Classical automorphic forms and representations of SL(2)

Bill Casselman
University of British Columbia
cass@math.ubc.ca

This essay will explain the relationship between classical automorphic forms and representations of $GL_2(\mathbb{R})$. The classical theory of automorphic forms, in spite of initial appearances, is about the group $GL_2$, not $SL_2$. The classical theory is concerned with functions on the upper half plane, which is acted on by fractional linear transformations in $GL_{2, \mathbb{R}}$, and it happens that the intersection of $GL_{2, \mathbb{R}}$ with $GL_2(\mathbb{Z})$ is $SL_2(\mathbb{Z})$. This accident is responsible for some mild confusion.


Contents

1. Automorphic forms
2. Representations on holomorphic functions
3. Identification with discrete series
4. References

1. Automorphic forms

A congruence subgroup $\Gamma$ of $GL_{2, \mathbb{R}}$ is a subgroup of

$$SL_2(\mathbb{Z}) = GL_2(\mathbb{Z}) \cap GL_{2, \mathbb{R}}$$

containing some principal congruence subgroup

$$\Gamma(N) = \{ \gamma \in SL_2(\mathbb{Z}) \mid \gamma \equiv I \ (\text{mod} \ N) \}$$

The region

$$D = \{ z = x + iy \mid |x| \leq 1/2, |z| \geq 1 \}$$

is a fundamental domain of $SL_2(\mathbb{Z})$ in $\mathcal{H}$, and some finite number of transforms of this region will be a fundamental domain for $\Gamma$. A function on $D$ is said to be of moderate growth if it is $O(y^M)$ for some $M > 0$, and rapidly decreasing if $O(y^{-M})$ as $y \to \infty$. Similarly for a fundamental domain of $\Gamma$.

An automorphic form of weight $m$ for $\Gamma$ is a holomorphic function $f(z)$ on $\mathcal{H}$ satisfying a condition of moderate growth on every fundamental domain such that:

$$f \mid [\gamma]_m = f \quad \text{or} \quad f(\gamma(z)) = j(g, z)^m f(z)$$

for all $\gamma$ in $\Gamma$. Following the previous section, define the corresponding function on $G = GL_{2, \mathbb{R}}$

$$\Phi_f(g) = f(g(i)) \det^{m/2}(g) j(g, i)^{-m}.$$

It is also invariant under $\Gamma$ and again in some sense of moderate growth. From the results of the previous section:

1.1. Theorem. The space of holomorphic forms of weight $m$ for $\Gamma$ is isomorphic to the space of smooth (in fact, necessarily real analytic) functions $\Phi$ on $\Gamma \backslash G$ of moderate growth such that
(a) $\Phi(gk) = \varepsilon^{-m}(k)\Phi(g)$ for $k$ in $SO_2$;
(b) $R_{x, \Phi} = 0$;
(c) $\Phi(gz) = \Phi(g)$ for $z$ in the connected component of the center of $G$.

The existence of an automorphic form of weight $m$ says something about the occurrence of a discrete series representation of $G$ in $C^\infty(\Gamma\backslash G)$, but not quite what might be naively expected. Conditions (a)–(c) mean that the representation of $G$ generated by $\Phi$ is a copy of the dual discrete series $\hat{\pi}_m$. Formally, this is related to Frobenius reciprocity, since a $\Gamma$-invariant vector in $\pi_m$ lies in

$$\text{Hom}_\Gamma(\pi_m, \mathbb{C})$$

which according to a kind of Frobenius reciprocity is formally the same as

$$\text{Hom}_{(g,K)}(\pi_m, C^\infty(\Gamma\backslash G))$$.

This formal analysis is not legal, because the $\Gamma$-invariant function is not in $\pi_m$ but in some larger space. However, this heuristic idea can be justified with some modest amount of trouble.

2. Representations on holomorphic functions

The space $\mathbb{P}^1(\mathbb{C})$ is by definition the quotient of $\mathbb{C}^2 - \{0\}$ by scalars. This may be identified with $\mathbb{C} \cup \{\infty\}$ via the map $(z, w) \mapsto z/w$. The copy of $\mathbb{C}$ in $\mathbb{P}^1(\mathbb{C})$ is the image of the pairs $(z, w)$ with $w \neq 0$, and there is a splitting over it taking $z$ to $(z, 1)$.

