$\mathrm{SL}_2(\mathbf{Z})$

KEITH CONRAD

1. Introduction

The group $\mathrm{SL}_2(\mathbf{Z})$, which lies discretely in $\mathrm{SL}_2(\mathbf{R})$, has a role somewhat like that of $\mathbf{Z}$ inside of $\mathbf{R}$. It is the most basic example of a discrete nonabelian group. Two particular elements in $\mathrm{SL}_2(\mathbf{Z})$ are

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

The matrix S has order 4 ($S^2 = -I_2$), while T has infinite order ($T^n = \left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)$) and $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ has order 6 ($(ST)^3 = -I_2$).

Theorem 1.1. *The matrices S and T generate $\mathrm{SL}_2(\mathbf{Z})$.*

After proving this theorem and running through a few quick consequences, we will look at subgroups of finite index in $\mathrm{SL}_2(\mathbf{Z})$.

2. Proof of Theorem 1.1

Let $G = \langle S, T\rangle$ be the subgroup of $\mathrm{SL}_2(\mathbf{Z})$ generated by S and T. We will give two proofs that $G = \mathrm{SL}_2(\mathbf{Z})$, one algebraic and the other geometric.

For the algebraic proof, we start by writing down the effect of S and T^n on a general matrix by multiplication from the left:

$$S\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}, \quad T^n\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+nc & b+nd \\ c & d \end{pmatrix}. \tag{2.1}$$

Now pick $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z})$. Suppose $c \neq 0$. If $|a| \geq |c|$, divide a by c: $a = cq + r$ with $0 \leq r < |c|$. By (2.1), $T^{-q}\gamma$ has upper left entry $a - qc = r$, which is smaller in absolute value than the lower left entry c in $T^{-q}\gamma$. Applying S switches these entries (with a sign change), and we can apply the division algorithm in $\mathbf{Z}$ again if the lower left entry is nonzero in order to find another power of T to multiply by on the left so the lower left entry has smaller absolute value than before. Eventually multiplication of γ on the left by enough copies of S and powers of T gives a matrix in $\mathrm{SL}_2(\mathbf{Z})$ with lower left entry 0. Such a matrix, since it is integral with determinant 1, has the form $\left(\begin{smallmatrix} \pm 1 & m \\ 0 & \pm 1 \end{smallmatrix}\right)$ for some $m \in \mathbf{Z}$ and common signs on the diagonal. This matrix is either T^m or $-T^{-m}$, so there is some $g \in G$ such that $g\gamma = \pm T^n$ for some $n \in \mathbf{Z}$. Since $T^n \in G$ and $S^2 = -I_2$, we have $\gamma = \pm g^{-1}T^n \in G$, so we are done.

In this algebraic proof, G acted on the set $\mathrm{SL}_2(\mathbf{Z})$ by left multiplication. For the geometric proof, we make $\mathrm{GL}_2^+(\mathbf{R})$ act on the upper half-plane $\mathfrak{h} = \{x+iy : y > 0\}$ by linear fractional transformations: for $\tau \in \mathfrak{h}$, define

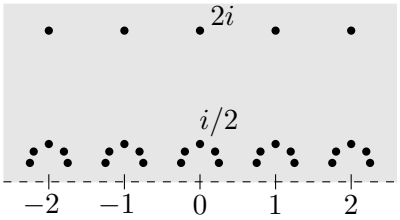
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\tau := \frac{a\tau + b}{c\tau + d}. \tag{2.2}$$

The reason (2.2) lies in $\mathfrak{h}$ follows from the imaginary part formula

$$\mathrm{Im}\left(\frac{a\tau+b}{c\tau+d}\right) = \frac{(ad-bc)\,\mathrm{Im}\,\tau}{|c\tau+d|^2}, \tag{2.3}$$

for $\tau \in \mathbf{C} - \{-d/c\}$ and real a, b, c, d. By this formula, which the reader can check as an exercise, if $\tau \in \mathfrak{h}$ and $ad - bc > 0$ then $(a\tau+b)/(c\tau+d) \in \mathfrak{h}$. To show (2.2) defines a (left) group action of $\mathrm{GL}_2^+(\mathbf{R})$ on $\mathfrak{h}$, check that $I_2\tau = \tau$ and $A(B\tau) = (AB)\tau$ for all A and B in $\mathrm{GL}_2^+(\mathbf{R})$. This action does not distinguish between matrices that differ by a sign (γ and $-\gamma$ act on $\mathfrak{h}$ in the same way), but this will not be a problem for the purpose of using this action to prove $G = \mathrm{SL}_2(\mathbf{Z})$ since $-I_2 = S^2 \in G$.

The key geometric idea is that when $\mathrm{SL}_2(\mathbf{Z})$ acts on a point in $\mathfrak{h}$, the orbit appears to accumulate towards the x-axis. This is illustrated by the picture below, which shows points in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of $2i$ (including $S(2i) = -1/(2i) = i/2$). It appears that the imaginary parts of points in the orbit never exceed 2.



With the picture in mind, pick $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ and set $\tau := \gamma(2i)$.

