_{n=1}^{\infty} \frac{a_f(n)}{n^s}.$$

It satisfies a functional equation. If we define

$$\Lambda(s, f) = \frac{N^{\frac{s}{2}}}{\pi^s} \Gamma\left(\frac{s + \epsilon + i\nu}{2}\right) \Gamma\left(\frac{s + \epsilon - i\nu}{2}\right) L(s, f)$$

where $\epsilon = 0$ or -1 depending on whether f is even or odd, then

$$\Lambda(s, f) = \pm \Lambda(1 - s, f).$$

Choose $F(x) = x^2 e^{-x}$.

Definition 19.5. $S_{Y,N} = \displaystyle\sum_{\substack{n=1 \\ (n,N)=1}}^{\infty} \lambda_f(n) F\left(\frac{n}{Y}\right).$

We have $S_{Y,N} \approx \sum_{n \ll Y} \lambda_f(n) F\left(\frac{n}{Y}\right)$. If $\lambda_f(n)$ are randomly chosen, we expect $S_{Y,N} \approx \sqrt{Y}$. However, Booker shows that modularity implies $S_{Y,N}$ is much smaller than $\sqrt{Y}$. I can now state the precise theorem.

Theorem 19.6 (Booker)**.** *Assume f is a Maass form for $\Gamma_0(N)$ with eigenvalue λ. Then*

$$\left|\frac{S_{Y,N}}{\sqrt{Y}}\right| < \prod_{\substack{p|N \\ p^2\nmid N}} (1+p) \cdot \left(\frac{N(\lambda+3)}{4288Y}\right)^2 \ll \frac{1}{Y^2}.$$

Corollary 19.7. *If Katz's conjecture is true, then $N(\lambda+3) > (18.3)\cdot 10^6$.*

I should mention that Venkatesh–Booker–Strömbergsson numerically computed the first 1000 Maass forms for $\mathrm{SL}(2,\mathbb{Z})$ (for each of these Maass forms, they computed the first 100 Fourier coefficients to 1000 decimal places). These Fourier coefficients seem to be irrational and transcendental.

Next time I will show how to use this kind of numerical modularity test in combination with Langlands' ideas, as suggested by Sarnak.

20. Lecture 20 (November 21, 2013)

20.1. **Booker's Theorems.** We will give a proof of the theorem by Booker. Let

$$f(x+iy) = \sum_{n\neq 0} A_f(n)\sqrt{y}K_{i\nu}(2\pi|n|y)e^{2\pi inx}$$

be a Maass newform on $\Gamma_0(N)$ with Laplace eigenvalue $\frac{1}{4}+\nu^2$. We have the L-function

$$L(s,f) = \sum \frac{A_f(n)}{n^s}$$

which satisfies the functional equation

$$\Lambda(s,f) = \left(\frac{\sqrt{N}}{\pi}\right)^s \Gamma\left(\frac{s+\epsilon+i\nu}{2}\right)\Gamma\left(\frac{s+\epsilon-i\nu}{2}\right)L(s,f) = \pm\Lambda(1-s,f)$$

where

$$\epsilon = \begin{cases} 1 & \text{if } f \text{ is even,} \\ -1 & \text{if } f \text{ is odd.} \end{cases}$$

We define

$$S_{Y,N} := \sum_{(n,N)=1} A_f(n)\left(\frac{n}{Y}\right)^2 e^{-\frac{n}{Y}}.$$

Theorem 20.1 (Booker)**.** *We have*

$$\frac{S_{Y,N}}{\sqrt{N}} < \prod_{\substack{p|N \\ p^2\nmid N}} (1+p) \cdot \left(\frac{N\left(\frac{1}{4}+\nu^2+3\right)}{42.88Y}\right)^2.$$

This implies that for $Y\to\infty$,

$$\frac{S_{Y,N}}{\sqrt{N}} \ll \frac{1}{Y^2}.$$

On the other hand, if $A_f(n)$ are random, we expect

$$\frac{S_{Y,N}}{N} = O(1)$$

as $Y \to \infty$. So this gives a numerical test for modularity.

The proof uses standard methods in analytic number theory.