The group $GL_2(\mathbb{C})$ acts on $\mathbb{C}^2$ in the standard way:

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \omega = \begin{bmatrix} z \\ w \end{bmatrix} \mapsto g\omega = \begin{bmatrix} az + bw \\ cz + dw \end{bmatrix}.$$ 

and the corresponding action on $\mathbb{P}^1(\mathbb{C})$ is by fractional linear transformations:

$$g \begin{bmatrix} z \\ 1 \end{bmatrix} = (cz + d) \begin{bmatrix} (az + b)/(cz + d) \\ 1 \end{bmatrix} = j(g, z) \begin{bmatrix} g(z) \\ 1 \end{bmatrix}$$ if $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Let $Z$ be the space of complex pairs $(z, w)$ with $w \neq 0$ and $\text{IM}(z/w) > 0$ (that is to say, $z/w$ is real, with positive imaginary part). The map $(z, w) \mapsto z/w$ specifies it as a fibre space over the upper half plane $\mathcal{H}$ with fibre equal to $\mathbb{C}^\times$. As I have already mentioned, this fibring is split by the inverse map $z \mapsto (z, 1)$, and $\mathcal{H}$ may be identified with its image in $Z$.

If $g$ is real, we have

$$\frac{az + bw}{cz + dw} = \frac{a(z/w) + b}{c(z/w) + d}$$

$$= \frac{a\zeta + b}{c\zeta + d} \quad (\zeta = z/w)$$

$$\text{IM} \left( \frac{a\zeta + b}{c\zeta + d} \right) = \frac{1}{2i} \left( \frac{a\zeta + b}{c\zeta + d} - \frac{a\overline{\zeta} + b}{c\overline{\zeta} + d} \right)$$

$$= \frac{ad - bc}{|c\zeta + d|^2} \text{IM}(\zeta).$$

Therefore the group $G = GL_2^\text{pos}(\mathbb{R})$ takes $Z$ into itself. Since it commutes with scalar multiplication it preserves the fibring over $\mathcal{H}$. The traditional action of $G$ on $\mathcal{H}$ is derived from this, and the automorphy
factor $j(g,z)$ measures the extent to which the splitting fails to be $G$-covariant. That $j(g,z)$ measures an obstruction explains why the equation

$$j(gh,z) = j(g,h(z))j(h,z)$$

is reminiscent of a cocycle relation.

The group $G$ acts also on the space of smooth functions on $Z$:

$$F \mapsto L_g F, \quad [L_g F](\omega) = F(g^{-1}\omega)$$

is the associated right action. The space $Z$ is an open subset of $\mathbb{C} \times \mathbb{C}$, hence possesses an inherited complex structure. Since $G$ preserves this complex structure, it takes the subspace of holomorphic functions into itself.

Since $G$ commutes with scalar multiplication it leaves stable the subspace $C^\infty_m(Z)$ of smooth functions of weight $m$—those $F$ such that

$$F(\lambda z, \lambda w) = \lambda^{-m} F(z, w)$$

as well as its subspace $C^\text{hol}_m(Z)$ of holomorphic functions. In effect, the functions in $C^\infty_m(Z)$ are smooth sections of a holomorphic bundle over $H$.

The group $G$ acts transitively on $Z$, and the isotropy subgroup of $(i, 1)$ is just $I$, so $Z$ is in fact a principal homogeneous space over $G$. That is to say, it is $G$-isomorphic to $G$ given a choice of base point. The stabilizer of any fibre of the projection onto $H$ is a copy of $\mathbb{C}^\times$ in $G$. In particular, the fibre over $i$ in $H$ may be identified with the copy of $\mathbb{C}^\times$ given by the embedding of $\mathbb{C}$ into $M_2(\mathbb{R})$

$$\iota: a + ib \mapsto \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

since if $\lambda = a + bi$ then

$$i(\lambda) \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} ai - b \\ bi + a \end{bmatrix}$$

$$= (a + bi) \begin{bmatrix} ai - b/(a + bi) \\ 1 \end{bmatrix}$$

$$= (a + bi) \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$= \lambda \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Choosing a base point in $Z$ allows us to identify functions on $Z$ with functions on $G$. For each smooth $F$ on $Z$, therefore, let $\Phi_F$ be the corresponding function on $G$:

$$\Phi_F(g) = F\left( g \begin{bmatrix} i \\ 1 \end{bmatrix} \right).$$

This follows immediately from the equation above:

2.1. Lemma. The map $F \mapsto \Phi_F$ identifies $C^\infty_m(Z)$ with the space of smooth $\Phi$ satisfying the condition

$$\Phi(g \cdot i(\lambda)) = \lambda^{-m} \Phi(g)$$

for $\lambda$ in $\mathbb{C}^\times$.

