UNIVERSITY OF CHICAGO REU 2007: GROUPS OF 2×2 MATRICES

COURSE BY URI BADER
JULY 16-27, 2007
NOTES TYPED BY SPENCER DOWDALL

1. Linear Transformations

Let K be a field (such as $\mathbb{R}$, $\mathbb{C}$, $\ldots$), and let $V = K^2 = \{\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right) : s,t \in K\} = \{se_1 + te_2 : s,t \in K\}$ (where $e_1 = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), e_2 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$).

Definition 1.1. A *linear transformation* is a map $T : V \to V$ such that

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

for all $x, y \in V$ and all $\alpha, \beta \in K$.

Claim 1.2. *T is determined by the four scalars $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ where $T(e_1) = ae_1 + ce_2$, $T(e_2) = be_1 + de_2$*

Proof. For the vector $\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right) = se_1 + te_2 \in V$, we have

$$\begin{aligned} T\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right) &= T(se_1 + te_2) = sT(e_1) + tT(e_2) = s(ae_1 + ce_2) + t(be_1 + de_2) \\ &= (as + bt)e_1 + (cs + dt)e_2 = \left(\begin{smallmatrix} as+bt \\ cs+dt \end{smallmatrix}\right) =: \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} s \\ t \end{smallmatrix}\right) \end{aligned}$$

$\square$

Matrix mutliplication is given by the composition rule: If $T : V \to V$ and $T' : V \to V$ are linear transformations corresponding to $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$, then $T \circ T'$ is the linear transformation corresponding to

$$\left(\begin{smallmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{smallmatrix}\right) =: \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$$

In particular, matrix multiplication is associative and non-commutative. (Why? Because we know this for linear transformations). Also, if $0 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is the zero matrix and $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ is the identity matrix, then

$$\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) \quad \text{and} \quad \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$$

Definition 1.3. A matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is *invertible* if there exists a matrix $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ such that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) = I$.

Claim 1.4. *The matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invertible if and only if* $\det\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = ad - bc \neq 0$

Proof. Firstly, if $ad - bc \neq 0$ then one easily checks that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^{-1} = \frac{1}{ad-bd}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$ so that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is indeed invertible. Conversely, if $ad - bc = 0$ then $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} d \\ -c \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$. Thus, if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$, then

$$\left(\begin{smallmatrix} d \\ -c \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} d \\ -c \end{smallmatrix}\right) = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} d \\ -c \end{smallmatrix}\right) = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \implies c, d = 0$$

Similarly, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} -b \\ a \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) \implies a, b = 0$. This is a contradiction since $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is not invertible. ☐

Exercise. Show that $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invertible as a matrix if and only if the corresponding linear transformation $T : V \to V$ is one-to-one and onto.

Definition 1.5. The above shows that the set of invertible matrices over K forms a group. This group is called the *General Linear Group* and is denoted by $GL_2(K)$.

Remark. $GL_2(\mathbb{R}) \subset \mathcal{M}_{2\times 2}(\mathbb{R}) \simeq \mathbb{R}^4$ ($\mathcal{M}_{2\times 2}(K)$ is the set of 2×2 matrices over K). The compliment of this subset is the inverse image of $\{0\}$ under the determinant map $\det : \mathcal{M}_{2\times 2}(\mathbb{R}) \to \mathbb{R}$; so $GL_2(\mathbb{R}) = \det^{-1}(\mathbb{R} - \{0\})$ which is an open set. Thus $GL_2(\mathbb{R})$ is not only an algebraic object (group) but is also a topological and geometrical object.

2. The Projective Line

Definition 2.1. Let $\mathbb{P}(V)$ denote the set of all linear lines in V. (A linear line in V is a straight line through the origin, i.e. a linear subspace of dimension 1; such lines should not be confused with affine lines which need not pass through the origin). Thus

$$\mathbb{P}(V) = \{\text{all linear lines in } V\} = \{\text{all 1-dimensional linear subspaces of } V\}$$

To each nonzero vector we may associate the unique line containing it and the origin, thus we have a map $V - \{0\} \to \mathbb{P}(V)$, $x \mapsto \{\lambda x : \lambda \in K\}$. In this way, two nonzero vectors define the same line iff they are scalar multiplies of eachother; whence $\mathbb{P}(V) \simeq V - \{0\}/\sim$ where $u \sim v$ if there exists a scalar $\lambda \in K$ such that $\lambda u = v$ (this is obviously an equivalence relation).

Definition 2.2. $\mathbb{P}^1(K) := \mathbb{P}(K^2)$ is called the *projective line over* K. (It is called a line because it is (in some sense) a one-dimensional object. In higher deminsions we set $\mathbb{P}^n(K) = \mathbb{P}(K^{n+1})$).

If $v \in V - \{0\}$, then let $[v] = \{u \in V - \{0\} : u \sim v\}$ be the equivalence class of all nonzero vectors equivalent to v (i.e., all nonzero vectors on the same linear line as v). Then

$$\mathbb{P}^1(K) = \{[\begin{smallmatrix} s \\ t \end{smallmatrix}] : s, t \in K\} = \{[\begin{smallmatrix} 1 \\ t \end{smallmatrix}] : t \in K\} \cup \{[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}]\} \cong K \cup \{\infty\} = \hat{K}$$

(think "one-point compactification").

Exercise. Convince yourself that $\mathbb{P}^1(\mathbb{C}) = \hat{\mathbb{C}}$ is homeomorphic (that is, topologically equivalent) to the two-sphere $\mathbb{S}^2$.

Note that the projective line is homogeneous: it looks the same in every direction (the point $\{\infty\}$ is not special). Observe that the action of $GL_2(K)$ on K^2 preserves lines (i.e., it sends lines to lines). Therefore this action decsends to an action of $GL_2(K)$ on $\mathbb{P}^1(K)$.

Claim 2.3. *The action of $GL_2(K)$ on $\mathbb{P}^1(K)$ is three-transitive (but not four-transitive), that is, every (ordered) triple of distinct lines can be mapped to any other (ordered) triple of disinct lines by some element of $GL_2(K)$.*

The proof of this is given below, but try to prove it yourself first as an exercise. (Hint: consider the triple $0, 1, \infty \in \hat{K} = \mathbb{P}^1(K)$. Show that every triple of distinct lines can be mapped to it.)

