Compactification in low dimension

Tri Nguyen - University of Alberta

April 29, 2025

In this expository note, I will try to explain explicitly how to compactify $\Gamma\backslash\mathbb{H}$ by adding points in two ways.

1 Some preparations

We will always denote Γ a subgroup of the group $SL_2(\mathbb{Z})$ of finite index, and this group acts on the upper half complex plane $\mathbb{H}$ by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \circ z := \frac{az+b}{cz+d}$$

When z tends to infinity, we have

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

so we add the rational line to define the action of this group at ∞. In particular, we consider the set

$$\overline{\mathbb{H}} = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$$

Note that on the projective rational line, we define the action to be the multiplication of a 2×2 matrix with a 2×1 vector. Then under this action, we have the following lemma

Lemma 1. *$SL_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$.*

Proof. For each point in $\mathbb{P}^1(\mathbb{Q})$, we can choose the representative to be of the form $[a : b]$, where $\gcd(a, b) = 1$. Then there exists $x, y \in \mathbb{Z}$ such that

$$ax - by = 1$$

Thus we get the following equality

$$\begin{bmatrix} b & a \\ -x & y \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This implies any points in $\mathbb{P}^1(\mathbb{Q})$ can be moved to $[0 : 1]$, and thus the action is transitive. □

Corollary 2. *If Γ is a subgroup of finite index in $\mathrm{SL}_2(\mathbb{Z})$ then $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ has only finite orbits.*

2 Compactification of $\Gamma\backslash\mathbb{H}$ by adding points.

We introduce a topology on $\overline{\mathbb{H}}$. For the usual upper half plane, the topology is the usual metric topology on $\mathbb{C}$, and we only define the system of the neighborhood of $r \in \mathbb{P}^1(\mathbb{Q})$.

Let $S(c,\omega)$ be the circle that touches the real line at $\omega = p/q$ and has the radius $\frac{c}{2q^2}$. Then the collection of circles $D(c,\omega) = \bigcup_{0<c'\leqslant c} S(c',\omega)$ is called *Farey disk*. Let $c \to 0$, these disks define a neighborhood of ω. The Farey disks at ∞ are defined to be the region

$$D(T,\infty) = \{z : \Im z \geqslant T\}$$

It can be checked easily that the matrix under inversion, the Farey disk at ∞ is mapped to $D(1/T,0)$. In general, if $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma \circ \infty = \omega$ then $D(T,\infty)$ is mapped to $D(1/T,\omega)$. With the above topology on the extended upper half plane, we could show that

Lemma 3. $\Gamma\backslash\overline{\mathbb{H}}$ *is a compact set.*

The proof is taken from [1] and I rewrite it here for completeness.

Proof. We first prove for the case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. It is well known that the quotient space $\Gamma\backslash\mathbb{H}$ is identical to the set

$$\mathcal{F} = \{z \in \mathbb{H} : \Re z \in [-1/2, 1/2), |z| \geqslant 1 \text{ and } |z| > 1 \text{ if } \Re z > 0\}$$

By lemma 1, the projective rational line "shrinks" to a point under the action of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, Thus we can identify $\Gamma\backslash\overline{\mathbb{H}}$ with the set $\tilde{\mathcal{F}} = \mathcal{F} \cup \{\infty\}$. Consider an open cover $\{U_i\}_{i\in I}$ of $\tilde{\mathcal{F}}$ and the natural projection $\pi\colon \overline{\mathbb{H}} \to \tilde{\mathbb{F}}$. Then the set $\{\pi^{-1}(U_i)\}_{i\in I}$ forms an open cover of $\overline{\mathbb{H}}$. There must be an index i_0 such that $\pi^{-1}(U_{i_0})$ contains a neighborhood of ∞, namely contains a Farey disk $D(T,\infty)$ for some $T > 0$. Since $\overline{\mathcal{F}} - D(T,\infty)$ is a compact set, its image under π is compact, hence it can be covered by $U_{i_1}, \ldots, U_{i_m}$. Altogether, $\tilde{\mathcal{F}}$ admits a finite subcover $U_{i_0}, \ldots, U_{i_m}$.
Now we proceed to the general case. Note that

