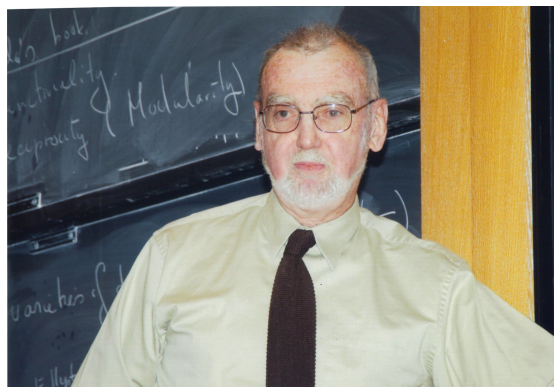# MATH808F - Modular Forms

Taught by `Jingren Chi`
Notes taken by `Haoran Li`
2020 Fall

Department of Mathematics
University of Maryland

Contents

1 Overview

Definition 1.1. G is a Lie group, $K \leq G$ is a closed subgroup, $X = G/K$ is then a homogeneous space with transitive left G-action, $\Gamma \leq G$ is a discrete subgroup. The so called *automorphic functions* are $\mathbb{C}$-valued functions f on X such that

$$f(\gamma \cdot x) = f(x), \quad \forall x \in X, \gamma \in \Gamma \tag{1.1}$$

Loosely speaking, *automorphic forms* (for Γ) on X are automorphic functions that are also eigenfunctions for invariant differential operators on X (+ some technical growth conditions when necessary)

Question 1.2. How to decompose automorphic functions into sums (or integrals) of automorphic forms

Example 1.3. $\Gamma = \mathbb{Z}$, $X = G = \mathbb{R}$, automorphic functions are functions on $\mathbb{R}/\mathbb{Z} = \mathbb{T}$, automorphic forms are $e^{2\pi i n x}$, $n \in \mathbb{Z}$. Fourier analysis tells us $L^2(\mathbb{R}/\mathbb{Z}) = \widehat{\bigoplus}_{n \in \mathbb{Z}} \mathbb{C} e^{2\pi i n x}$

Example 1.4. $G = \mathrm{SL}_2(\mathbb{R})$, $K = \mathrm{SO}(2)$, $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ is a finite index subgroup, $G/K = \mathcal{H} = \{\mathrm{Im}\, z > 0\}$ is the Poincaré upper half plane. G-invariant differential operators on $\mathcal{H}$ are polynomials with constant coefficients of the hyperbolic Laplacian $\Delta = -y^2 \left(\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} \right)$, Examples of automorphic forms in this setting: Maass forms. Γ are sometimes called "modular groups", the corresponding automorphic forms on $\mathcal{H}$ are called *modular forms*

Note. $\mathcal{H}$ has the structure of a complex manifold, it is natural to look for holomorphic automorphic forms

Example 1.5.
$$j(z) = q^{-1} + 744 + 196884q + 21493760q^2 + \cdots$$

Where $q = e^{2\pi i z}$, $z \in \mathcal{H}$, is invariant under $\mathrm{SL}_2(\mathbb{Z})$, hence a modular form

Definition 1.6. G induces a right action on $\mathbb{C}(X)$ by $(f \cdot g)(x) = f(gx)$, (1.1) becomes $f \cdot \gamma = f$, $\forall \gamma \in \Gamma$. More generally, we can allow a nontrivial *automorphy factor* $(f \cdot_c g) = c_g(x)f(gx)$, $\forall g \in G$, here $c_g : X \to \mathbb{C}^\times$

Exercise 1.7. For the action to be well-defined, the family of functions c_g must satisfy $c_{g_1 g_2}(x) = c_{g_2}(x)c_{g_1}(g_2 x)$, so called cocycle condition, $\forall g_1, g_2 \in G$, $x \in X$

Exercise 1.8. For $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, denote $j(g, z) = cz + d$, $G = \mathrm{SL}(2, \mathbb{R})$ acting on $\mathcal{H}$ by $g \cdot z = \dfrac{az + b}{cz + d}$. For $k \in \mathbb{Z}$, we consider the automorphy factor $c_g(z) = (cz + d)^{-k}$. Show c_g satisfies the cocycle condition

Definition 1.9. Then we get an action $(f \cdot_k g)(z) = (cz + d)^{-k} f\left(\dfrac{az + b}{cz + d} \right)$, $z \in \mathcal{H}$. For a modular group $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$, holomorphic function f on $\mathcal{H}$ is called a *holomorphic modular form* of *weight* k and *level* Γ (one may also need to add some boundness condition) if $f \cdot_k \gamma = f$, $\forall \gamma \in \Gamma$ which is equivalent to $f\left(\dfrac{az + b}{cz + d} \right) = (cz + d)^k f(z)$

