Introduction to Modular Forms

General outline

- $SL_2(\mathbf{Z})$, fundamental domains and S and T generate $SL_2(\mathbf{Z})$.
- Definition of modular forms and how for each function we get an infinte number of functions.
- Remarks: e.g. k even, $\mathcal{M}(\Gamma)$ is a vector space, $\mathcal{M}^0(R)$ is a field, and necessary and sufficient conditions for something to be a modular form.
- Eisenstein series
- Definition of modular forms
- Zeroes formula for modular forms
- Dimensional characterizations of modular forms

Things to keep on the board

- $S.z$, and $T.z$
- $\text{Im}(\gamma.z) = |cz+d|^{-2}\text{Im}(z)$
- Picture of the fundamental domain

Definition of $SL_2(\mathbf{Z})$ and its action.
We define

$$SL_2(\mathbf{Z}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : a, b, c, d \in \mathbf{Z}, ad - bc = 1 \right\}.$$

We denote this set by Γ.

We also define an action . on $\mathbf{C} \cup \infty$ by $SL_2(\mathbf{Z})$ by

$$g.z = \frac{az + b}{cz + d}$$

where $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and

$$g.\infty = \lim_{z \to \infty} \frac{az + b}{cz + d} = \frac{a}{c}.$$

The result of this action is called a *fractional linear transformation*. Two important examples are:

$$T.z = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right).z = z + 1$$
$$S.z = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right).z = \frac{-1}{z}$$

$SL_2(\mathbf{Z})$ acts on the complex upper half-plane

Let $H = \{z \in \mathbf{C} : \mathrm{Im}(z) > 0\}$. Note that the action . preserves H; i.e., if $z \in H$ then $g.z \in H$:

$$\begin{aligned}
\mathrm{Im}(g.z) &= \mathrm{Im}\left(\frac{az+b}{cz+d}\right) \\
&= \mathrm{Im}\left(\frac{(az+b)(c\bar{z}+d)}{|cz+d|^2}\right) \\
&= |cz+d|^{-2}\mathrm{Im}(adz+bc\bar{z}) \\
&= |cz+d|^{-2}(ad-bc)\mathrm{Im}(z) \\
&= |cz+d|^{-2}\mathrm{Im}(z)
\end{aligned}$$

Notice, though, that this is in H since the coefficient is positive.

Γ-equivalence and Fundamental Domains

An action breaks a set into congruence classes (i.e., orbits). We say, for $z_1, z_2 \in H$, that $z_1 \sim z_2$ if there exists $g \in \Gamma$ such that $g.z_1 = z_2$. The relation $\sim$ is called Γ-equivalence.

Let F be a closed region of H. We say that F is a *fundamental domain* if

1. every $z \in H$ is Γ-equivalent to a point in F
2. no two distinct points in the interior of the F are Γ-equivalent.

A fundamental domain for Γ.

Theorem A fundamental domain for Γ is

$$F = \left\{ z \in H : \frac{-1}{2} \le \mathrm{Re}(z) \le \frac{1}{2} \right\}.$$

The proof will work something like this: Take any $z \in H$. Move it over until it's in the strip. If it's in the unit circle, apply S if it's not, we're fine. Take any two elements in F they can't be related by any power of T since T moves everything over 1 (the width of F) and S moves things inside the unit circle. You might notice this would prove the claim only if S and T generate Γ, which we'll see is true.

Proof. Let $z \in H$ and let $\Gamma' = \langle S, T \rangle$. If $\gamma \in \Gamma'$ then as we saw before

$$\mathrm{Im}\gamma.z = |cz + d|^{-2}\mathrm{Im}(z).$$

We claim that there is some γ for which this is maximal.

The numbers $|cz+d|$ have a lower bound that is not 0. Notice that $cz+d$ is an element of the lattice spanned by z and 1. We can always fit a unit disk around the origin that doesn't hit any lattice points. Since the $cz+d$ have a lower bound, the claim is true.

Replacing this γ with $T^j\gamma$ for some j, we can assume, WLOG, that $\gamma.z$ has real part between $-1/2$ and $1/2$. But then if $\gamma.z \notin F$, i.e., $|\gamma.z| < 1$, then we'd have

$$\mathrm{Im}(S.\gamma.z) = \frac{\mathrm{Im}(\gamma.z)}{|\gamma.z|^2} > \mathrm{Im}(\gamma.z),$$

contradicting our choice of γ.

Now we prove that no two points in the interior of F are Γ-equivalent. Let $z_1, z_2 \in F$ (note this is more than we need). WLOG suppose $\mathrm{Im}(z_2) \ge \mathrm{Im}(z_1)$. By what we've just proved, there exists a $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ such that $\gamma.z_1 = z_2$.

