

IDEAL CLASSES AND SL_2

KEITH CONRAD

1. INTRODUCTION

A standard group action in complex analysis is the action of $\mathrm{GL}_2(\mathbf{C})$ on the Riemann sphere $\mathbf{C} \cup \{\infty\}$ by linear fractional transformations (Möbius transformations):

$$(1.1) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d}.$$

We need to allow the value ∞ since $cz + d$ might be 0. (If that happens, $az + b \neq 0$ since $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible.) When $z = \infty$, the value of (1.1) is $a/c \in \mathbf{C} \cup \{\infty\}$.

It is easy to see this action of $\mathrm{GL}_2(\mathbf{C})$ on the Riemann sphere is transitive (that is, there is one orbit): for every $a \in \mathbf{C}$,

$$(1.2) \quad \begin{pmatrix} a & a-1 \\ 1 & 1 \end{pmatrix} \infty = a,$$

so the orbit of ∞ passes through all points. In fact, since $\begin{pmatrix} a & a-1 \\ 1 & 1 \end{pmatrix}$ has determinant 1, the action of $\mathrm{SL}_2(\mathbf{C})$ on $\mathbf{C} \cup \{\infty\}$ is transitive.

However, the action of $\mathrm{SL}_2(\mathbf{R})$ on the Riemann sphere is not transitive. The reason is the formula for imaginary parts under a real linear fractional transformation:

$$\mathrm{Im} \left(\frac{az + b}{cz + d} \right) = \frac{(ad - bc) \mathrm{Im}(z)}{|cz + d|^2}$$

when $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbf{R})$. Thus, z and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} z$ have the same imaginary part when $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has determinant 1. The action of $\mathrm{SL}_2(\mathbf{R})$ on the Riemann sphere has three orbits: $\mathbf{R} \cup \{\infty\}$, the upper half-plane $\mathfrak{h} = \{x + iy : y > 0\}$, and the lower half-plane. To see that the action of $\mathrm{SL}_2(\mathbf{R})$ on \mathfrak{h} is transitive, pick $x + iy$ with $y > 0$. Then

$$\begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix} i = x + iy,$$

and the matrix here is in $\mathrm{SL}_2(\mathbf{R})$. (This action of $\mathrm{SL}_2(\mathbf{R})$ on the upper half-plane is essentially one of the models for the isometries of the hyperbolic plane.)

The action (1.1) makes sense with \mathbf{C} replaced by any field K , and gives a transitive group action of $\mathrm{GL}_2(K)$ on the set $K \cup \{\infty\}$. Just as over the complex numbers, the formula (1.2) shows the action of $\mathrm{SL}_2(K)$ on $K \cup \{\infty\}$ is transitive.

Now take K to be a number field, and replace the group $\mathrm{SL}_2(K)$ with the subgroup $\mathrm{SL}_2(\mathcal{O}_K)$. We ask: how many orbits are there for the action of the group $\mathrm{SL}_2(\mathcal{O}_K)$ on $K \cup \{\infty\}$?

Theorem 1.1. *For a number field K , the number of orbits for $\mathrm{SL}_2(\mathcal{O}_K)$ on $K \cup \{\infty\}$ is the class number of K .*

Therefore there are finitely many orbits, and moreover this finiteness is a non-trivial statement!

In Section 2, we will prove $\mathrm{SL}_2(\mathcal{O}_K)$ acts transitively on $K \cup \{\infty\}$ if and only if K has class number 1. This is the simplest case of Theorem 1.1. As preparation for the general case, in Section 3 we will change our language from $K \cup \{\infty\}$ to the projective line over K , whose relevance (among other things) is that it removes the peculiar status of ∞ . (It seems useful to treat the special case of class number 1 without mentioning the projective line, if only to underscore what it is one is gaining by using the projective line in the general case.) In Section 4 we prove Theorem 1.1 in general. This theorem is particularly important for totally real K (in the context of Hilbert modular forms [1, pp. 36–38], [2, pp. 7–8]), but it holds for any number field K .

As a further illustration of the link between SL_2 and classical number theory, we show in an appendix that the Euclidean algorithm on \mathbf{Z} is more or less equivalent to the group $\mathrm{SL}_2(\mathbf{Z})$ being generated by the matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

The prerequisites we need about number fields are: in any number field all fractional ideals are invertible, and any fractional ideal has two generators. That only two generators are needed for fractional ideals in a number field appears as an exercise in several introductory algebraic number theory books, but it may seem like an isolated fact in such books (I thought so when I first saw it!). Its use in the proof of Theorem 1.1 shows it is not.