Proof. We have the test function $h(x) = x^2 e^{-x}$, which is equal to

$$h(x) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \Gamma(s+2) x^{-s} ds$$

by Mellin inversion, because

$$\widetilde{h}(s) = \int_0^\infty x^2 e^{-x} x^s \frac{dx}{x} = \Gamma(s+2).$$

This implies that

$$\begin{aligned}
S_{Y,N} &= \sum_{(n,N)=1} A_f(n) h\left(\frac{n}{Y}\right) \\
&= \sum_{(n,N)=1} A_f(n) \cdot \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \Gamma(s+2) \left(\frac{Y}{n}\right)^s ds \\
&= \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \Gamma(s+2) L(s,f) \prod_{\substack{p|N \\ p^2 \nmid N}} \left(1 \pm p^{-s-\frac{1}{2}}\right) Y^s ds \\
&= \pm \frac{1}{2\pi i} \int_{-\frac{3}{2}-i\infty}^{-\frac{3}{2}+i\infty} \Gamma(s+2) Y^s \left(\frac{\sqrt{N}}{\pi}\right)^{1-2s} \frac{\Gamma\left(\frac{1-s+\epsilon+i\nu}{2}\right)\Gamma\left(\frac{1-s+\epsilon-i\nu}{2}\right)}{\Gamma\left(\frac{s+\epsilon+i\nu}{2}\right)\Gamma\left(\frac{s+\epsilon-i\nu}{2}\right)} L(1-s,f) \prod_{\substack{p|N \\ p^2 \nmid N}} \left(1 \pm p^{-s-\frac{1}{2}}\right) ds
\end{aligned}$$

where we used the functional equation, and shifted the line of integration because the L-function of a cusp form is holomorphic.

Recall Stirling's formula: for $s = \sigma + it$ where σ is a fixed real number, we have

$$\Gamma(\sigma + it) \sim c|t|^{\sigma - \frac{1}{2}} e^{-\frac{\pi}{2}|t|}$$

as $|t| \to \infty$, where c is a constant. This implies that in the integral above, $\Gamma(s+2)$ has exponential decay as $|t| \to \infty$, and

$$\frac{\Gamma\left(\frac{1-s+\epsilon+i\nu}{2}\right)\Gamma\left(\frac{1-s+\epsilon-i\nu}{2}\right)}{\Gamma\left(\frac{s+\epsilon+i\nu}{2}\right)\Gamma\left(\frac{s+\epsilon-i\nu}{2}\right)}$$

has polynomial decay in $|t|$. If we compute the constants carefully we will get Booker's theorem

$$\frac{S_{Y,N}}{\sqrt{Y}} \leq c_{N,f} Y^{-2}.$$

□

In the appendix, Booker proves the following

Theorem 20.2 (Booker)**.** *Suppose $0 \leq \alpha_p \leq \pi$ are arbitrary real numbers (angles) for primes $p < P$ where $P \to \infty$. Then there exists $c > 0$ such that for all $\epsilon > 0$, there exists a Maass cusp newform f of level 1 such that $\Delta f = \lambda f$ with*

$$\lambda < e^{\frac{cP^2}{\log P} \cdot \frac{\log(1+\frac{1}{\epsilon})}{\epsilon}}$$

whose Fourier coefficients $A_f(p) = 2\cos\theta_p$ $(0 \le \theta_p \le \pi)$ *satisfy*

$$|\theta_p - \alpha_p| < \epsilon$$

for all $p \le P$.

This implies we cannot disprove Katz's conjecture by numerical calculations! However, if we know that the coefficients are algebraic (e.g. modular forms), once we know they are close enough we can prove things by numerical computations.

The proof of this theorem uses the trace formula. I think it will be interesting to generalize this theorem adelically.

20.2. **Selberg Trace Formula for Holomorphic Forms.** Let us talk about the Selberg trace formula for holomorphic forms. An interesting proof for $\Gamma = \mathrm{SL}(2,\mathbb{Z})$ was found by Zagier in 1975, and it was generalized to level N by H. Cohen and Oèsterlé.

For $z, z' \in \mathfrak{h}$, we will construct a kernel $h(z,z')$ which is assumed to be a holomorphic form of weight k in both z and z' independently. We define the convolution

$$f * h(z') := \int_{\Gamma\backslash\mathfrak{h}} f(z)\overline{h(z, -\overline{z'})}(\operatorname{Im} z)^k \frac{dxdy}{y^2}.$$

Zagier chooses the kernel function to be

$$h_m(z,z') := \sum_{ad-bc=m} (czz' + dz' + az + b)^{-k} = \sum_{ad-bc=m} (cz+d)^{-k}\left(z' + \frac{az+b}{cz+d}\right)^{-k},$$

which converges absolutely for $k \ge 4$. This kernel function simplifies Selberg's original proof of his trace formula, but only works when there is a holomorphic structure.