The group $G$ acts on this by means of the left regular representation, compatibly with its action on $C^\infty_m(Z)$. 

Next, I’ll show how to characterize the image of $C^\text{hol}(\mathbb{Z})$, with respect to $\Phi$. Before I state the main result, I’ll recall some elementary facts of complex analysis. A smooth $\mathbb{C}$-valued function $f = u(x, y) + iv(x, y)$ on an open subset of $\mathbb{C}$ is holomorphic if and only if the real Jacobian matrix of $f$

$$
\begin{pmatrix}
\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\
\frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}
\end{pmatrix}
$$

considered as a map from $\mathbb{R}^2$ to itself lies in the image of $\mathbb{C}$ in $M_2(\mathbb{R})$. (Since this image generically coincides with the group of orientation-preserving similitudes, this means precisely that it is conformal.) This condition is equivalent to the Cauchy-Riemann equations

$$
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.
$$

Holomorphicity may also be expressed by the single equation

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

where

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

When $f$ is holomorphic, its complex derivative is

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

The notation is designed so that for an arbitrary smooth function

$$
df = \frac{\partial f}{\partial z} dz + \frac{\partial f}{\partial \overline{z}} d\overline{z}
$$

where $dz = dx + idy$.

The Lie algebra of $G$ has as basis:

$$
\zeta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
\alpha = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},
$$

$$
= (1/2)(\alpha + \zeta)
$$

$$
\kappa = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},
$$

$$
\nu_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
$$

In addition, we shall find useful these:

$$
\eta = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},
$$

$$
x_+ = \begin{bmatrix} 1 & -i \\ -i & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}
$$

$$
= \alpha - i(2\nu_+ + \kappa)
$$

$$
= (\alpha + \zeta) - \zeta - i(2\nu_+ + \kappa)
$$

$$
= (2\eta - \zeta) - i(2\nu_+ + \kappa).
$$
The significance of \( x_+ \) and its conjugate \( x_- \) is that
\[
[\kappa, x_\pm] = \pm 2ix_\pm.
\]

Now for \( F \) in \( C^\infty(\mathcal{Z}) \), define a function \( f \) on the upper half-plane \( \mathcal{H} \) to be its restriction to its image in \( \mathcal{Z} \):
\[
f(z) = F(z, 1).
\]

2.2. **Lemma.** For \( F \) in \( C_m^\infty(\mathcal{Z}) \) we have
\[
R_{x_+}\Phi_F(p) = -4iy \frac{\partial f(z)}{\partial z}
\]
if
\[
p = \begin{bmatrix} y & x \\ 0 & 1 \end{bmatrix}, \quad p(i) = z = x + iy.
\]

**Proof.** Since \( \kappa \) and \( \zeta \) are in the Lie algebra of \( \iota(\mathbb{C}^\times) \):
\[
R_\kappa F = -miF, \quad R_\zeta F = -mF.
\]

But then
\[
R_{x_+} F(p) = (R_\alpha - 2iR_{\nu_+} - iR_\kappa) F(p)
\]
\[
= (2R_\eta - R_\zeta - 2iR_{\nu_+} - iR_\kappa) F(p)
\]
\[
= (2R_\eta - 2iR_{\nu_+} + m - i(-mi)) F(p)
\]
\[
= (2R_\eta - 2iR_{\nu_+}) F(p)
\]
Now I apply the basic formula \( R_X f(g) = [L_X g^{-1}, f](g) \) to get
\[
(2R_\eta - 2iR_{\nu_+}) F(p) = (2L_{\eta p\nu^{-1}} - 2iL_{\nu_+ p^{-1}}) F(p).
\]

But
\[
L_\eta = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}
\]
and
\[
p\eta^{-1} = y\eta - x\nu_+
\]
so
\[
(2L_{\eta p\nu^{-1}} - 2iL_{\nu_+ p^{-1}}) F(p) = 2y \frac{\partial f}{\partial y} - 2iy \frac{\partial f}{\partial x}
\]
\[
= -2iy \left( \frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right)
\]
\[
= -4iy \frac{\partial f}{\partial z}.
\]

This has as consequence:

2.3. **Proposition.** The image of \( C_m^{\text{hol}}(\mathcal{Z}) \) under the map \( F \mapsto \Phi_F \) is the space \( C_m^{\text{hol}}(G) \) of all smooth functions \( \Phi \) in \( C^\infty(G) \) such that
(a) \( \Phi(g \cdot \iota(\lambda)) = \lambda^{-m} \Phi(g) \) for all \( \lambda \) in \( \mathbb{C}^\times \);  
(b) \( R_{x^+} \Phi = 0 \).

