Automorphic Forms and Automorphic Representations

Wee Teck Gan

References:

- A. Borel, *Automorphic forms on* $SL_2(\mathbf{R})$
- D. Bump, *Automorphic forms and representations*
- S. Gelbart, *Automorphic forms on adele groups*

Introduction

About ninety years ago, Ramanujan considered the following power series of q

$$\Delta(q) = q \cdot \prod_{n \geq 1} (1 - q^n)^{24}.$$

Expanding this out formally, we have:

$$\begin{aligned} \Delta(q) &= \sum_{n > 0} \tau(n) q^n \\ &= q - 24q^2 + \ldots \end{aligned}$$

Ramanujan made a number of conjectures about the coefficients $\tau(n)$. These conjectures have turned out to be very influential. They say:

- τ is multiplicative, i.e. if m and n are relatively prime, then

$$\tau(mn) = \tau(m) \cdot \tau(n)$$

Moreover,

$$\tau(p^{k+1}) = \tau(p)\tau(p^k) - p^{11}\tau(p^{k-1}).$$

- For all primes p, $|\tau(p)| \leq 2 \cdot p^{11/2}$. This implies that for any n,

$$|\tau(n)| \leq C_\epsilon \cdot n^{11/2+\epsilon}$$

for any ϵ.

The first conjecture was proved by Mordell (around 1920), while the second by Deligne (around 1970).

These conjectures led to the theory of modular forms. We shall begin with a brief description of the basic results in this theory, and then give a reformulation using representation theory. This reformulation leads to a vast generalization of the theory.

Contents:

- Classical modular forms

- Automorphic forms on real groups

- Automorphic representations of adele groups

- Eisenstein series

- Multiplicity One for $GL(n)$

Classical Modular Forms

Upper Half Plane

Let

$$\mathbf{H} = \{x + iy : x \in \mathbf{R}, y > 0\}$$

be the upper half plane. It is a homogeneous space for $SL_2(\mathbf{R})$ under the action:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z \mapsto \frac{az+b}{cz+d}.$$

In fact, this defines an action of $GL_2(\mathbf{R})^+$ on $\mathbf{H}$ with center acting trivially. Moreover, it extends to an action on $\mathbf{H}^* = \mathbf{H} \cup \mathbf{R} \cup \{\infty\}$. The action on $\mathbf{R} \cup \{\infty\}$ is also transitive. The stabilizer of ∞ is the Borel subgroup B of upper triangular matrices.

The stabilizer of $i = \sqrt{-1}$ is

$$K = SO(2) = \{g \in SL_2(\mathbf{R}) : g^t g = 1\}$$

which is a maximal compact subgroup of $SL_2(\mathbf{R})$. So we have:

$$\mathbf{H} \cong SL_2(\mathbf{R})/K.$$

There is an $SL_2(\mathbf{R})$-invariant measure on $\mathbf{H}$, namely

$$\frac{dx\,dy}{y^2}.$$

This is invariant under $SL_2(\mathbf{R})$ because

$$Im(gz) = \frac{Im(z)}{|cz+d|^2}.$$

Standard notations

- $B =$ the Borel subgroup of upper triangular matrices

- $N =$ the unipotent radical of $B = \{\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}\}$

- $T =$ the group of diagonal matrices

- $A = \{\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} : a > 0\}$

- $K = SO(2)$, the maximal compact.

Arithmetic Subgroups

If Γ is an arithmetic subgroup of $SL_2(\mathbf{R})$, then Γ acts on $\mathbf{H}$ in a properly discontinuous fashion. The quotient $\Gamma\backslash\mathbf{H} = \mathbf{\Gamma\backslash SL_2(R)/K}$ possesses a fundamental domain $\mathcal{F}$ and has finite volume (since $\Gamma\backslash SL_2(\mathbf{R})$ has finite volume).

An example of Γ is $SL_2(\mathbf{Z})$. A fundamental domain for the discrete subgroup $SL_2(\mathbf{Z})$ is:

$$\mathcal{F} = \{z = x + iy : |x| <\leq 1/2, |z| \geq 1\}$$

It looks like:

Cusps

As a Riemann surface, $\mathcal{F}$ is a punctured sphere. It has a natural compactification: by adding the point $i\infty$. This extra point is called the **cusp** at infinity.

More formally, a point $x \in \mathbf{R} \cup \{\infty\}$ is cuspidal for Γ if the stabilizer of x in Γ contains non-trivial unipotent elements. For the purpose of this lecture, our Γ is always contained in $SL_2(\mathbf{Q})$, in which case the cuspidal points are simply $\mathbf{Q} \cup \{\infty\}$.

A **cusp of** Γ is a Γ-orbit in $\mathbf{Q} \cup \{\infty\}$. Because $SL_2(\mathbf{Z})$ acts transtively on $\mathbf{Q}\cup\infty = SL_2(\mathbf{Q})/B(\mathbf{Q})$, there is one cusp when $\Gamma = SL_2(\mathbf{Z})$.

More generally, the number of cusps of Γ is $\#\Gamma\backslash SL_2(\mathbf{Q})/B(\mathbf{Q})$ (which is finite), and $\Gamma\backslash\mathbf{H}$ can be compactified by adding these cusps:

$$\overline{\Gamma\backslash\mathbf{H}} = \Gamma\backslash(\mathbf{H} \cup \mathbf{Q} \cup \{\infty\}).$$

Siegel sets

It is often useful to have a set which covers the fundamental domain, but is easier to describe. One such example is the **Siegel set** associated to the cusp $i\infty$:

$$\mathfrak{S}_{c,d} = \{x + iy : |x| < c, y > d\}$$

For $\Gamma = SL_2(\mathbf{Z})$, if c is large enough and d is small enough, this set will cover $\mathcal{F}$. The volume of $\mathfrak{S}_{c,d}$ (with respect to the invariant measure) is easily computed and seen to be finite. This shows that $\mathcal{F}$ has finite volume.

Hecke congruence subgroups

Another example of arithmetic group is the Hecke congruence subgroup:

$$\Gamma_0(N) = \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : c \equiv 0 \mod N \}.$$

Here the fundamental domain may have more than one cusp. For example, when $N = 2$, the fundamental domain, as shown below, has 3 cusps.

In this case, the fundamental domain cannot be covered by a single Siegel set. One needs 3 Siegel sets, one for each cusp.

In the following, Γ shall mean $\Gamma_0(N)$ for some N.

Modular Forms

Definition: A holomorphic modular form for Γ is a holomorphic function f on $\mathbf{H}$ satisfying some extra properties:

- (automorphy) for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$,

  $$f(\gamma z) = (cz + d)^k \cdot f(z),$$

  where k is a positive integer.

- (holomorphy) f is holomorphic at every cusp of Γ.

One motivation for studying these type of functions comes from the theory of elliptic curves, but we will not go into this here.

Notations: For $g \in GL_2(\mathbf{R})^+$ and $z \in \mathbf{H}$, set

$$j(g, z) = (cz + d) \cdot det(g)^{-1/2}$$

$$(f|_k g)(z) = j(g, z)^{-k} \cdot f(gz).$$

Then the automorphy condition can be expressed as:

$$f|_k \gamma = f$$

for any $\gamma \in \Gamma$.

The integer k is called the **weight** of f, whereas if $\Gamma = \Gamma_0(N)$, then N is the **level** of f.

Observe that f is necessarily 0 if k is odd (because $-1 \in \Gamma$).

Next we want to explain more precisely the holomorphy condition.

Fourier Expansion

Because the element

$$t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

lies in Γ, we have:

$$f(z+1) = f(z).$$

Thus f is really a function on the strip

$$\{x + iy : -1/2 \leq x < 1/2, y > 0\}.$$

The map $z \mapsto q = e^{2\pi i z}$ sends this strip onto the punctured open unit disc, and sends the cusp $i\infty$ to 0.

The function f gives rise to a holomorphic function $\tilde{f}(q)$ on the punctured disc, and $\tilde{f}(q)$ has a Laurent expansion about 0:

$$\tilde{f}(q) = \sum_n a_n q^n.$$

By "f is holomorphic at the cusp $i\infty$", we mean that the singularity at 0 is removable, so that $a_n = 0$ if $n < 0$.

Thus we can expand f as a Fourier series:

$$f(z) = \sum_{n \geq 0} a_n(f) e^{2\pi i n z}.$$

The numbers $\{a_n(f)\}$ are the Fourier coefficients of f (at the cusp $i\infty$).

One has analogous Fourier expansion at every cusp of Γ. More precisely, any cusp $a \in \mathbf{Q}$ can be moved to ∞ by an element γ of $SL_2(\mathbf{Z})$, and one can consider the Fourier expansion of $f|_k\gamma$ at $i\infty$ as above.

CUSP FORMS: f is a cusp form (or is cuspidal) if f vanishes at every cusp.

Thus f is cuspidal iff the zeroth Fourier coefficient $a_0(f)$ in the Fourier expansion of f at every cusp is zero.

Moderate growth and rapid decrease

In the presence of holomorphy of f on $\mathbf{H}$, the holomorphy condition at a cusp is implied by a weaker assumption, namely that of moderate growth.

We say that f is of **moderate growth** at the cusp $i\infty$ if there exists n such that

$$|f(x+iy)| \leq C \cdot y^n$$

as $y \to \infty$ with z in a Siegel set for $i\infty$ (i.e. as $y \to \infty$ with x bounded).

Say that f is **rapidly decreasing** at $i\infty$ if, for any k, there exists C_k such that

$$|f(x+iy)| \leq C_k y^{-k}$$

as $y \to \infty$ in a Siegel set for $i\infty$.

Because $|e^{2\pi ikz}| = e^{-2\pi ky}$, we see that if f is holomorphic on $\mathbf{H}$ and satisfies automorphy condition, then

- f is of moderate growth at $i\infty$ iff f is holomorphic at $i\infty$.

- f is rapidly decreasing at $i\infty$ iff f vanishes at $i\infty$.

Natural Question: Suppose we are given a q-series $f(z) = \sum_{n\geq 0} a_n q^n$ which converges for $z \in \mathbf{H}$. Under what conditions can we conclude that $f(z)$ is a modular form, with respect to Γ? Such a result is called a **converse theorem**. We will come to this type of result later.

Remarks: We know that Γ is finitely generated. For example, $SL_2(\mathbf{Z})$ is generated by the following two elements:

$$t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad\qquad w = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

So to check that a given Fourier series is modular with respect to Γ, it suffices to check the transformation property for w. But this is not apparent from the Fourier series at all!

Finite Dimensionality

Let $M_k(N)$ denote the space of modular forms of weight k and level N, and let $S_k(N)$ be the subspace of cusp forms. These spaces are finite-dimensional. For example:

$$dim M_k(1) = \begin{cases} \frac{k}{12} + 1, \text{ if } k \neq 2 \mod 12, \\ \frac{k}{12}, \text{ if } k = 2 \mod 12. \end{cases}$$

Ring structure

Observe that if $f_i \in M_{k_i}(N)$, then $f_1 \cdot f_2 \in M_{k_1+k_2}(N)$. Thus $\bigoplus_k M_k(N)$ has a ring structure. Moreover, if one of f_i's is cuspidal, so is $f_1 \cdot f_2$.

Examples

(i) **Eisenstein series E_{2k}.**

$$E_{2k}(z) = \frac{1}{2\zeta(2k)} \sum_{m,n\in\mathbf{Z},(m,n)\neq(0,0)} (mz+n)^{-2k}.$$

This converges absolutely for all $z \in \mathbf{H}$ provided that $k \geq 2$. We claim that this is a modular form of weight $2k$ and level 1. Given $g \in SL_2(\mathbf{Z})$, we have:

$$(E_{2k}|_{2k}g)(z) = (cz+d)^{-2k} \cdot \sum \left(m(\frac{az+b}{cz+d}) + n \right)^{-2k}$$

$$= \sum \left((ma+nc)z + (mb+nd)\right)^{-2k}.$$

Since $g \in SL_2(\mathbf{Z})$, $(m,n) \mapsto (ma+nc, mb+nd)$ is a bijection of $\mathbf{Z}^2$ onto itself. Thus the last expression is equal to E_{2k}. This shows the automorphy condition.

For the holomorphy at $i\infty$, we need to find the Fourier expansion of E_{2k}. It turns out that

$$E_{2k}(z) = 1 + \frac{(-1)^k 4k}{B_{2k}} \cdot \sum_{n \geq 1} \sigma_{2k-1}(n) q^n$$

where B_{2k} is the $2k$-th Bernoulli number and $\sigma_{2k-1}(n)$ is the sum of the $(2k-1)$-powers of the divisors of n.

As a corollary, we see that for $k > 2$ even,

$$M_k(1) = \mathbf{C} \cdot E_k \oplus S_k(1)$$

$$dim M_k(1) = dim S_k(1) + 1.$$

So $\dim S_k(1) = 0$ if $k < 12$ and $\dim S_{12}(1) = 1$.

(ii) **Ramanujan Δ function.** The easiest way to construct cusp forms is to use linear combinations of Eisenstein series. For example, let us set

$$f(z) = E_4(z)^3 - E_6(z)^2.$$

One computes to see that its Fourier expansion looks like:

$$f(z) = \frac{1}{1728}(q + 24q^2 + ...);$$

so it is indeed a cusp form of weight 12 and level 1.

It turns out that this cusp form is precisely equal to a multiple of the Ramanujan Δ-function:

$$q \prod_{n \geq 1} (1 - q^n)^{24} = q + 24q^2 + ...$$

Ramanujan showed that this is an element of $S_{12}(1)$. Because this latter space has dimension 1, and so this is equal to our definition given above.

(iii) **Theta functions.** Another source of modular forms is the so-called theta functions. Let $A = (a_{ij})$ be a positive definite symmetric matrix with integer entries. Assume for simplicity that A has determinant 1 and A is even, i.e. $x^t Ax$ is even for all $x \in \mathbf{Z}^k$. Define a quadratic form by

$$Q(x) = \frac{1}{2} x^t Ax$$

and consider the series

$$\theta_Q(z) = \sum_{x \in \mathbf{Z}^k} q^{Q(x)} =$$

$$= \sum_{n \geq 0} a_n(Q) q^n$$

where

$$a_n(Q) = \#\{x \in \mathbf{Z}^k : Q(x) = n\}.$$

Then $\theta_Q(z)$ is a modular form of weight $k/2$ and level 1. This is a consequence of the Poisson summation formula.

Bounds on Fourier Coefficients

We have the following trivial bound of Hecke.

Proposition: If $f(z) = \sum_{n>0} a_n(f) e^{2\pi i n z}$ is a cusp form of weight k, then

$$|a_n(f)| \leq C_f \cdot n^{k/2}.$$

Proof: Consider the function $F(z) = y^{k/2}|f(z)|$, if $z = x + iy$. Then $F(\gamma z) = F(z)$ for any $\gamma \in \Gamma$. Thus, F is a continuous function on $\Gamma \backslash \mathbf{H}$. Since f is cuspidal, $F(z)$ tends to 0 as z approaches the cusps. Thus, F is bounded and we have:

$$|f(x+iy)| \leq C_f \cdot y^{-k/2}.$$

Now, for fixed y,

$$|a_n(f) \cdot e^{2\pi n y}| = |\int_0^1 e^{-2\pi i n x} \cdot f(z)\, dx| \leq$$

$$\leq \int_0^1 |f(x+iy)|\,dx.$$

Thus

$$|a_n(f)| \leq C_f e^{-2\pi n y} y^{-k/2}$$

and putting $y = 1/n$ gives the result.

Ramanujan-Petersson Conjecture: Let f be a cuspidal Hecke eigenform form of weight k. Then for all primes p,

$$|a_p(f)| \leq 2 \cdot p^{\frac{k-1}{2}}.$$

This implies that

$$|a_n(f)| \leq C_\epsilon \cdot n^{\frac{k-1}{2}+\epsilon}$$

for any $\epsilon > 0$.

This conjecture has been proved by Deligne in 1971 as a consequence of his proof of the Weil conjectures. This bound on Fourier coefficients have many number theoretic applications.

Petersson Inner Product

The space $S_k(N)$ comes equipped with a natural inner product:

$$\langle f_1, f_2 \rangle_k = \int_{\Gamma \backslash \mathbf{H}} f_1(z)\overline{f_2(z)} y^k \cdot \frac{dx\,dy}{y^2}.$$

This is convergent because $y^{k/2} f_i(z)$ tends to zero at the cusps. It remains convergent as long as one of the functions is cuspidal. Thus it makes sense to take the inner product of a cusp form with any modular form.

L-functions and Hecke theory

Given a cusp form of weight k,

$$f(z) = \sum_{n \geq 0} a_n(f) e^{2\pi i n z},$$

one may consider the Dirichlet series:

$$L(f, s) = \sum_{n \geq 1} \frac{a_n(f)}{n^s}.$$

This converges for $Re(s) > \frac{k}{2} + 1$, because $|a_n(f)| = O(n^{k/2})$. It is related to f by a Mellin transform. Indeed,

$$\int_0^\infty f(iy) \cdot y^s \cdot \frac{dy}{y}$$

$$= \int_0^\infty \sum_{n \geq 1} a_n(f) e^{-2\pi n y} y^s \cdot \frac{dy}{y}$$

$$= \sum_n a_n(f) \cdot \int_0^\infty e^{-t} (2\pi n)^{-s} t^s \cdot \frac{dt}{t} \quad (t = 2\pi n y)$$

$$= (2\pi)^{-s}\Gamma(s) \cdot \sum_n \frac{a_n(f)}{n^s}$$

where

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \cdot \frac{dt}{t}$$

is the so-called Gamma function, which satisfies $\Gamma(z+1) = z\Gamma(z)$ and $\Gamma(n+1) = n!$.

Here is an important theorem about the L-functions attached to cusp forms (for simplicity, we state it for level 1 forms only):

Theorem: Let f be a cusp form of weight k of level 1. Then we have:

(i) (analytic continuation) $L(s, f)$ extends to an entire function on $\mathbf{C}$.

(ii) (functional equation) The function $\Lambda(s, f) = (2\pi)^{-s}\Gamma(s)L(s, f)$ satisfies the functional equation

$$\Lambda(s, f) = (-1)^{k/2}\Lambda(k - s, f).$$

(iii) (boundedness in vertical strips) $\Lambda(s,f)$ is bounded in vertical strips.

proof: When $Re(s)$ is large, we have:

$$\Lambda(s,f) = \int_0^\infty f(iy) \cdot y^s \cdot \frac{dy}{y}.$$

But the RHS is convergent for all s (and thus gives (i)). This is because:

- $f(iy)$ is exponentially decreasing in y as $y \to \infty$, since f is cuspidal.