Theorem 20.3 (Zagier)**.** *Define* $C_k = \frac{(-1)^{\frac{k}{2}}\pi}{2^{k-3}(k-1)}$, *and let* $\mathcal{S}_k$ *be the space of holomorphic weight* k *cusp forms for* $\mathrm{SL}(2,\mathbb{Z})$.

- *The function* $C_k^{-1} m^{k-1} h_m(z,z')$ *satisfies*
$$f * h_m(z') = C_k m^{-k+1}(T_m f)(z')$$
where $T_m : \mathcal{S}_k \to \mathcal{S}_k$ *is the Hecke operator, i.e.* $C_k^{-1} m^{k-1} h_m(z,z')$ *is the kernel for* T_m.
- *Let* $f_1, \cdots, f_r$ *be an eigenbasis of* $\mathcal{S}_k$ *with* $T_m f_i = a^i(m) f_i$. *Then the trace of* T_m *is equal to*
$$\operatorname{Tr}(T(m)) = \sum_{i=1}^r a^i(m) = \frac{m^{k-1}}{C_k}\int_{\Gamma\backslash\mathfrak{h}} h_m(z, -\overline{z})(\operatorname{Im} z)^k \frac{dxdy}{y^2}.$$

Proof. We want to prove

$$\frac{m^{k-1}}{C_k} h_m(z,z') = \sum_{i=1}^r a^i(m)\frac{f_i(z)f_i(z')}{\langle f_i, f_i\rangle}.$$

This is the spectral expansion into a basis $f_1, \cdots, f_r$ of $\mathcal{S}_k$. This would then imply

$$\int_{\Gamma\backslash\mathfrak{h}} \frac{m^{k-1}}{C_k} h(z, -\overline{z})(\operatorname{Im} z)^k \frac{dxdy}{y^2} = \sum a^i(m).$$

When $m = 1$, $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$. Then the convolution is

$$\begin{aligned} f * h_1(z) &= \int_{\Gamma\backslash\mathfrak{h}} \sum_{\gamma\in\mathrm{SL}(2,\mathbb{Z})} (-z' + \gamma\overline{z})^{-k} f(\gamma z)(\mathrm{Im}\,\gamma z)^k \frac{dxdy}{y^2} \\ &= 2\int_0^{\infty}\int_{-\infty}^{\infty} (x - iy - z')^{-k} f(x+iy) y^{k-2} dxdy \\ &= 2\int_0^{\infty} \frac{2\pi i}{(k-1)!} f^{(k-1)}(2iy + z')dy \\ &= C_k f(z'). \end{aligned}$$

So we have proved the theorem when $m = 1$. We can easily generalize it to $m > 1$. □

We now state the

Theorem 20.4 (Selberg trace formula of T_m for holomorphic forms of weight k for $\mathrm{SL}(2, \mathbb{Z})$)**.** *We have*

$$\mathrm{Tr}(T(m)) = -\frac{1}{2}\sum_{t=-\infty}^{\infty} P_k(t, m)H(4m - t^2) - \frac{1}{2}\sum_{dd'=m} \min(d, d')^k.$$

Here

$$P_k(t, m) = \frac{\rho^{k-1} - \overline{\rho}^{k-1}}{\rho - \overline{\rho}}$$

where $\rho + \overline{\rho} = t$ *and* $\rho \cdot \overline{\rho} = m$, *and*

$$H(n) = \begin{cases} 0 & \textit{if } n < 0, \\ -\frac{1}{12} & \textit{if } n = 0, \\ \textit{class number} = \#\{ax^2 + bxy + cy^2 : b^2 - 4ac = -n\} & \textit{if } n > 0, \end{cases}$$

where we have to count forms equivalent to $x^2 + y^2$ *(resp.* $x^2 + xy + y^2$*) with multiplicity* $\frac{1}{2}$ *(resp.* $\frac{1}{3}$*).*

Zagier gives the following table:

n	0	3	4	7	8	11	12	15
$H(n)$	$-\frac{1}{12}$	$\frac{1}{3}$	$\frac{1}{2}$	1	1	1	$\frac{4}{3}$	2

Example 20.5. Let $k = 4$. There are no cusp forms of weight 4 for $\mathrm{SL}(2, \mathbb{Z})$, so $\mathrm{Tr}(T_m) = 0$. On the other hand,

$$\sum_{t=-\infty}^{\infty} (t^2 - m)H(4m - t^2) = -5H(20) - 8H(19) - 2H(16) + 8H(11) + 22H(4) = -10 - 8 - 3 + 8 + 11 = -2$$

and

$$\sum \min(d, d')^3 = 1^3 + 1^3 = 2.$$

In fact, when $k = 2, 4, 6, 8, 10$, $\mathrm{Tr}(T_m) = 0$ so we get an identity between class numbers and divisor sums. But those relations were discovered previously by Kronecker. When $k \geq 12$, the trace involves cusp forms and the formula becomes more interesting.

