DECOMPOSING $\mathrm{SL}_2(\mathbf{R})$

KEITH CONRAD

1. Introduction

The group $\mathrm{SL}_2(\mathbf{R})$ is not easy to visualize: it naturally lies in $\mathrm{M}_2(\mathbf{R})$, which is 4-dimensional (the entries of a variable 2×2 real matrix are 4 free parameters). We will derive a product decomposition for $\mathrm{SL}_2(\mathbf{R})$ and use it to get a concrete image of $\mathrm{SL}_2(\mathbf{R})$.

Inside $\mathrm{SL}_2(\mathbf{R})$ are the following three subgroups:

$$K = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right\}, \quad A = \left\{ \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} : r > 0 \right\}, \quad N = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \right\}.$$

Theorem 1.1. *We have a decomposition* $\mathrm{SL}_2(\mathbf{R}) = KAN$*: every* $g \in \mathrm{SL}_2(\mathbf{R})$ *has a unique representation as* $g = kan$ *where* $k \in K$*,* $a \in A$*, and* $n \in N$*.*

This formula $\mathrm{SL}_2(\mathbf{R}) = KAN$ is called the *Iwasawa decomposition* of the group. Don't confuse the use of a in Theorem 1.1 as the label for a matrix in A with a as a real number in the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. The distinction should always be clear from the context. Since $\mathrm{SL}_2(\mathbf{R})$ is defined by the single equation $ad - bc = 1$ inside of $\mathrm{M}_2(\mathbf{R})$, it is a manifold of dimension $4 - 1 = 3$. The subgroups K, A, and N are each 1-dimensional ($K \cong S^1$, $A \cong \mathbf{R}_{>0}$, and $N \cong \mathbf{R}$), and Theorem 1.1 shows they fully account for the 3 dimensions of $\mathrm{SL}_2(\mathbf{R})$.

The subgroups in the Iwasawa decomposition are related to conjugacy classes. We will see that a matrix in $\mathrm{SL}_2(\mathbf{R})$ is, up to sign, conjugate to a matrix in K, A, or N.

2. Iwasawa decomposition

To derive the Iwasawa decomposition of $\mathrm{SL}_2(\mathbf{R})$ we will use an action of this group on bases in $\mathbf{R}^2$.

For $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{R})$, apply it to the standard basis e_1, e_2. The vectors

$$ge_1 = \begin{pmatrix} a \\ c \end{pmatrix}, \quad ge_2 = \begin{pmatrix} b \\ d \end{pmatrix}$$

are also a basis of $\mathbf{R}^2$. We will pass from this new basis of $\mathbf{R}^2$ back to the standard basis e_1, e_2 of $\mathbf{R}^2$ by a sequence of transformations in $\mathrm{SL}_2(\mathbf{R})$ that amounts to something like the Gram-Schmidt process (which turns any basis of $\mathbf{R}^n$ into an orthonormal basis of $\mathbf{R}^n$).

Let θ be the angle *from* the positive x-axis *to* ge_1. Let ρ_θ be the counterclockwise rotation of the plane around the origin by θ, so $\rho_{-\theta}(ge_1)$ is on the positive x-axis. Because $\det g$ is positive, the ordered pair of vectors (ge_1, ge_2) has the same orientation as the ordered pair (e_1, e_2), so $\rho_{-\theta}(ge_2)$ is in the upper (rather than lower) half-plane.

Since $\rho_{-\theta}(ge_1)$ is a positive scalar multiple of e_1, we want to divide $\rho_{-\theta}(ge_1)$ by its length so it becomes e_1. Its length is $r = ||\rho_{-\theta}(ge_1)|| = ||ge_1|| = \sqrt{a^2 + c^2}$. Applying $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & 1/r \end{smallmatrix}\right)$ will have the desired effect $\rho_{-\theta}(ge_1) \mapsto e_1$, but this matrix doesn't have determinant 1. On the other hand, $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)$ also has the desired effect on $\rho_{-\theta}(ge_1)$ and has determinant 1. So apply the matrix $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)$ to $\mathbf{R}^2$. It sends $\rho_{-\theta}(ge_1)$ to $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)\rho_{-\theta}(ge_1) = e_1$. What does it do to $\rho_{-\theta}(ge_2)$? The vector $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)\rho_{-\theta}(ge_2)$ is in the upper half-plane (because $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)$ has positive determinant) and along with e_1 it forms two edges of a parallelogram with area 1 (because $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)$ has determinant ± 1). A parallelogram with area 1 having base e_1 must have height 1, so $\left(\begin{smallmatrix} 1/r & 0 \\ 0 & r \end{smallmatrix}\right)\rho_{-\theta}(ge_2) = \left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right)$ for some x.