- since f is modular with respect to w,

  $$f(iy) = (-1)^{k/2} y^{-k} f(i/y).$$

  So as $y \to 0$, $f(iy) \to 0$ faster than any power of y.

To see (ii), note that

$$\begin{aligned}
&\int_0^\infty f(iy) \cdot y^s \cdot \frac{dy}{y} \\
=&\int_0^\infty (-1)^{k/2} y^{-k} f(i/y)\, y^s \,\frac{dy}{y} \\
=&(-1)^{k/2} \cdot \int_0^\infty f(it)\, t^{k-s} \cdot \frac{dy}{y} \qquad\qquad (t = 1/y).
\end{aligned}$$

This proves (ii), and (iii) is clear.

The L-function of f turns out to be a very important invariant of f.

Hecke Operators

The theory of Hecke operators explains why the Fourier coefficients of certain modular forms are multiplicative functions. Let us assume that $\Gamma = SL_2(\mathbf{Z})$ for simplicity. For each positive integer n, we are going to define a Hecke operator T_n which is a linear operator on M_k preserving S_k.

A general construction: Let $g \in GL_2^+(\mathbf{Q})$ and write the double coset $\Gamma g \Gamma$ as a union of single cosets:

$$\Gamma g \Gamma = \bigcup_i \Gamma a_i.$$

Here the union is over a finite indexing set and the a_i's are in $GL_2^+(\mathbf{Q})$. We then set

$$f|_k[g] = \sum_i f|_k a_i.$$

This defines an operator $M_k(\Gamma) \to M_k(\Gamma)$. It is independent of the choice of the a_i's.

<u>**Hecke operator T_n:**</u> Now let $M(n)$ be the integral 2×2 matrices with determinant n By the theory of elementary divisors,

$$M(n) = \bigcup_{d|a, ad=n} \Gamma\, t(a,d)\, \Gamma$$

where

$$t(a,d) = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

We set

$$f|T_n = n^{k/2-1} \sum_{d|a, ad=n} f|_k[t(a,d)].$$

For example, T_p is simply the operator defined by the double coset $t(p,1)$. More explicitly, because

$$M(p) = \Gamma \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \Gamma$$
$$= \bigcup_{k=0}^{p-1} \Gamma \begin{pmatrix} 1 & k \\ 0 & p \end{pmatrix} \cup \Gamma \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix},$$

we have:

$$(f|T_p)(z) = p^{k-1} f(pz) + \frac{1}{p} \sum_{k=0}^{p-1} f(\frac{z+k}{p}).$$

Proposition:

- (Effects on Fourier coefficients) we have:
$$a_n(T_p f) = a_{pn}(f) + p^{k-1} a_{n/p}(f)$$
where the second summand is interpreted to be 0 if p does not divide n. More complicated formulas exist for T_n.

- T_n preserves S_k.

- T_n is self-adjoint with respect to the Petersson inner product:
$$\langle f_1, T_n f_2 \rangle = \langle T_n f_1, f_2 \rangle.$$

- if $(n, m) = 1$, then $T_n T_m = T_{nm} = T_{mn}$. Moreover,
$$T_p T_{p^r} = T_{p^{r+1}} + p^{k-1} T_{p^{r-1}}.$$

Thus we see that the linear span of the T_n's form an algebra and this algebra is generated by T_p's with p prime. Moreover, this algebra is commutative.

As consequences of the above properties, we have:

Corollary:

- The action of $\{T_n\}$ on S_k can be simultaneously diagonalized.

- If f is a cuspidal Hecke eigenform with eigenvalues λ_n for T_n, then

$$a_n(f) = \lambda_n \cdot a_1(f).$$

  Thus if f is non-zero, then $a_1(f) \neq 0$, and we can normalize it by scaling to make $a_1(f) = 1$.

- **Multiplicity One Theorem**: If f is a normalized cuspidal eigenform, then f is completely determined by its Hecke eigenvalues.

- If f is a normalized cuspidal eigenform, then the Fourier coefficients of f are multiplicative, and satisfy:

$$a_p a_{p^r} = a_{p^{r+1}} + p^{k-1} a_{p^{r-1}}.$$

Euler products

The fact that the Fourier coefficients of f (a normalized cuspidal eigenform) are multiplicative implies that $L(f,s)$ has an Euler product (when $Re(s)$ is large):

$$L(f,s) = \prod_p (\sum_k a_{p^k} p^{-ks})$$

$$= \prod_p \frac{1}{1 - a_p(f)p^{-s} + p^{k-1-2s}}.$$

For example, Δ is a normalized cuspidal Hecke eigenform, and this explains the first of Ramanujan's conjectures on the coeffcients $\tau(n)$.

We say that an L-function defined by a Dirichlet series in some right half plane is **nice** if it has **analytic continuation, satisfies appropriate functional equation, is bounded in vertical strips and possesses an Euler product**. Thus $L(s,f)$ is nice if f is a cuspidal Hecke eigenform.

Hecke operators for $\Gamma_0(N)$

The above theory of Hecke works beautifully for modular forms with respect to $SL_2(\mathbf{Z})$, but there are complications when $\Gamma = \Gamma_0(N)$ ($N > 1$). We can still define the operators T_n as before, the algebra is still commutative and generated by all the T_p 's. But T_n is self-adjoint only if $(n, N) = 1$. So we can only simultaneously diagonalize the actions of T_n with $(n, N) = 1$.

If f happens to be an eigenfunction for all T_n's, then $L(f, s)$ will still have an Euler product as before. But now, we have no guarantee that $S_k(N)$ has a basis of this type.

This shows that the theory of modular forms can be quite sensitive to the group Γ.

Old and new forms

Another complication is that the eigenvalues of the T_p's, with $(p, N) = 1$, do not separate the forms.

Example: $\Delta(z)$ and $\Delta(2z)$ are both elements of $S_{12}(2)$, but have same eigenvalues for T_n for all odd n.

This is not so surprising since both these functions are built out of a single function Δ. They are so-called old forms:

Definition: Suppose that $m \cdot n$ divides N, then for $f \in M_k(m)$, the function $f(nz)$ is an element of $M_k(mn)$ and thus of $M_k(N)$. The subspace of $M_k(N)$ spanned by elements of this type is the space of **old forms**. Its

orthogonal complement in $S_k(N)$ is the space of **new forms**.

This definition is due to Atkin-Lehner.

Results of Atkin-Lehner

What Atkin and Lehner showed is that when one restricts to the space of new forms, then the neat results for $SL_2(\mathbf{Z})$ are restored. Namely,

- the action of ALL T_n's can be simultaneously diagonalized, so that the L-function of a cuspidal Hecke eigen-newform has an Euler product.

- we have the **multiplicity one theorem**: the newforms can be distinguished from one another by their eigenvalues with respect to the T_p's with $(p, N) = 1$.

Converse theorems.

We have seen that the L-function of a cuspidal Hecke eigen-newform is nice. Hecke showed the following converse:

Theorem: Suppose that $|a_n| = O(n^r)$ for some r so that $L(s) = \sum_{n>0} a_n n^{-s}$ converges absolutely when $Re(s)$ is large. If $L(s)$ has analytic continuation, is bounded in vertical strips and satisfies the functional equation

$$\Lambda(s) = (-1)^{k/2}\Lambda(k-s)$$

then $f(z) = \sum_{n>0} a_n q^n$ is a cusp form of weight k and level 1. If $L(f,s)$ has Euler product, then f is a Hecke eigenform.

The role of the functional equation is that it allows us to deduce automorphy with respect to w.

How does one characterize the cusp forms of level N? Weil proved an analog of the above theorem, but with a crucial "twist". More precisely, instead of just requiring $L(s)$ to satisfy a single functional equation, his theorem requires functional equat

ions for twists of $L(s)$ by various Dirichlet characters χ, i.e. for

$$L_\chi(s) = \sum_n \frac{a_n \cdot \chi(n)}{n^s}.$$

The reason why this extra functional equations are needed is because $\Gamma_0(N)$ usually have more than 2 generators! For the precise statement of the theorem and a proof, see [Bump, Thm. 1.5.1].

Maass Forms

Maass introduced certain analogs of modular forms which are not holomorphic. More precisely, one says that a function f on $\mathbf{H}$ is a Maass form with respect to Γ with paramater s if

- f is smooth;
- $f(\gamma g) = f(g)$ for any $\gamma \in \Gamma$;
- f is of moderate growth at the cusps of Γ;
- f is an eigenfunction for the hyperbolic Laplacian

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

with eigenvalue $\frac{1}{4} - s^2$.

From this, one can show that f is in fact real-analytic.

Say that f is cuspidal if it vanishes at the cusps of Γ. Then f is rapidly decreasing at the cusps and thus is bounded on $\Gamma\backslash\mathbf{H}$. In particular, a Maass cusp form belongs to $L^2(\mathbf{H})$.

Nonholomorphic Eisenstein series

An example of a Maass form is a nonholomorphic Eisenstein series, constructed as follows. Consider the function

$$\phi(z) = y^{\frac{1}{2}+s}.$$

This satisfies:

$$\Delta\phi = \left(\frac{1}{4} - s^2\right)\phi.$$

Moreover, because

$$\phi(\gamma z) = \frac{\phi(z)}{|cz+d|^{2s+1}}$$

we see that $\phi(\gamma z) = \phi(z)$ if $\gamma \in N(\mathbf{Z})$. If we set

$$E(z,s) = \sum_{\gamma \in N(\mathbf{Z})\backslash SL_2(\mathbf{Z})} \phi(\gamma z),$$

then this converges absolutely if $Re(s) > 1/2$. In that case, $E(z,s)$ is $SL_2(\mathbf{Z})$-invariant and $\Delta E(z,s) = (1/2 - s^2)E(z,s)$. Note that $E(z,s)$ is not cuspidal.

Fourier expansion

Assume that $\Gamma = SL_2(\mathbf{Z})$ for simplicity. Because $f(z+1) = f(z)$, we also have a Fourier expansion of f at the cusp $i\infty$. But because f is not holomorphic, the Fourier expansion is not as clean:

$$f(x+iy) = \sum_{n=-\infty}^{\infty} \alpha_n(y,s)e^{2\pi inx},$$

for some functions $\alpha_n(y,s)$. These functions are not random, because $\alpha_n(y,s)e^{2\pi inx}$ is also an eigenfunction of Δ with eigenvalue $1/4 - s^2$.

For example,

$$\alpha_0(y) = ay^{1/2+s} + by^{1/2-s}.$$

Moreover, f is cuspidal iff $\alpha_0(y) = 0$.

In general, it turns out that

$$\alpha_n(y) = a_n \cdot \sqrt{y} \cdot K_s(2\pi|n|y)$$

where a_n is a constant, and K_s is the normalized Bessel function.

L-functions

As in the holomorphic case, one can develop a Hecke theory by attaching to f an L-function. The map

$$\iota : x + iy \mapsto -x + iy$$

gives an involution on the space of Maass forms. We say that f is even (respectively odd) if $\iota f = f$ (respectively $-f$), in which case $a_{-n}(f) = a_n(f)$ (resp. $-a_n(f)$).

For an even or odd f, we set

$$L(s, f) = \sum_{n \geq 1} a_n n^{-s}$$

which converges when $Re(s) > 3/2$, because $a_n = O(n^{1/2})$. Then one can show that $L(s, f)$ has analytic continuation and functional equations. Moreover, there are also actions of Hecke operators and if f is a Hecke eigenform, then $L(s, f)$ has Euler product.

Existence of cusp forms

It is not easy to show the existence of cuspidal Maass forms. Selberg used the **trace formula** to show that many such functions exist. In fact, the trace formula give a count of these cuspidal functions (the so-called **Weyl's law**). Perhaps this will be covered in Labesse's lectures.

There are, however, not many explicit constructions. One such construction is due to Maass, but let's not go into this here. He showed that there is a Maass form whose L-function is the same as that of a Hecke character of a real quadratic extension of $\mathbf{Q}$. This was the initial question which led to the discovery/definition of Maass forms. For more details, see [Bump, Thm. 1.9.1].

Selberg's conjecture

Because Δ is a positive definite operator (on $L^2(\Gamma\backslash\mathbf{H})$), the eigenvalue $\lambda = 1/4 - s^2$ is $>$ 0 if f is cuspidal. It implies that one of the following 2 situations occurs:

- s is real and $|s| < 1/2$, in which case, $\lambda < 1/4$;

- s is purely imaginary, in which case $\lambda \geq 1/4$.

Selberg proved that when $\Gamma = SL_2(\mathbf{Z})$, the first possibility never occurs, so that $\lambda \geq 1/4$ always. He conjectured that the same holds for any congruence subgroups.

Selberg's Conjecture:

For any congruence subgroup Γ, if f is cuspidal with respect to Γ with parameter s, then s is purely imaginary and $\lambda \geq 1/4$.

It is known that this is not true for some non-congruence Γ.

Later, when we reformulate everything in terms of representation theory, we shall see that the Selberg conjecture and the Ramanujan conjecture (on size of Fourier coefficients) are basically the same phenomenon.

Automorphic Forms on Real Groups

GOAL: to reformulate the theory of modular forms and Maass forms in a single framework, which is susceptible to generalization to general reductive groups.

A classical modular form f is a function on $SL_2(\mathbf{R})/K$, which is "quasi-invariant" on the left under Γ with respect to the factor of automorphy

$$j(g, z) = (cz + d) \cdot det(g)^{-1/2}.$$

One can generalize this to certain general group G (in place of SL_2), namely those real semisimple G such that the symmetric space G/K has a complex structure. In that case, G/K is a **hermitian symmetric domain**.

An example is the symplectic group $G = Sp_{2n}$, where

$$G/K = \{Z = X + iY \in M_n(\mathbf{C}) : Z^t = Z, Y > 0\}$$

is the so-called Siegel upper half space. In this case, one has the theory of **Siegel modular forms**, with

$$j(g, Z) = CZ + D, \qquad g \in Sp_{2n}(\mathbf{R}), \ Z \in G/K$$

if

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

However, it is not clear what is the analog of this for general groups, e.g. $G = SL_n$ ($n \geq 3$).

The key insight in the refomulation is to transform f to a function which is left-invariant by Γ and only "quasi-invariant" on the right by K. This can be achieved by setting

$$\phi_f(g) = (f|_k g)(i)$$

for $g \in SL_2(\mathbf{R})$ and k is the weight of f. Thus we will consider ϕ_f rather than f and for general G, an automorphic form will be a function on $\Gamma\backslash G$ for an arithmetic subgroup of G, satisfying some extra properties. We shall explain what are these properties. Then we shall explain how representation theory enters the picture.

Let's examine some properties of $\phi = \phi_f$.

Γ-invariance and K-finiteness

- ϕ is a smooth function.

- $\phi(\gamma g) = \phi(g)$ for any $\gamma \in \Gamma$. This is because
$$(f|\gamma g)(i) = ((f|\gamma)|g)(i) = (f|g)(i)$$
Thus f is a function on $\Gamma \backslash SL_2(\mathbf{R})$. (Note that $SL_2(\mathbf{R})$ acts on the space of functions on $\Gamma \backslash SL_2(\mathbf{R})$ by right translation).

- $\phi(gr_\theta) = e^{ik\theta} \cdot \phi(g)$ where
$$r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$
is a typical element in K. This is because:
$$(f|gr_\theta)(i) = (-i\sin\theta + \cos\theta)^{-k} \cdot (f|g)(i).$$
Thus the right K-translates of ϕ span a one-dimensional vector space. In particular, ϕ_f is right K-finite.

Holomorphy on $\mathbf{H}$

What does the holomorphy condition on $\mathbf{H}$ and the cusps translate to? To explain this, need some notations. To say that f is holomorphic on $\mathbf{H}$ means that

$$\frac{\partial f}{\partial \bar{z}} = 0$$

where

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

So we expect that this will translate to something like: "ϕ_f is killed by some differential operator".

Review on differential operators

One source of differential operators on smooth functions on $\Gamma\backslash SL_2(\mathbf{R})$ is the complexified Lie algebra $\mathfrak{sl}_2(\mathbf{C})$, acting by right infinitesimal translation: if $X \in \mathfrak{g}_0 = \mathfrak{sl}_2(\mathbf{R})$, then

$$(X\phi)(g) = \frac{d}{dt}\phi(g \cdot exp(tX))|_{t=0}.$$

This defines a left-invariant first-order differential operator on smooth functions on $SL_2(\mathbf{R})$. We extend this action to $\mathfrak{g} = \mathfrak{sl}_2(\mathbf{C})$ by linearity. The left-invariance of X implies that X preserves functions on $\Gamma\backslash G$.

The differential operator we need will arise in this way.

Iwasawa decomposition

$$SL_2(\mathbf{R}) = N \cdot A \cdot K \cong \mathbf{R} \times \mathbf{R}_+^\times \times S^1.$$

Explicitly,

$$g = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} \cdot \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Thus we can regard ϕ_f as a funciton of (x, y, θ):

$$\phi_f(x, y, \theta) = e^{ik\theta} y^{k/2} f(x + iy).$$

Lemma: f is holomorphic on $\mathbf{H}$ iff

$$L\phi_f = 0$$

where

$$L = -2iy\frac{\partial}{\partial \bar{z}} + \frac{i}{2}\frac{\partial}{\partial \theta}.$$

How to think about L

In the discussion on representations of $SL_2(\mathbf{R})$, we have come across the raising and lowering operator. Namely, the following 3 elements form an $\mathfrak{sl}_2$-triple:

$$H = i \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathfrak{k} = Lie(K) \otimes_{\mathbf{R}} \mathbf{C},$$

$$E = \frac{1}{2} \cdot \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \qquad F = \frac{1}{2} \cdot \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix}.$$

They satisfy:

$$[H, E] = 2E, \quad [H, F] = -2F, \quad [E, F] = H$$

Thus F lowers eigenvalues of H by 2, whereas E increases it by 2.

It turns out that if we think of F has a differential operator on functions of $SL_2(\mathbf{R})$, then

$$F = e^{-2i\theta} \cdot L.$$

Thus L is basically the lowering operator in the representation theory of $SL_2(\mathbf{R})$.

