ON THE ACTION OF THE FULL MODULAR GROUP

DAKSH DHEER
(THIRD YEAR, HANSRAJ COLLEGE, 2023)

Abstract. In this paper, we introduce the idea of groups acting topologically on spaces and give an algebraic description, alongside a geometric view, of the action of the full modular group $SL_2(\mathbb{Z})$ on the upper half complex plane. This notion lends itself to providing an alternate perspective for studying an abstract space by relating its topology with that of a coset space of a known group.

Keywords: *Group actions, Modular group, Special linear group, topological groups, orbit spaces, Upper-half complex plane, Fundamental domains*

1. Introduction

Definition 1.1. A group G is said to be a topological group if it is endowed with a topology such that the operation map $(x, y) \to x \cdot y$ and inversion map $x \to x^{-1}$ are continuous.

Definition 1.2. The (continuous) action of a group G on a topological space X is defined as follows:

For all elements $g \in G$, we have a continuous map $x \to g \cdot x \in X$ such that

- $1 \cdot x = x, \ \forall x \in X$
- $g \cdot (h \cdot x) = (gh) \cdot x, \quad \forall g, h \ \in G$

Clearly, $x \to g^{-1} \cdot x \in X$ is the inverse continuous map, thus each element $g \in G$ defines a homeomorphic automorphism of X.
We call X a G**-space** and say G *acts topologically* on X.

For each element $x \in X$, we define the *stabilizer of* x and *orbit* of x as follows:

Definition 1.3 (Stabilizer)**.**

$$G_x = stab_G(x) := \{g \in G : g \cdot x = x\} \subseteq G$$

Definition 1.4 (Orbit)**.**

$$Gx = orb_G(x) := \{g \cdot x : g \in G\} \subseteq X$$

If there is only one orbit, we say G acts *transitively* on X.
Given a group G acting on a space X, if $v \in orb_G(w)$, we say v and w are *G-equivalent.*

2. The Action of $SL_2(\mathbb{C})$

We define the special linear group of 2×2 matrices over the complex field $SL_2(\mathbb{C})$ as:

Definition 2.1.

$$SL_2(\mathbb{C}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{C} \text{ and } ad - bc = 1 \right\}$$

This group acts on the Riemann Sphere, i.e., $\mathbb{C} \cup \{\infty\}$:
For all elements $\gamma \in SL_2(\mathbb{C})$, we have a continuous map $z \to \gamma \cdot z$, where

$$\gamma \cdot z = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \frac{az+b}{cz+d}$$

Note that the image of z under this action is a Möbius transformation, thus it is well-defined at all points on the Riemann sphere. A complete description is given by:

$$\gamma \cdot z = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \begin{cases} \infty & z = \dfrac{d}{c} \text{ and } c \neq 0 \\ \dfrac{az+b}{cz+d} & z \neq \dfrac{d}{c} \text{ and } c \neq 0 \\ \dfrac{a}{c} & z = \infty \text{ and } c \neq 0 \\ \infty & z = \infty \text{ and } c = 0 \end{cases}$$

Now, consider the imaginary part of $\gamma \cdot z$:

$$\begin{aligned} \Im(\gamma \cdot z) = \Im\left(\frac{az+b}{cz+d}\right) &= \Im\left(\frac{(az+b)(c\bar{z}+d)}{|cz+d|^2}\right) \\ &= \frac{1}{|cz+d|^2}\Im(ac|z|^2 + adz + bc\bar{z} + bd) \\ &= \frac{ad-bc}{|cz+d|^2}\Im(z) = \frac{1}{|cz+d|^2}\Im(z) \end{aligned}$$

Hence, we have

$$\Im(\gamma \cdot z) = \frac{1}{|cz+d|^2}\Im(z) \tag{1}$$

We know that given any $z \in \mathbb{R}$, $\Im(z) = 0 \implies \Im(\gamma \cdot z) = 0$.
Furthermore, $\Im(z) < 0 \implies \Im(\gamma \cdot z) < 0$ and $\Im(z) > 0 \implies \Im(\gamma \cdot z) > 0$, thus the action of $SL_2(\mathbb{R})$ preserves:

- $\mathcal{H} := \{z \in \mathbb{C} : \Im(z) > 0\}$, which is the right hemisphere of the Riemann sphere.
- $\mathcal{H}^* := \{z \in \mathbb{C} : \Im(z) < 0\}$ which is the left hemisphere of the Riemann sphere.
- $\mathbb{R} \cup \{\infty\}$, which is the central longitude in the Riemann sphere.