Proof (Sketch). We have

$$\operatorname{Tr}(T(m)) = \frac{m^{k-1}}{C_k} \int_{\Gamma\backslash\mathfrak{h}} \sum_{ad-bc=m} \frac{y^k}{(c|z|^2 + d\overline{z} - az - b)^k} \frac{dxdy}{y^2}.$$

Zagier writes $\operatorname{Tr}(T(m)) = \sum_{t=-\infty}^{\infty} I(m,t)$ where

$$I(m,t) = \frac{m^{k-1}}{C_k} \int_{\Gamma\backslash\mathfrak{h}} \sum_{\substack{ad-bc=m \\ a+d=t}} \frac{y^k}{(c|z|^2 + d\overline{z} - az - b)^k} \frac{dxdy}{y^2}.$$

Then he proves that

$$\frac{1}{2}I(m,t) + I(m,-t) = \begin{cases} -\frac{1}{2}P_k(t,m)H(4m-t^2) & \text{if } t^2-4m<0, \\ \frac{k-1}{24}m^{\frac{k-2}{2}} - \frac{1}{4}m^{\frac{k-1}{2}} & \text{if } t^2-4m=0, \\ -\frac{1}{2}\left(\frac{|t|-u}{2}\right)^{k-1} & \text{if } t^2-4m=u^2, u>0, \\ 0 & \text{otherwise} \end{cases}$$

by elementary counting. This implies the trace formula. □

21. Lecture 21 (November 26, 2013)

21.1. **Jacquet–Langlands Correspondence.** Today I will talk about the Jacquet–Langlands correspondence. In 1970 Jacquet and Langlands published their book *Automorphic Forms on* GL(2). Ninety percent of that book was already known by work of the Russians — the notion of automorphic representations was due to Gelfand, and even the tensor product theorem was basically proved by Piatetski-Shapiro. Jacquet–Langlands gave new proofs, which weren't very surprising except when they gave the correspondence at the end of their book. The idea was to compare two different groups and obtain a matching. That was what inspired Langlands to make his conjecture in the end.

Let me first talk about quaternion algebras. Fix integers $q, r \geq 1$ which are co-prime and square-free.

Definition 21.1 (Quaternion algebra)**.** $D[r,q] := \{x_0 + x_1J_1 + x_2J_2 + x_3J_3 : x_0, x_1, x_2, x_3 \in \mathbb{Q}\}$, where J_1, J_2, J_3 are quaternions satisfying

$$J_1^2 = q, \qquad J_2^2 = r, \qquad J_3^2 = -rq,$$
$$J_1J_2 = -J_2J_1, J_1J_3 = -J_3J_1, J_2J_3 = -J_3J_2,$$
$$J_1J_2 = J_3, \quad J_2J_3 = -rJ_1, \ J_3J_1 = -qJ_2.$$

Example 21.2. When $r = q = -1$, we get Hamilton's quaternions.

Definition 21.3 (Conjugate)**.** If $x = x_0+x_1J_1+x_2J_2+x_3J_3$, then $\overline{x} = x_0-x_1J_1-x_2J_2-x_3J_3$.

Definition 21.4 (Norm)**.** $N(x) = x \cdot \overline{x} = x_0^2 - qx_1^2 - rx_2^2 + rqx_3^2$.

Definition 21.5 (Trace)**.** $\operatorname{Tr}(x) = x + \overline{x}$.

We can realize the quaternion algebra $D[q,r]$ as a matrix group inside $M(2\times 2, \mathbb{Q}(\sqrt{q}, \sqrt{r}))$ by

$$1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$J_1 \mapsto \begin{pmatrix} \sqrt{q} & 0 \\ 0 & -\sqrt{q} \end{pmatrix},$$
$$J_2 \mapsto \begin{pmatrix} 0 & \sqrt{r} \\ \sqrt{r} & 0 \end{pmatrix},$$
$$J_3 \mapsto \begin{pmatrix} 0 & \sqrt{rq} \\ -\sqrt{rq} & 0 \end{pmatrix}.$$

In general, this can be described by the following formula.