Holomorphy at cusp

Recall that holomorphy of f at the cusp $i\infty$ is implied by holomorphy of f on $\mathbf{H}$ and the fact that f does not grow too fast at $i\infty$:

$$|f(x+iy)| \leq C \cdot y^N$$

for some N, as $y \to \infty$ with x bounded. This condition translates to: as $y \to \infty$ with x bounded,

$$|\phi_f(g)| \leq C \cdot y^n$$

for some n (where g has coordinates (x, y, θ)). This last condition is a very important one. It is called the condition of **moderate growth** at the cusp $i\infty$. Of course, we also need to verify it for the other cusps.

We would like to formulate it in a "coordinate-free" manner, and we will do this for a general real reductive linear algebraic group G, say $G = GL_n$, Sp_{2n} or SO_n.

Norm functions

Since G is linear algebraic, we may choose an embedding

$$i : G \hookrightarrow GL_n.$$

For $g \in G$, define:

$$||g|| = max_{j,k}\{i(g)_{jk}, i(g^{-1})_{jk}\}$$

The norm function $|| - ||$ gives a measure of the size of elements of G. Though it depends on the choice of i, different choices lead to norm functions which are comparable. In other words, if $|| - ||_1$ and $|| - ||_2$ are 2 such norm functions, we have:

$$C \cdot ||g||_1^{1/r} \leq ||g||_2 \leq D \cdot ||g||_1^r$$

for some C, D and $r > 0$.

For our purpose, having such bounds is good enough. So the choice of i is not a serious one.

An important property of $||-||$ is:

- if K is a compact set, then there are constants C_K and D_K such that for any $g \in G$ and $k \in K$

$$C_K||g|| \leq ||gk|| \leq D_K||g||.$$

We write: $||g|| \asymp ||gk||$. Thus if $g = nak$ in the Iwasawa decomposition, then $||g||$ is more or less the same as $||na||$.

In the case, $G = SL_2$, if g has coordinates (x, y, θ), then

$$||g|| \asymp || \begin{pmatrix} y^{1/2} & xy^{-1/2} \\ 0 & y^{-1/2} \end{pmatrix} ||$$

$$= max\{y^{1/2}, y^{-1/2}, xy^{-1/2}\}.$$

Thus as $y \to \infty$ with x bounded,

$$||g|| \asymp y^{1/2}.$$

Moderate growth

In particular, the moderate growth condition for ϕ_f can be reformulated as:

Definition: A function ϕ on G is said to be of **moderate growth** if there is a constant n such that for all $g \in G$,

$$|\phi(g)| \leq C||g||^n$$

for some C.

This definition is independent of the choice of the norm function.

The advantage of this definition is that it simultaneously encompasses the condition at all the cusps! See [Borel, Prop. 5.11].

Cusp forms

A cusp form is defined by the vanishing of the zeroth Fourier coefficient at each cusp. At the cusp $i\infty$,

$$a_0(f) = \int_0^1 f(x+iy)\,dx \quad \text{for any } y.$$

We see that $a_0(f) = 0$ iff

$$\phi_N(g) := \int_{\mathbf{Z}\backslash\mathbf{R}} \phi_f\left(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g\right) dx = 0$$

for all g.

Recall that the cusps of Γ are in bijection with $\Gamma\backslash SL_2(\mathbf{Q})/B(\mathbf{Q})$. If x is a cuspidal point, its stabilizer in SL_2 is a Borel subgroup B_x defined over $\mathbf{Q}$. Then the zeroth coefficient of f at x vanishes iff

$$\int_{(\Gamma\cap N_x)\backslash N_x} \phi_f(ng)\,dn = 0.$$

Thus f **is cuspidal iff the above integral is** 0 **for any Borel subgroup defined over** $\mathbf{Q}$.

Of course, it suffices to check this for a set of representatives for the Γ-conjugacy classes of Borel $\mathbf{Q}$-subgroups.

We have noted before that a cuspidal f satisfies: for any k,

$$|f(x+iy)| \leq C_k y^{-k}.$$

as $y \to \infty$ with x bounded.

Definition: A function ϕ on $\Gamma\backslash G$ is **rapidly decreasing** on a Siegel set $\mathfrak{S}$ if, for any k,

$$|\phi(g)| \leq C_k \cdot ||g||^{-k}, \qquad g \in \mathfrak{S}.$$

One knows by reduction theory that $\Gamma\backslash G$ can be covered by finitely many Siegel sets. It is easy to see that if f is cuspidal, then ϕ is rapidly decreasing on each of these Siegel sets.

Now we have translated all the defining properties of a holomorphic modular form f on $\mathbf{H}$ to properties of ϕ_f. In other words, we have:

Proposition: The map $f \mapsto \phi_f$ defines an isomorphism of $M_k(\Gamma)$ to the space of smooth functions ϕ of $\Gamma \backslash SL_2(\mathbf{R})$ satisfying:

- ϕ is smooth;

- $\phi(gr_\theta) = e^{ik\theta}\phi(g)$;

- $F\phi = 0$ (F is lowering operator)

- ϕ is of moderate growth.

Moreover, the image of the space of cusp forms consists of those functions ϕ such that for ANY

Borel $\mathbf{Q}$-subgroup $B = T \cdot N$, the constant term ϕ_N along the unipotent radical N is zero. Further, the image of cusp forms is contained in $L^2(\Gamma \backslash SL_2(\mathbf{R}))$.

One has a similar proposition for the Maass forms, except that instead of $F\phi = 0$, we have a condition about the Laplace operator. We will come to this next.

Casimir Operator

The action of $\mathfrak{g} = \mathfrak{sl}_2(\mathbf{C})$ on smooth functions of $SL_2(\mathbf{R})$ as left-invariant differential operators extends to an action of the universal enveloping algebra $U\mathfrak{g}$. Those differential operators which are right-invariant as well form the center $Z\mathfrak{g}$ of $U\mathfrak{g}$. It is well-known that there is a canonical element in $Z(\mathfrak{g})$ (at least up to scaling) called the Casimir operator Δ. In the case of SL_2, one has:

$$\Delta = -\frac{1}{4}H^2 + \frac{1}{2}H - 2EF$$

and

$$Z(\mathfrak{g}) = \mathbf{C}[\Delta].$$

As a differential operator on $C^\infty(SL_2)$, we have:

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

Because Δ is bi-invariant, it acts on $C^\infty(\Gamma \backslash SL_2)$ as well as $C^\infty(SL_2/K)$.

- the action of Δ on a function on SL_2/K is the action of the hyperbolic Laplacian.

- If f is a holomorphic modular form, then
$$\Delta\phi_f = \frac{k}{2}(1 - \frac{k}{2})\phi_f.$$
This is because $H\phi_f = k\phi_f$ and $F\phi_f = 0$.

Proposition: The space of Maass forms is equal to the space of smooth functions which are right K-invariant, of moderate growth and satisfy
$$\Delta\phi = (\frac{1}{4} - s^2) \cdot \phi.$$

Thus we see that the theory of holomorphic modular forms and Maass forms can be subsumed in a single framework.

Passage from SL_2 to GL_2

We have yet to translate the Hecke operators from the classical picture to the new framework. For this purpose, it is convenient to pass from SL_2 to GL_2; for example, the Hecke operator T_p is defined by the diagonal matrix $diag(p, 1)$ which is not in SL_2.

There is nothing deep in this passage. It relies on the following identification:

$$\Gamma\backslash SL_2(\mathbf{R}) \cong Z(\mathbf{R})\Gamma'\backslash GL_2(\mathbf{R}).$$

Here, $\Gamma = \Gamma_0(N)$ and

$$\Gamma' = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbf{Z}) : c = 0 \mod N\}.$$

A function ϕ of $\Gamma\backslash SL_2(\mathbf{R})$ is thus naturally a function on $GL_2(\mathbf{R})$, or rather $PGL_2(\mathbf{R})$.

Now we can restate the last two propositions using GL_2 instead of SL_2.

Proposition: The map $f \mapsto \phi_f$ defines an isomorphism of $M_k(\Gamma)$ to the space $V_k(\Gamma')$ of functions ϕ of $Z(\mathbf{R})\Gamma' \backslash GL_2(\mathbf{R})$ satisfying:

- ϕ is smooth;
- $\phi(gr_\theta) = e^{ik\theta}\phi(g)$;
- $F\phi = 0$;
- ϕ is of moderate growth.

Similarly, the space of Maass forms with parameter s can be identified with the space of

ϕ on $Z(\mathbf{R})\Gamma'\backslash GL_2(\mathbf{R})$ which are smooth, right-K-invariant, of moderate growth and such that $\Delta\phi = (1/4 - s^2)\phi$.

Moreover, the image of the space of cusp forms consists of those functions ϕ such that for ANY Borel $\mathbf{Q}$-subgroup $B = T \cdot N$, the constant term ϕ_N along the unipotent radical N is zero. Moreover, the image of cusp forms is contained in $L^2(Z(\mathbf{R})\Gamma'\backslash GL_2(\mathbf{R}))$.

Observe that the maximal compact subgroup K' of $GL_2(\mathbf{R})$ is the orthogonal group O_2, which contains SO_2 with index 2. The span of the K'-translates of ϕ_f (for f holomorphic) is now 2-dimensional.

Hecke operators

Now we come to the Hecke operators. For $\alpha \in GL_2(\mathbf{Q})$, we have the Hecke operator T_h on the space of functions on $\Gamma \backslash GL_2(\mathbf{R})$ by:

$$(T_\alpha \phi)(g) = \sum_{i=1}^{r} \phi(a_i g)$$

if

$$\Gamma \alpha \Gamma = \bigcup_{i=1}^{r} \Gamma a_i$$

The definition is independent of the choice of representatives a_i. The reason that left Γ-invariance is preserved is that if $\gamma \in \Gamma$, then $\{\Gamma a_i \gamma\}$ is a permutation of $\{\Gamma a_i\}$.

This operator can be understood geometrically as a correspondence on $\Gamma' \backslash PGL_2$.

Let α_p denote the diagonal matrix $diag(p, 1)$. Earlier, we have defined an action of $\Gamma\alpha_p\Gamma$ on a modular form f:

$$T_{\alpha_p} f := f|_k[\alpha_p] = \sum_i f|_k a_i$$

if $\Gamma\alpha_p\Gamma = \bigcup_i \Gamma a_i$. This operator is basically the Hecke operator T_p:

$$T_p = p^{k/2-1} T_{\alpha_p}.$$

Proposition: The isomorphism $M_k(\Gamma) \longrightarrow V_k(\Gamma')$ is an isomorphism of Hecke modules, i.e. for any prime p,

$$\phi_{T_\alpha f} = T_\alpha \phi_f.$$

Proof:

$$\phi_{T_\alpha f}(g) = ((T_\alpha f)|_k g)(i) = ((\sum_j f|a_j)|g)(i)$$

$$\sum_j (f|(a_j g))(i) = (T_\alpha \phi_f)(g).$$

Automorphic forms

Let G be a reductive linear algebraic group defined over $\mathbf{Q}$, and let Γ be an arithmetic group. We shall assume for simplicity that $\Gamma \subset G(\mathbf{Q})$.

By an automorphic form on G with respect to an arithmetic group Γ, we mean a function ϕ on $\Gamma \backslash G(\mathbf{R})$ satisfying:

- ϕ is smooth
- ϕ is of moderate growth
- ϕ is right K-finite
- ϕ is $Z(\mathfrak{g})$-finite

$Z(\mathfrak{g})$-**finiteness**

Say that a function ϕ on G is $Z(\mathfrak{g})$-**finite** if dim($Z(\mathfrak{g})\phi$) is finite. Equivalently, if ϕ is annihilated by an ideal of finite codimension in $Z(\mathfrak{g})$.

In the case of SL_2 above, if ϕ is an eigenfunction for Δ with eigenvalue λ, then ϕ is annihilated by the ideal $(\Delta - \lambda)$ which is of codimension 1 in $\mathbf{C}[\Delta]$. Such ϕ's are in particular $Z(\mathfrak{g})$-finite.

Observe that if ϕ is $Z(\mathfrak{g})$-finite and $X \in Z(\mathfrak{g})$, then ϕ is killed by some polynomial in X. For if ϕ is killed by J (of finite codimension), then the kernel of

$$C[X] \longrightarrow Z(\mathfrak{g})/J$$

has finite codimension.

The space of automorphic forms

Let $\mathcal{A}(G, \Gamma)$ denote the space of automorphic forms on G. Also, if $\rho \subset \widehat{K}$ is a finite set of irreducible representations of K and J is an ideal of finite codimension in $Z(\mathfrak{g})$, then we let

- $\mathcal{A}(G, \Gamma, J)$ be the subspace of $\mathcal{A}(G, \Gamma)$ consisting of functions which are killed by J;

- $\mathcal{A}(G, \Gamma, J, \rho)$ be the subspace of $\mathcal{A}(G, \Gamma, J)$ consisting of functions ϕ such that the finite-dim representation of K generated by ϕ is supported on ρ. Shall see later that this space is finite-dim: this is a fundamental result of Harish-Chandra.

For example, when $G = SL_2$, $J = \langle \Delta - \frac{k}{2}(\frac{k}{2} - 1) \rangle$ and ρ consists of the single character $r_\theta \mapsto e^{ik\theta}$ of K, we have seen that

$$\mathcal{A}(G, \Gamma, J, \rho) \supset M_k(\Gamma).$$

Analytic Properties of automorphic forms

We are going to list some basic analytic properties of an automorphic form f. For this, we shall need some analytic inputs. Two of the most useful ones are:

- (**Elliptic regularity theorem**) If f is killed by an elliptic differential operator, then f is real analytic.

- (**Abundance of K-invariant test functions**) If f is a right K-finite and $Z(\mathfrak{g})$-finite function on G, then given any neighbourhood U of 1, one can find $\alpha \in C_c^\infty(G)$ such that

  - support of α is in U

  - $\alpha(kgk^{-1}) = \alpha(g)$ for any $k \in K$ and $g \in G$

– $f * \alpha = f$, where

$$(f * \alpha)(g) = \int_G f(gh^{-1})\alpha(h)\, dh.$$

The first result is a standard result in the theory of differential operators. The second one is actually best viewed in the framework of representation theory. It was proved by Harish-Chandra. We shall take these for granted, but see [Borel, Thm. 2.14].

Proposition: Let f be an automorphic form on $\Gamma\backslash G$. We have:

- f is real analytic.

- f is of **uniform** moderate growth, i.e. there exists a N such that for any $X \in U(\mathfrak{g})$,

$$|(Xf)(g)| \leq C_X ||g||^N.$$

Proof: For (i), we show that a K-finite, $Z(\mathfrak{g})$-finite function on $G(\mathbf{R})$ is real analytic.

We know that f is annihilated by some polynomial $P(\Delta)$ of the Casimir element Δ. Unfortunately, the Casimir element is not elliptic.

To create an elliptic operator, we let Δ_K be the Casimir element of the maximal compact K. Then $\Delta - 2\Delta_K$ is elliptic (but it is not an element of $Z(\mathfrak{g})$).

We claim however that f is killed by some polynomial in $\Delta - 2\Delta_K$.

Indeed, because f is K-finite, f is contained in a finite-dim K-invariant subspace, and because every finite-dim representation of K is semisimple, the action of Δ_K on this space can be diagonalized. So we can write: $f = f_1 + ... + f_r$ so that each f_i is an eigenfunction of Δ_K, say

$\Delta_K f_i = \lambda_i f_i$. Moreover, f_i is $Z(\mathfrak{g})$-finite as well, and is still killed by $P(\Delta)$.

If $P(\Delta) = \prod_j (\Delta - c_j)$, then it follows that f_i is killed by

$$P_i(\Delta - 2\Delta_K) = \prod_j (\Delta - 2\Delta_K + 2\lambda_i - c_j).$$

Taking a product of the P_i's gives the result.

For (ii), choose α as in the proposition, and note that

$$Xf = X(f * \alpha) = f * X\alpha.$$

Then

$$|Xf(g)| \leq \int_U |f(gh^{-1})| \cdot |X\alpha(h)| dh$$

$$\leq C \cdot ||g||^n \cdot ||X\alpha||_1.$$

<u>Cusp forms</u>

Definition: If f is automorphic, then f is **cuspidal** if for any parabolic $\mathbf{Q}$-subgroup $P = MN$ of G, we have

$$f_N(g) := \int_{(\Gamma\cap N)\backslash N} f(ng)dn = 0.$$

The function f_N on G is called the **constant term of f along N**.

Remark: The restriction of f_N to $M(\mathbf{R})$ is an automorphic form on M. This is not that trivial, namely one needs to deduce $Z(\mathfrak{m})$-finiteness from $Z(\mathfrak{g})$-finiteness.

To check for cuspidality, it suffices to check for a set of representatives for the Γ-orbits of **maximal** parabolic $\mathbf{Q}$-subgroups.

We let $\mathcal{A}_0(G, \Gamma)$ be the space of cusp forms.

Fourier coefficients: The constant term is but one Fourier coefficient of f along N. For any unitary character χ of N which is left-invariant under $\Gamma\backslash N$, we set:

$$f_{N,\chi}(g) = \int_{(\Gamma\cap N)\backslash N} f(ng)\cdot\overline{\chi(n)}\,dn.$$

This is the χ-th Fourier coefficient of f along N.

Fourier expansion for abelian N

If N is abelian, then

$$f(g) = \sum_{\chi} f_{N,\chi}(g)$$

so that f can be recovered from its Fourier coefficients along N. To see this, consider the function on $N(\mathbf{R})$:

$$\Phi_g(x) = f(xg).$$

It is in fact a function on

$$(\Gamma \cap N)\backslash N \cong (\mathbf{Z}\backslash\mathbf{R})^r.$$

So we can expand this in a Fourier series:

$$\Phi_g(x) = \sum_{\chi} a_{\chi}(g)\,\chi(x)$$

where

$$a_{\chi}(g) = \int_{(\Gamma\cap N)\backslash N} \overline{\chi(x)} \cdot f(xg)\,dx = f_{N,\chi}(g)$$

Putting $x = 1$ in the Fourier series gives the assertion.

An important estimate

The following is an important estimate:

Proposition: Let $P = MN$ be a **maximal** parabolic $\mathbf{Q}$-subgroup, and A the split component of its Levi subgroup M (so $M = M_0 \cdot A$ with M_0 having compact center and $A \cong \mathbf{R}_+^\times$ is in the center of M).

Then $f - f_N$ is rapidly decreasing on any Siegel set

$$\mathfrak{S} = \omega \cdot A_t \cdot K$$

where ω is a compact set of $N \cdot M_0$.