Remark 1.10. To unify these examples for $G = \mathrm{SL}_2(\mathbb{R})$ to acts on $\mathcal{H}$ and "get rid of" the automorphy factors, it is better to consider $\Gamma \backslash G$. The advantage is $\Gamma \backslash G$ has a large symmetry group coming from right multiplication of G, (whereas $\Gamma \backslash \mathcal{H}$ does not have so many automorphisms). The invariant differential operators on $\Gamma \backslash \mathcal{H}$ come from $Z(\mathfrak{g})$, then center of the universal enveloping algebra of $\mathrm{Lie}(G)$. The automorphic forms in the above examples all correspond to certain C^∞ functions on $\Gamma \backslash G$, their automorphy factors are determined by their behavior under right $K = \mathrm{SO}(2, \mathbb{R})$ action

Example 1.11. Classical Maass forms on $\Gamma\backslash\mathcal{H}$ correspond to certain right K-invariant functions on $\Gamma\backslash G$. The basic problem of decomposing automorphic functions motivates the more refined problem of decomposiing the right regular representation of G on $L^2(\Gamma\backslash G)$

Theorem 1.12. Assume $\Gamma\backslash G$ is compact (equivalently, $\Gamma\backslash\mathcal{H}$ is compact. Modular groups which unfortunately do not satisfy this assumption, is one of the difficulty of the subject), then

$$L^2(\Gamma\backslash G) = \bigoplus_\pi \pi \otimes \mathrm{Hom}_G(\pi, L^2(\Gamma\backslash G))$$
$$= \bigoplus_\pi \pi^{\oplus m_\pi}$$

π run over irreducible representations of G, $m_\pi = \dim\mathrm{Hom}_G(\pi, L^2(\Gamma\backslash G)) < \infty$. Each multiplicity space $\mathrm{Hom}_G(\pi, L^2(\Gamma\backslash G))$ can be identified with a space of certain automorphic forms (The automorphy factors, eigenvalues of Laplacian are determined by the G-representations π)

Remark 1.13. In general we only assume that $\Gamma\backslash\mathcal{H}$ has finite volume, then we still have a decomposition of a subspae of $L^2(\Gamma\backslash G)$ (the discrete spectrum) whose orthogonal complement (the continuous spectrum) can be analyzed using theory of Eisenstein series. This is not the end of the story! Now comes the (arguably) more interesting part: when $\Gamma \leq G$ is arithmetic (e.g. modular groups, groups coming from indefinite quaternion algebras over $\mathbb{Q}$), then we can decompse each multiplicity space $\mathrm{Hom}_G(\pi, L^2(\Gamma\backslash G))$ further under the action of a big algebra on $L^2(\Gamma\backslash G)$ commuting with the right regular G-representation, this is the so-called "Hecke algebra". Where does this extra symmetry come from? Let $N_G(\Gamma) = \{g \in G | g\Gamma g^{-1} = \Gamma\}$ be the normarlizer, then $N_G(\Gamma)$ acts on $\Gamma\backslash G$ by left multiplication (so obviously commute with right G-action). This action factors through the quotient group $\Gamma\backslash N_G(\Gamma)$ and also induces automorphisms of $\Gamma\backslash\mathcal{H}$. Thus we ge4t an action of $\Gamma\backslash N_G(\Gamma)$ on $L^2(\Gamma\backslash G)$ that commutes with right G-regular representations. So $\Gamma\backslash N_G(\Gamma)$ acts on the multiplicity spaces $\mathrm{Hom}_G(\pi, L^2(\Gamma\backslash G))$ and decompose it further. The group $\Gamma\backslash N_G(\Gamma)$ is small (finite if $\Gamma\backslash G$ is compact, not sure if only finite volume), so the resulting decomposition is not so interesting. However, the action of $\Gamma\backslash N_G(\Gamma)$ on $\Gamma\backslash\mathcal{H}$ (and $\Gamma\backslash G$) can be extended to certain correspondences on $\Gamma\backslash\mathcal{H}$ (and $\Gamma\backslash G$)