Definition 21.6.

$$\phi(x_0 + x_1 J_1 + x_2 J_3 + x_3 J_3) = \begin{pmatrix} x_0 + x_1\sqrt{q} & \sqrt{r}(x_2 + x_3\sqrt{q}) \\ \sqrt{r}(x_2 - x_3\sqrt{q}) & x_0 - x_1\sqrt{q} \end{pmatrix}.$$

Then $\phi : D[q,r] \overset{\cong}{\to} M'(2 \times 2, \mathbb{Q}(\sqrt{r}, \sqrt{q})) \subset M(2 \times 2, \mathbb{Q}(\sqrt{r}, \sqrt{q}))$ is an algebra homomorphism.

Definition 21.7. A subring $\mathcal{O} \subset D[q,r]$ is called an order if $1 \in \mathcal{O}$ and, in addition, $\mathcal{O}$ is a free $\mathbb{Z}$-module of rank 4.

Definition 21.8. Let $\mathcal{O}$ be an order of $D[q,r]$. Then we define

$$\mathrm{disc}(\mathcal{O}) := \det(\mathrm{Tr}[\xi_j, \xi_k])$$

where $\{\xi_j\}$ is a $\mathbb{Z}$-basis for $\mathcal{O}$.

Definition 21.9. Let $\mathcal{O}$ be an order of $D[q,r]$. We define $\Gamma_{\mathcal{O}} \subset \mathrm{SL}(2,\mathbb{R})$ to be the set

$$\{\phi(x) : x \in \mathcal{O}, N(x) = 1\}.$$

This will turn out to be a discrete subgroup. This was studied intensely by Eichler, who proved the following

Theorem 21.10 (Eichler)**.** *Assume $D = D[q,r]$ is a division algebra (i.e. for all $a, b \in D - \{0\}$, there exists a unique $x \in D$ such that $a = bx$ and a unique $y \in D$ such that $ay = b$.) Let $\mathcal{O} \subset D[q,r]$ be an order. Let $\mathfrak{h} = \{x + iy : x \in \mathbb{R}, y > 0\}$ be the upper half plane. Then $\Gamma_{\mathcal{O}}$ is a finitely generated Fuchsian group where* $\mathrm{Vol}(\Gamma_{\mathcal{O}}\backslash\mathfrak{h}) < \infty$ *and $\Gamma_{\mathcal{O}}\backslash\mathfrak{h}$ is compact.*

In other words, there are no cusps! Hecke had intensely studied modular forms for $\Gamma_0(N)$, and Eichler was trying to generalize the theory to (quaternion) groups whose quotients are compact. Jacquet and Langlands were aware of this, and they started to look at the trace formula for these groups. Recall that when the quotient is compact, the trace formula has no continuous spectrum and is very simple.

Now I can state the "naive" version of the Jacquet–Langlands correspondence. Let $D[q,r]$ be a division algebra over $\mathbb{Q}$. Let $\mathcal{O}$ be an order of $D[q,r]$. To each Maass form $\phi \in \mathcal{L}^2(\Gamma_{\mathcal{O}}\backslash\mathfrak{h})$ satisfying

$$\Delta\phi = \lambda\phi$$

where $\Delta = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$ is the Laplacian, there exists an integer $N \geq 1$ and a Maass form $\Phi \in \mathcal{L}^2(\Gamma_0(N)\backslash\mathfrak{h})$ where $\Delta\Phi = \lambda\Phi$.

This is worked out in Hejhal's book, following Eichler who proved some special cases of this. It is remarkable that there is such a correspondence, since the groups $\Gamma_{\mathcal{O}}$ have compact quotient and the theory of Hecke operators is very different from usual.

Jacquet–Langlands worked out the trace formulae on $\mathcal{L}^2(\Gamma_{\mathcal{O}}\backslash\mathfrak{h})$ and $\mathcal{L}^2(\Gamma_0(N)\backslash\mathfrak{h})$ and looked for matching. This led to Langlands' functoriality conjectures, endoscopy and beyond endoscopy. The proof was local and involved the "non-principal series".

Let me now talk about the more general case.

Definition 21.11 (Essentially square-integrable local representations)**.** Let F_v be a local field. An irreducible admissible representation of $\mathrm{GL}(2, F_v)$ is essentially square-integrable if it is a twist of a local representation which is not a principal series.

Theorem 21.12 (Jacquet–Langlands for GL(2))**.** *Put $G = \mathrm{GL}(2)$ and G' the unit group of a quaternion algebra D over a field F. There is a one-to-one correspondence*

$$\{\textit{Automorphic representations of } G'\} \longleftrightarrow \left\{\begin{matrix}\textit{Automorphic representations of } G \textit{ which} \\ \textit{are essentially square-integrable at every} \\ \textit{place where } D \textit{ ramifies}\end{matrix}\right\}.$$

Further, if π' is an automorphic representation of G' and π is the corresponding automorphic representation of G, then π_v is completely determined by π'_v where π' ramifies at v.