Here

$$A_t = \{a \in A : \delta(a) > t\}$$

where δ is the unique simple root occurring in N.

We sketch the proof, under the simplifying assumption that N is abelian. In the general case, one can find a filtration of N by normal subgroups whose successive quotients are abelian, and one can apply induction.

Basically the proof is by repeated integration by parts.

We have, by Fourier expansion:

$$(f - f_N)(g) = \sum_{\chi \neq 1} f_{N,\chi}(g).$$

So we need to estimate $f_{N,\chi}(g)$ for $g \in \mathfrak{S}$. In particular, we would be done if we can show: for any k,

$$|f_{N,\chi}(g)| \leq C_{\chi,k} ||g||^{-k},$$

with $\sum_\chi C_{\chi,k} < \infty$.

Let's examine $f_{N,\chi}(g)$.

Firstly, since g is the Siegel set $\mathfrak{S}$, there is no loss in assuming that $g = a \in A_t$ with t large. Then $||a|| \asymp max_\alpha \alpha(a)$, with α ranging over the roots in N. It is easy to see that there is a $c > 0$ so that for any root α in N,

$$\alpha(a) > ||a||^c.$$

Now we shall make another simplifying assumption. We suppose that we can choose a basis $\{X_\alpha\}$ of $Lie(N)$ such that

- each X_α is a root vector for A for the root α:

$$a \cdot X_\alpha = \alpha(a) \cdot X_\alpha$$

- under the natural identification of N with $Lie(N)$, $\Gamma \cap N$ is identified with the $\mathbf{Z}$-span of the X_α's. This gives an isomorphism

$$\Gamma \cap N \backslash N \cong \mathbf{Z}^r \backslash \mathbf{R}^r.$$

The assumption is satisfied if, for example, G is a Chevalley group defined over $\mathbf{Z}$ and P is a parabolic subgroup defined over $\mathbf{Z}$.

The non-trivial character χ is then of the form

$$\chi(\mathbf{x}) = e^{2\pi i \mathbf{k}\cdot\mathbf{x}}$$

for some $\mathbf{k} \in \mathbf{Z}^r$.

Now

$$f_{N,\chi}(a) = \int_{\mathbf{Z}^r \backslash \mathbf{R}^r} f(exp(\sum x_\alpha X_\alpha) \cdot a)\, e^{-2\pi i \mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}.$$

Choose β so that $|k_\beta|$ is maximum among the coordinates of $\mathbf{k}$. Then using integration by parts repeatedly, we have:

$$f_{N,\chi}(a) =$$

$$\left(\frac{-1}{2\pi i k_\beta}\right)^p \int e^{-2\pi i \mathbf{k}\cdot\mathbf{x}} . \frac{\partial^p}{\partial x_\beta^p} \Big(f(exp(\sum x_\alpha X_\alpha)\, a)\Big)\, d\mathbf{x}.$$

Now we have

$$\frac{\partial^p}{\partial x_\beta^p}\Big(f(exp(\sum x_\alpha X_\alpha)\, a)\Big)$$

$$= (X_\beta'^p f)(exp(\sum x_\alpha X_\alpha(\, a))$$

where $X_\beta' = Ad(a^{-1})(-X_\beta) = -\beta(a)^{-1} \cdot X_\beta$.

So

$$f_{N,\chi}(a) = \frac{1}{(2\pi i k_\beta)^p} \cdot \frac{1}{\beta(a)^p}\cdot$$

$$\int_{\mathbf{Z}^r\backslash\mathbf{R}^r} e^{-2\pi i \mathbf{k}\cdot\mathbf{x}} \cdot (X_\beta^p f)(exp(\sum x_\alpha X_\alpha)\, a)\, d\mathbf{x}.$$

Now using the fact that f is of uniform moderate growth, say of exponent N, we get a bound of the type:

$$|f_{N,\chi}(a)| \leq C_{f,p} \cdot \frac{1}{||\mathbf{k}||^p} \cdot ||a||^{N-cp}$$

where $||k|| = max_\alpha |k_\alpha| = |k_\beta|$.

Since p can be arbitrarily large, and $\sum_\mathbf{k} \frac{1}{||\mathbf{k}||^p}$ converges for large p, we are done.

Analytic properties of cusp forms

One consequence of the above estimate is:

Theorem: Suppose that G is semisimple.

(i) If f is a cusp form, then f is rapidly decreasing on any Siegel set for $\Gamma\backslash G$.

(ii) Conversely, suppose that f satisfies all the properties of an automorphic form, except for the condition of moderate growth, but suppose that $f_N = 0$ for all $P = MN$. Then the following are equivalent (G semisimple):

- f is of moderate growth
- f is bounded
- f is in $L^p(\Gamma\backslash G)$ for all $p \geq 1$
- f is in $L^p(\Gamma\backslash G)$ for some $p \geq 1$.

Proof: The only thing that remains to be proven is that if f is in L^p, then f is bounded. Choose a K-invariant $\alpha \in C_c^\infty(G)$ such that $f = f * \alpha$. Then

$$\begin{aligned} |f(g)| &= |(f * \alpha)(g)| \\ &\leq \int_G |f(gx^{-1})| \cdot |\alpha(x)| \, dx \\ &\leq ||l_{g^{-1}}(f)||_p \cdot ||\alpha||_q \\ &= ||f||_p \cdot ||\alpha||_q \end{aligned}$$

by the Holder inequality (with $p^{-1} + q^{-1} = 1$).

The $(\mathfrak{g}, K)$-module structure

We now consider the vector space $\mathcal{A}(G, \Gamma)$ as a whole. The main result is:

Theorem: $\mathcal{A}(G, \Gamma)$ is naturally a $(\mathfrak{g}, K)$-module.

Proof: It is easy to see that if $\mathcal{A}(G, \Gamma)$ is preserved by K. Further, the action of $\mathfrak{g}$ and K are compatible in the usual sense:

$$kXf = (Ad(k)X)kf.$$

Next we show that $\mathcal{A}(G)$ is invariant under the $\mathfrak{g}$-action. If $X \in \mathfrak{g}$, then it is again clear that Xf is smooth, left-invariant under Γ, $Z(\mathfrak{g})$-finite and right K-finite. The only thing left to check is the condition of moderate growth. But this follows from uniform moderate growth of f.

Thus we see the entrance of representation theory.

Remarks: Note that $\mathcal{A}(G, \Gamma)$ is not invariant under right translation by G. Indeed, the K-finiteness condition is not preserved.

What properties does this $(\mathfrak{g}, K)$-module have? The following is a **fundamental result of Harish-Chandra**:

Theorem: Fix an ideal J of finite codimension in $Z(\mathfrak{g})$. Then $\mathcal{A}(G, \Gamma, J)$ is an admissible $(\mathfrak{g}, K)$-submodule. Equivalently, if ρ is an irreducible representation of K, then $\mathcal{A}(G, \Gamma, J, \rho)$ is finite-dimensional.

This theorem has many applications. Let us list two of them.

Corollary: Any irreducible $(\mathfrak{g}, K)$-module π occurs as a submodule of $\mathcal{A}(G, \Gamma)$ with finite multiplicity, i.e.

$$\dim Hom_{\mathfrak{g},K}(\pi, \mathcal{A}(G, \Gamma)) < \infty.$$

Proof; π has an infinitesimal character; let J be its kernel. Fix a K-type ρ of π. Then

$$Hom(\pi, \mathcal{A}(G, \Gamma)) = Hom(\pi, \mathcal{A}(G, \Gamma, J)).$$

If this space is infinite dimensional, then ρ will occur infinitely often in $\mathcal{A}(G, \Gamma, J)$, contradicting the fact that $\mathcal{A}(G, \Gamma, J)$ is admissible.

Corollary: The space $\mathcal{A}_0(G, \Gamma)$ of cusp forms is a semisimple $(\mathfrak{g}, K)$-module, with each irreducible summand occurring with finite multiplicities.

Proof: By the theorem, $\mathcal{A}_0(G, \Gamma, J)$ is admissible and we know it is contained in $L^2(\Gamma\backslash G)$; so it is also unitarizable. It is a standard result in representation theory that a unitarizable, admissible $(\mathfrak{g}, K)$-module is semisimple. So $\mathcal{A}_0(G, \Gamma, J)$ is semisimple.

Because $\mathcal{A}_0(G, \Gamma)$ is the union of the $\mathcal{A}_0(G, \Gamma, J)$, an argument using Zorn's lemma shows that $\mathcal{A}_0(G, \Gamma)$ is also semisimple.

Hecke algebra

Besides the structure of a $(\mathfrak{g}, K)$-module, $\mathcal{A}(G, \Gamma)$ also possesses the action of Hecke operators. This is defined as before: if $\alpha \in G(\mathbf{Q})$ and

$$\Gamma\alpha\Gamma = \bigcup_{i=1}^{r} \Gamma a_i$$

then

$$(T_\alpha f)(g) = \sum_{i=1}^{r} f(a_i g).$$

We think of $\Gamma\alpha\Gamma$ as the characteristic function of this double coset. The **Hecke algebra for** Γ is the algebra of fuctions on $G(\mathbf{Q})$ which are bi-Γ-invariant and supported on finitely many Γ-double cosets. The multiplication is by convolution.

The above formula makes $\mathcal{A}(G, \Gamma)$ into a module for the Hecke algebra. Observe:

- the action of the Hecke algebra commutes with the action of $(\mathfrak{g}, K)$.

This is because the $(\mathfrak{g}, K)$-action is by right translation, whereas the action of a Hecke operator is a sum of left translation.

Thus, if π is an irreducible $(\mathfrak{g}, K)$-module, then the Hecke algebra acts on

$$Hom_{\mathfrak{g},K}(\pi, \mathcal{A}(G, \Gamma)).$$

Note that this Hom-space is finite dimensional, by the fundamental theorem of Harish-Chandra.

Classical modular forms

We have seen that a classical modular form f corresponds to certain automorphic form ϕ on SL_2. One can ask: what is the $(\mathfrak{g}, K)$-module generated by ϕ?

Now we have seen that ϕ is annihilated by the lowering operator F whereas the set

$$\{\phi, E\phi, E^2\phi, ...\}$$

are eigenfunctions are eigenfunctions of K with eigenvalues k, $k+2$,.... Moreover, the span of these is invariant a $(\mathfrak{g}, K)$-submodule.

Thus we conclude that ϕ generates the holomorphic discrete series π_k of minimal weight k, and

$$M_k(\Gamma) \cong Hom_{\mathfrak{g},K}(\pi_k, \mathcal{A}(G, \Gamma)).$$

This is an isomorphism of modules for the Hecke algebra.

Given $l \in Hom_{\mathfrak{g},K}(\pi_k, \mathcal{A}(G, \Gamma))$, the corresponding classical modular form is obtained by taking the lowest weight vector in $l(\pi_k)$ and then transforming it back to the upper half plane.

Similarly, if π_s is the principal series representation

$$\pi_s = Ind_B^{SL_2} \delta_B^{1/2+s},$$

then the space of Maass forms with respect to Γ with parameter s is isomorphic to

$$Hom_{\mathfrak{g},K}(\pi_s, \mathcal{A}(G, \Gamma))$$

in a Hecke equivariant fashion.

Selberg's conjecture again

We can now provide a representation theoretic interpretation of the **Selberg conjecture** for cuspidal Maass forms: $\lambda \geq 1/4$, or equivalently that s is purely imaginary.

Now s is purely imaginary iff π_s is a so-called **tempered** $(\mathfrak{g}, K)$-module. Thus Selberg's conjecture says that if the only π_s which can embed into $\mathcal{A}_0(G, \Gamma)$ are the tempered ones (if Γ is a congruence group).

As we shall see later, the theory of Eisenstein series shows that for most s, one can embed π_s into $\mathcal{A}(G, \Gamma)$.

Relation with $L^2(\Gamma\backslash G)$

In Labesse's lectures, one encounters the question of decomposing the unitary representation $L^2(\Gamma\backslash G(\mathbf{R}))$ of $G(\mathbf{R})$. This is of course a very natural question.

What does one know about this problem from Labesse's lectures?

One of the results discussed is that when $\Gamma\backslash G(\mathbf{R})$ is compact, then $L^2(\Gamma\backslash G(\mathbf{R}))$ decomposes into the direct sum of irreducible unitary representations, each occurring with finite multiplicity. In other words,

$$L^2(\Gamma\backslash G(\mathbf{R})) \cong \widehat{\bigoplus}_{\pi\in\widehat{G}} Hom_G(\pi, L^2(\Gamma\backslash G)) \otimes \pi$$

with

$$\dim Hom_G(\pi, L^2(\Gamma\backslash G)) < \infty.$$

More generally, even if $\Gamma\backslash G$ is not compact, the above result holds if we consider the space $L^2_0(\Gamma\backslash G)$ of *cuspidal* L^2-functions.

When $\Gamma\backslash G$ is non-compact, then $L^2(\Gamma\backslash G)$ will not decompose into a direct sum of irreducibles. Rather, there will be a part which decomposes as a direct sum (called the **discrete spectrum**) and a part which decomposes as a direct integral (called the **continuous spectrum**). Thus,

$$L^2(\Gamma\backslash G) = L^2_{disc}(\Gamma\backslash G) \oplus L^2_{cont}(\Gamma\backslash G).$$

For example, the space L^2_0 is contained in the discrete spectrum L^2_{disc}.

The (very non-trivial) theory of Eisenstein series shows that L^2_{cont} can be described in terms of the discrete spectrum of the Levi subgroups of G and thus can be understood inductively.

Thus **the fundamental problem in the study of $L^2(\Gamma\backslash G)$ is the decomposition of L^2_{disc}.**

What is the relation, if any, between the unitary representation $L^2(\Gamma\backslash G)$ and the $(\mathfrak{g}, K)$-module $\mathcal{A}(G, \Gamma)$?

Well, a priori, not much. These two spaces of functions are certainly different: an L^2-function is not necessarily an automorphic form (since it may not be smooth), and an automorphic form needs not be L^2 (for example, the Eisenstein series). So none of these spaces is contained in the other.

It turns out, however, that the two problems are very much related. Let us explain this.

We have seen that

$$\mathcal{A}_0(G, \Gamma) \subset L_0^2(\Gamma\backslash G).$$

In fact, $\mathcal{A}_0(G, \Gamma)$ is the subspace of smooth, K-finite and $Z(\mathfrak{g})$-finite vectors in the unitary representation $L_0^2(\Gamma\backslash G)$!

Representation theoretically, if

$$L^2_0(\Gamma\backslash G) \cong \widehat{\bigoplus_{\pi}} m_\pi \pi$$

(Hilbert direcct sum), then we have

$$\mathcal{A}(G,\Gamma) = \bigoplus_{\pi} m_\pi \pi_K$$

(algebraic direct sum) where π_K is the $(\mathfrak{g}, K)$-module underlying π. Note that this is slightly smaller than the $(\mathfrak{g}, K)$-module underlying the unitary representation $L^2_0(\Gamma\backslash G)$.

Thus the decomposition of L^2_0 into irreducible unitary representations is the same problem as the decomposition of $\mathcal{A}_0$ into irreducible $(\mathfrak{g}, K)$-modules.

More generally, if we consider the intersection

$$\mathcal{A}^2(G,\Gamma) := \mathcal{A}(G,\Gamma) \cap L^2(\Gamma\backslash G)$$

then $\mathcal{A}^2(G,\Gamma)$ (the space of square-integrable automorphic forms) is precisely the space of smooth K-finite, $Z(\mathfrak{g})$-finite vectors in the discrete spectrum $L^2_{disc}(\Gamma\backslash G)$.

Thus the decomposition of L^2_{disc} is the same as the problem of decomposing $\mathcal{A}^2(G,\Gamma)$.

This problem is one of the central problems in the theory of automorphic forms. It is far from being resolved.

Automorphic Representations of Adele Groups

We have defined the space $\mathcal{A}(G, \Gamma)$ of automorphic forms with respect to an arithmetic group Γ of G (a reductive linear algebraic group defined over $\mathbf{Q}$). We saw that $\mathcal{A}(G, \Gamma)$ has a commuting action of $(\mathfrak{g}, K)$ and the Hecke algebra $\mathcal{H}(G, \Gamma)$.

From this point of view, we saw that the classical modular forms correspond to different ways of embedding certain irreducible $(\mathfrak{g}, K)$-modules into $\mathcal{A}(G, \Gamma)$:

$$M_k(N) \cong Hom_{\mathfrak{g},K}(\pi_k, \mathcal{A}(PGL_2, \Gamma_0'(N)))$$

where π_k is discrete series of $PGL_2(\mathbf{R})$ with lowest weight k. Thus we are interested in how $\mathcal{A}(G, \Gamma)$ decomposes as a $(\mathfrak{g}, K) \times \mathcal{H}(G, \Gamma)$ module.

GOAL: to formulate the theory of automorphic forms using adelic language.

The reasons, among others, are:

- we want a theory that deals with $\mathcal{A}(G, \Gamma)$ for all choices of Γ simultaneously.

- we want a framework in which the roles of the $(\mathfrak{g}, K)$-action and the $\mathcal{H}(G, \Gamma)$-action are parallel, i.e. so that they are actions of the same kind.

- To describe the process of attaching an L-function to a classical modular form in terms of representation theory, it is cleanest to use the adelic framework, as demonstrated in Tate's thesis.

Adeles

Let F be a number field. Then one can associate a locally compact topological ring $\mathbf{A}_F$, called the adele ring of F. For concreteness, in these lectures, we shall just work with $\mathbf{Q}$.

Let's recall the definition of $\mathbf{A} = \mathbf{A}_{\mathbf{Q}}$:

$$\mathbf{A} \subset \mathbf{R} \times \prod_p \mathbf{Q}_p$$

consisting of those $x = (x_v)$ such that for almost all primes p,

$$x_p \in \mathbf{Z}_p.$$

It is clearly a subring of $\prod_v \mathbf{Q}_v$.

This construction is called the **restricted direct product**.

The ring $\mathbf{A}$ has a natural topology: a basis of open neighbourhoods at a point x consists of:

$$\prod_{v \in S} U_v \times \prod_{v \notin S} \mathbf{Z}_v$$

where S is a finite set of places of $\mathbf{Q}$, including the archimedean prime, and U_v is an open neighbourhood of x. In particular, $\mathbf{A}$ is a locally compact ring.