$$\overline{\mathbb{H}} = \mathrm{SL}_2(\mathbb{Z}) \circ \tilde{\mathcal{F}} = \bigcup \Gamma a_i \circ \tilde{\mathcal{F}}$$

by corollary 2. Then under the surjective map $\pi\colon \overline{\mathbb{H}} \to \Gamma\backslash\overline{\mathbb{H}}$, we have

$$\Gamma\backslash\overline{\mathbb{H}} = \bigcup \pi\left(\Gamma a_i \circ \tilde{\mathcal{F}}\right),$$

which shows that the set $Y(\Gamma) = \Gamma\backslash\overline{\mathbb{H}}$ is compact as it is the union of compact sets. □

The orbit of $\mathbb{P}^1(\mathbb{Q})$ under the action of Γ is called *cusps*. We have the obvious equality that

$$\Gamma\backslash\overline{\mathbb{H}} = \Gamma\backslash\mathbb{H} \cup \underbrace{\Gamma\backslash\mathbb{P}^1(\mathbb{Q})}_{\text{cusps}}$$

So in fact lemma 3 tells us that we only need to add a finite cusp to get a compact domain. That means we only need to consider the actions of Γ on the projective rational line. By the orbit-stabilizer theorem, we get the decomposition

$$\Gamma\backslash\bigcup_{\omega} D(c,\omega) = \bigcup \Gamma_{\omega_i}\backslash D(c,\omega_i)$$

where ω_i is the set of representative for the action of Γ on $\mathbb{P}^1(\mathbb{Q})$ and Γ_{ω_i} are the stabilizer of $\omega_i \in \Gamma$.
Again, since the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{Q})$ is transitive, for each $r \in \mathbb{P}^1(\mathbb{Q})$, there exists an element $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma \circ \infty = r$. So we have $\Gamma_r = \gamma\Gamma_\infty\gamma^{-1}$. Hence we only need to know the "shape" of the domain $\Gamma_\infty\backslash D(T,\infty)$. WLOG, we could assume $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, and hence

$$\Gamma_\infty = \left\{\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} : n \in \mathbb{Z}\right\}$$

Geometrically, $\Gamma_\infty\backslash D(T,\infty)$ is the strip $\{\Re z \in [-1/2, 1/2), \Im z \geqslant T\}$. But this is biholomorphic to a closed disk that misses a point on the boundary. So compactification is obtained by filling in the missing points to get finitely many compact disks.

3 Borel - Serre compactification of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$

We consider another compactification, by looking at the Farey disk $D(c,\omega)$ for fixed parameters c,ω. Then for any points $y \neq \omega$ on the Farey circle $S(c,\omega)$, we could connect y with ω by a unique geodesic in the upper half-plane.
These geodesics are either upper half circles that are orthogonal to the real line or the vertical line passing through ω. We thus can identify the Farey disks as follows

$$D(c,\omega) - \{\omega\} = X_{\infty,\omega} \times (0,c],$$

since a point θ on the Farey circles $S(c,\omega)$ is defined by its radius, up to a scaling of c, and the intersection of the geodesic $\overline{\theta\omega}$ with the real line. The uniqueness of the geodesics gives us a bijection between two sets. Here we let $X_{\infty,\omega} = \mathbb{P}^1(\mathbb{R}) - \{\omega\}$
How does the group Γ act on the set on the RHS set in the above identification? First, we look at the special case where $\omega = \infty$. In this case, the identification is

$$D(T,\infty) - \infty = X_{\infty,\infty} \times [T,\infty)$$

On the left, stabilizer subgroup Γ_∞ can be thought of as a subgroup of the group of translation, which leaves all the Farey circles $S(t,\infty)$ - which are the line $\{\Im z = t \geqslant T\}$ in this case - intact. Thus on the right-hand side, the action of Γ_∞ only affects the first coordinate. In general case, we need a lemma

Lemma 4. *If $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma \circ \infty = \omega$ then $\gamma \circ D(T,\infty) = D(1/T,\omega)$.*