This is essentially proved by the trace formula.

Let v be a place of F and F_v be its local field. Then there exists a unique quaternion algebra over F_v which is a division ring. Let K be an extension of F. Then $D_K = D \otimes_F K$ is again a quaternion algebra over K.

Definition 21.13. We say K splits D if $D_K \hookrightarrow M(2 \times 2, K)$.

Definition 21.14. (Ramified) D is ramified at v if F_v splits D. Otherwise it is unramified.

In the Jacquet–Langlands correspondence, the left hand side is really co-compact finite-dimensional representations with no parabolic conjugacy classes, but the right hand side has infinite-dimensional automorphic representations with parabolic conjugacy classes. This is something even the Russians didn't expect. Langlands was so struck by this correspondence that he kept thinking about this and was led to his entire program.

In their proof of the correspondence, Jacquet and Langlands show that if an irreducible automorphic representation of GL(2) has a supercuspidal local representation, then the contribution of the continuous spectrum to the trace formula is zero for suitable test functions. This idea was generalized by Deligne-Kazhdan leading to the simple trace formula (no continuous spectrum contribution).

22. Lecture 22 (December 5, 2013)

22.1. **Whittaker Transforms.** Today I will talk about the Whittaker functions on $\mathrm{GL}(n, \mathbb{R})$ at the archimedean place ∞.

Let $\mathfrak{h}^n$ be the generalized upper half plane defined by

$$\mathfrak{h}^n = \left\{ xy : x = \begin{pmatrix} 1 & & & & \\ & 1 & & x_{ij} & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix}, y = \begin{pmatrix} y_1 \cdots y_{n-1} & & & & \\ & \ddots & & & \\ & & y_1 y_2 & & \\ & & & y_1 & \\ & & & & 1 \end{pmatrix}, x_{ij} \in \mathbb{R}, y_j > 0 \right\}.$$

When $n = 2$, this is the classical upper half plane. $\mathrm{SL}(n, \mathbb{Z})$ acts on $\mathfrak{h}^n$ by left multiplication, and we have the Iwasawa decomposition

$$\mathrm{GL}(n, \mathbb{R})/(\mathrm{SO}(n, \mathbb{R}) \times \mathbb{R}^\times) \cong \mathfrak{h}^n.$$

We are interested in functions $\phi \in \mathcal{L}^2(\Gamma\backslash\mathfrak{h}^n)$, where $\Gamma \subset \mathrm{SL}(n, \mathbb{Z})$ is a congruence subgroup, satisfying:

- $\phi(\gamma z) = \phi(z)$ for all $\gamma \in \Gamma$ and $z \in \mathfrak{h}^n$;
- $\displaystyle\int_{\Gamma\backslash\mathfrak{h}^n} |\phi(z)|^2 d^\times z < \infty$;
- $D\phi = \lambda_D \phi$ for all invariant differential operators D in the center of the universal enveloping algebra, where $\lambda_D \in \mathbb{C}$.

Such a function ϕ is called an automorphic form.

Let $U_n = \left\{ \begin{pmatrix} 1 & * & * \\ & \ddots & * \\ & & 1 \end{pmatrix} \right\}$, and $\psi_{m_1,\cdots,m_{n-1}} : U_n \to \mathbb{C}^\times$ be the character defined by

$$\psi_{m_1,\cdots,m_{n-1}} \left(\begin{pmatrix} 1 & u_{1,2} & & & \\ & 1 & u_{2,3} & & \\ & & \ddots & \ddots & \\ & & & 1 & u_{n-1,n} \\ & & & & 1 \end{pmatrix} \right) = e^{2\pi i (m_1 u_{1,2} + \cdots + m_{n-1} u_{n-1,n})}.$$

We define the $(m_1, \cdots, m_{n-1})$-th Fourier coefficient of ϕ to be

$$\phi_{m_1,\cdots,m_{n-1}}(z) := \int_0^1 \cdots \int_0^1 \phi(uz) \overline{\psi_{m_1,\cdots,m_{n-1}}(u)} du$$

where $u = \begin{pmatrix} 1 & & u_{ij} \\ & \ddots & \\ & & 1 \end{pmatrix}$ and $du = \prod du_{ij}$. Then $\phi_{m_1,\cdots,m_{n-1}}$ satisfies the following properties:

- $\displaystyle\int_{\Gamma\backslash\mathfrak{h}^n} |\phi_{m_1,\cdots,m_{n-1}}(z)|^2 d^\times z < \infty$;
- $D\phi_{m_1,\cdots,m_{n-1}} = \lambda_D \phi_{m_1,\cdots,m_{n-1}}$ for all invariant differential operators D;
- $\phi_{m_1,\cdots,m_{n-1}}(uz) = \psi_{m_1,\cdots,m_{n-1}}(u) \phi_{m_1,\cdots,m_{n-1}}(z)$ for all $u \in U_n(\mathbb{R})$ and $z \in \mathfrak{h}^n$.