Alternatively, one can think of $\mathbf{A}$ as the inductive (or direct) limit of

$$\prod_{v \in S} \mathbf{Q}_v \times \prod_{v \notin S} \mathbf{Z}_v.$$

Observe that $\mathbf{Q} \hookrightarrow \mathbf{A}$ diagonally, so that $\mathbf{A}$ is a $\mathbf{Q}$-algebra. An important property is:

Theorem: $\mathbf{Q}$ is discrete in $\mathbf{A}$ with $\mathbf{Q}\backslash\mathbf{A}$ compact.

Thus, the situation of $\mathbf{Q} \subset \mathbf{A}$ is analogous to the situation of $\mathbf{Z} \subset \mathbf{R}$.

There are some variants of the above construction. If S is a finite set of places of $\mathbf{Q}$, we let:

$$\begin{cases} \mathbf{Q}_S = \prod_{v \in S} \mathbf{Q}_v \\ \mathbf{A}^S = \{(x_v) \in \prod_{v \notin S} \mathbf{Q}_v : x_v \in \mathbf{Z}_v \text{ for almost all } v\} \end{cases}$$

We call $\mathbf{A}^S$ the S-adeles. If S consists only of the place ∞, then we call $\mathbf{A}^S$ the finite adeles and denote it by $\mathbf{A}_f$.

The following is called the **strong approximation theorem for adeles:**

Theorem: Let S be a non-empty finite set of places of $\mathbf{Q}$. Then $\mathbf{Q}$ is dense in $\mathbf{A}^S$.

When S consists only of the archimedean place, this is the so-called Chinese remainder theorem.

Adele Groups

Let G be a linear algebraic group defined over $\mathbf{Q}$; the examples to keep in mind are the reductive groups GL_n and SL_n, as well as unipotent groups.

Note that G is an affine algebraic variety over $\mathbf{Q}$. One can thus consider the group $G(\mathbf{A})$ of adelic points of G. We simply call this the **adele group of** G. It is a locally compact group and we can give it a more concrete description as follows.

We consider the set of sequences (K_p) (indexed by primes) of open compact subgroups K_p of $G(\mathbf{Q}_p)$, and consider two such sequences (K_p) and (K'_p) to be equivalent if $K_p = K'_p$ for almost all p.

Now given a linear algebraic group G, one can associate an equivalent class of such sequences.

Namely, choose any embedding $i : G \hookrightarrow GL(V)$ where V is a $\mathbf{Q}$-vector space, and pick a lattice Λ in V. Then let K_p be the stabilizer of $\Lambda \otimes_{\mathbf{Z}} \mathbf{Z}_p$ in $G(\mathbf{Q}_p)$. The resulting sequence (K_p) will depend on (i, Λ) but different choices give equivalent sequences.

Let us pick one such sequence (K_p) from the equivalence class determined by G. Then it can be shown that:

$$G(\mathbf{A}) =$$

$$\{(g_v) \in \prod_v G(\mathbf{Q}_v) : g_v \in K_v \text{ for almost all } v\}.$$

This is independent of the choice of the sequence.

From this, one sees that if S is a finite set of primes, then

$$G(\mathbf{A}) = G(\mathbf{Q}_S) \times G(\mathbf{A}^S).$$

Moreover, a basis of neighbourhoods at 1 in $G(\mathbf{A}_f)$ consists of open compact subgroups $U = \prod_v U_v$ with $U_v = K_v$ for almost all v.

For almost all p, the open compact subgroup K_p is a so-called **hyperspecial maximal compact subgroup**. For example, when $G = GL_n$, $K_p = GL_n(\mathbf{Z}_p)$.

We can modify K_v at the remaining places (including ∞) and assume that they are special maximal compact subgroups. Then

$$K = \prod_v K_v$$

is a maximal compact subgroup of $G(\mathbf{A})$. We fix this K henceforth.

For example, when $G = GL_1$, then

$$GL_1(\mathbf{A}) =$$

$$\{x = (x_v) \in \prod_v \mathbf{Q}_v^\times,\, x_p \in \mathbf{Z}_p^\times \text{ for almost all } p\}.$$

This is the so-called idele group of $\mathbf{Q}$.

The quotient $G(\mathbf{Q})\backslash G(\mathbf{A})$

Because $\mathbf{Q} \subset \mathbf{A}$, we have:

$$G(\mathbf{Q}) \subset G(\mathbf{A}).$$

The situation of $G(\mathbf{Q}) \subset G(\mathbf{A})$ is entirely analogous to the situation of $SL_2(\mathbf{Z}) \subset SL_2(\mathbf{R})$. Indeed, we have:

- $G(\mathbf{Q})$ is a discrete subgroup of $G(\mathbf{A})$.

- $G(\mathbf{Q})\backslash G(\mathbf{A})$ has a fundamental domain which can be covered by a sufficiently large Siegel set (associated to any parabolic $\mathbf{Q}$-subgroup).

- $G(\mathbf{Q})\backslash G(\mathbf{A})$ has finite volujme if G is semisimple; it is compact if G is anisotropic.

Strong Approximation for G

The strong approximation theorem allows one to relate the adelic picture to the case of $\Gamma\backslash G(\mathbf{R})$.

Theorem: Assume that G is simply-connected and S is a finite set of places of $\mathbf{Q}$ such that $G(\mathbf{Q}_S)$ is not compact, then $G(\mathbf{Q})$ is dense in $G(\mathbf{A}^S)$.

Here is a reformulation. Given any open compact subgroup $U^S \subset G(\mathbf{A}^S)$, we have:

$$G(\mathbf{A}) = G(\mathbf{Q}) \cdot G(\mathbf{Q}_S) \cdot U^S.$$

A consequence of this is:

Corollary: Under the asumptions of the theorem, if we let $\Gamma = G(\mathbf{Q}) \cap U^S$, then

$$G(\mathbf{Q})\backslash G(\mathbf{A})/U^S \cong \Gamma\backslash G(\mathbf{Q}_S).$$

An example

As an example, consider the case when $G = SL_2$ and $S = \{\infty\}$. Then

$$SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})/U_f \cong \Gamma\backslash SL_2(\mathbf{R})$$

where U_f is any open compact subgroup of $G(\mathbf{A}_f)$ and $\Gamma = G(\mathbf{Q}) \cap U_f$.

Let's take $\mathbb{U}_f$ to be the group

$$K_0(N) = \prod_{p|N} I_p \cdot \prod_{(p,N)=1} SL_2(\mathbf{Z}_p)$$

where I_p is an Iwahori subgroup of $SL_2(\mathbf{Q}_p)$:

$$I_p = \{g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}_p) : c \equiv 0 \mod p\}.$$

Then it is clear that

$$\Gamma_0(N) = K_0(N) \cap SL_2(\mathbf{Q}).$$

So we have:

$$SL_2(\mathbf{Q})\backslash SL_2(\mathbf{A})/K_0(N) \cong \Gamma_0(N)\backslash SL_2(\mathbf{R}).$$

Passage from real to adele groups

The above consideration allows us to regard an automorphic form f on $\Gamma\backslash G(\mathbf{R})$ as a function on $G(\mathbf{Q})\backslash G(\mathbf{A})$, at least for certain Γ.

We say that Γ is a **congruence subgroup** of G if $\Gamma = G(\mathbf{Q}) \cap U_\Gamma$ for some open compact subgroup U_Γ of $G(\mathbf{A}_f)$.

Thus if Γ is congruence, and G satisfies strong approximation, we have:

$$\Gamma\backslash G(\mathbf{R}) \cong G(\mathbf{Q})\backslash G(\mathbf{A})/U_\Gamma$$

and we can regard an automorphic form on $\Gamma\backslash G(\mathbf{R})$ as a function on $G(\mathbf{Q})\backslash G(\mathbf{A})$ which is right-invariant under U_Γ.

Remarks: (i) In general, for any reductive G and any open compact $U \subset G(\mathbf{A}_f)$, we have

$$\# G(\mathbf{Q}) \backslash G(\mathbf{A}) / G(\mathbf{R}) U < \infty.$$

In this case, if $\{g_i \in G(\mathbf{A}_f)\}$ is a set of double coset representatives, then

$$G(\mathbf{Q}) \backslash G(\mathbf{A}) / U = \bigcup_i \Gamma_i \backslash G(\mathbf{R})$$

with $\Gamma_i = G(\mathbf{Q}) \cap g_i U g_i^{-1}$.

Thus, the passage between real and adele groups is not seriously affected by the lack of strong approximation.

(ii) When $G = GL_n$, we still have the decomposition:

$$GL_n(\mathbf{A}) = GL_n(\mathbf{Q}) \cdot GL_n(\mathbf{R}) \cdot U$$

for any open compact $U \subset GL_n(\mathbf{A}_f)$ on which the image of determinant is equal to $\prod_p \mathbb{Z}_p^\times$.

This is a consequence of strong approximation for SL_n and the fact that $\mathbf{Q}$ has class number 1. So it would not be true for a general number field.

In particular, when $U = K_0'(N)$ (the analog of $K_0(N)$ for GL_2), we have

$$\Gamma_0(N)\backslash SL_2(\mathbf{R}) \cong \Gamma_0'(N)Z(\mathbf{R})\backslash GL_2(\mathbf{R})$$

$$\cong Z(\mathbf{A})GL_2(\mathbf{Q})\backslash GL_2(\mathbf{A})/K_0'(N).$$

Properties of functions

When we regard an automorphic form on $\Gamma\backslash G(\mathbf{R})$ as a function f on $G(\mathbf{Q})\backslash G(\mathbf{A})$, the function f will inherit the properties of an automorphic form. Let us spell out some of these.

Definition:

- A function f on $G(\mathbf{A})$ is said to be **smooth** if it is C^∞ in its archimedean variable, and locally constant in the finite-adeles variable.

- f is K-**finite** if the right K-translates of K span a finite dimension vector space. Equivalently, f is K_∞-finite and is right-invariant under an open compact subgroup of $G(\mathbf{A}_f)$.

- If we fix a $\mathbf{Q}$-embedding $i : G \hookrightarrow GL_n$, we may define a **norm function** by

$$||g|| = \prod_v max_{j,k}\{|i(g)_{jk}|_v, |i(g^{-1})_{jk}|_v\}.$$

- f is said to be of **moderate growth** if there exists $n \geq 0$ and $C > 0$ such that

$$|f(g)| \leq C||g||^n$$

for all g.

- f is **rapidly decreasing on a Siegel set** $\mathfrak{S}$ if, for any $k > 0$, there exists $C_k > 0$ such that

$$|f(g)| \leq C_k \cdot ||g||^{-k}$$

for any $g \in \mathfrak{S}$.

Automorphic forms on adele groups

Let G be a reductive linear algebraic group over $\mathbf{Q}$.

Definition: A function f on $G(\mathbf{Q})\backslash G(\mathbf{A})$ is called an **automorphic form** if

- f is smooth
- f is right K-finite
- f is of moderate growth
- f is $Z(\mathfrak{g})$-finite.

We let $\mathcal{A}(G)$ denote the space of automorphic forms on G. This space contains $\mathcal{A}(G(\mathbf{R}), \Gamma)$ (for any congruence Γ) as the space of functions right-invariant under an open compact subgroup of $G(\mathbf{A}_f)$.

Cusp forms

Definition: An automrophic form f on G is called a **cusp form** if, for any parabolic $\mathbf{Q}$-subgroup $P = MN$ of G, the constant term

$$f_N(g) = \int_{N(\mathbf{Q})\backslash N(\mathbf{A})} f(ng)\, dn$$

is zero as a function on $G(\mathbf{A})$.

It suffices to check this vanishing on a set of representatives of G-conjugacy classes of maximal parabolic subgroups.

We let $\mathcal{A}_0(G)$ denote the space of cusp forms on G.

Analytic properties

The analytic properties of an automorphic form f follow immediately from those on real groups, using the passage from real to adele groups. For example, we have:

- f is real analytic when restricted to $G(\mathbf{R})$

- f is of uniform moderate growth

- if f is cuspidal, f is rapidly decreasing on a Siegel set defined using any parabolic $\mathbf{Q}$-subgroup. In particular, $f \in L^2(G(\mathbf{Q})\backslash G(\mathbf{A}))$ if G is semisimple.

Automorphic representations

The space $\mathcal{A}(G)$ possesses the structure of a $(\mathfrak{g}, K)$-module as before. In addition, for each prime p, the group $G(\mathbf{Q}_p)$ acts on $\mathcal{A}(G)$ by right translation. Thus, $\mathcal{A}(G)$ has the structure of a representation of

$$(\mathfrak{g}, K) \times G(\mathbf{A}_f).$$

Moreover, as a representation of $G(\mathbf{A}_f)$, it is a smooth representation.

We shall abuse terminology, and say that $\mathcal{A}(G)$ is a smooth representation of $G(\mathbf{A})$, even though $G(\mathbf{R})$ does not preserve $\mathcal{A}(G)$.

Definition: An irreducible smooth representation π of $G(\mathbf{A})$ is called an **automorphic representation** if π is a subquotient of $\mathcal{A}(G)$.

Admissibility

Theorem: An automorphic representation π is **admissible**, i.e. given any irreducible representation ρ of K, the multiplicity with which ρ occurs in π is finite.

An equivalent definition of "admissibility" is: for any open compact subgroup $U = \prod_p U_p \subset G(\mathbf{A}_f)$ and any irreducible representation ρ_∞ of K_∞, the subspace of vectors in π which are fixed by U and which is contained in the ρ_∞-isotypic subspace of π^U is finite-dimensional.

Proof: Suppose that $V_1 \subset V_2 \subset \mathcal{A}(G)$ are submodules with $V_2/V_1 \cong \pi$. We may assume that V_2 is generated over $G(\mathbf{A})$ by $f \in V_2 \setminus V_1$. Otherwise, we simply replace V_2 by the $G(\mathbf{A})$-submodule V_2' generated by f and V_1 by $V_1 \cap V_2'$.

If f is killed by an ideal J of finite codimension in $Z(\mathfrak{g})$, then V_2 is killed by J. Thus

$$V_2^U \subset \mathcal{A}(G, J)^U$$

and

$$\mathcal{A}(G, J)^U \cong \oplus_{i=1}^r \mathcal{A}(G(\mathbf{R}), \Gamma_i, J).$$

The RHS is an admissible $(\mathfrak{g}, K)$-module by a fundamental theorem of Harish-Chandra. This proves the theorem.

Restricted tensor product

We usually expect an irreducible representation of a direct product of groups G_i to be the tensor product of irreducible representations V_i of G_i. In the case of interest here, the adele group $G(\mathbf{A})$ is almost a direct product; it is a restricted direct product with respect to a family (K_p). It turns out that an irreducible admissible representation of $G(\mathbf{A})$ is almost a tensor product.

Definition: Suppose we have a family (W_v) of vector spaces, and for almost all v, we are given a non-zero vector $u_v^0 \in W_v$. The **restricted tensor product** $\otimes_v' W_v$ of the W_v's with respect to (u_v^0) is the inductive limit of $\{W_S = \otimes_{v \in S} W_v\}$, where for $S \subset S'$, one has $W_S \longrightarrow W_{S'}$ defined by

$$\otimes_{v \in S} u_v \mapsto (\otimes_{v \in S} u_v) \otimes (\otimes_{v \in S' \setminus S} u_v^0).$$

We think of $\otimes'_v W_v$ as the vector space generated by the elements

$$u = \otimes_v u_v \qquad \text{with } u_v = u_v^0 \text{ for almost all } v,$$

subject to the usual linearity conditions in the definition of the usual tensor product.

Now if each W_v is a representation of $G(\mathbf{Q}_v)$, and for almost all v, the distinguished vector u_v^0 is fixed by the maximal compact K_v, then the restricted tensor product inherits an action of $G(\mathbf{A})$: if $g = (g_v)$, then

$$g(\otimes_v u_v) = \otimes_v g_v u_v.$$

Because $g_v \in K_v$ and $u_v = u_v^0$ for almost all v, the resulting vector still has the property that almost all its local components are equal to the distinguished vector u_v^0.

Representations of adele groups

The following is a theorem of Flath:

Theorem: An irreducible admissible representation of $G(\mathbf{A})$ is a restricted tensor product of irreducible admissible representations π_v of $G(\mathbf{Q}_v)$ with respect to a family of vectors (u_v^0) such that $u_v^0 \in \pi_v^{K_v}$ for almost all v.

For the proof of this, see [Bump, §3.4].

Corollary: An automorphic representation π has a restricted tensor product decomposition: $\pi \cong \otimes'_v \pi_v$, where for almost all v, $\pi_v^{K_v} \neq 0$.

Unramified representations

Remarks: Note that if π_p is an irreducible admissible representation of $G(\mathbf{Q}_p)$, and K_p is a hyperspecial maximal compact subgroup of $G(\mathbf{Q}_p)$, then

$$\dim \pi_p^{K_p} \leq 1.$$

So the choice of u_p^0 is unique up to scaling.

We call an irreducible representation of $G(\mathbf{Q}_p)$ **unramified** or **spherical** with respect to K_p if $\dim \pi_p^{K_p} = 1$. These has been classified in Yu's lectures, using the **Satake isomorphism**.

We shall come back to this later, when we give a representation theoretic interpretation of the Ramanujan-Petersson conjecture.

Cuspidal automorphic representations

The space $\mathcal{A}_0(G)$ of cusp forms is clearly a submodule under $G(\mathbf{A})$. When G is reductive, with center Z, we usually specify a central character χ for $Z(\mathbf{A})$. Namely, if χ is a character of $Z(\mathbf{Q})\backslash Z(\mathbf{A})$, then we let $\mathcal{A}(G)_\chi$ be the subspace of automorphic forms f which satisfy:

$$f(zg) = \chi(z) \cdot f(g).$$

We let $\mathcal{A}_0(G)_\chi$ be the subspace of cuspidal functions in $\mathcal{A}(G)_\chi$. Then as in the case of $\mathcal{A}_0(G(\mathbf{R}), \Gamma)$ (with G semisimple), $\mathcal{A}_0(G)_\chi$ decomposes as the direct sum of irreducible representations of $G(\mathbf{A})$, each occurring with finite multiplicities.

Definition: A representation π of $G(\mathbf{A})$ is **cuspidal** if it occurs as a submodule of $\mathcal{A}_0(G)_\chi$.