Remark. The action of $GL_2(K)$ on $\mathbb{P}^1(K)$ has a kernel. (If a group G acts on the set X, then the *kernel* of the action is the set of those group elements $g \in G$ which act trivially on X.) The kernel of this action is exactly the subgroup of scalar matrices:

$$\text{kernel} = \{ \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right) : \lambda \in K\} = \text{Scalars}$$

(note that this subgroup is exactly the center of $GL_2(K)$).

Since the scalar matrices are the kernel of the action of $GL_2(K)$ on $\mathbb{P}^1(K)$, the action factors through an action of the quotient group $PGL_2(K) := GL_2(K)/\text{Scalars}$ on $\mathbb{P}^1(K)$ (note that scalar matrices are a normal subgroup of $GL_2(K)$). In fact, $PGL_2(K)$ acts *simply transitively* on the set of distinct ordered triples in $\mathbb{P}^1(K)$, that is, for every two ordered triples of distinct points in $\mathbb{P}^1(K)$ there is a unique element in $PGL_2(K)$ that maps the one to the other.

Exercise. These remarks are justified below, but you should try to prove them yourself first.

3. Möbius transformations: another point of view

Recall that $\mathbb{P}^1(K) = K \cup \{\infty\}$. In coordinates ($z \leftrightarrow \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right]$, $\infty \leftrightarrow \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$) the action of $GL_2(K)$ on $\mathbb{P}^1(K)$ is given by

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} az+b \\ cz+d \end{smallmatrix}\right] = \begin{bmatrix} \frac{az+b}{cz+d} \\ 1 \end{bmatrix}$$

On $K \subset \mathbb{P}^1(K)$ this is the mapping $z \mapsto \frac{az+b}{cz+d}$ (such mappings are called *Möbius transformations*). More precisely, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL_2(K)$ acts on $\mathbb{P}^1(K)$ by

$$z \mapsto \begin{cases} \infty & cz + d = 0 \\ \frac{a}{c} & z = \infty, c \neq 0 \\ \infty & z = \infty, c = 0 \\ \frac{az+b}{cz+d} & \text{else} \end{cases}$$

Exercise. Verify that composition of Möbius transformations corresponds to matrix multiplication.

Remark. If $\text{Möb}(K)$ is the group of Möbious transformations of K, then $\text{Möb}(K) \simeq PGL_2(K) \simeq \text{Gal}(K(z), K)$

We now give a proof the of Claim 2.3:

Proof. (of 2.3) We will show that for every triple of lines $L_1, L_2, L_3 \in \mathbb{P}^1(K)$, we can find a $g \in GL_2(K)$ such that $g\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \in L_1$, $g\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \in L_2$, and $g\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \in L_3$. Pick $u_i \in L_i$. $\{u_1, u_2\}$ is a basis, hence $u_3 = \alpha u_1 + \beta u_2$ for some $\alpha, \beta \in K$. Set $v_1 = \alpha u_1$, $v_2 = \beta u_2$, $v_3 = u_3 = \alpha u_1 + \beta u_2 = v_1 + v_2$. Let $g \in GL_2(K)$ be the unique element sending $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \mapsto v_1$ and $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \mapsto v_2$ (if in coordinates $v_i = \left(\begin{smallmatrix} v_{ix} \\ v_{iy} \end{smallmatrix}\right)$, then take $g = \left(\begin{smallmatrix} v_{1x} & v_{2x} \\ v_{1y} & v_{2y} \end{smallmatrix}\right)$). Then g satisfies $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \mapsto v_1 + v_2 = v_3$. Hence it will satisfy $0 \mapsto L_1$, $\infty \mapsto L_2$, and $1 \mapsto L_3$. □

We now justify the remark made earlier that the action of $PGL_2(K)$ on the set of ordered triples of distinct points in $\mathbb{P}^1(K)$ is simply transitive.

If a group G acts on a set X then the *stabilizer* of a point $x \in X$ is the set $\text{Stab}(x) = \{g \in G : gx = x\}$.

Claim 3.1. *Suppose that $x, y \in X$ are such that $y = gx$ for some $g \in G$ (such x and y are said to lie in the same* orbit*), Then*

$$Stab(y) = \{ghg^{-1} : h \in Stab(x)\} =: Stab(x)^g$$

Proof. If $h \in \mathrm{Stab}(x)$ then $(ghg^{-1})y = ghg^{-1}gx = gx = y$ so that $ghg^{-1} \in \mathrm{Stab}(y)$. Conversely, if $\gamma \in \mathrm{Stab}(y)$ then $g^{-1}\gamma g \in \mathrm{Stab}(x)$ so that $\gamma = g(g^{-1}\gamma g)g^{-1} \in \mathrm{Stab}(x)^g$. □

In the action of $GL_2(K)$ on ordered triples of distinct points in $\mathbb{P}^1(K)$, the stabilizer of $(0, 1, \infty)$ is the subgroup of scalar matrices. If (L_1, L_2, L_3) is any ordered triple, then $(L_1, L_2, L_3) = g(0, 1, \infty)$ for some $g \in GL_2(K)$. Whence 3.1 implies that $\mathrm{Stab}(L_1, L_2, L_3) = \{ghg^{-1} : h \in \mathrm{Scalars}\} = \mathrm{Scalars}$. Since the stabilzer of every triple is exactly the subgroup of scalar matrices, the quotient group $PGL_2(K)$ indeed acts simply transitively. In fact, there is a bijection of sets: $PGL_2(K) \simeq \{(\text{distint, ordered) triples in } \mathbb{P}^1(K)\}$.

4. Cross-Ratio

Observe that the mapping $z \mapsto \frac{(z-z_1)(z_2-z_3)}{(z-z_3)(z_2-z_1)}$ is a Möbius transformation that sends $(z_1, z_2, z_3) \mapsto (0, 1, \infty)$ (in our coordinates $z \leftrightarrow \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right]$ and $\infty \leftrightarrow \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$).