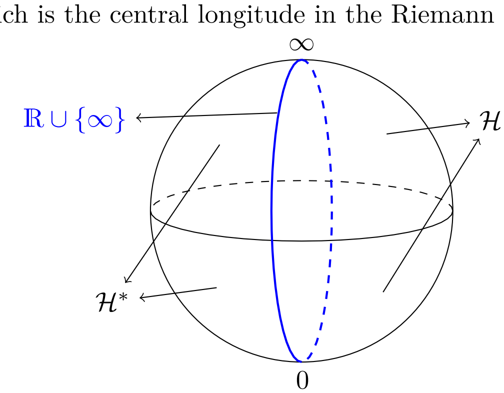

FIGURE 1. The Riemann Sphere

Hence, to study the action of $SL_2(\mathbb{Z})$ on $\mathbb{C} \cup \{\infty\}$, it is sufficient to study it on $\mathcal{H}^*$, $\mathcal{H}$ and $\mathbb{R} \cup \{\infty\}$.
On $\mathbb{R} \cup \{\infty\}$, the action just shuffles the real numbers, and the spaces $\mathcal{H}^*$ and $\mathcal{H}$ are homeomorphic. Hence, in this paper, we shall focus on the action of $SL_2(\mathbb{Z})$ on $\mathcal{H}$, which is the upper-half complex plane.

3. The Full Modular Group

We define the full modular group to be the special linear group of 2×2 matrices over the ring of integers with unit determinant, i.e.,

$$SL_2(\mathbb{Z}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z} \text{ and } ad - bc = 1 \right\}$$

Consider the set $\mathcal{D} := \{z \in \mathcal{H} : |\Re(z)| < \frac{1}{2}, |z| > 1\} \subset \mathcal{H}$.
We then have its closure as:

$$\overline{\mathcal{D}} := \{z \in \mathcal{H} : |\Re(z)| \leq \frac{1}{2}, |z| \geq 1\}$$

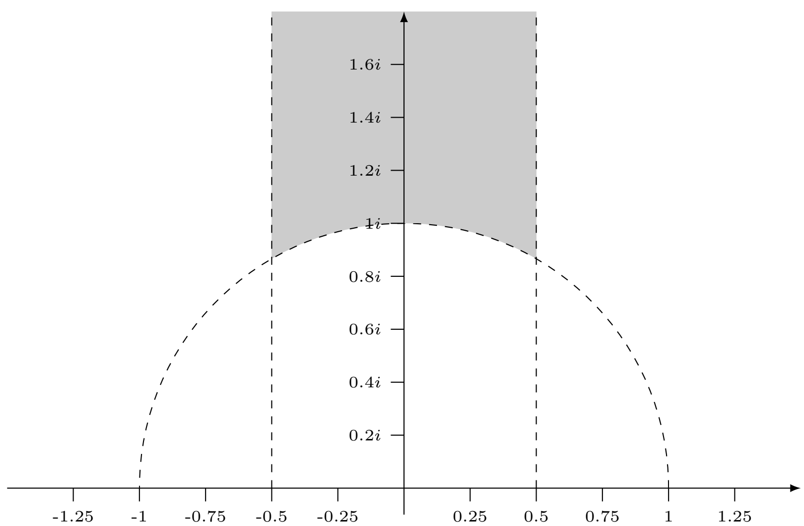

Figure 2. The region $\mathcal{D}$

Theorem 1.

(A) The group $SL_2(\mathbb{Z})$ *is generated by* $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ *and* $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, *i.e.,* $SL_2(\mathbb{Z}) = \langle S, T \rangle$.