Any horizontal shear transformation $\left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)$, which has determinant 1, fixes the x-axis and acts as a stretching along each horizontal line. Applying the horizontal shear transformation $\left(\begin{smallmatrix} 1 & -x \\ 0 & 1 \end{smallmatrix}\right)$ to $\mathbf{R}^2$ takes $\left(\begin{smallmatrix} x \\ 1 \end{smallmatrix}\right)$ to $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = e_2$ and fixes e_1. We have finally returned to the standard basis e_1, e_2 from the basis ge_1, ge_2 by a sequence of transformations in $\mathrm{SL}_2(\mathbf{R})$. Our overall composite transformation is

$$\begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/r & 0 \\ 0 & r \end{pmatrix} \rho_{-\theta},$$

so

$$\begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/r & 0 \\ 0 & r \end{pmatrix} \rho_{-\theta} g$$

sends e_1 to e_1 and e_2 to e_2. A linear transformation on $\mathbf{R}^2$ is determined by what it does to a basis, so

$$\begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/r & 0 \\ 0 & r \end{pmatrix} \rho_{-\theta} g = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Solving for g,

$$\begin{aligned} g &= \rho_\theta \begin{pmatrix} 1/r & 0 \\ 0 & r \end{pmatrix}^{-1} \begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \\ &\in KAN. \end{aligned}$$

Such an expression for g as a product kan with $k \in K$, $a \in A$, and $n \in N$ is called the Iwasawa decomposition of g.

To check this decomposition is unique, for any angle θ, $r > 0$, and $x \in \mathbf{R}$, set

$$g = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} r\cos\theta & xr\cos\theta - (1/r)\sin\theta \\ r\sin\theta & xr\sin\theta + (1/r)\cos\theta \end{pmatrix}.$$

If this is $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{R})$ then

$$r = \sqrt{a^2 + c^2} > 0, \quad \cos\theta = \frac{a}{r}, \quad \sin\theta = \frac{c}{r}, \tag{2.1}$$

and

$$x = \begin{cases} \frac{b + (1/r)\sin\theta}{r\cos\theta}, & \text{if } \cos\theta \neq 0, \\ \frac{d - (1/r)\cos\theta}{r\sin\theta}, & \text{if } \sin\theta \neq 0. \end{cases}$$

Substituting the formulas for $\cos\theta$ and $\sin\theta$ into the formula for x, and using $ad - bc = 1$, we obtain the uniform formula

$$x = \frac{ab + cd}{a^2 + c^2}. \tag{2.2}$$

All the parameters in the matrices making up the Iwasawa decomposition of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ are determined in (2.1) and (2.2), so the Iwasawa decomposition is unique. This completes the proof of Theorem 1.1. In an appendix we derive the Iwasawa decomposition using a different action of $\mathrm{SL}_2(\mathbf{R})$, on the upper half-plane.

The Iwasawa decomposition for $\mathrm{SL}_2(\mathbf{R})$ extends to higher dimensions: $\mathrm{SL}_n(\mathbf{R}) = KAN$ where $K = \mathrm{SO}_n(\mathbf{R}) = \{T \in \mathrm{GL}_n(\mathbf{R}) : TT^\top = I_n, \det T = 1\}$, A is the group of diagonal matrices with positive diagonal entries (and determinant 1) and N is the group of upper-triangular matrices with 1's along the main diagonal. While A and N are both isomorphic to $\mathbf{R}$ when $n = 2$, N becomes nonabelian for $n > 2$. The group K is compact, the group $A \cong (\mathbf{R}_{>0})^{n-1} \cong \mathbf{R}^{n-1}$ is abelian, and N is a nilpotent group. This explains the notation A and N, for abelian and nilpotent.

Since

$$\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & r^2x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}, \tag{2.3}$$

we can move any element of A past an element of N (on either side) at the cost of changing the element of N. Therefore $AN = NA$ is a subgroup of $\mathrm{SL}_2(\mathbf{R})$. Explicitly,

$$AN = \left\{\begin{pmatrix} y & x \\ 0 & 1/y \end{pmatrix} : y > 0, x \in \mathbf{R}\right\}. \tag{2.4}$$

The Iwasawa decomposition $KAN = K(AN)$ for $\mathrm{SL}_2(\mathbf{R})$ is the analogue of the polar decomposition $S^1 \times \mathbf{R}_{>0}$ for $\mathbf{C}^\times$.

In the Iwasawa decomposition, neither K nor AN (nor A or N) is normal in $\mathrm{SL}_2(\mathbf{R})$. For example, the conjugate of an element of K by $(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix})$ is usually not in K and the conjugate of an element of AN by $(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix})$ is usually not in AN. Because of the non-normality, it is not easy to describe the group operation in $\mathrm{SL}_2(\mathbf{R})$ in terms of its Iwasawa decomposition. This decomposition is important for other purposes, such as the following.