2. TRANSITIVITY AND CLASS NUMBER ONE

As an example of class number one, take $K = \mathbf{Q}$. We will show every rational number is in the $\mathrm{SL}_2(\mathbf{Z})$ -orbit of ∞ . Pick a rational number r , and write it in reduced form as $r = a/c$, so a and c are relatively prime integers. (If $r = 0$, use $a = 0$ and $c = 1$.) Since $(a, c) = 1$, we can solve the equation $ad - bc = 1$ in integers b and d , which means we get a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\mathrm{SL}_2(\mathbf{Z})$ whose first column is $\begin{pmatrix} a \\ c \end{pmatrix}$. This matrix sends ∞ to $a/c = r$.

Conversely, if we know by some independent means that the $\mathrm{SL}_2(\mathbf{Z})$ -action on $\mathbf{Q} \cup \{\infty\}$ is transitive, then for any rational number r we can find a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$ sending ∞ to r , so $r = a/c$. Since $ad - bc = 1$, a and c have no common factors, so we can write r as a ratio of relatively prime integers. Thus, the fact that the $\mathrm{SL}_2(\mathbf{Z})$ -action on $\mathbf{Q} \cup \{\infty\}$ is transitive is equivalent to the ability to write rational numbers in reduced form over \mathbf{Z} .

A similar argument shows the action of $\mathrm{SL}_2(\mathcal{O}_K)$ on $K \cup \{\infty\}$ is transitive if and only if every element of K can be written in ‘reduced form,’ i.e., as a ratio of relatively prime algebraic integers from \mathcal{O}_K .

Theorem 2.1. *Every element of K^\times has the form α/β where $(\alpha, \beta) = (1)$ in \mathcal{O}_K if and only if K has class number 1.*

Proof. If K has class number 1 then \mathcal{O}_K is a PID, so a UFD, so any ratio of nonzero elements of \mathcal{O}_K can be put in a reduced form.

Conversely, suppose each ratio of nonzero elements of \mathcal{O}_K can be put in reduced form. To show every ideal is principal, pick an ideal \mathfrak{a} . We may suppose $\mathfrak{a} \neq (0)$, so $\mathfrak{a} = (x, y)$ where x and y are in \mathcal{O}_K and neither is 0. By hypothesis we can write $x/y = \alpha/\beta$ where $(\alpha, \beta) = (1)$. Then $x\beta = y\alpha$, so $(x)(\beta) = (y)(\alpha)$. The ideals (α) and (β) are relatively prime, so $(\alpha)|(x)$. Thus $\alpha|x$, so $x = \alpha\gamma$ for some $\gamma \in \mathcal{O}_K$. Then $\alpha\gamma\beta = y\alpha$, so $y = \beta\gamma$. It follows that $\mathfrak{a} = (x, y) = (\alpha\gamma, \beta\gamma) = (\alpha, \beta)(\gamma) = (1)(\gamma) = (\gamma)$ is principal. \square

Thus, the number of orbits for $\mathrm{SL}_2(\mathcal{O}_K)$ on $K \cup \{\infty\}$ is 1 if and only if K has class number 1.

3. THE PROJECTIVE LINE

In this section, K is any field.

The set of numbers $K \cup \{\infty\}$ can be thought of as the possible slopes of different lines through the origin in K^2 . Rather than determine such lines by their slopes, we can determine such lines by naming a representative point (x, y) on the line, excluding $(0, 0)$ (which lies on all such lines). But we face the issue: when do two non-zero points (x, y) and (x', y') lie on the same line through the origin? Since a line through the origin is the set of scalar multiples of any non-zero point on that line, (x, y) and (x', y') lie on the same line through the origin when $(x', y') = \lambda(x, y)$ for some $\lambda \in K^\times$.

Definition 3.1. The *projective line* over K is the set of points in $K^2 - \{(0, 0)\}$ modulo scaling by K^\times . That is, we set $(x, y) \sim (x', y')$ if and only if there is some $\lambda \in K^\times$ such that $x' = \lambda x$ and $y' = \lambda y$. The equivalence classes for \sim form the projective line over K .