(B) Given any $z \in \mathcal{H}$, *there exists* $\gamma \in SL_2(\mathbb{Z})$ *such that* $\gamma \cdot z \in \overline{\mathcal{D}}$.

(C) Let z, w *be two distinct elements of* $\overline{\mathcal{D}}$ *such that* $w \in orb_{SL_2(\mathbb{Z})}(z)$. *Then, either*
- $\odot$ $\Re(z) = \pm\frac{1}{2}$ *and* $w = z \pm 1$, *i.e.,* $w \in T^{\pm 1} \cdot z$, *or*
- $\odot$ $|z| = 1$ *and* $w = -\frac{1}{z}$, *i.e.,* $w \in S \cdot z$

(D) Let $z \in \mathcal{H}$ such that $z \notin \{i, \rho = e^{i\frac{\pi}{3}}, -\overline{\rho} = e^{i\frac{2\pi}{3}}\}$. Then, $stab_{SL_2(\mathbb{Z})}(z) = \{\pm I\}$

Proof. (A) Let G denote $SL_2(\mathbb{Z})$. Note the fact that $\langle S, T\rangle \subseteq G$.
Observe that $T^n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ and that $T^n \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+nc & b+nd \\ c & d \end{pmatrix}$
Similarly, we have $S \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}$ and that $S^2 = -I$.
Now, let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$ be arbitrary.
Case (i) $c = 0$. Then,

$$\det \gamma = 1 \implies ad - b(0) = ad = 1 \implies a = d = \pm 1$$

Thus, we have $\gamma = \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = T^b$ or $S^2 \cdot T^b$, so $\gamma \in \langle S, T\rangle$.
Case (ii) $c \neq 0$.
Without loss of generality, assume $|a| \geq |c|$ (if this is not the case, then apply S to γ to make it so.)
By Division Algorithm, we have $a = cq + r$, $0 \leq r < |c|$. Now,

$$T^{-q} \cdot \gamma = \begin{pmatrix} a - qc & b - qd \\ c & d \end{pmatrix} = \begin{pmatrix} r & b - qd \\ c & d \end{pmatrix}$$

$$\text{Then, } S^2 \cdot T^{-q} \cdot \gamma = \begin{pmatrix} -c & -d \\ r & b - qd \end{pmatrix}$$

and now $|-c| > r$ so we may repeat the process.
Since all the entries of the matrix are integers, the process terminates when $c = 0$, and we revert to case (i). Thus, $\gamma \in \langle S, T\rangle \quad \forall \gamma \in G$, and so we are done.

(B) Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$ be fixed. Now, if $g' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in G$ such that $\Im(g \cdot z) \leq \Im(z)$, then, using equation (1), we have:

$$\frac{\Im(z)}{|cz+d|^2} \leq \frac{\Im(z)}{|c'z+d'|^2} \implies 0 \leq |c'z + d'|^2 \leq |cz+d|^2 \tag{2}$$

Since c, d are fixed and $c', d' \in \mathbb{Z}$, there are only finitely many choices for c', d' such that equation (2) holds. We choose $\gamma \in \langle S, T\rangle$ such that $\Im(\gamma \cdot z)$ is maximal. Now, $T^n \cdot z = z + n$, hence given any $\theta \in \mathcal{H}$, there exists an integer n such that $|\Re(T^n \cdot \theta)| \leq \frac{1}{2}$. Thus, $\exists\, n_0 \in \mathbb{Z}$ such that $|\Re(T^{n_0} \cdot \gamma \cdot z)| \leq \frac{1}{2}$.
If $|T^{n_0} \cdot \gamma \cdot z| \geq 1$, then we are done. Suppose not, i.e., $|T^{n_0} \cdot \gamma \cdot z| < 1$, then we have:

$$\begin{aligned}\Im(S\cdot T^{n_0}\cdot\gamma\cdot z) &= \Im\left(-\frac{1}{T^{n_0}\cdot\gamma\cdot z}\right) = \Im\left(-\frac{1}{T^{n_0}\cdot\gamma\cdot z}\cdot\frac{\overline{T^{n_0}\cdot\gamma\cdot z}}{\overline{T^{n_0}\cdot\gamma\cdot z}}\right)\\ &= -\frac{\Im(\overline{T^{n_0}\cdot\gamma\cdot z})}{|T^{n_0}\cdot\gamma\cdot z|^2} = \frac{\Im(T^{n_0}\cdot\gamma\cdot z)}{|T^{n_0}\cdot\gamma\cdot z|^2}\\ &= \frac{\Im(\gamma\cdot z)}{|T^{n_0}\cdot\gamma\cdot z|^2} > \Im(\gamma\cdot z) \qquad (\text{since } |T^{n_0}\cdot\gamma\cdot z|<1)\end{aligned}$$

But this contradicts the maximality of $\Im(\gamma\cdot z)$.
Thus, $\forall\ z\in\mathcal{H},\ \exists\ n_0\in\mathbb{Z}$ and $\gamma\in\langle S,T\rangle\subseteq G$ such that $T^{n_0}\cdot\gamma\cdot z\in\overline{\mathcal{D}}$.

(C) We have $z,w\in\overline{\mathcal{D}}$ such that $z\neq w$ and $w\in orb_{SL_2(\mathbb{Z})}(z)$. Hence, $w=\gamma\cdot z$ for some $\gamma\in G$.
We may assume, without loss of generality, that $\Im(\gamma\cdot z)\geq\Im(z)$. Thus, using equation (1), we obtain $\frac{\Im(z)}{|cz+d|^2}\geq\Im(z)\implies|cz+d|^2\leq 1$. Now, observe that

$$\begin{aligned}|cz+d|^2 &= (c\cdot\Re(z)+d)^2+(c\cdot\Im(z))^2\\ &= c^2\cdot\Re(z)^2+2cd\cdot\Re(z)+d^2+c^2\cdot\Im(z)^2\\ &= c^2|z|^2+2cd\cdot\Re(z)+d^2\\ &\geq c^2+2cd\cdot\Re(z)+d^2 \qquad (\text{since } z\in\overline{\mathcal{D}}\implies|z|\geq 1)\end{aligned}$$

Hence, we get

$$c^2+2cd\cdot\Re(z)+d^2\leq|cz+d|^2\leq 1 \tag{3}$$

We claim that $|c|<2$. If not, then

$$\begin{aligned}|cz+d|^2 &\geq 2^2+2(2)d\cdot\Re(z)+d^2\\ &\geq 2^2+2(2)d\cdot(\frac{1}{2})+d^2\\ &= 4+d(2+d)\geq 4+min\{d(2+d)\}\\ &= 4+(-1)=3 \qquad (\text{which contradicts the inequality (3)})\end{aligned}$$

Thus, $c=-1,0,$ or 1. We consider all cases separately.

<u>*Case (1)*</u>: $c=0$.
In this case, $d\neq 0$ (otherwise $\det\gamma=ad-bc=0$).
From (3), we have $1\geq 0^2+2(0)d\cdot\Re(z)+d^2\implies 1\geq d^2\implies d=\pm 1$.

Let $d=1$. Then, $\gamma=\begin{pmatrix}a&b\\0&1\end{pmatrix}$ and using the fact that $\det\gamma=1$, we obtain $a=1$.
Hence, $\gamma=\begin{pmatrix}1&b\\0&1\end{pmatrix}$, so $w=\gamma\cdot z=z+b$. Moreover, both $z,w\in\overline{\mathcal{D}}$, hence $\Re(z)=\pm\frac{1}{2}$ and $b=\pm 1$ and so $w\in T^{\pm 1}\cdot z$.
For $d=-1$, we would similarly obtain $w\in T^{\pm 1}\cdot z$.

Case (2): $c = 1$.
From (3), we have $1 \geq 1 + 2d \cdot \Re(z) + d^2$.
Clearly, $|d| < 2$, otherwise we get $1 \geq 1 + 2(2) \cdot \Re(z) + 2^2$, which is absurd.
Hence, $d = 0, 1, -1$.