Corollary 2.1. *As a topological space,* $\mathrm{SL}_2(\mathbf{R})$ *is homeomorphic to the inside of a solid torus.*

Proof. Let $f\colon K \times A \times N \to \mathrm{SL}_2(\mathbf{R})$ by $f(k, a, n) = kan$. This is continuous, and by Theorem 1.1 it is surjective. We can write down an inverse function using the computations at the end of the proof of Theorem 1.1. For $g = (\begin{smallmatrix} a & b \\ c & d \end{smallmatrix})$ in $\mathrm{SL}_2(\mathbf{R})$, define $r(g) = \sqrt{a^2 + c^2}$ and

$$k(g) = \begin{pmatrix} a/r(g) & -c/r(g) \\ c/r(g) & a/r(g) \end{pmatrix}, \quad a(g) = \begin{pmatrix} r(g) & 0 \\ 0 & 1/r(g) \end{pmatrix},$$

$$n(g) = \begin{pmatrix} 1 & (ab + cd)/(a^2 + c^2) \\ 0 & 1 \end{pmatrix}.$$

The function $g \mapsto (k(g), a(g), n(g))$ from $\mathrm{SL}_2(\mathbf{R})$ to $K \times A \times N$ is continuous and is an inverse to f.

Topologically, $K \cong S^1$, $A \cong \mathbf{R}_{>0} \cong \mathbf{R}$, and $N \cong \mathbf{R}$. Therefore topologically, $\mathrm{SL}_2(\mathbf{R}) \cong S^1 \times \mathbf{R}^2$. The plane $\mathbf{R}^2$ is homeomorphic to the open unit disc D by $v \mapsto v/\sqrt{1 + ||v||^2}$ (with inverse $w \mapsto w/\sqrt{1 - ||w||^2}$), where $||\cdot||$ is the usual length function on $\mathbf{R}^2$, so as a topological space $\mathrm{SL}_2(\mathbf{R})$ is homeomorphic to $S^1 \times D$, which is the inside of a solid torus.

As an alternate ending, on the decomposition $K \times A \times N \cong S^1 \times \mathbf{R}_{>0} \times \mathbf{R}$ treat the product $\mathbf{R}_{>0} \times \mathbf{R}$ as the right half plane $\{x + iy : x > 0\}$ and identify it with the open unit disc D by the Cayley transformation $z \mapsto (z - 1)/(z + 1)$. (Vertical lines in the half-plane are sent to circles inside D that are tangent to the unit circle at 1.) $\square$

Although the proof of Corollary 2.1 shows $\mathrm{SL}_2(\mathbf{R})$ and $S^1 \times \mathbf{R}^2$ are homeomorphic as topological spaces, they are not isomorphic as groups. Equivalently, the homeomorphism $K \times A \times N \to \mathrm{SL}_2(\mathbf{R})$ in Corollary 2.1 is not a group homomorphism.

The Iwasawa decomposition of a matrix in K, A, or N is the obvious one. For a lower triangular matrix $\left(\begin{smallmatrix} 1 & 0 \\ y & 1 \end{smallmatrix}\right)$, which is in none of these subgroups, the inverse map in the proof of Corollary 2.1 gives us the decomposition

$$\begin{pmatrix} 1/\sqrt{1+y^2} & -y/\sqrt{1+y^2} \\ y/\sqrt{1+y^2} & 1/\sqrt{1+y^2} \end{pmatrix} \begin{pmatrix} \sqrt{1+y^2} & 0 \\ 0 & 1/\sqrt{1+y^2} \end{pmatrix} \begin{pmatrix} 1 & y/(1+y^2) \\ 0 & 1 \end{pmatrix}.$$

Remark 2.2. The inside of a solid torus has a circle as a strong deformation retract, so the fundamental group of $\mathrm{SL}_2(\mathbf{R})$ is the same as that of a circle, namely $\mathbf{Z}$. In terms of $\mathrm{SL}_2(\mathbf{R})$ itself, one circle inside it is the subgroup K and a strong deformation retract to K can be written down using the Iwasawa decomposition (make a suitable path in the group from kan to k). From the connection between covering spaces and subgroups of the fundamental group, $\mathrm{SL}_2(\mathbf{R})$ admits a unique covering space of degree d for each positive integer d and the universal covering space of $\mathrm{SL}_2(\mathbf{R})$ is the inside of a solid cylinder $\mathbf{R} \times D$ (homeomorphic to $\mathbf{R}^3$). The degree-2 cover of $\mathrm{SL}(2, \mathbf{R})$ is an important group called the metaplectic group.