We denote the projective line over K by $\mathbf{P}^1(K)$. (Strictly speaking, the projective line over K is a richer geometric object than merely the set of equivalence classes $\mathbf{P}^1(K)$, but our definition will be adequate for our purposes.) The equivalence class of (x, y) in $\mathbf{P}^1(K)$ is denoted $[x, y]$ and is called a point of $\mathbf{P}^1(K)$. For instance, in $\mathbf{P}^1(\mathbf{R})$, $[2, 3] = [4, 8] = [1, 3/2]$. Provided $x \neq 0$, we have $[x, y] = [1, y/x]$, and $[1, a] = [1, b]$ if and only if $a = b$. We have $[0, y] = [x', y']$ if and only if $x' = 0$, and in this case $[0, y] = [0, 1]$. Thus, every point of $\mathbf{P}^1(K)$ equals $[1, y]$ for a unique $y \in K$ or is the point $[0, 1]$. By an analogous argument, every point of $\mathbf{P}^1(K)$ is $[x, 1]$ for a unique $x \in K$ or is the point $[1, 0]$. For the points $[x, y]$ with neither x nor y equal to 0, we can write them either as $[1, y/x]$ or $[x/y, 1]$. (To change between the two coordinates amounts to $t \leftrightarrow 1/t$ on K^\times .)

The passage from $[x, y]$ to the ratio y/x , with the exceptional case $x = 0$, corresponds to the idea of recovering a line's slope as a number in $K \cup \{\infty\}$. In other words, the correspondence between $\mathbf{P}^1(K)$ and $K \cup \{\infty\}$ comes about from

$$[x, y] \mapsto \begin{cases} y/x, & \text{if } x \neq 0, \\ \infty, & \text{if } x = 0. \end{cases}$$

Since $[x, y] = [x', y']$ if and only if (x, y) and (x', y') are non-zero scalar multiples, the ratio y/x (provided $x \neq 0$) is a well-defined number in terms of the point $[x, y]$ even though the coordinates x and y themselves are not uniquely determined from $[x, y]$.

We get another correspondence between $\mathbf{P}^1(K)$ and $K \cup \{\infty\}$ by associating $[x, y]$ to x/y or ∞ :

$$(3.1) \quad [x, y] \mapsto \begin{cases} x/y, & \text{if } y \neq 0, \\ \infty, & \text{if } y = 0. \end{cases}$$

Now we describe an action of $GL_2(K)$ on $\mathbf{P}^1(K)$ which corresponds to (1.1). For an invertible matrix $A \in GL_2(K)$, and a non-zero vector $v \in K^2$, the product Av is non-zero and

$$A(\lambda v) = \lambda Av$$

for any $\lambda \in K$. Therefore A sends all points on one line through the origin in K^2 to all points on another line through the origin in K^2 . (No such line collapses under A since A is invertible.) This means the usual action of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ on column vectors in K^2 lets us

define A as a transformation of $\mathbf{P}^1(K)$:

$$(3.2) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} \rightsquigarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} [x, y] := [ax+by, cx+dy].$$

When $y \neq 0$, let $z = x/y$. Then the element of $K \cup \{\infty\}$ corresponding by (3.1) to $[ax+by, cx+dy]$ is

$$\frac{ax+by}{cx+dy} = \frac{az+b}{cz+d},$$

interpreted as ∞ when the denominator is 0. Writing $[x, y]$ as $[z, 1]$, we see that the action of $\mathrm{GL}_2(K)$ on $K \cup \{\infty\}$ given by (1.1), with the peculiar role of ∞ , is the same as the action of $\mathrm{GL}_2(K)$ on $\mathbf{P}^1(K)$ given by the right side of (3.2). And now, in $\mathbf{P}^1(K)$, there is no more mysterious ∞ . Everything is homogeneous.

4. ORBITS AND IDEAL CLASSES

For $x, y \in K$, not both zero, we write $[x, y]$ for a point in $\mathbf{P}^1(K)$ and $(x, y) = x\mathcal{O}_K + y\mathcal{O}_K$ for a fractional ideal. Since every fractional ideal has two generators, (x, y) is a completely general fractional ideal as x and y vary (avoiding $x = y = 0$).

Now we are ready to prove Theorem 1.1 in general.

Proof. Step 1: If $[x, y]$ and $[u, v]$ are in the same $\mathrm{SL}_2(\mathcal{O}_K)$ -orbit, then the fractional ideals (x, y) and (u, v) are in the same ideal class.