The representation π_f

If f is a classical cuspidal Hecke eigenform on $\Gamma_0(N)$, we have seen that f gives rise to an automorphic form ϕ_f on $\Gamma_0'(N)\backslash PGL_2(\mathbf{R})$ which generates an irreducible $(\mathfrak{g}, K)$-module isomorphic to the discrete series representation of lowest weight k.

Now if we then transfer ϕ_f to a cusp form Φ_f on $PGL_2(\mathbf{Q})\backslash PGL_2(\mathbf{A})$, we can consider the subrepresentation π_f of $\mathcal{A}_0(PGL_2)$ generated by Φ_f. It turns out that this is an irreducible representation of $G(\mathbf{A})$ if f is a newform.

Thus a Hecke eigen-newform in $S_k(N)$ corresponds to a cuspidal representation of $PGL_2(\mathbf{A})$. Moreover, if $\pi_f \cong \otimes_v' \pi_v$, then π_p is unramified for all p not dividing N.

Basic questions

Having defined the notion of automorphic representations, some basic questions one can ask is:

- Given an irreducible admissible representation $\pi = \otimes'_v \pi_v$ of $G(\mathbf{A})$, can we decide if π is automorphic? When is it cuspidal?

- More generally, classify the automorphic representations of $G(\mathbf{A})$. One purpose of the Langlands program is to formulate an answer to this question.

- Construct some examples of automorphic representations. It turns out that there is a general method of constructing submodules of $\mathcal{A}(G)$. This is the theory of Eisenstein series. However, there is no known general method for constructing submodules of $\mathcal{A}_0(G)$.

Hecke algebra

On $\mathcal{A}(G(\mathbf{R}), \Gamma)$, we have the action of the Hecke algebra $\mathcal{H}(G, \Gamma)$. Under the isomorphism

$$\mathcal{A}(\Gamma \backslash G(\mathbf{R})) \cong \mathcal{A}(G)^{U_\Gamma},$$

what does the operator $T_\alpha = \Gamma \alpha \Gamma$ (with $\alpha \in G(\mathbf{Q})$) get translated to?

Let's call this new operator on the adelic side T'_α.

If $f \in \mathcal{A}(G)^{U_\Gamma}$, then the identification of f with a function on $\Gamma \backslash G(\mathbf{R})$ is simply given by restriction to $G(\mathbf{R})$. So we want to find T'_α such that

$$(T'_\alpha f)|_{G(\mathbf{R})} = T_\alpha (f|_{G(\mathbf{R})}).$$

Let us evaluate $T'_\alpha f$ at an element $g = (g_\infty, g_f) \in G(\mathbf{R}) \times G(\mathbf{A}_f)$. Because we are assuming that

$$G(\mathbf{A}_f) = G(\mathbf{Q}) \cdot U_\Gamma,$$

we can accordingly write

$$g_f = \gamma \cdot u.$$

Writing:

$$\Gamma\alpha\Gamma = \bigcup_i \Gamma a_i,$$

we compute

$$\begin{aligned} T'_\alpha f(g) &= T'_\alpha f(g_\infty, \gamma u) &&= T'_\alpha f(\gamma^{-1} g_\infty, 1) \\ &= \sum_i f(a_i \gamma^{-1} g_\infty, 1) &&= \sum_i f(g_\infty, \gamma a_i^{-1}) \\ &= \sum_i f(g_\infty, \gamma u u^{-1} a_i^{-1}) &&= \sum_i f(g_\infty, \gamma u a_i^{-1}) \\ &= \sum_i f(g_\infty, g_f a_i^{-1}) &&= \sum_i f(g a_i^{-1}) \end{aligned}$$

Note that because

$$\Gamma\alpha\Gamma = \bigcup_i \Gamma a_i \quad \text{in } G(\mathbf{Q})$$

we have

$$U_\Gamma \alpha U_\Gamma = \bigcup_i U_\Gamma a_i \quad \text{in } G(\mathbf{A}_f)$$

and so

$$U_\Gamma \alpha^{-1} U_\Gamma = \bigcup_i a_i^{-1} U_\Gamma.$$

We have now translated the action of $\mathcal{H}(G, \Gamma)$ to the adelic picture. Is the resulting operator something we have seen before?

Recollection from Yu's lectures

Recall from Yu's lectures that if V is a smooth representation of a locally profinite group G and $U \subset G$ is an open compact subgroup, then the map $V \mapsto V^U$ defines a functor from the category of smooth representatioons of G to the category of modules for the Hecke agebra $\mathcal{H}(G//U)$.

Recall that $\mathcal{H}(G//U)$ is the ring of functions in $C_c^\infty(G)$ which are bi-U-invariant, and the product is given by convolution of functions.

A basis for $\mathcal{H}(G//U)$ is given by the characteristic functions $f_\alpha = 1_{U\alpha U}$. The action of this on a vector in V^U is:

$$f_\alpha \cdot v = \int_G f_\alpha(g) \cdot gv\, dg$$

$$= \int_{U\alpha U} vdg = \sum_i a_i v$$

if $U\alpha U = \bigcup a_i U$ (and dg gives U volume 1).

Adelic Hecke algebras

We can apply the material from Yu's lectures to the smooth representation $\mathcal{A}(G)$ of $G(\mathbf{A}_f)$. Then the adelic Hecke algebra $\mathcal{H}(G(\mathbf{A}_f)//U_\Gamma)$ acts on $\mathcal{A}(G)^{U_\Gamma}$.

More explicitly, if $U_\Gamma \alpha^{-1} U_\Gamma = \bigcup_i a_i^{-1} U_\Gamma$, then the characteristic function of $U_\Gamma \alpha^{-1} U_\Gamma$ acts by

$$(T_\alpha f)(g) = \sum_i (a_i^{-1} f)(g) = \sum_i f(g a_i^{-1})$$

In conclusion, we see that the action of $\mathcal{H}(G,\Gamma)$ on $\mathcal{A}(G,\Gamma)$ gets translated to an action of the adelic Hecke algebra $\mathcal{H}(G(\mathbf{A}_f)//K_\Gamma)$ on $\mathcal{A}(G)^{K_\Gamma}$. This action of the adelic Hecke algebra arises from the smooth $G(\mathbf{A}_f)$-module structure on $\mathcal{A}(G)$. Note also that

$$\Gamma\backslash G(\mathbf{Q})/\Gamma \leftrightarrow U_\Gamma\backslash G(\mathbf{A}_f)/U_\Gamma.$$

Local Hecke algebras

Because $G(\mathbf{A}_f)$ is a restricted direct product, we have in fact

$$\mathcal{H}(G(\mathbf{A}_f)//U) \cong \otimes'_v \mathcal{H}(G(\mathbf{Q}_p)//U_p)$$

if $U = \prod_p U_p$. So the structure of $\mathcal{H}(G(\mathbf{A}_f)//U)$ is known once we understand the local Hecke algebras $\mathcal{H}(G(\mathbf{Q}_p)//U_p)$.

For almost all p, however, we know that $U_p = K_p$ is a hyperspecial maximal compact subgroup. In that case, the structure of the local Hecke algebra is known, by the **Satake isomorphism**. In particular, $\mathcal{H}(G(\mathbf{Q}_p)//K_p)$ is commutative and its irreducible modules are classified.

Because $V \mapsto V^{K_p}$ induces a bijection of irreduible unramified representations with irreducible modules of $\mathcal{H}(G(\mathbf{Q}_p)//K_p)$, we get in this way the classification of irreducible unramified representations of $G(\mathbf{Q}_p)$. We recall this classification next.

Classification of unramified representations

Let us assume for simplicity that G is a split group (e.g. $G = GL_n$). Let $B = T \cdot N$ be a Borel subgroup of G, with maximal torus T. So $T \cong (GL_1)^r$ and $T(\mathbf{Q}_p) \cong (\mathbf{Q}_p^\times)^r$. We let $W := N_G(T)/T$ be the Weyl group of G.

Let $\chi : T(\mathbf{Q}_p) \longrightarrow \mathbf{C}^\times$ be a (smooth) character of $T(\mathbf{Q}_p)$. We say that χ is an **unramifed** character if χ is trivial when restricted to $T(\mathbf{Z}_p) \cong (\mathbf{Z}_p^\times)^r$. If χ is unramified, then it is of the form

$$\chi(a_i, ..., a_r) = t_1^{ord_p(a_1)} \cdot ... \cdot t_r^{ord_p(a_r)}, \quad a_i \in \mathbf{Q}_p^\times$$

for some $s_i \in \mathbf{C}^\times$.

We may regard χ as a character of $B(\mathbf{Q}_p)$ using the projection $B(\mathbf{Q}_p) \to N(\mathbf{Q}_p)\backslash B(\mathbf{Q}_p) \cong T(\mathbf{Q}_p)$.

Given an unramified character χ of $T(\mathbf{Q}_p)$, we may form the induced representation

$$I_B(\chi) := Ind_{B(\mathbf{Q}_p)}^{G(\mathbf{Q}_p)} \delta_B^{1/2} \cdot \chi.$$

Here, δ_B is the modulus character of B, defined by:

$$\delta_B(b) = |det(Ad(b)|Lie(N))|_p.$$

We recall that the space of $I_B(\chi)$ is the subspace of $C^\infty(G(\mathbf{Q}_p))$ satisfying:

- $f(bg) = \delta(b)^{1/2} \cdot \chi(b) \cdot f(g)$ for any $b \in B(\mathbf{Q}_p)$ and $g \in G(\mathbf{Q}_p)$.

- f is right-invariant under some open compact subgroup U_f of $G(\mathbf{Q}_p)$.

Then $I_B(\chi)$ is an admissible representation of $G(\mathbf{Q}_p)$, possibly reducible. The representations $I_B(\chi)$ are called the **principal series representations**.

Because of the Iwasawa decomposition $G(\mathbf{Q}_p) = B(\mathbf{Q}_p) \cdot K_p$, an element f of $I_B(\chi)$ is completely determined by its restriction to K_p. Thus we see that

$$\dim I_B(\chi)^{K_p} = 1$$

and a vector in this 1-dimensional space is given by

$$f_0(bk) = \delta_B(b)^{1/2} \cdot \chi(b);$$

equivalently, $f_0|_{K_p}$ is the constant function 1.

Thus $I_B(\chi)$ has a unique irreducible subquotient π_χ with the property that $\pi_\chi^{K_p} \neq 0$.

Theorem: Any irreducible unramified representation of $G(\mathbf{Q}_p)$ is of the form π_χ for some unramfied character χ of $T(\mathbf{Q}_p)$.

The Weyl group W acts naturally on $T(\mathbf{Q}_p)$ and thus on the set of characters of $T(\mathbf{Q}_p)$: Namely, for $w \in W$,

$$(w\chi)(t) = \chi(w^{-1}tw).$$

Proposition: $\pi_\chi \cong \pi_{\chi'}$ iff $\chi = w\chi'$ for some $w \in W$.

Thus, the irreducible unramified representations are classified by W-orbits of unramified characters of $T(\mathbf{Q}_p)$.

Unitarizability: One may ask whether π_χ is unitarizable. When χ is a unitary character, then it is clear that $I_B(\chi)$ is unitarizable, and thus so is π_χ. Indeed, a $G(\mathbf{Q}_p)$-invariant inner product on $I_B(\chi)$ is:

$$\langle f_1, f_2 \rangle = \int_{K_p} f_1(k) \cdot \overline{f_2(k)}\, dk$$

Though this inner product is not $G(\mathbf{Q}_p)$-invariant if χ is not unitary, it may be possible to define an invariant inner product in some other ways.

At this point, it appears that the problem of determining the unitarizable unramified representations is not completely solved for all groups. Of course, it has been solved for GL_n some time ago.

The example of GL_2

Let us look at the example of GL_2. Then B is the group of upper triangular matrices, and

$$\delta_B \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = |a/d|_p.$$

The Weyl group W is isomorphic to the group S_2; the non-trivial element of W is represented by the matrix

$$w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

An unramified character is of the form

$$\chi_{t_1,t_2} \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} = t_1^{ord_p(a)} \cdot t_2^{ord_p(d)}.$$

Moreover, under the action of w, we see that

$$w\chi_{t_1,t_2} = \chi_{t_2,t_1}.$$

So the irreducible unramified representations of $GL_2(\mathbf{Q}_p)$ are parametrized by diagonal matrices

$$t_\chi = \begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix}$$

modulo the action of w. In other words, they are parametrized by *semisimple conjugacy clases* in $GL_2(\mathbf{C})$.

Observe that $\pi_{\chi_{t_1,t_2}}$ is a representation of the group $PGL_2(\mathbf{Q}_p)$ iff $t_1 t_2 = 1$.

Analogously, the irreducible unramified representations of $GL_n(\mathbf{Q}_p)$ are naturally parametrized by semisimple conjugacy classes in $GL_n(\mathbf{C})$. The semisimple class associated to an unramified representation π is called the **Satake parameter** of π.

Tempered representations

Let us recall the notion of a tempered representation of $G(\mathbf{Q}_p)$. Let π be a unitarizable representation of $G(\mathbf{Q}_p)$. Suppose that $\langle -, - \rangle$ is a $G(\mathbf{Q}_p)$-invariant inner product on π. Then by a matrix coefficient of π is a function on $G(\mathbf{Q}_p)$ of the form

$$f_{v_1,v_2}(g) = \langle gv_1, v_2 \rangle$$

with v_1 and v_2 in π.

Assume that π has a unitary central character, so that $|f_{v_1,v_2}(g)|$ is a functon on $Z(\mathbf{Q}_p)\backslash G(\mathbf{Q}_p)$.

Definition: Say that π is a **tempered representation** if the matrix coefficients of π lie in $L^{2+\epsilon}(Z(\mathbf{Q}_p)\backslash G(\mathbf{Q}_p))$ for any $\epsilon > 0$.

Proposition: The unramified representation π_χ is tempered iff χ is a unitary character.

We mentioned earlier that there may be other π_χ's which are unitarizable, but for which χ is not unitary. We call these other π_χ's the (unramified) **complementary series representations**. For example, the trivial representation of $G(\mathbf{Q}_p)$ is certainly not tempered.

Remarks: The notion of being tempered is a natural one. Indeed, if one considers the regular representation $L^2(G(\mathbf{Q}_p))$ (G semisimple, say), which decomposes into the sum of a discrete spectrum and a continuous one, then an irreducible unitarizable representation is tempered if and only if it occurs in the decomposition of $L^2(G(\mathbf{Q}_p))$.

Example: Let us describe the (unramified) complementary series of $PGL_2(\mathbf{Q}_p)$.

Proposition: $\pi_{\chi_{t_1,t_2}}$ is unitarizable and non-tempered iff $p^{-1/2} \leq t_1 \leq p^{1/2}$.

When $t_1 = p^{1/2}$ or $p^{-1/2}$, the corresponding unramified representation is the trivial representation.

Note that t_1 and t_1^{-1} parametrize the same representation.

Reformulating Ramanujan's conjecture

We are now in a position to reformulate the Ramanujan-Petersson conjecture in terms of representation theory. This reformulation is due to Satake.

We start with a cuspidal Hecke eigenform f of weight k for $SL_2(\mathbf{Z})$ with Fourier coefficients $\{a_n(f)\}$. The Ramanujan-Petersson conjecture says:

$$|a_p(f)| \leq 2 \cdot p^{(k-1)/2}.$$

Obviously, since $S_k(1)$ is finite dimensional, it suffices to prove this bound for a basis of $S_k(1)$. Recall that the action of the Hecke operators $\{T_n\}$ can be simultaneously diagonalized. So we have a natural basis of $S_k(1)$ consisting of Hecke eigenforms. We can further assume that these Hecke eigenforms are normalized, i.e. $a_1(f) = 1$.

Now assume that f is a normalized Hecke eigenform and suppose that

$$T_n f = \lambda_n(f) f.$$

Then we have seen that

$$a_n(f) = \lambda_n(f).$$

Thus, the Ramanujan-Petersson conjecture is equivalent to saying that the Hecke eigenvalues λ_p of T_p occurring in $S_k(1)$ satisfy

$$|\lambda_p| \leq 2p^{(k-1)/2}$$

Now we want to reformulate this on $\mathcal{A}(PGL_2(\mathbf{R}), \Gamma)$, wit $\Gamma = PGL_2(\mathbf{Z})$.

We saw that f gives rise to a function

$$\phi_f(g) := (f|_k g)(i)$$

on $SL_2(\mathbf{Z})\backslash SL_2(\mathbf{R}) \cong PGL_2(\mathbf{Z})\backslash PGL_2(\mathbf{R})$. On $\mathcal{A}(PGL_2, \Gamma)$, we have the Hecke algebra $\mathcal{H}(PGL_2, \Gamma)$ acting and we showed that:

$$\phi_{T_p f} = p^{k/2-1} \cdot T_{\alpha_p} \phi_f$$

where

$$\alpha_p = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus the Ramanujan-Petersson conjecture says the eigenvalue ν_p of ϕ_f with respect to T_{α_p} satisfies:

$$\nu_p = 2p^{1/2},$$

since $\lambda_p = p^{k/2-1}\nu_p$.

Now we pass to $\mathcal{A}(PGL_2)^{K_f}$ where $K_f = \prod_p PGL_2(\mathbf{Z}_p)$.

If we regard ϕ_f as an element Φ_f of $\mathcal{A}(PGL_2)^{K_f}$, then we saw that

$$T_{\alpha_p}\phi_f = T_{\alpha_p^{-1}}\Phi_f.$$

Here $T_{\alpha_p^{-1}}$ is the element in the adelic Hecke algebra corresponding to the double coset

$$K_f \begin{pmatrix} p^{-1} & 0 \\ 0 & 1 \end{pmatrix} K_f = K_f \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} K_f,$$

and is equal to T_{α_p}.

Further, under the decomposition

$$\mathcal{H}(PGL_2(\mathbf{A}_f)//K_f) = \otimes_p' \mathcal{H}(PGL_2(\mathbf{Q}_p)//K_p),$$

we see that T_{α_p} is supported only at the prime p.

Now suppose $\pi_f \subset \mathcal{A}_0(PGL_2)$ is the irreducible representation generated by Φ_f, and $\pi_f \cong \otimes_v' \pi_v$.

The action of T_{α_p} on $\pi_f^{K_f} = \mathbf{C} \cdot \Phi_f$ is simply the action of the characteristic functon t_p of $PGL_2(\mathbf{Z}_p)\alpha_p PGL_2(\mathbf{Z}_p)$ on $\pi_p^{K_p}$. So the eigenvalue of t_p on π_p is ν_p.