Subcase (1): $d = 0$.
We have $1 \geq |cz + d| \implies 1 \geq |z|$. But $z \in \overline{\mathcal{D}} \implies |z| \geq 1$, hence we get $|z| = 1$.
Now, $\det \gamma = 1 \implies ad - bc = -b = 1$, hence $\gamma = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$ and $w = \gamma \cdot z = \frac{az-1}{z} = a - \frac{1}{z} = a - \frac{\overline{z}}{|z|^2} = a - \overline{z}$.
Since $z, w \in \overline{\mathcal{D}}, |\Re(z)| \leq \frac{1}{2}$ and $|\Re(w)| \leq \frac{1}{2}$, i.e., $|a - \Re(z)| \leq \frac{1}{2}$ (since $\Re(z) = \Re(\overline{z})$.
Thus, from $|\Re(z)| \leq \frac{1}{2}$ and $|a - \Re(z)| \leq \frac{1}{2}$, we may conclude $a = -1, 0, 1$.
If $a = 0$, then $w = -\overline{z} = -\frac{1}{z}$ and so $w \in S \cdot z$.
If $a = 1$, then $w = 1 - \frac{1}{z}$ and also $w, z \in \overline{\mathcal{D}}$. The only such possibility is $w = -\overline{\rho}$ and $z = -\overline{\rho}$, which contradicts the fact that $z \neq w$. Hence, $a \neq 1$.
By a similar argument, one may also conclude that $a \neq -1$.

Subcase (2): $d = 1$.
We have $1 \geq 1 + 2 \cdot \Re(z) + 1 \implies \Re(z) \leq -\frac{1}{2} \implies \Re(z) = -\frac{1}{2}$.
Also, $1 \geq |z + 1|$. The only such possibility is $z = \rho$.
Now, $\det \gamma = 1 \implies ad - bc = a - b = 1$, hence we have $\gamma = \begin{pmatrix} a & a-1 \\ 1 & 1 \end{pmatrix}$. Thus, $w = \gamma \cdot z = \frac{a\rho + a - 1}{\rho + 1} = a - \frac{1}{\rho+1} = a - \frac{1}{-\overline{\rho}} = a + \rho$.
Since both $z, w \in \overline{\mathcal{D}}$, the only possibilities for a are $a = 1$ or $a = 0$.
If $a = 1$, then we get $z = \rho$ and $w = -\overline{\rho}$, which satisfies $|z| = 1$ and $w = -\frac{1}{z}$. If $a = 0$, then we get $z = \rho$ and $w = \rho$, which contradicts $z \neq w$.

Subcase (3): $d = -1$.
The same argument gives $z = -\overline{\rho}$ and $1 = -a - b$ and proceeding similarly, we get $a = 0$ or $a = -1$.
If $a = -1$, then we get $w = \rho$ and $z = -\overline{\rho}$, which satisfies $|z| = 1$ and $w = -\frac{1}{z}$. If $a = 0$, then we get $z = -\overline{\rho}$ and $w = -\overline{\rho}$, which contradicts $z \neq w$.

Case (3): $c = -1$.
Note that $-\gamma \cdot z = \frac{-az-b}{-cz-d} = \frac{az+b}{cz+d} = \gamma \cdot z$, hence we may replace γ by $-\gamma$ in the above cases and repeat the argument.

(D) Let $z \in \mathcal{H}$ such that $stab_{SL_2(\mathbb{Z})}(z) \neq \{\pm I\}$. Let $\gamma \in stab_{SL_2(\mathbb{Z})}(z)$ be arbitrary. Then $\gamma \cdot z = z$ and so $z \pm 1 = z$ or $-\frac{1}{z} = z$ [using part (B)]

If $z \pm 1 = z$, we get $z = \rho$ or $-\overline{\rho}$. If $-\frac{1}{z} = z$, we get $z = i$. (since $z \in \overline{\mathcal{D}}$)
Hence, $stab_{SL_2(\mathbb{Z})}(z) = \{\pm I\}$ $\forall$ $z \in \mathcal{H}$ such that $z \notin \{i, \rho = e^{i\frac{\pi}{3}}, -\overline{\rho} = e^{i\frac{2\pi}{3}}\}$.