Exercise. Write this in matrix form

Definition 4.1. (first definition) For distinct points $w, x, y, z \in \mathbb{P}^1(K)$, let $g \in GL_2(K)$ be such that $(gw, gx, gy) = (0, 1, \infty)$. The *cross-ratio* of w, x, y, z is defined to be the element $(w, x, y, z) := gz \in \mathbb{P}^1(K)$. Thus the corss-ratio is a mapping $CR : \{\text{distinct 4-tuples in } \mathbb{P}^1(K)\} \to \mathbb{P}^1(K)$. Observe that the cross-ratio is $GL_2(K)$-invariant: for $h \in GL_2(K)$, $(hw, hx, hy, hz) = (w, x, y, z)$.

Definition 4.2. (second definition: in coordinates) Representing elements of $\mathbb{P}^1(K)$ by elements in K (via $z \leftrightarrow \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right]$), we define a *cross-ratio* $CR : \{\text{distinct 4-tuples in } K\} \to K$ by $(z_1, z_2, z_3, z_4) \mapsto z$, where $(\left[\begin{smallmatrix} z_1 \\ 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} z_2 \\ 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} z_3 \\ 1 \end{smallmatrix}\right], \left[\begin{smallmatrix} z_4 \\ 1 \end{smallmatrix}\right]) = \left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right]$. This definition leads to an explicit formula:

$$(z_1, z_2, z_3, z_4) = \frac{(z_4 - z_1)(z_2 - z_3)}{(z_4 - z_3)(z_2 - z_1)}$$

Corollary 4.3. *The formula is invariant under Möbius transformations.*

Proof. Exercise. □

Theorem 4.4. *A self-map of $\mathbb{P}^1(K)$ that preserves the cross-ratio is from $PGL_2(K)$ (i.e. a Möbius transformation).*

Proof. Exercise. □

Claim 4.5. *Every Möbius transformation can be written as a composition of transformations of the following forms:*

$$z \mapsto az, \qquad z \mapsto z + b, \qquad z \mapsto \frac{-1}{z}$$

Proof. Consdier $z \mapsto \frac{az+b}{cz+d}$. If $c = 0$, then the map is

$$z \mapsto \tfrac{az+b}{d} = \tfrac{a}{d}z + \tfrac{b}{d} = (z \mapsto z + \tfrac{b}{d}) \circ (z \mapsto \tfrac{a}{d}z)$$

If $c \neq 0$, then

$$\tfrac{az+b}{cz+d} = \frac{\frac{a}{c}(cz+d)+(b-\frac{ad}{c})}{cz+d} = (b - \tfrac{ad}{c})\tfrac{1}{cz+d} + \tfrac{a}{c}$$

Therefore

$$z \mapsto \tfrac{az+b}{cz+d} = (z \mapsto z + \tfrac{a}{c}) \circ (z \mapsto (b - \tfrac{ad}{c})z) \circ (z \mapsto \tfrac{1}{z}) \circ (z \mapsto z + d) \circ (z \mapsto cz)$$

$\square$

Corollary 4.6. *This decomposistion of $PGL_2(K)$ can be used to give a decomposition of $GL_2(K)$, known as the* Bruhat decomposition of matrices.

When $K = \mathbb{C}$, the complex projective line $\mathbb{P}^1(\mathbb{C}) = \hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\} \simeq \mathbb{S}^2$ is called the *Riemann Sphere*. In this case the transformation $z \mapsto z + b$ is a translation by $b \in \mathbb{C}$, the transformation $z \mapsto \lambda z$ is scaling by $|\lambda|$ and a rotation by α where $\lambda = |\lambda| e^{i\alpha}$, and the transformation $z \mapsto \frac{-1}{z}$ is an inversion and reflection in the unit circle.

Definition 4.7. A *generalized circle* in $\hat{\mathbb{C}}$ is a Euclidean circle or line in $\mathbb{C}$. (Lines in $\mathbb{C}$ are thought of as circles through $\{\infty\} \in \hat{\mathbb{C}}$.) Therefore, any three distinct points in $\hat{\mathbb{C}}$ determine a unique generalized circle.

Theorem 4.8. *Let $w, x, y, z \in \mathbb{P}^1(\mathbb{C})$ be distinct. Then for their corss-ratio we have $(w, x, y, z) \in \mathbb{R}$ iff the four points w, x, y, z all lie on a generalized circle (i.e., iff z lies on the generalized circle determined by w, x, y).*

Proof. Exercise. (Hint: first prove the case $w, x, y \in \mathbb{P}^1(\mathbb{R})$, then generalize by 3-transitivity.) $\square$

Corollary 4.9. *Möbius transfromations take generalized circles to generalized circles. Moreover, Möb$(\mathbb{C})$ acts transitively on the set of generalized circles.*

Proof. This follows from 3-transitivity and Thm 4.8 since the cross-ratio is preserved by Möbius transformations. $\square$

Question. What is the stabilizer of a circle? (recall that since the action is transitive, the stabilizers of different circles will be conjugate subgroups in $Mob(\mathbb{C})$; thus it is enough to know the stabilizer of one circle).

Claim 4.10. *The stabilizer of $\mathbb{P}^1(\mathbb{R}) \subset \mathbb{P}^1(\mathbb{C})$ is exactly $PGL_2(\mathbb{R})$.*

Proof. It is clear that $PGL_2(\mathbb{R})$ stabilizes the circle $\mathbb{P}^1(\mathbb{R}) \subset \mathbb{P}^1(\mathbb{C})$. The fact that this is the whole stabilizer follows from 3-transitivity: if $g \in PGL_2(\mathbb{C})$ stabilizes $\mathbb{P}^1(\mathbb{R})$, then $g(0), g(1), g(\infty) \in \mathbb{P}^1(\mathbb{R})$. By 3-transitivity of the action of $PGL_2(\mathbb{R})$ on $\mathbb{P}^1(\mathbb{R})$, there is a unque $h \in PGL_2(\mathbb{R})$ with $h(0, 1, \infty) = (g(0), g(1), g(\infty))$. Since $PGL_2(\mathbb{C})$ acts *simply* 3-transitively on $\mathbb{P}^1(C)$, the the fact that g and $h \in PGL_2(\mathbb{R}) \subset PGL_2(\mathbb{C})$ agree on the three points 0, 1, and ∞ implies that they agree everywhere, whence $g = h \in PGL_2(\mathbb{R})$. $\square$

Considering the circle $\mathbb{P}^1(\mathbb{R}) \subset \mathbb{P}^1(\mathbb{C}) = \hat{C} \cong \mathbb{S}^2$ as the equator, the upper and lower hemispheres are the sets:

$$\text{upper hemisphere} = U = \{z \in \mathbb{C} : \text{Im} z > 0\}$$
$$\text{lower hemisphere} = L = \{z \in \mathbb{C} : \text{Im} z < 0\}$$

(so the points $\{i\}$ and $\{-i\}$ are the north and south poles).