Definition 1.14. Two discrete subgroups Γ_1, Γ_2 of G are *commensurable*, denoted $\Gamma_1 \approx \Gamma_2$, if their intersection $\Gamma_1 \cap \Gamma_2$ has fintie index in both of them. For $\Gamma \leq G$, let $\tilde\Gamma = \{g \in G | g\Gamma g^{-1} = \Gamma\}$ be the *commensurator* of Γ (this generalizes normalizer), elements in $\tilde\Gamma$ define correspondences on $\Gamma\backslash\mathcal{H}$ (and $\Gamma\backslash G$), which induces action of the convolution algebra $\mathbb{C}[\Gamma\backslash\tilde\Gamma/\Gamma]$ on $L^2(\Gamma\backslash G)$, and also on the cohomology of $\Gamma\backslash\mathcal{H}$. For modular groups Γ, we have $\tilde\Gamma = \mathrm{SL}_2(\mathbb{Q})$ which is large. For non-arithmetic groups Γ, $\tilde\Gamma/\Gamma$ is finite (This dichotomy between arithmetic and non-arithmetic cofinite volume subgroups follows from a general result of Margulis)

Remark 1.15. We will be mainly interested in congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$, i.e. subgroups that contain $\Gamma(N) = \ker(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}))$. In particular, such groups are modular, hence arithmetic. For each congruence subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$, we have G and $H_\Gamma = \mathbb{C}[\Gamma\backslash\tilde\Gamma/\Gamma]$ left and right acting on $L^2(\Gamma\backslash G)$, $\tilde\Gamma = \mathrm{SL}_2(\mathbb{Q})$. Put all these together (for the various congruence subgroups), $G = \mathrm{SL}_2(\mathbb{R})$ and $\varprojlim_\Gamma H_\Gamma = C_c^\infty(\mathrm{SL}_2(\mathbb{A}_f))$ left and right act on $\varinjlim_\Gamma L^2(\Gamma\backslash G) = L^2(\mathrm{SL}_2(\mathbb{Q})\backslash\mathrm{SL}_2(\mathbb{A}))$ ($\varinjlim_\Gamma \Gamma\backslash G = \mathrm{SL}_2(\mathbb{Q})\backslash\mathrm{SL}_2(\mathbb{A})$), decompose $L^2(\mathrm{SL}_2(\mathbb{Q})\backslash\mathrm{SL}_2(\mathbb{A}))$ into $\mathrm{SL}_2(\mathbb{A})$ representations, the irreducible summands are L^2-automorphic representations (Actually, we'll work with GL_2 instead, which is technically simpler). For (nice) irreducible representaion π of $\mathrm{GL}_2(\mathbb{A})$, Jacquet-Langlands associate an Euler prduct $L(s, \pi) = \prod_p L_p(s, \pi)$, (at least formally, may have convergence issues). This is done using tensor product theorem, which says roughly $\pi = \bigotimes_p \pi_p$ (restricted tensor product), π_p is the irreducible representation of $\mathrm{GL}_2(\mathbb{Q}_p)$, $L_p(s, \pi)$ is defined using only the factor π_p. Whether π occurs in decomposition of $L^2(\mathrm{GL}_2(\mathbb{Q}) \cdot Z(\mathbb{A}))\backslash\mathrm{GL}_2(\mathbb{A}))$ can be determined by analytic properties of $L(s, \pi)$. This is basically the converse theorem. If π occurs as a direct summand, then $\dim\mathrm{Hom}_{\mathrm{GL}_2(\mathbb{A})}(\pi, L^2(\mathrm{GL}_2(\mathbb{Q}) \cdot Z(\mathbb{A}))\backslash\mathrm{GL}_2(\mathbb{A}))) = 1$ (Multiplicity one theorem)

2 Upper half plane

Definition 2.1. $\mathcal{H} = \mathcal{H}^+ = \{\text{Im}(z) > 0\}$, $\mathcal{H}^- = \{\text{Im}(z) < 0\}$ are the upper and lower half planes, $\mathcal{H}^\pm = \mathbb{C} - \mathbb{R} = \mathbb{CP}^1 - \mathbb{RP}^1$

$$\text{PSL}_2(\mathbb{R}) = \text{SL}_2(\mathbb{R})/Z = \text{GL}_2^+(\mathbb{R})/Z \leq \text{GL}_2(\mathbb{R})/Z = \text{PGL}_2(\mathbb{R})$$

is a subgroup of index 2, $\text{PGL}_2(\mathbb{R})$ has two connected components, $\text{PSL}_2(\mathbb{R})$ is its identity component

$$\text{PSL}_2(\mathbb{C}) = \text{SL}_2(\mathbb{C})/Z = \text{GL}_2(\mathbb{C})/Z = \text{PGL}_2(\mathbb{C})$$