Since $\mathrm{Im}(z_2) \ge \mathrm{Im}(z_1)$, then $|cz_1 + d| \le 1$.

$Im(z_2) = Im(\gamma.z_1) = |cz+d|^{-2}Im(z)$ which implies $|cz+d|^{-2} \ge 1$ which finally implies $|cz+d| \le 1$.

Since $z_1 \in F$ we see that $|c| < 2$

Say that $c = 2 + k$ for $k \in \mathbf{Z}^+$ and $z_1 = x + iy$. Then $|cz+d| = \sqrt{((2+k)x+d)^2 + (2+k)^2y^2}$ can be shown to be bigger than 1 (using the fact that $x^2 + y^2 > 1$), a contradiction.

So, we have just four cases:

1. $c = \pm 1, d = 0$
2. $c = 0, d = \pm 1$
3. $c = \pm 1, d = \pm 1$
4. $c = \pm 1, d = \mp 1$

For case 1 we have:

$$\begin{pmatrix} a & b \\ \pm 1 & 0 \end{pmatrix}.$$

Since the matrix is in $SL_2(\mathbf{Z})$, we have $b = \mp 1$. By a computation we get

$$\begin{pmatrix} a & \mp 1 \\ \pm 1 & 0 \end{pmatrix} = \pm T^a S.$$

For case 2 we have:

$$\begin{pmatrix} a & b \\ 0 & \pm 1 \end{pmatrix}$$

This implies that $a = \pm 1$. So we have

$$\begin{pmatrix} \pm 1 & b \\ 0 & \pm 1 \end{pmatrix} = \pm T^b.$$

For case 3 we have:

$$\begin{pmatrix} a & b \\ \pm 1 & \pm 1 \end{pmatrix}$$

This implies that $b = a - 1$ so by some computations we get:

$$\begin{aligned} \begin{pmatrix} a & a-1 \\ \pm 1 & \pm 1 \end{pmatrix} &= \begin{pmatrix} \pm 1 & a \\ 0 & \pm 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & \pm 1 \end{pmatrix} \\ &= \begin{pmatrix} \pm 1 & a \\ 0 & \pm 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \pm 1 & 1 \\ 0 & \pm 1 \end{pmatrix} \\ &= \pm T^a S T \end{aligned}$$

In case 4 we have, by an identical (almost) computation to case 3, that:

$$\begin{pmatrix} a & b \\ \pm 1 & \mp 1 \end{pmatrix} = \pm T^a S T^{-1}$$

In all four of these cases there are two subcases: $a = 0$ or $a = \pm 1$. It can't be more than one since the strip is only one unit wide. If $a = \pm 1$ this means that z_1, z_2 are on the boundary, so they don't enter into the fundamental domain. If $a = 0$ either $z_1 = z_2$ or (as in case 2) z_1 and z_2 are on the circular part of the boundary.

Various corollaries

Denote by Γ_z the set $\{\gamma \in \Gamma : \gamma.z = z\}$.

Corollary If $z \in F$ then $\Gamma_z = \pm I$ except

1. if $z = i$, then $\Gamma_z = \pm\{I, S\}$
2. if $z = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$, then $\Gamma_z = \{I, ST, (ST)^2\}$
3. if $z = \dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$, then $\Gamma_z = \{I, TS, (TS)^2\}$

Corollary The group $\bar{\Gamma} = \mathrm{SL}_2(\mathbf{Z})/\pm I$ is generated by S and T.

Proof Let $\Gamma' = \langle S, T\rangle$. Let z be any point on the interior of F. Let $g \in \Gamma$. Consider the point $g.z \in H$. We've already shown that there exists a $\gamma \in \Gamma'$ such that $\gamma.(g.z) = z$. But since z is in the interior, we know that $\Gamma'_z = \pm I$ so, in particular, $\gamma \cdot g = \pm I$. I.e., $g = \pm\gamma^{-1}$. This shows that (up to a sign) any $g \in \Gamma$ is also in Γ' (since γ^{-1} is definitely in Γ').

Finally, we get to Modular forms

- Let $f(z)$ be meromorphic on H, k be an integer.
  - *Meromorphic* means that the function is analytic except maybe for some poles.
- Further suppose $f(z)$ satisfies the relation:

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

  for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z})$.
- Suppose even further that $f(z)$ is meromorphic at infinity (i.e., $f(z) = \sum_{n\in\mathbf{Z}} a_n q^n$ has at most finitely many nonzero a_n when $n < 0$.
- Suppose furthest that $f(z)$ is actually holomorphic on H and at infinity (i.e., $a_n = 0$ when $n < 0$).
  - *Holomorphic* means analytic everywhere.
- Then $f(z)$ is a modular form of weight k for Γ; the set of these is denoted $\mathcal{M}_k(\Gamma)$.
- If we further have a_0, then $f(z)$ is called a *cusp form of weight k for* Γ. The set of functions is denoted $S_k(\Gamma)$.