Corollary 4.11. *$PGL_2(\mathbb{R})$ preserves the decomposition $\mathbb{P}^1(\mathbb{C}) = \mathbb{P}^1(\mathbb{R}) \cup U \cup L$.*

Question. What is the stabilizer of the upper-hemisphere in $PGL_2(\mathbb{R})$?

Answer. By the corollary, it suffices to consider whether the image of $i \in U$ lies in U or L. If $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in PGL_2(\mathbb{R})$, then

$$\operatorname{Im}\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) i\right) = \operatorname{Im}\left(\frac{ai+b}{ci+d}\right) = \operatorname{Im}\left(\frac{(ai+b)(d-ci)}{c^2+d^2}\right) = \frac{\det\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}{c^2+d^2} \in U \iff \det\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) > 0$$

Whence $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in PGL_2(\mathbb{R})$ stabilizes the upper-hemisphere iff $\det\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) > 0$.

Definition 4.12. Let $GL_2^+(\mathbb{R}) := \{g \in GL_2(\mathbb{R}) : \det g > 0\}$ be the subgroup of invertible matrices with positve determinant. This is an index-2 subgroup of $GL_2(\mathbb{R})$. The *Special Linear group* (over $\mathbb{R}$) is the subgroup $SL_2(\mathbb{R}) := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL_2(\mathbb{R}) : \det\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = 1\}$ of invertible matrices with determinant equal to 1. Whence we have the inclusion of subgroups $SL_2(\mathbb{R}) \leqslant GL_2^+(\mathbb{R}) \leqslant GL_2(\mathbb{R})$. The images of these subgroups under the natural quotient map $GL_2(\mathbb{R}) \to PGL_2(\mathbb{R})$ are denoted by $PSL_2(\mathbb{R})$ and $PGL_2^+(\mathbb{R})$.

Exercise. Show that $PSL_2(\mathbb{R}) = PGL_2^+(\mathbb{R})$ and that $PGL_2^+(\mathbb{R})$ is an index-2 subgroup of $PGL_2(\mathbb{R})$. Whence we have the diagram

$$\begin{tikzcd}
SL_2(\mathbb{R}) \arrow[d, "q"] & \leqslant & GL_2^+(\mathbb{R}) \arrow[d, "q"] & \overset{\text{index-2}}{\leqslant} & GL_2(\mathbb{R}) \arrow[d, "q"] \\
PSL_2(\mathbb{R}) & = & PGL_2^+(\mathbb{R}) & \overset{\text{index-2}}{\leqslant} & PGL_2(\mathbb{R})
\end{tikzcd}$$

Observations about the action of $PSL_2(\mathbb{R})$:

(i) $PSL_2(\mathbb{R})$ acts on the upper hemisphere $\{z : \operatorname{Im}(z) > 0\}$ and on the circle $\mathbb{P}^1(\mathbb{R})$.

(ii) The action of $PSL_2(\mathbb{R})$ is 2-transitive on $\mathbb{P}^1(\mathbb{R})$. (For two lines $L_1, L_2 \in \mathbb{P}^1(\mathbb{R})$, choose two nonzero vectors $v_i \in L_i$. Rescale v_1 so that the matrix having v_1 and v_2 as its columns has determinant 1. Then this matrix sends e_i to v_i and hence $\infty \mapsto L_1$, $0 \mapsto L_2$.

(iii) The action of $PSL_2(\mathbb{R})$ is not 3-transitive on $\mathbb{P}^1(\mathbb{R})$. Rather, the action on (distinct, ordered) triples has two orbits. (Given three points $x, y, z \in \mathbb{P}^1(\mathbb{R})$, one may map x and y to 0 and ∞ by 2-transitivity. An easy calculation shows that the stabilizer of the pair $(0, \infty)$ is diagonal matrices, i.e. matrices of the form $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{smallmatrix}\right)$. Such matrices act on the third point z by $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{smallmatrix}\right)\left[\begin{smallmatrix} z \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} \lambda z \\ \lambda^{-1} \end{smallmatrix}\right] = \left[\begin{smallmatrix} \lambda^2 z \\ 1 \end{smallmatrix}\right]$. Whence the third point z can be scaled by any positive real number, but its sign cannot be changed. Therefore the orbits of the tirples $(0, \infty, 1)$ and $(0, \infty, -1)$ are distinct.)

(iv) The action of $PSL_2(\mathbb{R})$ on the upper hemisphere $U = \{z : \operatorname{Im}(z) > 0\}$ is transitive. (By choosing the appropriate $a, b \in \mathbb{R}$, the matrix $\left(\begin{smallmatrix} a & ba^{-1} \\ 0 & a^{-1} \end{smallmatrix}\right)$ will map the point $\left[\begin{smallmatrix} i \\ 1 \end{smallmatrix}\right] \leftrightarrow i \in U$ to any desired point in U.)

(v) The stabilizer of i is $\left\{\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right) : \theta \in \mathbb{R}\right\} = SO_2(\mathbb{R})$ (exercise: verify this by a calculation). Identifying $i \in \hat{\mathbb{C}} \simeq \mathbb{S}^2$ with the north pole and $\mathbb{P}^1(\mathbb{R}) \subset \hat{\mathbb{C}}$ with the equator, such matrices rotate the sphere an angle of θ about the axis through i and $-i$. Be careful! This action of $SO_2(\mathbb{R})$ on $\hat{\mathbb{C}}$

by Möbius transformations is not linear and should not be confused with the usual (linear) action of $SO_2(\mathbb{R})$ on $\mathbb{R}^2 \simeq \mathbb{C}$ by rotation.