Assume lemma 4 with the note that $\Gamma_\omega = \gamma\Gamma_\infty\gamma^{-1}$, we conclude that the action of Γ_ω only affects $X_{\infty,\omega}$ for all $\omega \in \mathbb{P}^1(\mathbb{Q})$. Since $\Gamma_\omega\backslash X_{\infty,\omega}$ is a circle, it is compact. Hence we can compactify the quotient space $\Gamma_\omega\backslash D(c,\omega) - \{\omega\}$ as

$$\Gamma_\omega\backslash D(c,\omega) - \{\omega\} \hookrightarrow \Gamma_\omega\backslash X_{\infty,\omega} \times [0,c]$$

As in section 1, we only need to compactify finitely many such quotient spaces and get the compactification of $\Gamma\backslash\mathbb{H}$.
Now we give a proof of lemma 4

Proof. Assume $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ is an element that sends ∞ to $\omega = \frac{p}{q}$. Let $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix}.$$

Thus we must have $a = p, c = q$ and b, c are integers such that $aq - cp = 1$. A Farey circle in the neighborhood of ∞ is, in fact, a line $S(T,\infty) = \{\Im z = T\}$, and this line is mapped to a circle tangent to the real line. Direct calculation shows that, for $z = x + iT$

$$\Im(\gamma \circ z) = \frac{\Im z}{|cz+d|^2} = \frac{T}{(cx+d)^2 + c^2T^2} \leqslant \frac{1}{q^2T}$$

The equality happens if $x = -d/c$. Since this $\gamma \circ z$ is a point on the circle tangent to the real line at p/q and has the largest distance to the real line, the segment connect p/q and $\gamma \circ z$ must be the diameter of the image circle. In particular, the radius of the image circle is $\frac{1}{2Tq^2}$. Lemma 4 follows immediately. $\square$

The above process can be applied to finitely many Fareye disks as in section 2 to get a compactification of $\Gamma\backslash\mathbb{H}$.

4 Equivalent definitions of semi-stability

4.1 In two dimension

In this section, we give two definitions of semi-stability in $\mathbb{H}$, and show that they are essentially the same. We will also compute the semi-stable locus in $\mathbb{H}$.

We first introduce some terminology: for each complex number $z \in \mathbb{H}$, we assign to it a lattice in $\mathbb{C}$,where the lattice is spanned by two vectors $\{z, 1\}$. By identifying $\mathbb{C} \cong \mathbb{R}^2$, we can compute the volume of the fundamental domain given by this lattice is $y = \Im(z)$. Then we scale two vectors in the basis to get a unit lattice.

Following this process, each $z \in \mathbb{H}$ is assigned with a unique unit lattice, namely $\Gamma_z = \operatorname{span}_{\mathbb{Z}} \{a, z/a\}$, where $a = \Im(z)$. Now we are ready to define semi-stability. Furthermore, to each lattice, we assigned to it a plot in the following way:

- We start with the point $(0, 0)$ in the plane.
- Let u be the shortest vector in the lattice Γ_z, we highlight the point $(1, \log |u|)$ in the plane.
- Finally, we attached the point $(2, \log(vol(A)))$, where A is the fundamental domain of Γ_z.
- We connect consecutive points by line segments. The union of these line segments is called *profile* of the lattice.

In our setting, since we already normalize all lattices to unit lattice, the final point is in fact $(2, 0)$.

Definition 1. *The lattice assigned to the number $z \in \mathbb{H}$ is call semi-stable if and only if the point $(1, \log |u|)$ lies above the x-axis.*

Before giving the second definition of semi-stablility, we will try to find the semi-stable locus using this definition. First, we restrict our attention to the fundamental domain

$$\mathcal{F} = \{z \in \mathbb{H} : \Re z \in [-1/2, 1/2), |z| \geqslant 1 \text{ and } |z| > 1 \text{ if } \Re z > 0\}$$

For each point $z \in \mathcal{F}$, it is easily to see that the shortest vector in the corresponding lattice is a. Thus the 1-dim point is $(1, \log(a)) = (1, -\log(y)/2)$. By definition, Γ_z is semi-stable iff $\log(y) \leqslant 0$, i.e. $y \geqslant 1$. To find the semi-stable locus in the whole upper half plane, we need the following result