Definition 22.1. A Whittaker function for Γ acting on $\mathfrak{h}^n$ is a function $W : \mathfrak{h}^n \to \mathbb{C}$ satisfying:

(1) $\displaystyle\int_{\Gamma\backslash\mathfrak{h}^n} |W(z)|^2 d^\times z < \infty$;
(2) $DW = \lambda_D W$ for all invariant differential operators D;
(3) $W(uz) = \psi_{m_1,\cdots,m_{n-1}}(u)W(z)$ for all $u \in U_n(\mathbb{R})$ and $z \in \mathfrak{h}^n$.

Condition (2) is a second-order differential equation, so there are two solutions, but only one of them will be in $\mathcal{L}^2$, i.e. condition (1). Thus the Whittaker function is unique if it exists.

Theorem 22.2 (Shalika, Multiplicity one)**.** *Fix a character $\psi_{m_1,\cdots,m_{n-1}}$. There exists (up to a constant multiple) a unique Whittaker function W satisfying (1), (2) and (3).*

The proof of this is quite difficult. Diaconu and I found a simple proof for GL(3). Our idea was to take the Mellin transform of this Whittaker function.

Multiplicity one is important because it gives a unique Fourier coefficient at every place, which can then be used to construct L-functions.

We state the following theorem for $\Gamma = \mathrm{SL}(n,\mathbb{Z})$ and $n \geq 3$ only. There is a more complicated statement for general congruence subgroups.

Theorem 22.3 (Piatetski-Shapiro–Shalika)**.** *Let $\phi \in \mathcal{L}^2(\Gamma\backslash\mathfrak{h}^n)$ be an automorphic form. Let $W(z)$ be a fixed Whittaker function associated to $\psi_{1,\cdots,1}$ for Γ acting on $\mathfrak{h}^n$. Then*

$$\phi(z) = \sum_{\gamma\in U_{n-1}(\mathbb{Z})\backslash \mathrm{SL}(n-1,\mathbb{Z})} \sum_{m_1=1}^{\infty} \cdots \sum_{m_{n-2}=1}^{\infty} \sum_{m_{n-1}\neq 0} A(m_1,\cdots,m_{n-1}) W\left(M \begin{pmatrix} \gamma & \\ & 1 \end{pmatrix} z\right)$$

where $M =$.

A proof of this can be found in my book *Automorphic Forms and L-functions for the Group GL(n,R)*. An adelic proof is given in Cogdell's notes, and details are given in my book with Hundley.

When $m = 2$, the first sum does not appear and the expansion is

$$\phi(z) = \sum_{m\neq 0} A(m)W(mz)$$

where $z = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}$ and $W(z) = \sqrt{y}K_{i\nu}(2\pi y)e^{2\pi i x}$ is the K-Bessel function.

Now we talk about Jacquet's Whittaker functions, which he constructed in his thesis for the p-adic places but the construction also works at the archimedean places. The key point is that $I_{s_1,\cdots,s_{n-1}}(z) := y_1^{s_1} y_2^{s_2} \cdots y_{n-1}^{s_{n-1}}$ (where $s_1,\cdots,s_{n-1} \in \mathbb{C}$) is always an eigenfunction of all invariant differential operators D. For example, when $n = 2$, we have

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

and

$$Dy^s = s(1-s)y^s.$$

If we simply define

$$W(z) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} I_s(uz)\overline{\psi(u)}du,$$

then this does not converge! But assuming it does, we have formally

$$\begin{aligned} W(u_1 z) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} I_s(uu_1 z)\overline{\psi(u)} du \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} I_s(uz)\overline{\psi(uu_1^{-1})} du \\ &= \psi(u_1)W(z) \end{aligned}$$

for $u_1 \in U_n$ and $z \in \mathfrak{h}^n$.