Definition 2.2. Consider natural projection $\mathbb{C}^2 - \{0\} \rightarrow \mathbb{CP}^1$, $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \mapsto \dfrac{z_1}{z_2}$, the standard action of $\text{GL}_2(\mathbb{C})$ on $\mathbb{C}^2 - \{0\}$ by matrix multiplication, which induces an action on $\mathbb{CP}^1$ by *fractional linear transformation*. Since scalar matrices act trivially, this induces an action of $\text{PGL}_2(\mathbb{C})$ on $\mathbb{CP}^1$

Fact 2.3.   1. Under this action, $\text{PGL}_2(\mathbb{C})$ is identified with the holomorphic automorphism group of $\mathbb{CP}^1$, also algebraic automorphism group

2. For any three distinct points $z_1, z_2, z_3 \in \mathbb{CP}^1$, there exists a unique $g \in \text{PGL}_2(\mathbb{C})$ such that $gz_1 = 0$, $gz_2 = 1$, $gz_3 = \infty$. So any non scalar matrix has at most two fixed points on $\mathbb{CP}^1$

Lemma 2.4.   1. $\text{PSL}_2(\mathbb{R})$ has three orbits on $\mathbb{CP}^1$: $\mathcal{H}, \mathcal{H}^-, \mathbb{RP}^1$

2. $\text{PGL}_2(\mathbb{R})$ has two orbits on $\mathbb{CP}^1$: $\mathcal{H}^\pm, \mathbb{RP}^1$

3. $\text{PSL}_2(\mathbb{R})$ is the group of holomorphic automorphisms of $\mathcal{H}$

Proof. If $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \text{GL}_2(\mathbb{R})$, then

$$\begin{aligned}
\text{Im}(gz) &= \text{Im}\,\frac{(az+b)(c\bar{z}+d)}{|cz+d|^2} \\
&= \text{Im}\,\frac{ac|z|^2 + adz + bc\bar{z} + bd}{|cz+d|^2} \\
&= \frac{(ad-bc)\,\text{Im}\,z}{|cz+d|^2} \\
&= \frac{\det(g)}{|cz+d|^2}\,\text{Im}\,z
\end{aligned}$$

So $\text{PSL}_2(\mathbb{R})$ preserves $\mathcal{H}, \mathcal{H}^-, \mathbb{RP}^1$. While $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in \text{PGL}_2(\mathbb{R})$ interchanges $\mathcal{H}$ and $\mathcal{H}^-$

$$\begin{bmatrix} y^{\frac{1}{2}} & xy^{-\frac{1}{2}} \\ 0 & y^{-\frac{1}{2}} \end{bmatrix} \cdot i = x + yi$$

For any $(x, y) \in \mathcal{H}$, thus $\text{PSL}_2(\mathbb{R})$ acts transitively on $\mathcal{H}$. For 3. first use Cayley transformation $\begin{bmatrix} 0 & -i \\ 1 & i \end{bmatrix}$ which induces an isomorphism $\mathcal{H} \rightarrow \mathbb{D}$, then use Schwartz lemma to determine $\text{Aut}(\mathbb{D})$, and then translate back to $\mathcal{H}$ $\qquad\square$

Exercise 2.5. The stabilizer of i in $\text{SL}_2(\mathbb{R})$ is

$$\text{SO}(2) = \left\{ \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \middle| \theta \in \mathbb{R} \right\}$$

So the stabilier of i in $\text{PSL}_2(\mathbb{R})$ is $\text{SO}(2)/\{\pm I\} \cong \text{SO}(2)$. $\mathcal{H} \cong \text{SL}_2(\mathbb{R})/\text{SO}(2) \cong \text{PSL}_2(\mathbb{R})/\text{SO}(2)$ is a homogeneous space

Exercise 2.6. $g^*(dx^2+dy^2) = |cz+d|^{-4}(dx^2+dy^2)$. Hence $g^*(y^{-2}(dx^2+dy^2)) = (y \circ g)^2 g^*(dx^2+dy^2) = \text{Im}(gz)^{-2}|cz+d|^{-4}(dx^2+dy^2) = y^{-2}(dx^2+dy^2)$

Definition 2.7. The *hyperbolic metric* on $\mathcal{H}$ is $\dfrac{dx^2+dy^2}{y^2}$. Then $\mathcal{H}$ becomes a model of hyperbolic plane: a two dimensional simply connected Riemannian manifold with constant Gaussian curvature -1. $\text{PSL}_2(\mathbb{R})$ are isometries on $\mathcal{H}$

Proposition 2.8. $\text{PSL}_2(\mathbb{R}) = \text{Isom}^+(\mathcal{H})$, the group of orientation preserving isometries. The group of isometries $\text{Isom}(\mathcal{H})$ is genenrated by $\text{Isom}^+(\mathcal{H})$ and reflection $z \mapsto -\bar{z}$