Lemma 5. *If Γ_z is semi-stable, then so is the lattice $\Gamma_{g \circ z}$, where $g \in \mathrm{SL}_2(\mathbb{Z})$.*

Proof. If we denote $L_z = \operatorname{span}_{\mathbb{Z}} \{1, z\}$, then $L_{\gamma \circ z} = cL_z$ for some complex number c. Indeed, we just need to check for γ be an inversion or translation, since these two transformations generate $\mathrm{SL}_2(\mathbb{Z})$, but this is easy. Now let $c = re^{it}$. Multiplying by e^{it} doesn't change the length, hence doesn't change the semi-stability. Multiplying by a positive number r will shift $(1, \log |u|)$ to $(1, \log |u| + \log r)$ and $(2, \log(vol(A)))$ to $(2, \log(vol(A)) + 2 \log r)$. I think in 2 dimensional case, c is 1

The line segment d connecting origin with the final point intersect the line $x = 1$ at $(1, \log(vol(A)) + \log r)$. By the semi-stability of the original lattice the point $(1, \log |u| + \log r)$ is above the line segment d. □

From this lemma, we could see that the semi-stable locus is the complement of the Farey balls in the upper half plane.

Now we give the second definition of semi-stability. First we note that $\mathbb{H}$ can be identified with the set $\mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$. Using Iwasawa's decomposition, we could identify $z = x + iy$ with the pair $(a(z), n(z))$ where

$$a(z) = \begin{bmatrix} y^{-1/2} & 0 \\ 0 & y^{1/2} \end{bmatrix} \quad n(z) = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$$

Then we can define a map

$$H_B\colon \mathbb{H} \to \mathfrak{sl}_2, \quad z \mapsto \log(a(z))H,$$

where $H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then we define a unique linear map $\alpha\colon \mathbb{R}H \to \mathbb{C}$ such that $\alpha(H) = 2$. Set $\rho = \alpha/2$. For each $T = -kH$ where $k > 0$, we define the degree of instability of $x \in \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$ as follows

$$\deg^T_{\text{inst}}(x) := \min_{\gamma\in\Gamma/\Gamma\cap B} \langle \rho, H_B(x\gamma) - T\rangle$$

In this particular case, the RHS of the above definition can be simplified as

$$\deg^T_{\text{inst}}(x) = \min_{\gamma\in\Gamma} \log(a(x\gamma) + k)$$

For the case $k = 0$, the point x is call semi-stable iff $\deg^T_{\text{inst}}(x) \geqslant 0$. Since the minimum is always achieved, this implies $\frac{-1}{2}\log(\Im(x\gamma)) \geqslant 0$, i.e. $\Im(x\gamma) \leqslant 1$ for all γ.
A key observation is that the minimum of $a(z)$ is achieved inside the Siegel's set, which is exactly the fundamental domain $\mathcal{F}$ in this case. This implies the semi-stable locus inside the fudamental domain is the intersection of $\mathcal{F}$ with $\{\Im(z) < 1\}$. So the semi-stable locus is the union of orbits who has a representative in this part.

4.2 In higher dimension

We start with the standard basis $\{e_i\}_{i=1}^n$ in the Euclidean space $\mathbb{R}^n$. Then for each $g \in \mathrm{SL_n}(\mathbb{R})$, we can assign g to the lattice spanned by $\{ge_i\}_{i=1}^n$. By normalizing, we can choose g so that the fundamental domain corresponding to this lattice has unit volume and denoted by Γ_g. for each $1 \leqslant m \leqslant n-1$, pick the sublattice M of the smallest volume and assign it to the point $(m, \log vol(M))$, and connecting to consecutive points by a line segment. boundary of the convex hull of this plot is called **canonical plot**. The analogue of semi-stability in higher dimension is

Definition 2. *For each point g, if the canonical plot of the lattice Γ_g contains exactly a line connecting the origin with $(n, 0)$.*

We can defined a function H_B similarly to the function is previous section, then the degree of instability can be defined by computing the following function:

$$\deg_{\text{inst}}(g) := \min_{P\in ParSt, \gamma\in\mathrm{SL}(\mathbb{Q})/P(\mathbb{Q})} \langle \rho_P, H_B(g\gamma)\rangle$$