Observations about the action of Möb($\mathbb{C}$):

(i) The subgroup of Möb($\mathbb{C}$) that fixes the origin and the unit circle is $\{z \mapsto e^{2\pi i\alpha}z : \alpha \in [0,1]\}$. (Clearly such maps fix the origin and unit circle. Use 3-transitivity to show that all such maps are of this form.)

(ii) The action of the Möbius group is conformal (i.e. it preserves angles). (This follows from the complex-differentiability of Möbius transformations: the action of the differential is by scalar multiplication, which is conformal).

5. Hyperbolic Space

Definition 5.1. Consider the unit disk $\mathbb{D} = \{z \in \hat{\mathbb{C}} : |z| < 1\}$. A *hyperbolic line* in $\mathbb{D}$ is a (generalized) circle that intersects the unit circle $\mathbb{S}^1 = \partial\mathbb{D}$ in right angles.

By 3-transitivity, there is an element $g \in$ Möb($\mathbb{C}$) sending the circle $\mathbb{P}^1(\mathbb{R})$ to the unit circle $\partial\mathbb{D}$ and the upper-hemisphere $\{\mathrm{Im} z > 0\}$ to the unit disk $\mathbb{D}$. It follows that the conjugate subgroup $\mathcal{D} = g(PSL_2(\mathbb{R}))g^{-1}$ is the subgroup of Möb($\mathbb{C}$) which preserves $\mathbb{D}$. The action of $\mathcal{D}$ on $\mathbb{D}$ is therefore conformal, 2-transitive on $\partial\mathbb{D}$, 1-transitive on $\mathbb{D}$, and sends hyperbolic-lines to hyperbolic lines.

Claim 5.2. *For every two (distinct) points in $\overline{\mathbb{D}}$, there is a unique hyperbolic line containing them.*

Proof. This is immediate for a pair of antipodal points on $\partial\mathbb{D}$ as diameters are clearly hyperbolic lines. The claim thus follows for arbitrary pairs on $\partial\mathbb{D}$ by the 2-transitivity of the confomal action of $\mathcal{D}$ on $\partial\mathbb{D}$. Similarly, the claim is clear for pairs of the form $(0, y)$ with $y \in \overline{\mathbb{D}}$ as the diameter containing y again suffices. Therefore, the result holds for abitrary pairs by the 1-transitivity of the action of $\mathcal{D}$ on $\mathbb{D}$. ☐

Claim 5.3. *The stabilizer of $\partial\mathbb{D}$ in Möb($\mathbb{C}$) acts transitively on the set of hyperbolic lines.*

Proof. This follows from the 3-transitivity of the action of Möb($\mathbb{C}$) on the unit circle and the bijective correspondence between hyperbolic lines and pairs of points on the unit circle. ☐

We now adjust our ordering convention for the cross ratio: For $z_0, z_1, z_2, z_3 \in \hat{\mathbb{C}}$, set their cross ratio to be $(z_0, z_1, z_2, z_3) = \frac{(z_2-z_0)(z_3-z_1)}{(z_1-z_0)(z_3-z_2)}$. (Note: this is the standard convention for the cross-ratio; it is a permutation of the definition given earlier).

Definition 5.4. For $x, y \in \mathbb{D}$, let L be the unique hyberbolic line through x and y, and let s, t be the intersections of L with the unit circle $\partial\mathbb{D}$ (labeled so that the points appear in the order (s, x, y, t) on L). Define a metric on $\mathbb{D}$ by setting

$$d(x, y) = \frac{1}{2}\left|\log(s, x, y, t)\right|$$

(recall that by Thm 4.8 the cross-ratio of four points on the same circle is a real number).

Note that there are some symetries in the cross-ratio (s, x, y, t) obtained by permuting the vaiables. Exercise: Find all of the possibilities for the cross-ratio of four distinct points.

Remark. d is invariant under the subgroup $\mathcal{D}$ of Möb($\mathbb{C}$) preserving $\mathbb{D}$.

Observation. Let $z_0, z_1, \ldots, z_4$ be any 5 points on the same circle. Then

$$\begin{aligned}(z_0, z_1, z_3, z_4) =& \frac{(z_3 - z_0)(z_4 - z_1)}{(z_1 - z_0)(z_4 - z_3)} = \frac{(z_2 - z_0)(z_4 - z_1)}{(z_1 - z_0)(z_4 - z_2)} \cdot \frac{(z_3 - z_0)(z_4 - z_2)}{(z_2 - z_0)(z_4 - z_3)} \\ =& (z_0, z_1, z_2, z_4) \cdot (z_0, z_2, z_3, z_4)\end{aligned}$$

Corollary 5.5. *Hyperbolic lines are geodesics in the metric space* $(\mathbb{D}, d)$. *(A* geodesic *is an isometric embedding (copy) of* $\mathbb{R}$ *into a metric space.)*

Theorem 5.6. *d is a metric on $\mathbb{D}$.*

Proof. The only non-trivial thing to check is the triangle inequality. We sketch a proof of this: Let $x, y, z \in \mathbb{D}$ and consider the (hyperbolic) triangle $\Delta \subset \mathbb{D}$ having these points as vertices. Let a, b, c denote the lengths of these sides so that $d(x, z) = a$, $d(x, y) = b$, and $d(y, z) = c$, and let A, B, C be the cooresponding hyperbolic lines comprising the sides of Δ. We must show $a \leq b + c$.

Let P be the hyperbolic line through y and perpendicular to A. and let w be the intersection point of A and P. If $a' = d(x, w)$ and $a'' = d(w, z)$, then, by the above observation, we have $a = a' + a''$. Thus it suffices to prove $a' \leq b$ and $a'' \leq c$, so we have reduced to the case of a right triangle.