$\square$

Remark. The reader may verify via similar calculations that $stab_{SL_2(\mathbb{Z})}(i) = \langle S \rangle$, $stab_{SL_2(\mathbb{Z})}(\rho) = \langle S \cdot T \rangle$ and $stab_{SL_2(\mathbb{Z})}(-\overline{\rho}) = \langle T \cdot S \rangle$.

Definition 3.1. Let G be a group acting on a topological space X.
A **fundamental domain** F for the action of G is a (closed) subset, such that exactly one point of each orbit is contained in (the interior of) F.

Remark. Theorem (1) shows that $\overline{\mathcal{D}}$ is a fundamental domain for $SL_2(\mathbb{Z})$. The key geometric idea that motivated us to consider this region is that is that points in $\mathcal{H}$ tend to accumulate towards the real axis under the action of $SL_2(\mathbb{Z})$.

Fundamental domains of group actions are of interest to us because they provide a concrete description of the orbit space that one can visualize. In our case, the group $SL_2(\mathbb{Z})$ acts on $\mathcal{H}$, and instead of studying where each point in the upper-half complex plane goes under the action, we focus on the set $\overline{\mathcal{D}} \subset \mathcal{H}$. With $\overline{\mathcal{D}}$, we identify a particular region of $\mathcal{H}$ which gives us the set of orbit representatives for this action.

In other words, the theorem proves that there is a one-one correspondence between the set $\overline{\mathcal{D}}$ and the orbit space of $SL_2(\mathbb{Z})$, denoted by $SL_2(\mathbb{Z})\backslash\mathcal{H}$. Thus, we may consider the coset space as a topological space (see Theorem (2)).
Part (B) of the theorem ensures that every $z \in \mathcal{H}$ has a representative $\gamma \in SL_2(\mathbb{Z})$ which acts on it and shifts it inside $\overline{\mathcal{D}}$. Part (C) specifies that no two points in (the interior of) share the same orbit, and if two points do share an orbit, they lie on the boundary, ensuring uniqueness of representatives. Furthermore, the reader may verify the fact that all stabilizers of a group action are described by the stabilizers of elements in a fundamental domain[1], which is accomplished by part (D) and the remark given after the theorem.

The fundamental domain and its distinct translates are known as *ideal triangles* - they have two endpoints in $\mathcal{H}$ and one outside $\mathcal{H}$ (either a rational number on the real axis or $i\,\infty$), and they are each bounded by three sides. Ideal triangles increase indefinitely in number as we get closer to the real axis and only intersect along boundary curves.

Note. Every ideal triangle is a fundamental domain for the given action on $\mathcal{H}$. It is interesting to see that even though $\overline{\mathcal{D}}$ itself is unbounded, several of its translates are bounded. Thus, there exist both compact and non-compact fundamental domains for this action.
The image given below displays $\overline{\mathcal{D}}$ in blue, along with its translates by several different matrices from the full modular group.

---

[1]**Hint**: Remember that $\overline{\mathcal{D}}$ may be identified with $SL_2(\mathbb{Z})\backslash\mathcal{H}$ and if $s^{'} = g \cdot s$ for some $g \in G$, then $stab_G(s^{'}) = g \cdot stab_G(s) \cdot g^{-1}$.