Proof. We have already seen $\text{PSL}_2(\mathbb{R}) = \text{Hol}(\mathcal{H})$, the group of holomorphic automorphisms and $\text{PSL}_2(\mathbb{R}) \leq \text{Isom}^+(\mathcal{H})$, but $\text{Isom}^+(\mathcal{H}) \leq \text{Hol}(\mathcal{H})$ since orientation preserving conformal maps are holomorphic $\qquad \square$

Fact 2.9. The geodesics on $\mathcal{H}$ are semi-circles othogonal to the real axis and half-lines orthogonal to the real axis, see [Miyake, Lemma 1.4.1]
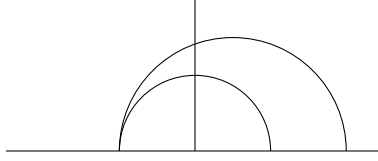
The hyperbolic metric induces a volume form $d\mu = \dfrac{dx \wedge dy}{y^2}$, and the hyperbolic Laplace operator $\Delta = -y^{-2}\left(\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}\right)$

Exercise 2.10. $d\mu, \Delta$ are invariant under $\text{PSL}_2(\mathbb{R})$ action (since the action preserves the metric)

Theorem 2.11 (Classification of motions). $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \text{SL}_2(\mathbb{R})$ and $g \neq \pm I$, then g has one or two fixed points on $\mathbb{CP}^1$

Proof.

Case 1: $c = 0$

       case i: $a = d = \pm 1$, so $b \neq 0$, g is a translation on $\mathbb{CP}^1$, ∞ is the only fixed point

       case ii: $a \neq d$, g is a linear function on $\mathbb{CP}^1$, $\infty, \dfrac{b}{d-a} \in \mathbb{R}$ are the two fixed points

Case 2: $c \neq 0$, then $\infty \mapsto \dfrac{a}{c}$ is not fixed

$$\frac{az+b}{cz+d} = z \Rightarrow z = \frac{a-d \pm \sqrt{(a+d)^2-4}}{2c}$$

       case i: $|a+d| = 2$, the only one fixed point is $\dfrac{a-d}{2c} \in \mathbb{R}$

       case ii: $|a+d| > 2$, there are two fixed points

       case iii: $|a+d| < 2$, there are two fixed points in $\mathcal{H}, \mathcal{H}^-$ and are conjugate of each other

In summary, there are three kinds of non-identity fractional linear transformation

1. Parabolic: When $|\text{tr}\, g| = 2$, only one fixed point, which is on $\mathbb{RP}^1$

2. Hyperbolic: When $|\text{tr}\, g| > 2$, two fixed points, both in $\mathbb{RP}^1$

3. Elliptic: When $|\text{tr}\, g| < 2$, two fixed points, one in $\mathcal{H}$, the other one in $\mathcal{H}^-$

$\square$

Example 2.12. Translation $\begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} : z \mapsto z + b$ is a parabolic motion. In general, parabolic elements move points along *horocycles*, i.e. horiontal lines or circles tangent to the x-axis. One can view horocycles as circles in $\mathbb{CP}^1 = S^2$ that are tangent to $\mathbb{RP}^1 = S^1$. $\mathrm{PSL}_2(\mathbb{R})$ action takes horocycles to horocycles and acts transitiviely on the set of horocycles. For any horizontal horocycle (say $\mathrm{Im}\, z = 1$), its stabilizer is $U = \left\{ \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \right\}$, identified with its image in $\mathrm{PSL}_2(\mathbb{R})$. Hence the set of horocycles can be identified with $\mathrm{PSL}_2(\mathbb{R})/U \cong (\mathbb{R}^2 - \{0\})/\{\pm I\}$ (note that $SL_2(\mathbb{R})/U \cong \mathbb{R}^2 - \{0\}$)

Example 2.13. $g = \begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix} : z \mapsto a^2 z$ is a hyperbolic motion, fixing $0, \infty$. In general, hyperbolic element moves points along *hypercycles*, i.e. intersections of circles in $\mathbb{CP}^1$ passing through the fixed points on $\mathbb{RP}^1$ with $\mathcal{H}$

Example 2.14. $g = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ is a elliptic motion, moving points along circles with *hyperbolic center* i, fixes i, induces counter-clockwise rotation of angle 2θ on the tangent space at i

Remark 2.15. Elliptic motions may have finite order ($\theta = \frac{\pi}{n}, n \in \mathbb{Z}$), parabolic and hyperbolic motions have infinite orders

References

[1] *A First Course in Modular Forms* - Fred Diamond, Jerry Shurman

Index