That the Iwasawa decomposition gives us a picture of $\mathrm{SL}_2(\mathbf{R})$ is a striking geometric application. Here is an algebraic application (whose punchline is the corollary).

Theorem 2.3. *The only continuous homomorphism* $\mathrm{SL}_2(\mathbf{R}) \to \mathbf{R}$ *is the trivial homomorphism.*

Proof. Let $f \colon \mathrm{SL}_2(\mathbf{R}) \to \mathbf{R}$ be a continuous homomorphism. Then

$$f(kan) = f(k) + f(a) + f(n).$$

We will show f is trivial on K, A, and N, and thus f is trivial on $KAN = \mathrm{SL}_2(\mathbf{R})$.

Since $K \cong S^1$, the elements of finite order in K are dense. Since $\mathbf{R}$ has no elements of finite order except 0, f is trivial on a dense subset of K and thus is trivial on K by continuity. (As an alternate argument, since K is a compact group so is $f(K)$, and the only compact subgroup of $\mathbf{R}$ is $\{0\}$.)

Now we look at f on A and N. Since $A \cong \mathbf{R}_{>0} \cong \mathbf{R}$ by $\left(\begin{smallmatrix} r & 0 \\ 0 & 1/r \end{smallmatrix}\right) \mapsto \log r$ and $N \cong \mathbf{R}$ by $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right) \mapsto x$, both algebraically and topologically, describing the continuous homomorphisms from A and N to $\mathbf{R}$ is the same as describing the continuous homomorphisms from $\mathbf{R}$ to $\mathbf{R}$. Any continuous homomorphism $\mathbf{R} \to \mathbf{R}$ has the form $x \mapsto tx$ for some real number t (see where 1 goes, call that t, and then appeal to the denseness of $\mathbf{Q}$ in $\mathbf{R}$). Therefore

$$f\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} = t \log r, \quad f\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = t'x$$

for some t and t'. Applying f to both sides of (2.3),

$$t \log r + t'x = t'r^2x + t \log r,$$

so $t'x = t'r^2x$ for all $r > 0$ and $x \in \mathbf{R}$. Thus $t' = 0$ (*e.g.*, take $x = 1$ and $r = 2$ to see this.) This shows f is trivial on N.

It remains to show f is trivial on A. For this we appeal to the conjugation relation

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1/r & 0 \\ 0 & r \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}^{-1}$$

Applying f, we get $f\left(\begin{smallmatrix} r & 0 \\ 0 & 1/r \end{smallmatrix}\right) = -f\left(\begin{smallmatrix} r & 0 \\ 0 & 1/r \end{smallmatrix}\right)$, so $f\left(\begin{smallmatrix} r & 0 \\ 0 & 1/r \end{smallmatrix}\right) = 0$. $\square$

Corollary 2.4. *Any continuous homomorphism* $\mathrm{SL}_2(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ *has image in* $\mathrm{SL}_n(\mathbf{R})$.

Proof. Let $f\colon \mathrm{SL}_2(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ be a continuous homomorphism. Composing f with the determinant map $\mathrm{GL}_n(\mathbf{R}) \to \mathbf{R}^\times$ gives a continuous homomorphism $\det \circ f\colon \mathrm{SL}_2(\mathbf{R}) \to \mathbf{R}^\times$. Since $\mathrm{SL}_2(\mathbf{R})$ is connected (Corollary 2.1), its image under $\det \circ f$ is a connected subgroup of $\mathbf{R}^\times$, and hence lies in $\mathbf{R}_{>0}$. As $\mathbf{R}_{>0} \cong \mathbf{R}$ both topologically and algebraically, $\det \circ f$ is trivial by Theorem 2.3. Thus $\det(f(g)) = 1$ for all $g \in \mathrm{SL}_2(\mathbf{R})$, so $f(g) \in \mathrm{SL}_n(\mathbf{R})$ for all $g \in \mathrm{SL}_2(\mathbf{R})$. $\square$

Example 2.5. We will construct a continuous homomorphism $\mathrm{GL}_2(\mathbf{R}) \to \mathrm{GL}_3(\mathbf{R})$ and see its restriction to $\mathrm{SL}_2(\mathbf{R})$ has values in $\mathrm{SL}_3(\mathbf{R})$.