Being in the same orbit means

$$(4.1) \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda u \\ \lambda v \end{pmatrix}$$

for some $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathcal{O}_K)$ and $\lambda \in K^\times$. Thus

$$\begin{aligned} ax+by &= \lambda u, \\ cx+dy &= \lambda v, \end{aligned}$$

so we have an inclusion of \mathcal{O}_K -modules $(\lambda u, \lambda v) \subset (x, y)$. Multiplying both sides of (4.1) by the inverse $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \in \mathrm{SL}_2(\mathcal{O}_K)$ gives the reverse inclusion, so $(x, y) = (\lambda u, \lambda v) = \lambda(u, v)$.

As far as Step 1 is concerned, the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ could have been in $\mathrm{GL}_2(\mathcal{O}_K)$ rather than $\mathrm{SL}_2(\mathcal{O}_K)$.

Step 2: If (x, y) and (u, v) are in the same ideal class, then the points $[x, y]$ and $[u, v]$ in $\mathbf{P}^1(K)$ are in the same $\mathrm{SL}_2(\mathcal{O}_K)$ -orbit.

Write $(x, y) = \lambda(u, v) = (\lambda u, \lambda v)$ for some $\lambda \in K^\times$. We want to show $[x, y]$ and $[u, v]$ are in the same orbit of $\mathrm{SL}_2(\mathcal{O}_K)$. Since $[\lambda u, \lambda v] = [u, v]$ in $\mathbf{P}^1(K)$, we may take $\lambda = 1$, *i.e.*, we may assume the fractional ideals (x, y) and (u, v) are equal. In other words, we are aiming at a relation between pairs of generators for the same fractional ideal.

Let $\mathfrak{a} = (x, y)$. The inverse ideal \mathfrak{a}^{-1} has two generators, say $\mathfrak{a}^{-1} = (r, s)$. From the equation $(1) = (x, y)(r, s) = (xr, xs, yr, ys)$, there are $\alpha, \beta, \gamma, \delta \in \mathcal{O}_K$ such that

$$\begin{aligned} 1 &= \alpha xr + \beta xs + \gamma ry + \delta ys \\ &= (\alpha r + \beta s)x + (\gamma r + \delta s)y. \end{aligned}$$

Note $y' := \alpha r + \beta s$ and $x' := -(\gamma r + \delta s)$ are in \mathfrak{a}^{-1} . Thus, we can form a matrix $M = \begin{pmatrix} x & x' \\ y & y' \end{pmatrix}$ in $M_2(K)$ with determinant $xy' - yx' = 1$ where the second column has entries in \mathfrak{a}^{-1} .

Similarly, there is a matrix $N = \begin{pmatrix} u & u' \\ v & v' \end{pmatrix} \in \mathrm{M}_2(K)$ with determinant 1 where the second column has entries in \mathfrak{a}^{-1} .

Since $x', y', u', v' \in \mathfrak{a}^{-1}$, the product

$$MN^{-1} = \begin{pmatrix} x & x' \\ y & y' \end{pmatrix} \begin{pmatrix} v' & -u' \\ -v & u \end{pmatrix}$$

has determinant 1 and entries in \mathcal{O}_K . Therefore MN^{-1} is in $\mathrm{SL}_2(K) \cap \mathrm{M}_2(\mathcal{O}_K) = \mathrm{SL}_2(\mathcal{O}_K)$.

As $M[1, 0] = [x, y]$ and $N[1, 0] = [u, v]$, $MN^{-1}[u, v] = M[1, 0] = [x, y]$, so $[x, y]$ and $[u, v]$ are in the same $\mathrm{SL}_2(\mathcal{O}_K)$ -orbit of $\mathbf{P}^1(K)$. \square

Our bijection between $\mathrm{SL}_2(\mathcal{O}_K)$ -orbits and ideal classes of K associates the identity ideal class ($x = 1, y = 0$) with the orbit of $[1, 0] = \infty$ in $\mathbf{P}^1(K)$.

Everything we have done here carries over to a general Dedekind domain, with identical proofs. We will just state the result.