Definition 3. *A point g is semi-stable if $\deg_{inst}(g) \geqslant 0$.*

To give a feeling how to compute the degree of instability of g, we look at the Iwasawa decomposition and consider $n = 3$. Then

$$g = kan,$$

where $k \in \mathrm{SO}(3, \mathbb{R})$ and a is the the diagonal matrix, n is unipotent upper triangular matrix. In particular

$$a = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \quad \text{and} \quad n = \begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix}$$

The value of $\deg_{\text{inst}}(g)$ will depend on which parabolic subgroup we are evaluating at. But in short, we will have the following system of inequalities

$$\begin{cases} \log(s_1) \geqslant 0 \\ \log(s_1) + \log(s_2) \geqslant 0 \end{cases}$$

Now we introduce some setting to understand what semi-stability means in term of canonical plot. In $\mathbb{R}^3$, we choose a specific set of linearly independent vector $\Delta = \{\alpha_1, \alpha_2\}$, where

$$\alpha_i = e_{i+1} - e_i, \quad i = 1, 2$$

For each $j = \overline{1, n}$, we defined the weight λ_j such that

$$\lambda_j \cdot \alpha_i = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

Clearly we have λ_n is orthogonal to Δ. Now consider the vector

$$\omega = \sum_{i=1}^{n} \log a_g \cdot e_1 = (s_1, \ldots, s_n)$$

Since multiplying by k doesn't change the volume of the lattice, we can assume that $g = an$. The following equations are easily verified:

Lemma 6. *We have the following identities:*

1. $\alpha_i \cdot \omega = s_{i+1} - s_i$.
2. $\lambda_1, \ldots, \lambda_n$ *is a basis of* $\mathbb{R}^n$.
3. $\nu_\Delta(\lambda_i) = \lambda_i - (i/n)\lambda_n$, *where* μ_Δ *is the projection on the subspace spanned by* Δ.

Let $\mathcal{C} = \{v \in \mathbb{R}^3 \mid \alpha_i \cdot v > 0 \text{ for all } \alpha_i\}$. This defines a cone in $\mathbb{R}^n$. We also denote

- $V_\Delta = \{v \in \mathbb{R}^n \mid \alpha_i \cdot v = 0\}$.
- $V_\Delta^\Delta = \{v \in \mathbb{R}^n \mid \text{ The nearest point to } v \text{ lies on the faces } V_\Delta\}$

In particular, we have a rather nice description of V_Δ^Δ in term of $\alpha_i's$ and λ_n.

Lemma 7.

$$V_\Delta^\Delta = \left\{ \sum_{i=1}^{n-1} c_i\alpha_i + c_n\lambda_n \mid c_i \leqslant 0, \quad \forall i \leqslant n-1 \right\}$$

Now for each point g, we assign to this point the point ω and we defined its **profile** to be the profile polygon that moves from $x = i$ to $x = i+1$ along a segment of slope s_i. Assume that this profile passes through n points (i, y_i), we have a relation

$$y_0 = 0, y_i - y_{i-1} = s_i$$

which implies

$$y_i = s_1 + \ldots + s_i = -\lambda_i \cdot \omega$$

By the lemma 7, V_Δ^Δ are spanned by λ_n and α_i's, which means for it consists of the points v such that $\nu_\Delta(\lambda_i) \cdot v \leqslant 0$ for $i = \overline{1, n-1}$ by lemma 6. In term of y_i's, it can be rephrased as $y_i - (i/n)y_n$ (note that $y_i = -\lambda_i \cdot \omega$). But in term of plot, this is just saying that the point (i, y_i) lies above the point (n, y_n) for all n. Note that in our setting, $y_n = 0$. So V_Δ^Δ is exactly the set of semi-stable lattice points. Note that we showed above $y_i/i \geqslant y_n/n = 0$, which implies that a point g corresponding to a lattice point lies inside V_Δ^Δ must satisfies $s_1 + \ldots + s_i \geqslant 0$, for $i \leqslant n-1$. For $n = 3$, we recover the previous definition.

References

[1] Diamond, F., & Shurman, J. M. (2005). A first course in modular forms (Vol. 228, pp. xvi-436). New York: Springer.