Let $V = \mathbf{R}x^2 + \mathbf{R}xy + \mathbf{R}y^2$ be the vector space of homogeneous polynomials in x and y of degree 2. It is 3-dimensional, with basis x^2, xy, y^2. For a matrix $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{R})$ and $Q(x,y) \in V$, let $(gQ)(x,y) = Q(ax+cy, bx+dy)$. A calculation shows $g_1(g_2Q) = (g_1g_2)(Q)$, so $\mathrm{GL}_2(\mathbf{R})$ acts on V. This action is a linear change of variables on V given by the entries of 2×2 matrices. Since $g(Q+Q') = g(Q) + g(Q')$ and $g(sQ) = sg(Q)$ for $s \in \mathbf{R}$, the action of g on V is a linear transformation (necessarily invertible, since the action of g^{-1} on V is its inverse). Using the basis x^2, xy, y^2 of V, we can compute a matrix representation of g on V: we have

$$g(x^2) = (ax+cy)^2 = a^2x^2 + 2acxy + c^2y^2,$$
$$g(xy) = (ax+cy)(bx+dy) = abx^2 + (ad+bc)xy + cdy^2,$$

and

$$g(y^2) = (bx+dy)^2 = b^2x^2 + 2bdxy + d^2y^2,$$

so the matrix of g with respect to the basis x^2, xy, y^2 is

$$\begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix}.$$

Call this matrix $f(g)$, so $f\colon \mathrm{GL}_2(\mathbf{R}) \to \mathrm{GL}_3(\mathbf{R})$ is a homomorphism and the formula for $f(g)$ shows f is continuous. By a calculation, $\det(f(g)) = (ad-bc)^3 = (\deg g)^3$, so when g has determinant 1 so does $f(g)$.

Example 2.5 can be generalized. For any integer $n \geq 1$, the space $V_n = \bigoplus_{i=0}^n \mathbf{R}x^{n-i}y^i$ of homogeneous 2-variable polynomials of degree n has dimension $n+1$ and $\mathrm{GL}_2(\mathbf{R})$ acts on this space by linear changes of variables. The restriction of this action to $\mathrm{SL}_2(\mathbf{R})$ on V_n accounts for essentially all "interesting" actions of $\mathrm{SL}_2(\mathbf{R})$ on finite-dimensional vector spaces.

Theorem 2.6. *The homomorphism* $\mathrm{SL}_2(\mathbf{Z}[x,y]) \to \mathrm{SL}_2(\mathbf{Z}[x,y]/(x^2+y^2-1))$ *is not surjective.*

Proof. Let $\overline{x}$ and $\overline{y}$ be the cosets of x and y in $\mathbf{Z}[x,y]/(x^2+y^2-1)$, so $\overline{x}^2 + \overline{y}^2 = 1$. An obvious matrix in $\mathrm{SL}_2(\mathbf{Z}[x,y]/(x^2+y^2-1))$ is

$$\begin{pmatrix} \overline{x} & \overline{y} \\ -\overline{y} & \overline{x} \end{pmatrix}. \tag{2.5}$$

Suppose there is a matrix

$$A(x,y) = \begin{pmatrix} a(x,y) & b(x,y) \\ c(x,y) & d(x,y) \end{pmatrix}$$

in $\mathrm{SL}_2(\mathbf{Z}[x,y])$ which reduces to the matrix (2.5) in $\mathrm{SL}_2(\mathbf{Z}[x,y]/(x^2+y^2-1))$.

A continuous map $S^1 \to \mathrm{SL}_2(\mathbf{R})$ is given by $(u,v) \mapsto \left(\begin{smallmatrix} u & v \\ -v & u \end{smallmatrix}\right) = A(u,v)$. The image is a loop in $\mathrm{SL}_2(\mathbf{R})$ which generates $\pi_1(\mathrm{SL}_2(\mathbf{R}))$. Let $h\colon [0,1] \times S^1 \to \mathrm{SL}_2(\mathbf{R})$ by $h(t,u,v) = A(tu,tv)$. Then $h(0,u,v)$ is a constant map and $h(1,u,v)$ is the standard loop in $\mathrm{SL}_2(\mathbf{R})$, so we can shrink the loop to a point continuously. This is false, so there is no such map. $\square$

3. Conjugacy Classes

The conjugacy class of a matrix in $\mathrm{SL}_2(\mathbf{R})$ is nearly determined by its eigenvalues, but we have to be a little bit careful so we don't confuse conjugacy in $\mathrm{SL}_2(\mathbf{R})$ with conjugacy in the larger group $\mathrm{GL}_2(\mathbf{R})$. For example, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and its inverse $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$ are conjugate in $\mathrm{GL}_2(\mathbf{R})$ by $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, whose determinant is -1. These two matrices are not conjugate in $\mathrm{SL}_2(\mathbf{R})$, since any $\mathrm{SL}_2(\mathbf{R})$-conjugate of $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ has a positive upper right entry, by an explicit calculation.