By applying a Möbius transformation, we may assume that $x = 0$ so that the hyperbolic lines A and B connconnecting x to w and y are diameters (i.e. straight Euclidian lines). Therefore, if $\tilde{a}$ and $\tilde{b}$ are the Euclidean distances from x to w and y, it suffices to check that $\tilde{a} \leq \tilde{b}$. The hyperbolic line P from w to y is now a circular arc perpendicular to the diameter A. Draw the striaght Euclidean line E through w perpendicular to A (so E is tangent to P at w). Since (looking at things Euclideanly) the line P is curving away from the origin x, we see that E intersects B closer to x than does P; call this Euclidean distnace $\tilde{b}'$ so that $\tilde{b}' \leq \tilde{b}$. Finally, the lines A, B, E form a Euclidean right triangle with hypotenuse of length $\tilde{b}'$. As $\tilde{a}$ is the length of one leg of this triangle, Euclidean geometry implies that $\tilde{a} \leq \tilde{b}' \leq \tilde{b}$ as desired. $\square$

Corollary 5.7. $(\mathbb{D}, d)$ *is a metric space and the subgroup* $\mathcal{D} \leqslant PGL_2(\mathbb{C})$ *preserving* $\mathbb{D}$ *acts on it by isometries. Indeed, the same holds for any other circle in* $\hat{\mathbb{C}}$, *in particular, for the upper half plane* $\{z : Im(z) > 0\}$ *on which* $PSL_2(\mathbb{R})$ *acts by isometries.*

Exercise. Show that $PSL_2(\mathbb{R}) \leqslant$ Isom(Upper half-plane) is an index-2 subgroup.

Definition 5.8. The metric space $(\mathbb{D}, d)$ is called the *hyperbolic plane*, and is denoted by $\mathbb{H}^2$ (sometimes $\mathbb{H}^2(\mathbb{R})$).

6. The group $PSL_2(\mathbb{Z})$

The goal in this section is to look at images of $PSL_2(\mathbb{Z})$ in $PSL_2(\mathbb{R})$ and understand the corresponding actions on $\mathbb{H}^2$. To this end, we study the space of unimodular lattices.

What are all of the basis of $\mathbb{C}$? Each such is given by a map (of abelian groups) $\mathbb{Z}^2 \xrightarrow{\varphi} \mathbb{C}$ such that $\varphi\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $\varphi\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ are not on the same line. Considering $\mathbb{C}$ as a 2-dimensional vector space, all such maps are linear, so we may set

$\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C}) = \{\varphi : \mathbb{Z}^2 \to \mathbb{C} : \det(\varphi) > 0\}$ to be the space of all such maps with positive determinant.

We will now mod out by some obvious equivalences on this space: There is an obvious left-action of $\mathbb{C}^*$ on $\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})$ by post-multiplication by a scalar: $\lambda \cdot \varphi = \lambda\varphi$. Similarly, there is a right-action of $\mathrm{Aut}^+(\mathbb{Z}^2) = SL_2(\mathbb{Z})$ on $\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})$ by pre-composition: $\varphi \cdot \alpha = \varphi \circ \alpha$. Since these two actions correspond to post- and pre-composition respectively ($\lambda \cdot \varphi \cdot \alpha = \mathbb{Z}^2 \xrightarrow{\alpha} \mathbb{Z}^2 \xrightarrow{\varphi} \mathbb{C} \xrightarrow{\lambda} \mathbb{C}$), the associativity of composition of maps implies that the two actions commute with each other

(Note: For R a ring,

$$GL_2(R) = \{2 \times 2 \text{ invertible matrices over } R\} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : ad - bc \in R^*\}$$

(where R^* denotes the set of units of R). Therefore $GL_2(\mathbb{Z}) \neq \mathcal{M}_{2\times 2} \cap GL_2(\mathbb{R})$; but rather $GL_2(\mathbb{Z}) = \mathcal{M}_{2\times 2} \cap \{\det = \pm 1\}$. Whence $SL_2(\mathbb{Z})$ is an index-2 subgroup of $GL_2(\mathbb{Z})$.)

There is a natural correspondence: $\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})/C^* \to \mathbb{H}^2$ (in upper half-plan model) given by $\varphi \mapsto \varphi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \varphi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)^{-1}$; this is clearly a bijection becaues for every $z \in \mathbb{H}^2$, there is a $\varphi \in \mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})$ with $\varphi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = z$ and $\varphi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = 1$. Since the actions of $SL_2(\mathbb{Z}$ and $\mathbb{C}^*$ on $\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})$ commute, we get an induced action of $SL_2(\mathbb{Z})$ on the quotient $\mathrm{Map}^+(\mathbb{Z}^2, C)/C^* \cong \mathbb{H}^2$. We also have the natural action of $SL_2(\mathbb{Z})$ on $\mathbb{H}^2$ by Möbius transformations. These two actions are related very simply:

Claim 6.1. *The action of $SL_2(\mathbb{Z})$ on the quotient $Map^+(\mathbb{Z}^2, \mathbb{C})/C^* \cong \mathbb{H}^2$ agrees with the usual action on $\mathbb{H}^2$ by Möbius transformations; i.e. the correspondence $Map^+(\mathbb{Z}^2, \mathbb{C})/C^* \to \mathbb{H}^2$ is an an $SL_2(\mathbb{Z})$-equivariant bijection.*

Remark. ${}_{GL_2(\mathbb{Z})}\backslash^{\mathrm{Map}^+(\mathbb{Z}^2, C)} \cong {}_{SL_2(\mathbb{Z})}\backslash^{\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})}$ can be identified with the space of lattices in $\mathbb{C}$, or with the space of "torus group quotients" of $\mathbb{C}$. In fact the quotient

$$MV = {}_{PSL_2(\mathbb{Z})}\backslash^{\mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})}/_{\mathbb{C}^*}$$

is the space of "complex structures on the torus"; also called the *modular variety*. Accordingly, $PSL_2(\mathbb{Z})$ is also called the *modular group*.

Question. What does the space MV look like?

To answer this, we will study/find a *fundamental domain* for the action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}^2$ (a *fundamental domain* is a subset of $\mathbb{H}^2$ that meets each (generic) orbit in exactly one point). (Note: Although $PSL_2(\mathbb{R})$ acts transitively on $\mathbb{H}^2$, the subgroup $PSL_2(\mathbb{Z}) \leqslant PSL_2(\mathbb{R})$ is discrete and therefore cannot act transitively.)