A function \( F \) in \( C^\infty_m(\mathbb{Z}) \) is determined by its restriction \( f \) to \( \mathcal{H} \):

\[
F(z, w) = w^{-m} F(z/w, 1) = w^{-m} f(z/w).
\]

How does the action of \( G \) on \( F \) translate to an action of \( G \) on \( f \)? If \( F \) in \( C^{\text{hol}}_m(\mathbb{Z}) \) restricts to \( f \) then

\[
\left[ L_g F \right](z, 1) = F(az + b, cz + d) \quad \text{where} \quad g^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}
\]

\[
= (cz + d)^{-m} F((az + b)/(cz + d), 1)
= (cz + d)^{-m} f \left( \frac{az + b}{cz + d} \right) = j(g^{-1}, z)^{-m} f(g^{-1}(z)).
\]

For \( z \) in \( \mathcal{H} \) and \( g \) in \( G \), let

\[
\left[ f \right|(g)_m(z) = j(g, z)^{-m} f(g(z)).
\]

This defines a right action of \( G \), whereas \( f \mapsto f \mid (g^{-1})_m \) defines a left action.

**2.4. Proposition.** The following diagram is commutative:

\[
C^{\text{hol}}_m(\mathbb{Z}) \rightarrow C^{\text{hol}}_m(G)
\]

\[
\downarrow
\]

\[
C^{\text{hol}}_m(\mathcal{H})
\]

The arrows are linear isomorphisms of \( G \)-spaces, if the action of \( G \) on the lower space is

\[
f \mapsto \left[ f \right|(g)_m.
\]

The map from \( C^{\text{hol}}(\mathcal{H}) \) to \( C^{\text{hol}}_m(G) \) takes \( f \) to

\[
f(g(i)) j(g, i)^{-m}.
\]

It will be convenient in the rest of this lecture to modify things slightly. Let

\[
\overline{\Phi}_F = \Phi_F(g) \cdot \det^{-m/2}(g).
\]

The functions \( \overline{\Phi}_F \) are now characterized by the conditions (a) \( \overline{\Phi}(g \cdot \iota(\lambda)) = (\lambda/|\lambda|)^{-m} \Phi(g) \) and (b) \( R_{x^+} \Phi = 0 \). In particular, \( \overline{\Phi} \) is fixed by the connected component of the center of \( G \). To go with this modification, redefine

\[
\left[ f \right|(g)_m(z) = \det^{m/2}(g)(cz + d)^{-m} f(g(z)).
\]

Here, the connected component of the scalars also acts trivially. The left regular action of \( G \) on the functions \( \overline{\Phi} \) and this new one on functions on \( \mathcal{H} \) are compatible. This normalization is just one of several in the literature. Each has its own charms, but for us invariance under the positive real scalars turns out to be most convenient. Keep in mind that the embedding of \( \text{SL}_2(\mathbb{R}) \) into \( \text{GL}_2(\mathbb{R}) \) induces an isomorphism of it with the quotient \( \text{GL}_2^{\text{pos}}(\mathbb{R})/\mathbb{R}^{\text{pos}}. \)
The explicit relation between $f$ and $\Phi$ is

$$\Phi_F(g) = f(g(i)) \det^{m/2}(g) j(g,i)^{-m}.$$ 

Instead of working with $GL_2^{\text{pos}}(\mathbb{R})$ we can work with $SL_2(\mathbb{R})$.

2.5. Proposition. The map from $C^\text{hol}(\mathcal{H})$ to $C^\infty(SL_2(\mathbb{R}))$ taking

$$f \mapsto f(g(i)) j(g,i)^{-m}$$

is an $SL_2(\mathbb{R})$ isomorphism of $C^\text{hol}(\mathcal{H})$ with the space of smooth functions $\Phi$ on $SL_2(\mathbb{R})$ such that

$$\Phi(g \cdot i(\lambda)) = \lambda^{-m} \Phi(g) \text{ for } \lambda \text{ such that } |\lambda| = 1$$

$$R_x \Phi = 0.$$ 

The image of $\{\lambda | |\lambda| = 1\}$ is the maximal compact subgroup

$$SO(2) = \left\{ \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \right\}$$

in $SL_2(\mathbb{R})$.