Theorem 3.1. *Let $T \in \mathrm{SL}_2(\mathbf{R})$. If $(\mathrm{Tr}\, T)^2 > 4$ then T is conjugate to a unique matrix of the form $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$ with $|\lambda| > 1$. If $(\mathrm{Tr}\, T)^2 = 4$ then T is conjugate to exactly one of $\pm I_2$, $\pm\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, or $\pm\left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$. If $(\mathrm{Tr}\, T)^2 < 4$ then T is conjugate to a unique matrix of the form $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$ other than $\pm I_2$.*

Proof. For $T \in \mathrm{SL}_2(\mathbf{R})$, its eigenvalues are roots of its characteristic polynomial, which is $X^2 - tX + 1$, where $t = \mathrm{Tr}(T)$. The nature of the eigenvalues of T are determined by the discriminant of this polynomial, $t^2 - 4$: two distinct real eigenvalues if $t^2 > 4$, a repeated eigenvalue if $t^2 = 4$, and two complex conjugate eigenvalues if $t^2 < 4$. We will find a representative for the conjugacy class of T based on the sign of $t^2 - 4$. Of course matrices with different t's are not conjugate.

In what follows, if v and w are vectors in $\mathbf{R}^2$ whose specific coordinates are not important to make explicit, we will write $([v]\ [w])$ for the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. This matrix is invertible when v and w are linearly independent.

Suppose $t^2 > 4$. Then T has distinct real eigenvalues λ and $1/\lambda$. Let v and v' be eigenvectors in $\mathbf{R}^2$ for these eigenvalues: $Tv = \lambda v$ and $Tv' = (1/\lambda)v'$. In coordinates from the basis v and v', T is represented by $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$, so T is conjugate to $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$ by the 2×2 matrix $([v]\ [v'])$. Scaling v' keeps it as an eigenvector of T, and by a suitable nonzero scaling the matrix $([v]\ [v'])$ has determinant 1. Therefore T is conjugate to $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{R})$. We did not specify an ordering of the eigenvalues, so $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1/\lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$ have to be conjugate to each other in $\mathrm{SL}_2(\mathbf{R})$. Explicitly, $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1/\lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$. Conjugate matrices have the same eigenvalues, so $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$ is conjugate to $\left(\begin{smallmatrix} \mu & 0 \\ 0 & 1/\mu \end{smallmatrix}\right)$ only when μ equals λ or $1/\lambda$. We can therefore pin down a representative for the conjugacy class of T in $\mathrm{SL}_2(\mathbf{R})$ as $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{smallmatrix}\right)$ with $|\lambda| > 1$.

Now suppose $t^2 = 4$. The roots of $X^2 - tX + 1$ are both 1 (if $t = 2$) or both -1 (if $t = -2$). Let $\lambda = \pm 1$ be the eigenvalue for T. Extend v to a basis $\{v, v'\}$ of $\mathbf{R}^2$. Scaling v', we may assume the matrix $([v]\ [v'])$ has determinant 1. Conjugating T by this matrix expresses it in the basis v and v' as $\left(\begin{smallmatrix} \lambda & x \\ 0 & y \end{smallmatrix}\right)$. Since the determinant is 1, $y = 1/\lambda = \lambda = \pm 1$. Therefore T is conjugate in $\mathrm{SL}_2(\mathbf{R})$ to a matrix of the form $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} -1 & x \\ 0 & -1 \end{smallmatrix}\right)$. If $x = 0$ these matrices are $\pm I_2$, which are in their own conjugacy class. The formulas

$$\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}^{-1} = \begin{pmatrix} 1 & r^2 \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -r^2 \\ 0 & 1 \end{pmatrix}$$

show $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ is conjugate to either $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$, depending on the sign of x. Similarly $\left(\begin{smallmatrix} -1 & x \\ 0 & -1 \end{smallmatrix}\right)$ is conjugate to either $\left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} -1 & -1 \\ 0 & -1 \end{smallmatrix}\right)$.

The four matrices $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} -1 & -1 \\ 0 & -1 \end{smallmatrix}\right)$ are nonconjugate in $\mathrm{SL}_2(\mathbf{R})$, *e.g.*, an $\mathrm{SL}_2(\mathbf{R})$-conjugate of $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ looks like $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ with x a perfect square. Other cases are left to the reader.