Historical remarks

1. You may wonder where the $(cz+d)^k$ in the definition comes from. Surprisingly it comes from geometry. A fundamental domain for a modular form can be made into a manifold by adding a point at infinity and gluing sides together. Geometers are interested in functions that are invariant on the manifold acted on by Γ. Poincaré created this functions he called θ-Fuchsian that popped out a scalar when the argument was acted on by a γ. Realizing that if he could two different θ-Fuchsian functions to spit out the same constant, taking their quotient, he would have the function he desired. It so happened that the function he created spat out $(cz+d)^k$. This history is due to Stillwell in his edition of Poincaré's papers.

2. Another version has as its, in brief, the first ingredient the classical definition of the Weierstraß $\wp$-function expanded in terms of partial fractions and the consequent differential equation $(y')^2 = 4x^3 - 60G_4(x) - 140G_6(x)$. Here the quantities $G_k(\Lambda) = \sum' \omega^{-k}$ are Eisenstein series with the sum taken over all periods $\omega \neq 0$ in the period lattice Λ. Note that a typical ω looks like $n\omega_1 + m\omega_2$ where ω_1 and ω_2 are the primitive periods of $\wp$ and the (n, m) run through all ordered pairs of integers not both zero. It's convenient to multiply the G_k by ω_1^k and to set $z = \omega_2/\omega_1$ (and still to call these things G_k; in effect rescale so that $\omega_1 = 1$). It's now straightforward to see (because the transformation essentially only permutes the periods)

that the G_{2k} are invariant under a transformation $z \mapsto (az+b)/cz+d)$, with $ad-bc=1$, precisely up to a factor $(cz+d)^{2k}$. Moreover, general considerations, which are little more than linear algebra, lead one to see that all modular forms are in fact polynomials in G_4 and G_6.

Remarks that can be easily verified:

1. k is even.

   - Note that if $\gamma = -I$ then $f(\gamma.z) = f(z)$, which in turn implies $f(\gamma.z) = (-1)^k f(z)$ which can't hold if k is odd.

2. For $\Gamma = \mathrm{SL}_2(\mathbf{Z})$ if the relation holds for S and T it holds for all elements of Γ.

   - Note first that $\dfrac{d\gamma.z}{dz} = (cz+d)^{-2}$. So, we can rewrite the relation as $(\dfrac{d\gamma.z}{dz})^{k/2} f(\gamma.z) = f(z)$. Notice this means that $(dz)^{k/2} f(z)$ is invariant if we multiply it by γ. So, we conclude if the relation holds for γ_1 and for γ_2, it holds for $\gamma_1\gamma_2$.

3. $\mathcal{M}_k(\Gamma)$ is a complex vector space, trivially. Moreover, take a form of weight k_1 and a form of weight k_2, their product is a form of weight $k_1 + k_2$.

Examples of modular forms

Example 1 Let k be an even integer greater than 2. For $z \in H$, let

$$G_k(z) = \sum_{m,n} \frac{1}{(mz+n)^k}$$

where the sum is over all pairs of integers m, n.

Because $k \geq 4$ the sum is absolutely convergent on all compact subsets of H. Hence $G_k(z)$ is holomorphic on H. Further, note that $G_k(z) = G_k(z+1)$ since the sum is over pairs of integers. Finally, note that

$$\begin{aligned} G_k(-\frac{1}{z}) &= \sum_{m,n} \frac{1}{(mz+n)^k} \\ &= \sum_{m,n} \frac{1}{\left(\frac{m-nz}{-z}\right)^k} \\ &= \sum_{m,n} \frac{(-z)^k}{(m-nz)^k} \\ &= (-z)^k G_k(z) \end{aligned}$$

Sometimes the Fourier coefficients are of arithmetic (and hence number theoretic) interest. For instance,

Theorem Let $k > 2$ be an even integer. Then the Eisenstein series has Fourier expansion:

$$G_k(z) = 2\zeta(k)\left(1 - \frac{2k}{B_k}\sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n\right)$$

where

- $q = e^{2\pi i z}$,
- the Bernoulli numbers B_k are defined by
$$\frac{x}{x-1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!},$$
- and the function σ_{k-1} is the divisor function of n and is defined by
$$\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}.$$

Example 2 The *discriminant modular form* $\Delta(z)$ is defined to be a linear combination of the normalized Eisenstein series $E_k(z) = \frac{1}{2\zeta(k)} G_k(z)$. Let $g_2(z) = \frac{4}{3}\pi^4 E_4(z)$ and $g_3(z) = \frac{8}{27}\pi^6 E_6(z)$. Define

$$\Delta(z) = g_2(z)^3 - 27 g_3(z)^2.$$

Since $\mathcal{M}(\Gamma)$ is a complex vector space, we see that $\Delta(z)$ is an element of $\mathcal{M}(\Gamma)$. Note that since both $E_6(z)$ and E_4 have $a_0 = 1$, Δ has $a_0 = 0$ so it is a cusp form of weight 12.