Theorem 4.1. *Let R be a Dedekind domain and F be its fraction field. The orbits for $\mathrm{SL}_2(R)$ on $F \cup \{\infty\}$ are in bijection with the ideal class group of R . In particular, $\mathrm{SL}_2(R)$ acts transitively on $F \cup \{\infty\}$ if and only if R is a PID.*

APPENDIX A. GENERATORS FOR $\mathrm{SL}_2(\mathbf{Z})$

There are two important matrices in $\mathrm{SL}_2(\mathbf{Z})$:

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

It is left to the reader to check that $S^2 = -I_2$, so S has order 4, while $T^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ for any $k \in \mathbf{Z}$, so T has infinite order.

Theorem A.1. *The group $\mathrm{SL}_2(\mathbf{Z})$ is generated by S and T .*

Proof. As the proof will reveal, this theorem is the Euclidean algorithm in disguise.

First we check how S and any power of T change the entries in a matrix. Verify that

$$S \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix},$$

and

$$T^k \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + ck & b + dk \\ c & d \end{pmatrix}.$$

Thus, up to a sign change, multiplying by S on the left interchanges the rows. Multiplying by a power of T on the left adds a multiple of the second row to the first row and does not change the second row. Given a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\mathrm{SL}_2(\mathbf{Z})$, we can carry out the Euclidean algorithm on a and c by using left multiplication by S and powers of T . We use the power of T to carry out the division (if $a = cq + r$, use $k = -q$) and use S to interchange the roles of a and c to guarantee that the larger of the two numbers (in absolute value) is in the upper-left corner. (Multiplication by S will cause a sign change, but this has no serious effect on the algorithm.)

Since $ad - bc = 1$, a and c are relatively prime, so the last step of Euclid's algorithm will have a remainder of 1. This means, after suitable multiplication by S 's and T 's, we will have transformed the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ into a matrix with first column $\begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 0 \\ \pm 1 \end{pmatrix}$. Left-multiplying by S interchanges the rows up to a sign, so we can suppose the first column is $\begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}$. Any matrix of the form $\begin{pmatrix} 1 & x \\ 0 & y \end{pmatrix}$ in $\mathrm{SL}_2(\mathbf{Z})$ must have $y = 1$ (the determinant is 1),

and then it is $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = T^x$. A matrix $\begin{pmatrix} -1 & x \\ 0 & y \end{pmatrix}$ in $\mathrm{SL}_2(\mathbf{Z})$ must have $y = -1$, so the matrix is $\begin{pmatrix} -1 & x \\ 0 & -1 \end{pmatrix} = (-I_2)T^{-x}$. Since $-I_2 = S^2$, we can finally unwind and express our original matrix in terms of S 's and T 's. \square

Example A.2. Take $A = \begin{pmatrix} 26 & 7 \\ 11 & 3 \end{pmatrix}$. Since $26 = 11 \cdot 2 + 4$, we want to subtract $11 \cdot 2$ from 26:

$$T^{-2}A = \begin{pmatrix} 4 & 1 \\ 11 & 3 \end{pmatrix}.$$

Now we want to switch the roles of 4 and 11. Multiply by S :

$$ST^{-2}A = \begin{pmatrix} -11 & -3 \\ 4 & 1 \end{pmatrix}.$$

Dividing -11 by 4, we have $-11 = 4 \cdot (-3) + 1$, so we want to add $4 \cdot 3$ to -11 . Multiply by T^3 :

$$T^3ST^{-2}A = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}.$$

Once again, multiply by S to switch the entries of the first column (up to sign):

$$ST^3ST^{-2}A = \begin{pmatrix} -4 & -1 \\ 1 & 0 \end{pmatrix}.$$

Our final division is: $-4 = 1(-4) + 0$. We want to add 4 to -4 , so multiply by T^4 :

$$T^4ST^3ST^{-2}A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = S.$$

Left-multiplying by the inverses of all the S 's and T 's on the left side, we obtain

$$A = T^2S^{-1}T^{-3}S^{-1}T^{-4}S.$$

Since $S^4 = I_2$, we can write S^{-1} as S^3 if we wish to use a positive exponent on S . However, a similar idea does not apply to the negative powers of T .

Remark A.3. Since $ST = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$ has order 6, we can write $\mathrm{SL}_2(\mathbf{Z}) = \langle S, ST \rangle$, which is a generating set of elements with finite order.

REFERENCES

- [1] E. B. Freitag, "Hilbert Modular Forms," Springer-Verlag, 1990.
- [2] P. B. Garrett, "Holomorphic Hilbert Modular Forms," Wadsworth & Brooks/Cole, 1990.