Finally, suppose $t^2 < 4$. Now T has complex conjugate eigenvalues which, by the quadratic formula for $X^2 - tX + 1$, are of absolute value 1 and are not ± 1 (since $t \neq \pm 2$). We can write the eigenvalues as $e^{\pm i\theta}$, with $\sin\theta \neq 0$. Pick an eigenvector v in $\mathbf{C}^2$ such that $Tv = e^{i\theta}v$. Since $e^{i\theta}$ is not real, $v \notin \mathbf{R}^2$. Let $\overline{v}$ be the vector with coordinates that are complex conjugate to those of v, so $T\overline{v} = e^{-i\theta}\overline{v}$. Then $v + \overline{v}$ and $i(v - \overline{v})$ are in $\mathbf{R}^2$, with

$$T(v + \overline{v}) = (\cos\theta)(v + \overline{v}) + (\sin\theta)i(v - \overline{v})$$

and

$$T(v - \overline{v}) = -(\sin\theta)(v + \overline{v}) + (\cos\theta)i(v - \overline{v}).$$

Therefore conjugating T by the (invertible) real matrix $([v + \overline{v}]\ [i(v - \overline{v})])$ turns T into $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. We don't know the determinant of $([v + \overline{v}]\ [i(v - \overline{v})])$, but scaling v by a real number (and $\overline{v}$ by the same amount, to keep it conjugate) can give this conjugating matrix determinant ± 1. If the determinant is 1 then T is conjugate to $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{R})$. If the determinant is -1, then reverse the order of the columns in the conjugating marix to give it determinant 1 and then T is conjugate to $\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right) = \left(\begin{smallmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{R})$. Two matrices $\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{smallmatrix}\right)$ in K can be conjugate only when $\varphi = \pm\theta \bmod 2\pi\mathbf{Z}$, by looking at eigenvalues, and a direct calculation shows the $\mathrm{SL}_2(\mathbf{R})$-conjugate of $\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)$ never equals $\left(\begin{smallmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{smallmatrix}\right)$ unless $\sin\theta = 0$, but we are in a case when $\sin\theta \neq 0$. $\square$

When $T \in \mathrm{SL}_2(\mathbf{R})$ satisfies $\mathrm{Tr}(T)^2 > 4$ we say T is *hyperbolic*, when $(\mathrm{Tr}\,T)^2 = 4$ we say T is *parabolic*, and when $(\mathrm{Tr}\,T)^2 < 4$ we say T is *elliptic*. This terminology is borrowed from the shape of a plane conic $ax^2 + bxy + cy^2 = 1$ in terms of its discriminant $d = b^2 - 4ac$: it is a hyperbola when $d > 0$, a parabola when $d = 0$, and an ellipse when $d < 0$. Up to sign, the hyperbolic conjugacy classes in $\mathrm{SL}_2(\mathbf{R})$ are represented by matrices in A (besides I_2), the elliptic conjugacy classes are represented by matrices in K (besides $\pm I_2$), and the parabolic conjugacy classes are represented by matrices in N.

Appendix A. Acting on the Upper Half-Plane

We will use an action of $\mathrm{SL}_2(\mathbf{R})$ on the upper half-plane $\mathfrak{h}$ to obtain the Iwasawa decomposition of $\mathrm{SL}_2(\mathbf{R})$ in a more efficient manner than the first proof that used an action on bases of $\mathbf{R}^2$.

For $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{R})$ and a non-real complex number z, set

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = \frac{az + b}{cz + d} \in \mathbf{C} - \mathbf{R}. \tag{A.1}$$

By a calculation left to the reader, $g_1(g_2(z)) = (g_1g_2)(z)$ for g_1 and g_2 in $\mathrm{GL}_2(\mathbf{R})$, and

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

Therefore when $ad-bc>0$, z and $(az+b)/(cz+d)$ have the same sign for their imaginary parts. In particular, if $g \in \mathrm{SL}_2(\mathbf{R})$ and z is in the upper half-plane then so is $g(z)$, so (A.1) is an action of the group $\mathrm{SL}_2(\mathbf{R})$ on the set $\mathfrak{h}$. This action has one orbit since we can get anywhere in $\mathfrak{h}$ from i using $\mathrm{SL}_2(\mathbf{R})$:

$$\begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}(i) = x+iy. \tag{A.2}$$

Notice that the matrix used here to send i to $x+iy$ is in the subgroup AN (see (2.4)).

Let's determine the stabilizer of i. Saying $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(i) = i$ is equivalent to $(ai+b)/(ci+d) = i$, so $ai+b = di-c$. Therefore $d=a$ and $b=-c$, so $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & -c \\ c & a \end{smallmatrix}\right)$ with $a^2+c^2=1$. We can therefore write $a=\cos\theta$ and $c=\sin\theta$, which shows the stabilizer of i is the set of matrices $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. This is the subgroup K (so $\mathfrak{h}$ can be viewed as the coset space $\mathrm{SL}_2(\mathbf{R})/K$ on which $\mathrm{SL}_2(\mathbf{R})$ acts by left multiplication).