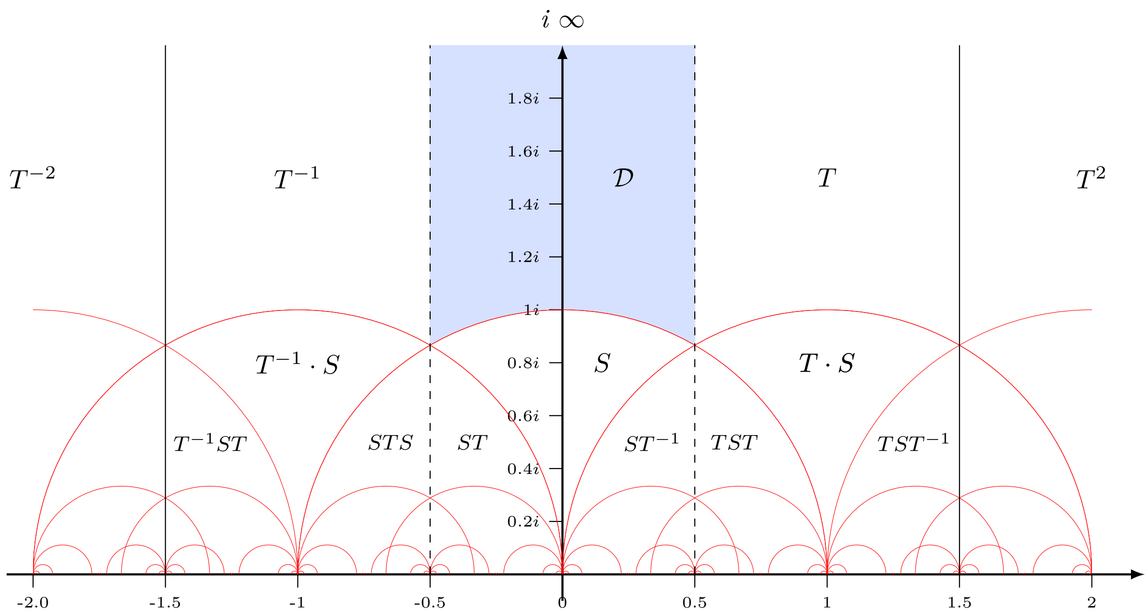

Figure 3. A geometric view

4. Generalisation

We now present a general version of Theorem (1) without proof:

Theorem 2. *Let G act topologically on a space X via a transitive action. If X is a locally compact Hausdorff space, and G is a locally compact group with countable basis, then for any given $x \in X$, the coset space $G\backslash G_x$ is homeomorphic to X under the correspondence $g \cdot G_x \to g \cdot x$.*

This theorem plays a crucial role in the higher dimensional theory of modular forms, which is beyond the scope of this paper. The key idea is that the transitivity of the action allows us to view X as a coset space for any $x \in X$. Thus, functions on X can be thought of as functions on G which are constant on G_x, thereby allowing an alternate perspective for study.
The advantage of this is that since G is a topological group, the coset space $G\backslash G_x$ inherits the quotient topology from the topology of G. Hence, with this action, we are identifying the topology of X with that of G, allowing us to go from an abstract topological space X, which had no visible connection to G except that G acted on it, to a coset space of a known group.
With respect to the action of $SL_2(\mathbb{R})$, the reader may verify that $\mathcal{H}$ is homeomorphic to $SL_2(\mathbb{R})\backslash stab(i) = SL_2(\mathbb{R})\backslash\langle S\rangle$. [2]

[2]Hint: Consider the map $\gamma \cdot stab(i)x \to \gamma \cdot i$.

5. Conclusion

In this paper, we have briefly explored the sagacious idea of identifying an abstract topological space with a coset space via a concrete example. This idea is at the heart of the Langlands program, a modern mathematics collective project that aims to establish connections between harmonic analysis and analytic number theory - two distinct fields of mathematics - via automorphic forms and Galois groups. Some refer to it as a kind of grand unified theory of mathematics.

Aside from that, the idea is charming in its own right. It allows to jump back and forth between abstraction and concreteness, and brings out an unexpectedly beautiful connection between groups and topological spaces: a magnificent display of the profundity and allure that is synonymous with mathematics.

References

[1] M. Ram Murty, Michael Dewar, Hester Graves, *Problems in the Theory of Modular Forms*, Springer Singapore, 2016. ISBN 9789811026515.
[2] Article. Fundamental domains for $SL_2(\mathbb{Z})$ and Γ_θ.
[3] Article. $SL_2(\mathbb{Z})$.
[4] Article. The Action of $SL_2(\mathbb{Z})$ on the Upper-half Complex Plane