3. Identification with discrete series

Let $\pi_m$ be the left representation of $G = GL_2^{\text{pos}}(\mathbb{R})$ on the space of holomorphic functions on $\mathcal{H}$ taking $f$ to $f \mid [g^{-1}]_m$. This is equivalent to the left representation on $C^\text{hol}(\mathbb{Z})$. This is trivial on the connected component of the scalar matrices, hence may be identified with a representation of $SL_2(\mathbb{R})$. In this section we’ll see that for $m > 1$ it is an incarnation of the discrete series $D^+_m$ and for $m = 1$ one of the so-called 'limit' of discrete series.

Let $K$ be the maximal compact subgroup $SO_2$ of $G$. Recall the character

$$\varepsilon: \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \mapsto c + is.$$ 

The restriction of $\pi_m$ to $K$ will decompose into a sum of one-dimensional eigenspaces with respect to $K$.

3.1. Proposition. The eigenfunctions of $K$ in $\pi_m$ are the functions

$$f_p(z) = \left( \frac{z - i}{z + i} \right)^p \left( \frac{1}{z + i} \right)^m$$

on $\mathcal{H}$ with $p \geq 0$. For $k$ in $K$

$$\pi_m(k)f_p = \varepsilon^{m+2p}(k)f_p.$$ 

Proof. The action of $G$ on $\mathcal{Z}$ is inherited from a linear one. It maps linear functions to linear functions, polynomials to polynomials. Linear functions correspond to row vectors, since

$$pz + qw = \begin{bmatrix} p & q \end{bmatrix} \begin{bmatrix} z \\ w \end{bmatrix}.$$
So $L_g$ takes the linear function $pz + qw$ to $L_g([p \ q])$, with
\[
[L_g([p \ q]) \left( \begin{bmatrix} z \\ w \end{bmatrix} \right)] = [p \ q]g^{-1} \left( \begin{bmatrix} z \\ w \end{bmatrix} \right)
\]
so
\[
[p \ q] \mapsto [p \ q]g^{-1}.
\]
The linear functions corresponding to $[1, \pm i]$ are eigenfunctions of $K$ in this representation, since
\[
\begin{bmatrix} 1, \pm i \end{bmatrix} \begin{bmatrix} c & -s \\ s & c \end{bmatrix} = (c \pm is)[1, \pm i]
\]
\[
L_k(z \pm iw) = \epsilon^{\mp 2}(z \pm iw)
\]
\[
L_k(z - iw)^p(z + iw)^q = \epsilon^{p-q}(k)(z - iw)^p(z + iw)^q.
\]
Since $1/(z - iw)$ has a singularity on $\mathcal{Z}$, the function $(z - iw)^p(z + iw)^q$ will be holomorphic on $\mathcal{Z}$ only if $p \geq 0$, and it then will lie in $C^2_{\text{hol}}(\mathcal{Z})$ if and only if $-m = p + q$. Restricting to the image of $H$ in $\mathcal{Z}$ concludes.

3.2. Corollary. The restriction of $\pi_m$ to $K$ is the direct sum of characters $\epsilon^{m+2p}$ for $p = 0, 1, \ldots$

When $m = 0$ the representation $\pi_m$ contains the trivial representation on constants. More generally, when $m \leq 0$ the representation $\pi_m$ contains the unique irreducible representation of dimension $|m| + 1$.

From now on in this essay, I shall assume that $m > 0$.

We know already that the measure $dx \ dy/y^2$ is $G$-invariant on $H$. Here is a useful generalization of this:

3.3. Proposition. The measure $y^{m-2} \ dx \ dy$ is $G$-invariant for $\pi_m$.

Proof. What this means is that for $f$ in $C^\infty_c(H)$ and $f_\ast = \pi_m(g)f$ we have
\[
\int_{H} |f_\ast(z)|^2 y^m \ dx \ dy = \int_{H} |f(z)|^2 y^m \ dx \ dy.
\]

We may as well assume that $g$ is in $\text{SL}_2(\mathbb{R})$. Recall that
\[
y(g(z)) = \frac{y(z)}{|cz + d|^2} = y(z)j(g(z))^{-2}.
\]
The first integrand is therefore
\[
|f(g(z))|^2 |j(g(z))|^{2m} y^m(z) = \int |f(g(z))|^2 |j(g(z))|^{2m} y^m(g(z)) \ dx \ dy.
\]
so the result follows from the invariance of $dx \ dy/y^2$.

The function $(z - i)^p/(z + i)^p$ is bounded on $H$ and $1/(z + i)^m$ is square-integrable with respect to the measure $y^{m-2} \ dx \ dy$ for $m > 1$. It follows that the representation $\pi_m$ is square-integrable for $m > 1$—i.e. is in fact a discrete summand of $L^2(\text{SL}_2(\mathbb{R}))$.

4. References