Definition 22.4. Let $s = (s_1, \cdots, s_{n-1})$. Define

$$W_s(z) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} I_s(w_0 uz)\overline{\psi(u)} du$$

where $w_0 = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}$, $u = \begin{pmatrix} 1 & & u_{ij} \\ & \ddots & \\ & & 1 \end{pmatrix}$ and $du = \prod du_{ij}$.

Putting in the long element w_0 of the Weyl group guarantees convergence.

Example 22.5 ($n = 2$)**.** Let $z = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix}$. Then

$$\begin{aligned} W_s(z) &= \int_{-\infty}^{\infty} I_s\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} uz\right) e^{-2\pi i ux} du \\ &= \int_{-\infty}^{\infty} I_s\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y & x+\nu \\ 0 & 1 \end{pmatrix}\right) e^{-2\pi i \nu x} d\nu \\ &= \int_{-\infty}^{\infty} \left(\frac{y}{(x+\nu)^2 + y^2}\right)^s e^{-2\pi i \nu x} d\nu \end{aligned}$$

is essentially the K-Bessel function $\sqrt{y}K_{it}(y)e^{2\pi i x}$, which converges absolutely for $\mathrm{Re}(s) > 1$.

It turns out that if we use other elements of the Weyl group, we get the degenerate Whittaker functions.

Definition 22.6 (Jacquet's Whittaker function)**.** Fix $n \geq 2$. Let $\nu = (\nu_1, \cdots, \nu_{n-1}) \in \mathbb{C}^{n-1}$. Define

$$I_\nu(z) = \prod_{i=1}^{n-1}\prod_{j=1}^{n-1} y_i^{b_{ij}\nu_j}$$

where $z = xy \in \mathfrak{h}^n$ and $b_{ij} = \begin{cases} ij & \text{if } i + j \leq n, \\ (n-i)(n-j) & \text{otherwise.} \end{cases}$ Let $v_{j,k} = \sum_{i=0}^{j-1} \frac{n\nu_{n-k+i} - 1}{2}$. We define

$$W(z) = \prod_{j=1}^{n-1} \prod_{j \leq k \leq n-1} \frac{\Gamma(\frac{1}{2} + v_{j,k})}{\pi^{\frac{1}{2} + v_{j,k}}} \int_{U_n(\mathbb{R})} I_\nu(w_0 uz)\overline{\psi(u)} d^\times u.$$

Writing $\nu_j = \frac{1}{n} + it_j$, we set $t = (t_1, \cdots, t_{n-1})$.

For $z = xy = \begin{pmatrix} 1 & & & \\ & 1 & x_{ij} & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} y_1 \cdots y_{n-1} & & & & \\ & y_1 \cdots y_{n-2} & & & \\ & & \ddots & & \\ & & & y_1 & \\ & & & & 1 \end{pmatrix}$, we define $d^\times z = d^\times x \cdot d^\times y$, where

$$d^\times x = \prod dx_{ij}$$

and

$$d^\times y = \prod_{k=1}^{n-1} y_k^{-k(n-k)} \frac{dy_k}{y_k}.$$

Definition 22.7 (Whittaker transform)**.** Let $f : \mathbb{R}_+^{n-1} \to \mathbb{C}$. We define the Whittaker transform

$$f^{\#}(t) := \int_{\mathbb{R}_+^{n-1}} f(y) W_{it}(y) d^\times y \tag{*}$$

where $t = (t_1, \cdots, t_{n-1})$.

Theorem 22.8 (Kontorovich–Goldfeld, Whittaker transform inversion formula)**.** *Let $f : \mathbb{R}_+^{n-1} \to \mathbb{C}$ be smooth of compact support and $f^{\#} : \mathbb{R}_+^{n-1} \to \mathbb{C}$ be as in (*). Then*

$$f(y) = \frac{1}{\pi^{n-1}} \int_{\mathbb{R}^{n-1}} f^{\#}(t) W_{-it}(y) \frac{dt}{\prod_{1 \leq k \neq l \leq n} \Gamma(\frac{\alpha_k \cdots \alpha_l}{2})}$$

where α_i are defined by

$$\frac{k(n-k)}{2} + \sum_{l=1}^{n-k} \frac{\alpha_l}{2} = \sum_{k=1}^{n-1} b_{kl} \nu_l$$

and

$$\alpha_n = -\sum_{k=1}^{n-1} \alpha_k.$$

Here $\frac{dt}{\prod_{1 \leq k \neq l \leq n} \Gamma(\frac{\alpha_k \cdots \alpha_l}{2})}$ is the Plancherel measure for GL(n), whose existence was proved by Harish-Chandra for all reductive Lie groups.