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).

- $\phi(g r_\theta) = e^{ik\theta} \cdot \phi(g)$ where
$$r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$
is a typical element in K. This is because:
$$(f|g r_\theta)(i) = (-i\sin\theta + \cos\theta)^{-k} \cdot (f|g)(i).$$
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

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

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

$\underline{Z(\mathfrak{g})\text{-finiteness}}$

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

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

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

Cusp forms

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,

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

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

$$L_0^2(\Gamma\backslash G) \cong \widehat{\bigoplus_{\pi}} m_\pi \pi$$

(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 $L_0^2(\Gamma\backslash G)$.

Thus the decomposition of L_0^2 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.