The zeroes for $f \in \mathcal{M}_k(\Gamma)$.

Theorem Let $f(z) \in \mathcal{M}(\Gamma)$. For $P \in H$ let $v_p(f)$ be the order of the zero (or minus the order of the pole) of f at P. Let $v_\infty(f)$ be the index of the first nonvanishing term in the Fourier expansion of f. Then

$$v_\infty(f) + \frac{1}{2}v_i(f) + \frac{1}{3}v_\omega(f) + \sum_{P \in \Gamma/H, P \neq i,\omega} v_P(f) = \frac{k}{12}.$$

(Here $\omega = 1/2 + \dfrac{\sqrt{-3}}{2}$.)

Sketch of proof Apply the residue theorem to the fundamental domain with loops around the zeroes and cut off at some T bigger than all the poles and zeroes of f. Say P is on the LHS of the contour at is a zero, Q is on the circular part and is a zero, and that it has zeroes at ω (hence $T.\omega$) and i. The residue theorem says

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}dz = \sum_{P \in \Gamma/H, P \neq i,\omega} v_P(f).$$

Integrating over the contour, the top part gives up $-v_\infty(f)$, around each ω gives us $-v_\omega(f)/6$ and integrating over the i gives us $-v_i(f)/2$. Finally the integral over the rest of the unit circle gives us $k/12$.

Corollary Let k be even.

1. The only modular forms of weight 0 for Γ are constants.

   Proof Let c be any value taken by f. Then $f(z) - c$ has a zero. This means one of the terms on the LHS is nonzero, but the RHS is zero. This means that $f(z) = c$ is the zero function.

2. $\mathcal{M}_k(\Gamma) = 0$ if $k < 0$ or $k = 2$.

   Proof There is no way for the number of zeroes to be less than 1/3.

3. $\mathcal{M}_k(\Gamma)$ is one-dimensional and generated by $E_k(z)$ for $k = 4, 6, 8, 10, 14$.

   Proof For $k = 4$, we have $v_\omega(f) = 1$. For $k = 6$, we have $v_i(f) = 1$. For $k = 8$, we must have $v_\omega(f) = 2$. For $k = 10$ we have $v_i(f) = v_\omega(f) = 1$. For $k = 14$ we have $v_\omega(f) = 2$, $v_i(f) = 1$. Let $f_1, f_2 \in \mathcal{M}_k(\Gamma)$. They have the same zeroes, so f_1/f_2 is also a modular function of weight 0, hence a constant. Picking $f_2 = E_k$, we see that we're done.

4. $\mathcal{S}_k(\Gamma) = 0$ if $k < 12$ or $k = 14$, $\mathcal{S}_{12}(\Gamma)$ is one dimensional and generated by Δ, $\mathcal{S}_k(\Gamma) = \Delta\mathcal{M}_{k-12}(\Gamma)$.

   Proof For f a cusp form we have that $v_\infty(f) > 0$. Notice that when $k = 12$ and $f = \Delta$ we see that ∞ is the only zero. This means that f/Δ is also a modular form of weigth $k - 12$.

5. $\mathcal{M}_k(\Gamma) = \mathcal{S}_k(\Gamma) \oplus \mathbf{C}E_k$ for $k > 2$.

   Proof E_k does not vanish at infinity, but we can find a c so that $f - cE_k$ does vanish at ∞, i.e., so that $f - cE_k$ is a cusp form g and hence $f = cE_k + g$.

Corollary to the Corollary Any $f \in \mathcal{M}_k(\Gamma)$ can be written in the form

$$f(z) = \sum_{4i+6j=k} c_{i,j} E_4(z)^i E_6(z)^j.$$

We use induction. For $k = 4, 6, 8, 10, 14$ we get that $\mathcal{M}_k(\Gamma)$ is spanned by $E_4, E_6, E_4^2, E_4E_6, E_4^2E_6$ respectively. For $f \in \mathcal{M}_k(\Gamma)$ we can find a c so that $f - cE_4^iE_6^j$ is a cusp form. From before we get $f - cE_4^iE_6^j = f_1\Delta$. So

$$f = cE_4^iE_6^j + f_1\Delta = cE_4^iE_6^j + \frac{(2\pi)^{12}}{1728}(E_4^3 - E_6^2)f_1$$

where $f_1 \in \mathcal{M}_{k-12}(\Gamma)$. Apply the induction hypothesis to f_1 and we're done.