Claim 6.2. *Let $F = \{z \in \mathbb{H}^2 : |z| \geq 1\} \cap \{z \in \mathbb{H}^2 : -\frac{1}{2} \leq Im(z) \leq \frac{1}{2}\}$. Then F is a fundamental domain for the action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}^2$; more precisely, every orbit of $PSL_2(\mathbb{Z})$ in $\mathbb{H}^2$ meets F in exactly one interior point, exactly two boundary points, or the point i.*

Proof. (sketch) Consider a lattice $\Lambda = \varphi(\mathbb{Z}^2)$ for $\varphi \in \mathrm{Map}^+(\mathbb{Z}^2, \mathbb{C})$. Up to the action of $\mathbb{C}^*$, we may assume that $\varphi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = 1$. Up to the $SL_2(\mathbb{Z})$ action, we may assume that (i) $\varphi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ is of shortest length in $\varphi(\mathbb{Z}^2)$ and that (ii) $|\mathrm{Re}(\varphi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right))| \leq \frac{1}{2}$; whence $x = \varphi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \in F$. Since $\{\varphi\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \varphi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\}$ forms a basis of $\varphi(\mathbb{Z}^2)$, it follows that (iii) $\mathrm{Im}(x)$ is minimal in $\varphi(\mathbb{Z}^2)$. (Note that facts (i)-(iii) uniquely determine a vector v such that $\{1, v\}$ is a basis.) Therefore, to the lattice Λ we associate the unique point $x = \varphi\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \in \Lambda$ such that $x \in F$ and $\mathrm{Im}(x)$ is minimal in Λ. Conversely, given an interior point $y \in F$ we associate the lattice with basis $\{1, y\}$. Thus we have a bijection:

$$\begin{tikzcd} \text{lattices} \arrow[rr, bend left, "\text{the above argument}"] & & \text{(interior) points } x \in F \arrow[ll, bend left, "\text{lattice generated by } \{1,x\}"] \end{tikzcd}$$

However, this is not a bijection on ∂F where we have the identifications $z \leftrightarrow z+1$ for $z \in F \cap \{\mathrm{Re}(z) = -\frac{1}{2}\}$ (corresonding to the action by the group element $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in PSL_2(\mathbb{Z})$), and $z \leftrightarrow -\frac{1}{z}$ for $z \in F \cap \{|z| = 1\}$ (corresponding to the action by the group element $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \in PSL_2(\mathbb{Z})$). Making these identifications, we see that the modular variety (recall that $MV =$ the quotient space ${}_{PSL_2(\mathbb{Z})}\backslash^{\mathrm{Map}^+(\mathbb{Z}^2,\mathbb{C})}/_{\mathbb{C}^*}$) is shaped like a "triangular pillowcase" with cusps at i, $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, and ∞. $\square$

This picture of the modular variety should help us understand the algebraic structure of $PSL_2(\mathbb{Z})$.

7. The Structure of $PSL_2(\mathbb{Z})$

Consider the elements $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right), T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in PSL_2(\mathbb{Z})$ corresponding to the Möbius transformations $S(z) = -\frac{1}{z}$, and $T(z) = z + 1$.

Claim 7.1. *The elements S and T generate $PSL_2(\mathbb{Z})$.*

Proof. (sketch) Observe that the stabilizer of $2i \in F$ is trivial in $PSL_2(\mathbb{Z})$. (Exercise: show this. Idea: if $\frac{a(2i)+b}{c(2i)+d} = 2i$ with $a, b, c, c \in \mathbb{Z}$, then conclude $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{smallmatrix}\right)$.) Now, consider $g \in PSL_2(\mathbb{Z})$ and let $y = g(2i)$. The goal is to use S and T to send y into F, i.e. to find a word w in S and T such that $wy \in F$. Since $\mathrm{Orbit}_{2i} \cap F = \{2i\}$, we have that $2i = wy = (wg)2i$ which implies that $g = w^{-1}$ is a word in S and T since the stabilizer of $2i$ is trivial.

We now describe the algorithm to construct the word w:

(i) If $\mathrm{Im}(y) > 1$, then $T^n y \in F$ for some $n \in \mathbb{Z}$
(ii) If $\mathrm{Im}(y) < 1$, then we apply T^n to get either into F or into the open hemisphere $\{|z| < 1\}$. In the first case we are done, in the second, apply S to increase $\mathrm{Im}(y)$.
(iii) Repeat.

We need only check that $\mathrm{Im}(y)$ increases by a sufficient amount each time so that we always eventually obtain $\mathrm{Im}(y) > 1$ (i.e., so that the algorithm terminates). This is made more precise by the following compactness arbument: If I start at $y = a + bi$, then when I apply S and T as in the above algorithm, I always end up in $\{z : |z| \leq \frac{1}{b}\} \cap \{z : \mathrm{Im}(z) \geq b\}$, which is a compact set. This compact set can only intersect finitely many copies of the fundamental domain (in the tiling of $\mathbb{H}^2$ by the images of F under S and T), so the algorithm must terminate. $\square$

Corollary 7.2. *Set $U = ST$. Then $PSL_2(\mathbb{Z})$ is generated by S and U.*

Observe that $S(z) = -\frac{1}{z} \leftrightarrow \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $U(z) = \frac{-1}{z+1} \leftrightarrow \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$. One easily checks that $S^2 = 1 = U^3$.

Consider the group G given by the presentation $G = \langle \tilde{S}, \tilde{U} : \tilde{S}^2 = \tilde{U}^3 = 1 \rangle$. Thus G is the (universal) group generated by the two elements $\tilde{S}$, $\tilde{U}$ under the relations $\tilde{S} = 1 = \tilde{U}^3$. (Note that this is the same as the free product: $G = C_2 * C_3$, where C_n is the cyclic group of order n.)

Such a group can be constructed as follows: Take the free group $F(\tilde{S}, \tilde{U})$ on the two symbols $\tilde{S}$ and $\tilde{U}$. That is

$$F(\tilde{S}, \tilde{U}) = \{\text{all reduced words in } \tilde{S}, \tilde{U}, \tilde{S}^{-1}, \tilde{U}^{-1}\}$$

(*reduced* means that we cancel all terms $\tilde{S}\tilde{S}^{-1}, \tilde{S}^{-1}\tilde{S}, \tilde{U}\tilde{U}^{-1}, \tilde{U}^{-1}\tilde{U}$). Note that this F is a functor from the category of sets to the category of groups. It is left adjoint to the forgetful functor U from the category of groups to the category of sets.