Now we are ready to derive the Iwasawa decomposition. For $g \in \mathrm{SL}_2(\mathbf{R})$, write $g(i) = x+iy \in \mathfrak{h}$. Using (A.2),

$$g(i) = x+iy = \begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}(i) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{pmatrix}(i) \in NA(i). \tag{A.3}$$

Since g acts on i in the same way as an element of NA, and the stabilizer of i is K, $g \in NAK$. Thus $\mathrm{SL}_2(\mathbf{R}) = NAK$. Applying inversion to this decomposition, $\mathrm{SL}_2(\mathbf{R}) = KAN$. That settles the existence of the Iwasawa decomposition.

To prove uniqueness, assume $nak = n'a'k'$. Applying both sides to i, k and k' fix i so $na(i) = n'a'(i)$. For $n = \left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ and $a = \left(\begin{smallmatrix} r & 0 \\ 0 & 1/r \end{smallmatrix}\right)$, $na = \left(\begin{smallmatrix} r & x/r \\ 0 & 1/r \end{smallmatrix}\right)$, so $na(i) = x + r^2 i$. In particular, knowing $na(i)$ tells us the parameters determining n and a. Hence $n=n'$ and $a=a'$, so $k=k'$.

The upper half-plane action of $\mathrm{SL}_2(\mathbf{R})$ leads in a second way to the formulas (2.1) and (2.2) for the matrix entries in the factors of the Iwasawa decomposition for $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{R})$. (We have just proved anew the existence and uniqueness of this decomposition.) Write, as in Section 2,

$$g = kan = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} r & 0 \\ 0 & 1/r \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}. \tag{A.4}$$

We want to determine the entries of these matrices in terms of the entries of g. We will work with $g^{-1} = n^{-1}a^{-1}k^{-1}$ since the $\mathrm{SL}_2(\mathbf{R})$-action on $\mathfrak{h}$ leads to the decomposition NAK rather than KAN:

$$g^{-1}(i) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}(i) = \frac{di-b}{-ci+a} = -\frac{ab+cd}{a^2+c^2} + \frac{1}{a^2+c^2}i.$$

Writing this as $u+iv$, from (A.3) (with g^{-1} in place of g and $u+iv$ in place of $x+iy$) we get

$$n^{-1} = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -(ab+cd)/(a^2+c^2) \\ 0 & 1 \end{pmatrix}$$

and

$$a^{-1} = \begin{pmatrix} \sqrt{v} & 0 \\ 0 & 1/\sqrt{v} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{a^2+c^2} & 0 \\ 0 & \sqrt{a^2+c^2} \end{pmatrix},$$

so

$$n = \begin{pmatrix} 1 & (ab+cd)/(a^2+c^2) \\ 0 & 1 \end{pmatrix}, \quad a = \begin{pmatrix} \sqrt{a^2+c^2} & 0 \\ 0 & 1/\sqrt{a^2+c^2} \end{pmatrix}.$$

Since $g = kan$,

$$k = gn^{-1}a^{-1} = \begin{pmatrix} a/\sqrt{a^2+c^2} & -c/\sqrt{a^2+c^2} \\ c/\sqrt{a^2+c^2} & a/\sqrt{a^2+c^2} \end{pmatrix}.$$

The formulas for the entries of k, a, and n match those in (2.1) and (2.2).

It is interesting to compare the role of the group K in the geometry of $\mathbf{R}^2$ and $\mathfrak{h}$. As a transformation of $\mathbf{R}^2$, an element of K is a rotation around the origin. This is an isometry of $\mathbf{R}^2$ using the Euclidean metric, and the K-orbit of a nonzero vector in $\mathbf{R}^2$ is the circle which passes through that vector and is centered at the origin. As a transformation of $\mathfrak{h}$, an element of K is a rotation around i relative to the hyperbolic metric on $\mathfrak{h}$. This is a hyperbolic isometry of $\mathfrak{h}$, and the K-orbit of a point in $\mathfrak{h}$ is the circle through that point which is centered at i relative to the hyperbolic metric.

The conjugacy class of a matrix $T \in \mathrm{SL}_2(\mathbf{R})$ was determined in Theorem 3.1 in terms of $(\mathrm{Tr}\, T)^2 - 4$, which is the discriminant of the characteristic polynomial of T. The sign of this quantity tells us whether T has real or non-real eigenvectors. The difference $(\mathrm{Tr}\, T)^2 - 4$ is also relevant to the action of T on the upper half-plane, with fixed points replacing eigenvectors. When $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, the fixed-point condition $T(z) = z$ is equivalent to $az + b = (cz+d)z$, which says $cz^2 + (d-a)z - b = 0$. The discriminant of this equation, which tells us the number of real roots, is

$$(d-a)^2 + 4bc = d^2 - 2da + a^2 + 4(ad-1) = (a+d)^2 - 4 = (\mathrm{Tr}\, T)^2 - 4.$$