Now the local Hecke algebra $\mathcal{H}(PGL_2(\mathbf{Q}_p)//K_p)$ is generated as an algebra by t_p. So the unramified representation π_p is completely determined by the eigenvalue of t_p on $\pi_p^{K_p}$. So the Satake parameter s_p of π_p is completely determined by the eigenvalue ν_p associated to Φ_f.

So we ask:

What is the relation between the Hecke eigenvalue ν_p and the Satake parameter s_p of π_p?

Proposition: $p^{1/2} \cdot Trace(s_p) = \nu_p$.

Proof: If $s_p = diag(t_1, t_2)$, then as we have seen,

$$\pi_p = I_B(\chi_{t_1,t_2})$$

If f_0 is the K_p-fixed vector in $I_B(\chi)$ with $f_0(1) = 1$, then

$$\begin{aligned}\nu_p &= (t_p f_0)(1) \\ &= \sum_{r=0}^{p-1} f_0 \begin{pmatrix} 1 & r \\ 0 & p \end{pmatrix} + f_0 \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \\ &= p^{1/2}(t_2 + t_1) \\ &= p^{1/2} \cdot Trace(s_p).\end{aligned}$$

Corollary: The Ramanujan-Petersson conjecture for $f \in S_k(1)$ is equivalent to saying that $\pi_{f,p}$ is tempered for all p.

Proof: Since π_p is necessarily unitarizable, π_p is either tempered or in the complementary series. If the Satake parameter is $s_p = diag(t, t^{-1})$, then π_p is complementary series iff

$$p^{-1/2} \leq t \leq p^{1/2}, \quad \text{but } t \neq 1,$$

which is equivalent to

$$|Trace(s_p)| > 2.$$

Remarks: Recall that the Selberg conjecture for the eigenvalues of the hyperbolic Laplacian on a Maass form is equivalent to saying that $\pi_{f,\infty}$ is tempered. Thus we have shown that the Selberg conjecture and the Ramanujan conjecture can be unified in a single statement in terms of the representation π_f.

Ramanujan Conjecture for GL_n

Let $\pi = \otimes_v' \pi_v$ be a cuspidal automorphic representation of GL_n with unitary central character. Then for each v, π_v is tempered.

In the conference next week, you will probably hear some progress towards this conjecture.

<u>Eisenstein Series</u>

GOAL: We will discuss a standard construction of automorphic representations: the theory of Eisenstein series.

Let $P = M \cdot N$ be a parabolic subgroup of G and let π be an automorphic representation of $M(\mathbf{A})$. Consider the induced representation

$$I_P(\pi) = Ind_{P(\mathbf{A})}^{G(\mathbf{A})} \delta_P^{1/2} \cdot \pi$$

Then the theory of Eisenstein series gives rise to a $G(\mathbf{A})$-intertwining map

$$E : I_P(\pi) \longrightarrow \mathcal{A}(G),$$

thus giving us concrete examples of automorphic representations.

Unfortunately, this construction does not provide cuspidal representations; indeed the image of E is orthogonal to $\mathcal{A}_0(G)$.

However, the automorphic forms in the image of E turn out to be very useful. For example, the theory of Eisenstein series is one of the most important tool we have for understanding the properties of automorphic L-functions. Moreover, they are necessary for the spectral decomposition of $L^2(G(\mathbf{Q})\backslash G(\mathbf{A}))$, as we saw in Labesse's lectures.

Parabolic induction

Let us recall the notion of parabolic induction. Let P be a parabolic $\mathbf{Q}$-subgroup of G, with Levi decomposition $P = M \cdot N$. We shall assume that P is a **maximal** parabolic.

The reason for this assumption is that the Eisenstein series will then be a function on $\mathbf{C}$ (taking values in $\mathcal{A}(G)$). If P is not maximal, then the Eisenstein series will be a function on $\mathbf{C}^n$ ($n \geq 2$), in which case results are harder to state.

(The standard example to keep in mind is the case of $B \subset PGL_2$).

The group $P(\mathbf{A})$ is not unimodular. So P has a modulus character δ_P. This character is trivial on $N(\mathbf{A})$ and its value on $M(\mathbf{A})$ is given by

$$\delta_P(m) = |\,\mathrm{det}(Ad(m)|Lie(N)(\mathbf{A}))|.$$

Given an abstract representation σ of $M(\mathbf{A})$, we can inflate (or pullback) σ to a representation of $P(\mathbf{A})$. Then one has the induced representation:

$$I_P(\sigma) := Ind_{P(\mathbf{A})}^{G(\mathbf{A})} \delta_P^{1/2} \cdot \sigma$$

More generally, we set

$$I_P(\sigma, s) := Ind_{P(\mathbf{A})}^{G(\mathbf{A})} \delta_P^{1/2+s} \cdot \sigma$$

Recall that the vector space for $I_P(\sigma, s)$ is the set of smooth functions

$$f : G(\mathbf{A}) \longrightarrow V_\sigma$$

such that

- $f(pg) = \delta_P(p)^{1/2+s} \cdot \sigma(p)(f(g))$

- f is right K-finite.

The action of $G(\mathbf{A})$ on $I_P(\sigma, s)$ is by right translation.

(In the standard example, σ is a character χ of $T(\mathbf{A}) \cong \mathbf{A}^\times$).

Flat sections

Because of Iwasawa decomposition $G(\mathbf{A}) = P(\mathbf{A}) \cdot K$, an element in $I_P(\sigma, s)$ is determined by its restriction to K. Indeed, restriction to K gives an isomorphism of vector spaces from $I_P(\sigma, s)$ to the space of smooth K-finite functions

$$f : K \to V_\sigma$$

satisfying

$$f(mk) = \sigma(m)(f(k)), \text{ for all } m \in M(\mathbf{A}) \cap K.$$

Given a function f in the latter space, we can extend it to an element $f_s \in I_P(\sigma, s)$. The family $\{f_s : s \in \mathbf{C}\}$ is called a **flat section**: the restriction of f_s to K is independent of s (it is equal to the f we started with). Sometimes, people also call it a **standard section**.

The case of automorphic σ

We shall assume that σ is a irreducible submodule of $\mathcal{A}(M)$ and so $V_\sigma \subset \mathcal{A}(M)$.

(In the standard example, σ is a character of the idele class group $T(\mathbf{Q})\backslash T(\mathbf{A}) = \mathbf{Q}^\times\backslash\mathbf{A}^\times$).

In this case, we can realize $I_P(\sigma, s)$ as $\mathbf{C}$-valued functions, rather than functions valued in V_σ. Indeed, if $f_s \in I_P(\sigma, s)$, then set

$$\tilde{f}_s(g) = [f_s(g)](1).$$

The function $\tilde{f}$ satisfies:

$$\tilde{f}_s(nmg) = f_s(g)(m).$$

In particular,

$$\tilde{f}_s : N(\mathbf{A})M(\mathbf{Q})\backslash G(\mathbf{A}) \longrightarrow \mathbf{C}.$$

(In the standard example, there is no need for this, because f_s is already $\mathbf{C}$-valued).

This $\tilde{f}_s$ has the property that for any $k \in K$, the function

$$m \mapsto \tilde{f}_s(mk)$$

is an element of $V_\sigma \subset \mathcal{A}(M)$. Moreover, if f_s is a flat section, then the element $\tilde{f}_s(-k) \in \mathcal{A}(M)$ is independent of s.

Formation of Eisenstein series

We want to make an automorphic form on $G(\mathbf{A})$ out of $\tilde{f}$. Since $\tilde{f}$ is only left-invariant under $P(\mathbf{Q})$ but not $G(\mathbf{Q})$, the easiest way to do this is to average over $P(\mathbf{Q})\backslash G(\mathbf{Q})$.

More precisely, let $f_s \in I_P(\sigma, s)$ be a flat section, whose restriction to K is a function f. We define the following function on $G(\mathbf{A})$:

$$E(f, s, g) = \sum_{\gamma \in P(\mathbf{Q})\backslash G(\mathbf{Q})} \tilde{f}_s(\gamma g)$$

Formally, this function on $G(\mathbf{A})$ is left-invariant under $G(\mathbf{Q})$. But we need to address convergence.

Convergence and properties

Here is the result on convergence:

Proposition:

(i) There exists $c > 0$ such that the above sum converges absolutely for any f and g when $Re(s) > c$. The convergence is locally uniform in g.

(ii) The function $E(f, s, g)$ is an automorphic form on G.

(iii) For fixed s, the map

$$f \mapsto E(f, s, -)$$

is a $G(\mathbf{A})$-equivariant map

$$I_P(\sigma, s) \longrightarrow \mathcal{A}(G).$$

(iv) In the half plane $Re(s) > c$, the function $s \mapsto E(f, s, g)$ (with f and g fixed) is holomorphic.

<u>Constant term of Eisenstein series</u>

The main result in the theory of Eisenstein series is the meromorphic continuation of $E(f, s, g)$ to $s \in \mathbf{C}$. A important ingredient in the proof of meromorphic continuation is the computation of the constant term of $E(f, s, g)$ along N.

Simplifying assumption:

Assume that σ is **cuspidal** and P is conjugate to its opposite parabolic. (This is automatic in the standard example).

Let $W_M = N_G(M)/M$. Because P is maximal parabolic, $W_M \cong S_2$, and we let w be the nontrivial element in W_M. We have:

Proposition: Assume $Re(s) >> 0$. Then

$$E_N(f, s, g) = \tilde{f}_s(g) + M_w(\sigma, s)\tilde{f}_s(g)$$

where $M_w(\sigma, s)\tilde{f}_s(g)$ is defined for $Re(s) >> 0$ by the absolutely convergent integral

$$M_w(\sigma, s)\tilde{f}(g) = \int_{N(\mathbf{A})} \tilde{f}_s(wng)\, dn.$$

The map $\tilde{f}_s \mapsto M_w(\sigma, s)\tilde{f}_s$ is a $G(\mathbf{A})$-equivariant map

$$I_P(\sigma, s) \longrightarrow I_P(w \cdot \sigma, -s).$$

Remarks: The operator $M_w(\sigma, s)$ is called a **standard intertwining operator**. It appears naturally in the constant term of the Eisenstein series, and is intricately connected with the properties of the Eisenstein series. In the course of proving that $E(f, s, g)$ has meromorphic continuation, one proves simultaneously that $M_w(\sigma, s)$ has meromorphic continuation. Observe that if $f = \prod_v f_v$, then $M_w(\sigma, s)$ factors into the product of **local intertwining operators** $M_{w,v}(\sigma_v, s)$. Thus this operator can often be analyzed locally.

The standard example

Let us compute the constant term in our standard example.

We have the Bruhat decomposition

$$G(\mathbf{Q}) = B(\mathbf{Q}) \bigcup B(\mathbf{Q}) w N(\mathbf{Q}).$$

So

$$B(\mathbf{Q})\backslash G(\mathbf{Q}) \leftrightarrow \{1, wN(\mathbf{Q})\}$$

and we have:

$$\sum_{\gamma \in B(\mathbf{Q})\backslash G(\mathbf{Q})} f_s(\gamma g) = f_s(g) + \sum_{\gamma \in N(\mathbf{Q})} f_s(w\gamma g).$$

This is what the previous proposition asserts.

Now we compute:

$$E_N(f,s,g)$$

$$=\int_{N(\mathbf{Q})\backslash N(\mathbf{A})} E(f,s,ng)\,dn$$

$$=\int_{N(\mathbf{Q})\backslash N(\mathbf{A})} \sum_{\gamma\in B(\mathbf{Q})\backslash G(\mathbf{Q})} f_s(\gamma ng)\,dn$$

$$=\int_{N(\mathbf{Q})\backslash N(\mathbf{A})} f_s(ng)\,dn +$$

$$\int_{N(\mathbf{Q})\backslash N(\mathbf{A})} \sum_{\gamma\in N(\mathbf{Q})} f_s(w\gamma ng)\,dn$$

$$= f_s(g) + \int_{N(\mathbf{A})} f_s(wng)\,dn$$

Meromorphic continuation

The following is the first main result of the theory.

Theorem:

(i) The function $s \mapsto E(f, s, g)$ can be continued to a meromorphic function on $\mathbf{C}$.

(ii) At a point s_0 where $E(f, s, g)$ is holomorphic for all g, the function $E(f, s_0, g)$ of g is an automorphic form.

(iii) At a point s_0 where $E(f, s, g)$ is holomorphic for all f and g, the map $f \mapsto E(f, s_0, -)$ is a $G(\mathbf{A})$-equivariant map of $I_P(\sigma, s_0)$ to $\mathcal{A}(G)$.

(iv) For any s_0, there is a constant N such that

$$\inf_{f,g}\{ord_{s=s_0} E(f, s, g)\} = -N.$$

Laurent expansion

The last part of the theorem says that one has a Laurent expansion about the point s_0:

$$E(f, s, g) =$$

$$\frac{a_{-N}(f, s_0, g)}{(s - s_0)^N} + \frac{a_{-(N-1)}(f, s_0, g)}{(s - s_0)^{N-1}} + \dots.$$

with $a_{-N}(f, s_0, g) \neq 0$ for some f and g.

Parts (ii) and (iii) can be extended to all $s_0 \in \mathbf{C}$, provided we use the **leading term** of the Laurent expansion at the point s_0.

More precisely, we have:

Theorem:

(i) For any i, the function $a_i(f, s_0, g)$ is an automorphic form on G.

(ii) The map $f \mapsto a_{-N}(f, s_0, -)$ is a $G(\mathbf{A})$-equivariant map $I_P(\sigma, s_0) \to \mathcal{A}(G)$.

Remarks: In part (ii), if we had used a_i, with $i \neq -N$, the map will not be $G(\mathbf{A})$-equivariant. One can only say that the composite map

$$a_i : I_P(\sigma, s_0) \to \mathcal{A}(G) \to \mathcal{A}(G)/\langle Im(a_k) : k < i\rangle$$

is equivariant.

Global intertwining operators

In the course of proving the meromorphic continuation of Eisenstein series, one also proves:

Theorem:

(i) The intertwining operator $M_w(\sigma, s)\tilde{f}_s$ has a meromorphic continuation to all of $\mathbf{C}$.

(ii) At each $s_0 \in \mathbf{C}$, the order of poles is bounded (as f varies).

(iii) the leading term of the Laurent expansion o $M_w(\sigma, s)$ at s_0 is an intertwining operator $I_P(\sigma, s) \to I_P(w \cdot \sigma, -s)$.

(iv) $M_w(w \cdot \sigma, -s) \circ M_w(\sigma, s) = id$.

As a consequence, the identity

$$E_N(f, s, g) = \tilde{f}_s(g) + M_w(\sigma, s)\tilde{f}_s(g)$$

holds for all $s \in \mathbf{C}$.

Functional equation

A consequence of the formula for the constant term E_N is that we have a functional equation for the Eisenstein series:

Theorem:

$$E(f, s, g) = E(M_w(s)f, -s, g).$$

Proof: Both sides have the same constant terms along (using the formula for the constant term E_N as well as (iv) of the last theorem). Moreover, under our simplifying assumption, the constant term along other parabolic subgroups are zero.

So the difference of the two sides is a cusp form. But as we shall see next, each Eisenstein series is orthogonal to the cusp forms. This shows that the difference is zero.

Orthogonality to cusp forms

The Eisenstein series $E(f,s,g)$ is orthogonal to the space $\mathcal{A}_0(G)$ of cusp forms. Suppose that $\phi \in \mathcal{A}_0(G)$. Then for $Re(s)$ large,

$$\int_{G(\mathbf{Q})\backslash G(\mathbf{A})} E(f,s,g)\cdot\phi(g)\,dg$$

$$=\int_{G(\mathbf{Q})\backslash G(\mathbf{A})}\left(\sum_{\gamma\in P(\mathbf{Q})\backslash G(\mathbf{Q})}\tilde{f}_s(\gamma g)\right)\cdot\phi(g)\,dg$$

$$=\int_{P(\mathbf{Q})\backslash G(\mathbf{A})}\tilde{f}_s(g)\cdot\phi(g)\,dg$$

$$=\int_{N(\mathbf{A})M(\mathbf{Q})\backslash G(\mathbf{A})}\tilde{f}_s(g)\cdot\left(\int_{N(\mathbf{Q})\backslash N(\mathbf{A})}\phi(ng)\,dn\right)\,dg$$

$$=0$$

The result for general s follows by meromorphic continuation.

Analytic behaviour of constant term

Clearly, if $E(f, s, -)$ has a pole of order k at s_0, then the order of pole at s_0 of the constant term $E_N(f, s, -)$ is $\leq k$. In fact, we have

Theorem: The order of pole of $E(f, s, -)$ at any s_0 is the same as that of the constant term $E_N(f, s, -)$.

Proof: If not, the leading term $a_{-k}(f, s, -)$ of the Laurent expansion of $E(f, s, -)$ at s_0 will be a non-zero cusp form, contradicting the fact that $a_{-k}(f, s, -)$ is orthogonal to all cusp forms.

Remarks: This result is of great importance in practice. It says that to decide the behaviour of E at s, it suffices to examine the behaviour of E_N at s. But we have a formula for E_N in terms of the intertwining operator, which is an Euler product of local intertwining operators.

It turns out that one can calculate these local intertwining operators explicitly for almost all p. This is the so-called **Gindikin-Karpelevich formula**. The answer can be expressed (as shown by Langlands) in terms of certain local L-functions of π_p.

Thus, we see that the analytic properties of E are controlled by those of appropriate L-functions.

This relation can be exploited both ways. Sometimes, one uses properties of the L-functions to deduce, for example, that E has a pole somewhere. On the other hand, one may also use the analytic properties of E to deduce analytic properties of the L-functions. This is the technique used in two of the standard approaches of studying L-functions: the Rankin-Selberg method and the Langlands-Shahidi method.

Spectral decomposition

For the application to the spectral decomposition of $L^2(G(\mathbf{Q})\backslash G(\mathbf{A}))$, it is necesary to know the following result:

Theorem: The Eisenstein series $E(f,s,g)$ is holomorphic on $Re(s) = 0$ (i.e. the so-called unitary axis). Moreover, it has only finitely many poles in $Re(s) > 0$.

The first statement is needed because in the spectral decomposition, one needs to consider integral of the type:

$$\int_{Re(s)=0} \phi(s) \cdot E(f,s,g)\, ds.$$

Moreover, this integral is obtained by a contour shift from the same integral over $Re(s) = c$ for some large c. From the second statement, we see that we may pick up finitely many residues of $E(f,s,g)$ during the contour shift, and these contribute to the so-called residual spectrum.