Let N be the minimal normal subgroup of $F(\tilde{S}, \tilde{U})$ containing the elements $\tilde{S}^2$ and $\tilde{U}^3$ (e.g, take N to be the intersection of all such subgroups). The presentation $\langle \tilde{S}, \tilde{U} | \tilde{S}^2 = \tilde{U}^3 = 1 \rangle$ is defined to be the quotient by N:

$$G = \langle \tilde{S}, \tilde{U} | \tilde{S}^2 = \tilde{U}^3 = 1 \rangle := F(\tilde{S}, \tilde{U})/N$$

The groups $F(\tilde{S}, \tilde{U})$ and G satisfy the following "universal" properties:

(i) For every group Γ, and any two elements $s, u \in \Gamma$, there exists a unique homomorphism $\Theta : F(\tilde{S}, \tilde{U}) \to \Gamma$ such that $\tilde{S} \mapsto s$ and $\tilde{U} \mapsto u$. Moreover, this map is onto iff Γ is generated by s and u.

(ii) If in Γ we have $s^2 = u^3 = 1$, then Θ factors through the quotient group G:

$$\begin{tikzcd} F(\tilde{S}, \tilde{U}) \arrow[d] \arrow[dr, "\Theta"] & \\ G = F(\tilde{S}, \tilde{U})/N \arrow[r, dashed, "\exists!\theta"] & \Gamma \end{tikzcd}$$

and gives a unique map $\theta : G \to \Gamma$ such that $\tilde{S} \mapsto s$ and $\tilde{U} \mapsto u$.

Corollary 7.3. *There is a unique (surjective) map* $\Psi : G = \langle \tilde{S}, \tilde{U} | \tilde{S}^2 = \tilde{U}^3 = 1 \rangle \to PSL_2(\mathbb{Z})$ *such that* $\Psi(\tilde{S}) = S$ *and* $\Psi(\tilde{U}) = U$.

Theorem 7.4. *The natrual map* $\Psi : G \to PSL_2(\mathbb{Z})$ *is an isomorphism.*

Proof. It remains to prove that Ψ is injective.

First, let's consider the action of $PLS_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{R})$. Consder $\mathbb{S}^1 \cong \mathbb{P}^1(\mathbb{R}) \subset \mathbb{P}^1(\mathbb{C}) \cong \mathbb{S}^2$. If we let $0, \infty \in \mathbb{P}^1(\mathbb{C})$ be the south and north poles of $\mathbb{S}^2$ respectively, then $\mathbb{P}^1(\mathbb{R})$ is a great meridian of the sphere. Then $\mathbb{P}^1(\mathbb{R}) - \{0, \infty\} = L' \cup R'$ where L' and R' are the two (open) circular arcs from 0 to ∞, labelled such that $-1 \in L'$ and $1 \in R'$ (i.e. the left and right halfs of the circle $\mathbb{P}^1(\mathbb{R})$). Let $L = \overline{L'} = L \cup \{0, \infty\}$ and, similarly, $R = \overline{R'} = R' \cup \{0, \infty\}$. We now consider the images of L and R under the maps S, U, and U^2:

$S(z) = -\frac{1}{z}$. Thus S maps $0 \mapsto \infty$, $\infty \mapsto 0$, $1 \mapsto -1$, and $-1 \mapsto 1$; whence $S(L) \subset R$. $U(z) = \frac{-1}{z+1}$. Thus U maps $0 \mapsto -1$, $\infty \mapsto 0$, $1 \mapsto -\frac{1}{2}$, and $-1 \mapsto \infty$; whence $U(R) \subset L$. Finally, $U^2(z) = \frac{z+1}{-z}$. Thus U^2 maps $0 \mapsto \infty$, $\infty \mapsto -1$, $1 \mapsto -2$, and $-1 \mapsto 0$; whence $U^2(R) \subset L$ as well.

The point is that S sends the left half to the right half, whereas U and U^2 send the right half to the left half. Therefore, in their action on $\mathbb{P}^1(\mathbb{R})$, we may imgine that S is playing ping-pong against the U's.

We now prove the claim: If Ψ is not injective, then there is an element $w \in G$ with $\Psi(w) = 1 \in PSL_2(\mathbb{Z})$. As $w \in \langle \tilde{S}, \tilde{U} | \tilde{S}^2 = \tilde{U}^3 = 1 \rangle$, w can be written as a reduced word in $\tilde{S}$, $\tilde{U}$, and $\tilde{U}^2 = \tilde{U}^{-1}$ in which no two $\tilde{U}$ terms next appear to each other. As $\Psi(gwg^{-1}) = 1$ for all $g \in G$, by conjugating by an appropriate $g \in G$, we may assume that w starts and ends with $\tilde{U}$ and $\tilde{U}^{-1}$ (i.e. that $\tilde{S}$ is neither the first nor last letter in w), hence $w = \tilde{U}^{\varepsilon_1}\tilde{S}\tilde{U}^{\varepsilon_2}\tilde{S}\cdots\tilde{S}\tilde{U}^{\varepsilon_n}$ for some n with $\varepsilon_i = \pm 1$ for each i. Then $U^{\varepsilon_1}SU^{\varepsilon_2}S\cdots SU^{\varepsilon_n} = \Psi(w) = 1$ acts as the identity on $\mathbb{P}^1(\mathbb{R})$. However, this contradicts the ping-pong game that S and the U's are playing against each other: since there are n U letters and $n-1$ S letters, we see that $\Psi(w)$ maps R into L. In particular have $1 = \mathrm{Id}_{\mathbb{P}^1(\mathbb{R})}(1) = (\Psi(w))(1) \in L$, which is a contradiction. Therefore, there is no such $w \in G$ and Ψ must be injective. $\square$

Remark. People take this ping-pong idea to great extents; there is even a whole abstract formulation and a "ping-pong lemma" and such. Indeed, by utilizing the ping-pong lemma, Jacques Tits proved the following:

Theorem 7.5. *(Tit's Alternative) Over any filed K, for any dimension n, and any group $\Gamma < GL_n(K)$, either Γ contains a copy of $\mathbb{F}_2$ (the free group on 2 generators), or Γ is virtualy solvable (i.e. contains a solvable subgroup of finite index).*

Exercise. Prove, using the same techniques, that $SL_2(\mathbb{Z}) \cong \langle a, b | a^4 = b^6 = 1 \rangle$.