The standard example

We illustrate the above in the case when $G = PGL_2$ and σ is the trivial character.

In this case, the Gindikin-Karpelevich formula gives:

$$M_w(s)f_v^0 = \frac{\zeta_v(2s)}{\zeta_v(2s+1)} \cdot f_v^0.$$

So we see that the constant term is:

$$E_N(f,s,g) = f_s(g) +$$

$$\prod_{v \in S} M_{w,v}(s) f_v(g_v) \cdot \frac{\zeta^S(2s)}{\zeta^S(2s+1)} f^{0,S}(g^S).$$

The local intertwining operators are defined by absolutely convergent integral when $Re(s) > 0$, and so has no poles there. Thus, one sees that in $Re(s) > 0$, there is precisely only one pole of order 1, namely at $s = 1/2$.

The residue there turns out to be a constant function.

A theorem of Langlands

The above discussion shows that the process of parabolic induction sends cuspidal automorphic representations on M to automorphic representations of G. In other words, if σ is cuspidal, then every irreducible constituent of $I_P(\sigma)$ is automorphic. Langlands showed the converse to this:

Proposition:

If π is an automorphic representation of G, then there exists a parabolic subgroup $P = MN$ and a cuspidal representation σ of M such that π is a constituent of $I_P(\sigma)$.

This theorem shows that cuspidal representations are the fundamental objects in the theory of automorphic forms, in the sense that every other automorphic representation is built out of them by parabolic induction.

[Compare this with the representation theory of p-adic groups. There, the basic objects are the supercuspidal representations and every irreducible smooth representation is a constituent of some $I_P(\sigma)$ with σ supercuspidal. Moreover, the pair (M, σ) is unique up to conjugacy.]

The theorem of Langlands above does not claim that the pair (M, σ) is unique up to conjugacy. For GL_n, this is in fact true: it is a non-trivial theorem of Jacquet-Shalika. In general, however, it is false!

For example, for the group $PGSp_4$, it was shown by Waldspurger that there are cuspidal representations π which are abstractly isomorphic to a constituent of some $I_P(\sigma)$ with σ cuspidal on $M = GL_2$. These are the so-called **CAP** representations.

Multiplicity One Theorem for GL_n

Last time, we saw a general construction of automorphic representations using Eisenstein series. This does not produce cuspidal representations. In fact, there are no known general methods which produce embeddings of a representation into $\mathcal{A}_0(G)$.

However, one can prove some results about the structure of $\mathcal{A}_0(G)$ in the case when $G = GL_n$. The goal of this lecture is to prove one such result:

Multiplicity One theorem

The multiplicity $m_0(\pi)$ of an irreducible representation π of $GL_n(\mathbf{A})$ in $\mathcal{A}_0(G)$ is ≤ 1.

Note that this theorem does not tell us which π has $m_0(\pi) = 1$.

Whittaker-Fourier coefficients

The proof of the multiplicity one theorem has two ingredients, one of which is global and the other local. We begin by explaining these 2 ingredients.

Let f be an automorphic form on $G = GL_n$. If $N \subset G$ is a unipotent subgroup, say the unipotent radical of a parabolic subgroup, one can consider the Fourier coefficients of f along N.

Namely, if χ is a unitary character of $N(\mathbf{A})$ which is trivial on $N(\mathbf{Q})$, we have

$$f_{N,\chi}(g) = \int_{N(\mathbf{Q})\backslash N(\mathbf{A})} \overline{\chi(n)} \cdot f(ng)\, dn$$

Note that if N is abelian, then we have:

$$f(g) = \sum_{\chi} f_{N,\chi}(g).$$

In general, when N is not abelian, the expansion on the RHS is only equal to $f_{[N,N]}$.

We apply the above to the unipotent radical N of the Borel subgroup B of upper triangular matrices

Definition: A character χ of $N(\mathbf{A})$ is **generic** if the stabilizer of χ in $T(\mathbf{A})$ is the center $Z(\mathbf{A})$ of $GL_n(\mathbf{A})$. An equivalent definition is that χ is non-trivial when restricted to every simple root subgroup in N.

Examples:

(i) When $G = GL_2$, a generic character of $N(\mathbf{Q})\backslash N(\mathbf{A})$ just means a non-trivial character of $\mathbf{Q}\backslash\mathbf{A}$. If we fix a character ψ of $F\backslash\mathbf{A}$, then all others are of the form

$$\chi_\lambda(x) = \psi(\lambda x)$$

for some $\lambda \in \mathbf{Q}$.

(ii) When $G = GL_3$, a character of $N(\mathbf{A})$ trivial on $N(\mathbf{Q})$ has the form

$$\chi_{\lambda_1,\lambda_2}\begin{pmatrix} 1 & a_1 & * \\ 0 & 1 & a_2 \\ 0 & 0 & 1 \end{pmatrix} = \psi(\lambda_1 a_1 + \lambda_2 a_2)$$

for some λ_1 and $\lambda_2 \in \mathbf{Q}$.

Saying that $\chi_{\lambda_1,\lambda_2}$ is generic means that λ_1 and λ_2 are non-zero.

Definition: A **Whittaker-Fourier coefficient** of f is a Fourier coefficient $f_{N,\chi}$ with χ generic. (N is unipotent radical of Borel).

We note:

Lemma: The group $Z(\mathbf{Q})\backslash T(\mathbf{Q})$ acts simply transitively on the generic characters of $N(\mathbf{A})$ trivial on $N(\mathbf{Q})$.

We let χ_0 be the generic character whose restriction to each simple root space is equal to ψ.

Observe that if $t \cdot \chi = \chi'$ with $t \in T(\mathbf{Q})$, then

$$f_{N,\chi'}(g) = f_{N,\chi}(t^{-1}g).$$

Thus, we see that $f_{N,\chi_0} \neq 0$ iff $f_{N,\chi} \neq 0$ for all generic χ.

Definition: A representation $\pi \subset \mathcal{A}(G)$ is said to be **globally generic** if there exists $f \in \pi$ whose Fourier-Whittaker coefficient $f_{N,\chi} \neq 0$ for some (and hence all) generic characters χ.

An equivalent reformulation is as follows. Consider the linear map

$$l_\chi : \mathcal{A}(G) \longrightarrow \mathbf{C}$$

defined by

$$l_\chi(f) = f_{N,\chi}(1)$$

with χ generic. Then π is globally generic iff $l_\chi \neq 0$ when restricted to π.

The example of GL_2

Suppose $G = GL_2$ and $\pi \subset \mathcal{A}_0(G)$ is an irreducible cuspidal representation.

Claim: π is globally generic.

Proof: Take any non-zero $f \in \pi$. Then we have the expansion

$$f(g) = \sum_{\chi} f_{N,\chi}(g)$$

Since f cuspidal, $f_N = 0$. So some $f_{N,\chi} \neq 0$.

In this case, the following are equivalent:

(i) π is globally generic;

(ii) $\pi = \otimes_v \pi_v$ is infinite-dimensional;

(iii) π_v is infinite-dimensional for all v.

Whittaker functionals

One can define the notion of a "generic representation" locally.

Let π_v be a representation of $G(\mathbf{Q}_v)$ and let

$$\chi_v : N(\mathbf{Q}_v) \longrightarrow \mathbf{C}$$

be a generic unitary character.

Definition: Let p be a finite prime. Then π_p is an **abstractly generic** representation if, given any generic χ_p, there is a non-zero linear functional $l_p : \pi_p \to \mathbf{C}$ such that

$$l_p(n \cdot v) = \chi_p(n) \cdot l_p(v)$$

for all $n \in N(\mathbf{Q}_p)$ and $v \in \pi_p$. Such a functional is called a local **Whittaker functional**.

Archimedean case

One can make the same definition at the infinite prime, except for one subtlety: π_∞ is a $(\mathfrak{g}, K)$-module and $N(\mathbf{R})$ does not act on π_∞.

Recall that given an admissible $(\mathfrak{g}, K)$-module V_K, there exists a continuous representation of $G(\mathbf{R})$ on a Hilbert space H whose underlying $(\mathfrak{g}, K)$-module is V_K. However, this Hilbert representation is not unique.

Regardless, for each such H, we can consider the space of smooth vectors H^∞, and equip this space with the smooth topology, making it into a Frechet space with continuous $G(\mathbf{R})$-action.

The following is an amazing result of Casselman-Wallach:

Theorem: The $G(\mathbf{R})$-representation H^∞ is independent of the choice of H.

We call this the canonical globalization of V_K.

Let π be an admissible $(\mathfrak{g}, K)$-module and let π^∞ be its canonical Frechet globalization.

Definition: Say that π is **abstractly generic** if π^∞ has a non-zero continuous Whittaker functional.

Note that such a functional is non-zero when restricted from π^∞ to π (by density of K-finite vectors).

Now let $\pi = \otimes_v \pi_v$ be an irreducible admissible representation of $G(\mathbf{A})$, one says that π is an **abstractly generic** representation if each of its local components π_v is abstractly generic.

The two ingredients

We can now state the two ingredients needed for the proof of the Multiplicity One Theorem.

Theorem A (Global genericity):

Let $\pi \subset \mathcal{A}(G)$ be an irreducible cuspidal representation. Then π is globally generic.

Theorem B (Local uniqueness of Whittaker functionals):

Let π_v be an irreducible smooth representation of $G(\mathbf{Q}_v)$. Then the space of (continuous) Whittaker fuctional on π_v is at most 1-dimensional.

We remark that while Theorem B is still true for an arbitrary (quasi-split) group, Theorem A is only valid for GL_n.

Proof of Multiplicity One

Assume Theorems A and B. We need to show that for any ireducible admissible representation π of $G(\mathbf{A})$,

$$\dim Hom_{G(\mathbf{A})}(\pi, \mathcal{A}_0(G)) \leq 1.$$

Let χ be a generic character of $N(\mathbf{A})$ trivial on $N(\mathbf{Q})$. Recall we have the map

$$l_\chi : \mathcal{A}(G) \longrightarrow \mathbf{C}_\chi$$

given by

$$l_\chi(\phi) = \int_{N(\mathbf{Q})\backslash N(\mathbf{A})} \overline{\chi(n)} \cdot \phi(n)\, dn.$$

Now we have a map

$$Hom_{G(\mathbf{A})}(\pi, \mathcal{A}_0(G)) \longrightarrow Hom_{N(\mathbf{A})}(\pi, \mathbf{C}_\chi)$$

given by $f \mapsto l_\chi \circ f$.

By Theorem A, this map is injective!

So it suffices to show that the RHS has dimension ≤ 1.

The generic character χ is of the form $\prod_v \chi_v$ for generic characters χ_v of $N(\mathbf{Q}_v)$.

Now if $L \in Hom_{N(\mathbf{A})}(\pi, \mathbf{C}_\chi)$ is non-zero, then for each v,

$$\dim Hom_{N(F_v)}(\pi_v, \mathbf{C}_{\chi_v}) \neq 0,$$

i.e. π is abstractly generic. By Theorem B, the above dimenson is 1, and for almost all v, a non-zero local functional l_v is non-zero on $\pi_v^{K_v}$.

Let us choose $l_v \neq 0$ so that for almost all v, $l_v(u_v^0) = 1$, where u_v^0 is the distinguished K_v-fixed vector in π_v. Then one has, for some constant c,

$$L(u) = c \cdot \prod_v l_v(u_v) \qquad \text{for any } u = \otimes_v u_v.$$

This shows that

$$\dim Hom_{N(\mathbf{A})}(\pi, \mathbf{C}_\chi) = 1$$

as desired.

Proof of Theorem A

Recall that we have shown Theorem A for GL_2, for we can express $f \in \pi$ as:

$$f(g) = \sum_{\chi} f_{N,\chi}(g) = \sum_{\gamma \in \mathbf{Q}^\times} f_{N,\chi_0}\left(\begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix} g\right).$$

Note that we have this expansion for any smooth function on $B(\mathbf{Q})\backslash GL_2(\mathbf{A})$.

Foir $n > 2$, because N is non-abelian, it appears that the Whittaker-Fourier coefficients $f_{N,\chi}$ only determines $f_{[N,N]}$. This is why Theorem A is not trivial.

However, we shall show:

Proposition: We have the expansion

$$f(g) = \sum_{\gamma \in N_{n-1}(\mathbf{Q}) \backslash GL_{n-1}(\mathbf{Q})} f_{N,\chi_0}\left(\begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix} g \right).$$

Here N_{n-1} is the unipotent radical of the Borel subgroup of GL_{n-1}.

Clearly, this proposition implies Theorem A. The proof of the proposition makes use of the so-called **mirabolic subgroup** P_n of GL_n (which is something specific to GL_n):

$$P_n = \{ \begin{pmatrix} g_{n-1} & * \\ 0 & 1 \end{pmatrix} : g_{n-1} \in GL_{n-1} \}.$$

It has a decomposition $P_n = GL_{n-1} \cdot U_n$.

One proves inductively the following statement.:

(*) Suppose that f is a smooth function on $P_n(\mathbf{Q})\backslash GL_n(\mathbf{A})$ whose constant terms along any standard unipotent subgroup $U \subset P_n$ vanishes. Then f has the expansion of the proposition.

For the purpose of this proof, we say that such an f is cuspidal. The proof proceeds as follows:

- Expand f along U_n:

$$f(g) = \sum_{\lambda} = f_{U_n,\lambda}(g).$$

  Then $f_{U_n} = 0$ since f cuspidal. Also $GL_{n-1}(\mathbf{Q})$ acts transitively on the non-trivial λ's. One such λ is $\lambda_0 = \chi_0|_{U_n}$. Its stabilizer in $GL_{n-1}(\mathbf{Q})$ is precisely $P_{n-1}(\mathbf{Q})$.

  So we have the preliminary expansion:

$$f(g) = \sum_{\gamma \in P_{n-1}(\mathbf{Q})\backslash GL_{n-1}(\mathbf{Q})} f_{U_n,\lambda_0}(\gamma g).$$

- Now let

$$\Phi_g = f_{U_n,\lambda_0}(-g)|_{GL_{n-1}}.$$

It is easy to see that Φ is smooth cuspidal as a function on $P_{n-1}(\mathbf{Q})\backslash GL_{n-1}(\mathbf{A})$. So we can apply induction hypothesis to Φ:

$$\Phi_g(h) = \sum_{\delta \in N_{n-2}\backslash GL_{n-2}} (\Phi_g)_{N_{n-1},\chi_0'}(\delta h)$$

where $\chi_0' = \chi_0|_{N_{n-1}}$.

- Finally, we substitute this expansion for Φ_g into the preliminary expansion for f. The double sum can be collapsed into a single sum over $N_{n-1}(\mathbf{Q})\backslash GL_{n-1}(\mathbf{Q})$. Also, one observes that the summand is given, for $\gamma \in P_{n-1}(\mathbf{Q})\backslash GL_{n-1}(\mathbf{Q})$, by:

$$(\Phi_g)_{N_{n-1},\chi_0'}(\gamma)$$

$$= \int_{N_{n-1}(\mathbf{Q})\backslash N_{n-1}(\mathbf{A})} \overline{\chi_0(n')} \cdot \Phi_g(n'\gamma)\, dn'$$

$$= \int_{N_{n-1}(\mathbf{Q})\backslash N_{n-1}(\mathbf{A})} \overline{\chi_0(n')} \cdot f_{U_n,\lambda_0}(n'\gamma g)\, dn'$$

$$= \int_{N_{n-1}(\mathbf{Q})\backslash N_{n-1}(\mathbf{A})} \int_{U_n(\mathbf{Q})\backslash U_n(\mathbf{A})} \overline{\chi_0(u)} \cdot \overline{\chi_0(n')} \cdot$$

$$f(un'\gamma g)\, du\, dn'$$

Because $N_n = U_n \cdot N_{n-1}$, this last double integral can be combined into the single integral

$$\int_{N_n(\mathbf{Q})\backslash N_n(\mathbf{A})} \overline{\chi_0(n)} \cdot f(n\gamma g)\, dn = f_{N_n,\chi_0}(\gamma g).$$

This is the desired result.

Strong Multiplicity One

In fact, for GL_n, a stronger result is true. Namely,

Theorem: (Rigidity)

Let π_1 and π_2 be irreducible cuspidal representations of GL_n. Assume that for almost all v, $\pi_{1,v}$ and $\pi_{2,v}$ are isomorphic as abstract representations. Then π_1 and π_2 are isomorphic as abstract representations.

The proof of this theorem, due to Jacquet-Shalika, proceeds by using L-functions.

Corollary: (Strong multiplicity one)

If $\pi_1, \pi_2 \subset \mathcal{A}_0(G)$ are such that $\pi_{1,v}$ and $\pi_{2,v}$ are isomorphic for almost all v, then $\pi_1 = \pi_2$ as subspaces of $\mathcal{A}_0(G)$.

The corollary follows from the theorem above and multiplicity one theorem.

Other groups

Finally, let us comment on $\mathcal{A}_0(G)$ for other groups G. It turns out that in general, the multiplicity one theorem is false, though we have:

Theorem (Ramakrishnan, 2000) The multiplicity one theorem is true for $G = SL_2$

The proof of this theorem is highly non-trivial; this shows once again that GL_2 is much easier to handle than SL_2. For example, strong multiplicity one is not true for SL_2. What about SL_n for $n \geq 3$?

Theorem (Blasius, 1994) For $G = SL_n$, $n \geq 3$, the multiplicity one theorem is false.

Does one have an explanation for this?

Generalized Ramanujan conjecture

Recall that we have formulated the Ramanujan conjecture for GL_n as: if π is cuspidal, then π is tempered.

Now one might make the same conjecture for general G. However, it turns out to be false! The first examples of such non-tempered cuspidal representations were discovered in the 70's by Saito-Kurokawa and Howe-Piatetski-Shapiro on the group Sp_4. However, these counterexamples to the naive generalization of Ramanujan's conjecture are non-generic. Thus, one is led to the following conjecture:

Generalized Ramanujan Conjecture: Let π be cuspidal and generic. Then π is tempered.

In fact, there is a deep conjecture of Arthur which extends the generalized Ramanujan conjecture. It explains the extent of the failure of

Ramanujan conjecture for general representations in the discrete spectrum of $L^2(G(\mathbf{Q})\backslash G(\mathbf{A}))$ as well as the presence of multiplicities in the discrete spectrum.