For $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in G, so $ad - bc = 1$, (2.3) tells us

$$\mathrm{Im}(g\tau) = \frac{\mathrm{Im}\,\tau}{|c\tau+d|^2}.$$

Write τ as $x + yi$. Then in the denominator

$$|c\tau+d|^2 = (cx+d)^2 + (cy)^2,$$

since $y \neq 0$ there are only finitely many integers c and d with $|c\tau+d|$ less than a given bound. Here τ is not changing but c and d are. Therefore $\mathrm{Im}(g\tau)$ has a *maximum* possible value as g runs over G (with τ fixed), so there is some $g_0 \in G$ such that

$$\mathrm{Im}(g\tau) \leq \mathrm{Im}(g_0\tau)$$

for all $g \in G$.

Since $Sg_0 \in G$, the maximality property defining g_0 implies $\mathrm{Im}((Sg_0)\tau) \leq \mathrm{Im}(g_0\tau)$, so (2.3) with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = S$ gives us

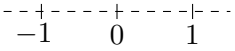
$$\mathrm{Im}(S(g_0\tau)) = \frac{\mathrm{Im}(g_0\tau)}{|g_0\tau|^2} \leq \mathrm{Im}(g_0\tau).$$

Therefore $|g_0\tau|^2 \geq 1$, so $|g_0\tau| \geq 1$. Since $\mathrm{Im}(T^n g_0\tau) = \mathrm{Im}(g_0\tau)$ and $T^n g_0 \in G$, replacing $g_0\tau$ with $T^n g_0\tau$ and running through the argument again shows $|T^n g_0\tau| \geq 1$ for all $n \in \mathbf{Z}$.

Applying T (or T^{-1}) to $g_0\tau$ adjusts its real part by 1 (or -1) without affecting the imaginary part. For some n, $T^n g_0\tau$ has real part between $-1/2$ and $1/2$. Using this power of T, we've shown that $\tau = \gamma(2i)$ has an element of its G-orbit in the set

$$\mathcal{F} = \{\tau \in \mathfrak{h} : |\,\mathrm{Re}(\tau)| \leq 1/2, |\tau| \geq 1\}. \tag{2.4}$$

See the picture below. Note $\mathrm{Im}\,\tau \geq \sqrt{3}/2 > 1/2$ for all $\tau \in \mathcal{F}$.



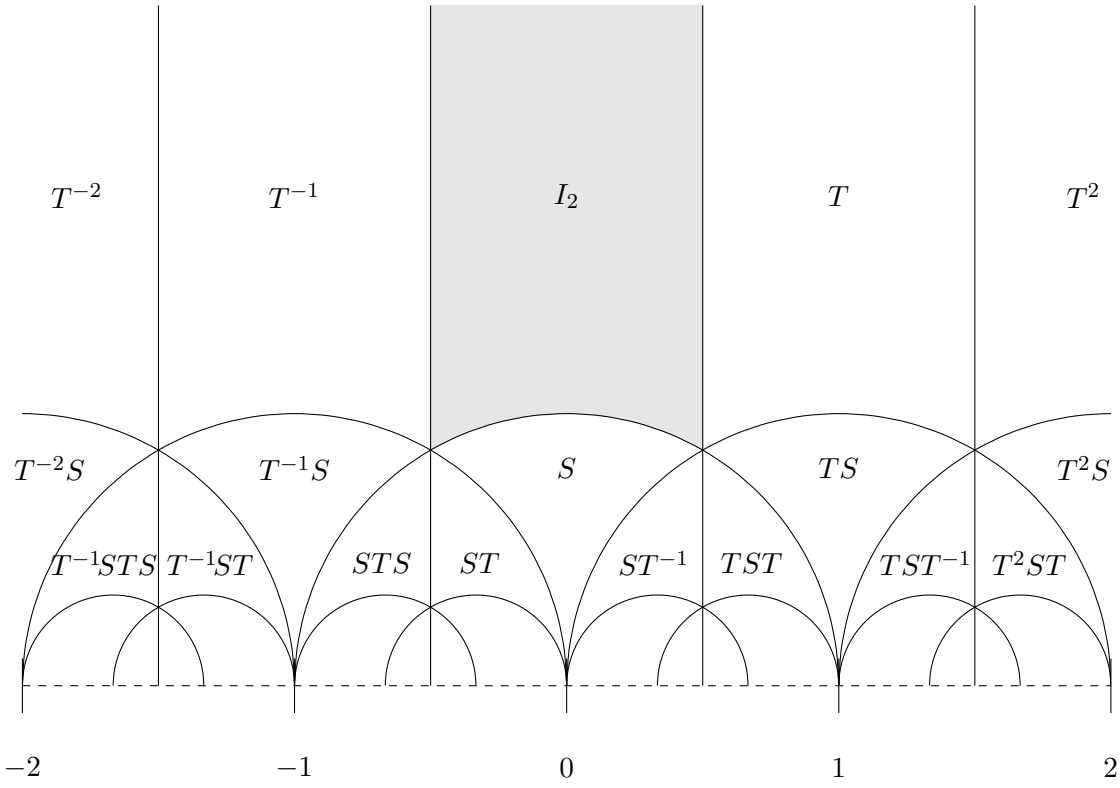
For γ in $\mathrm{SL}_2(\mathbf{Z})$ we showed there is $g \in G$ such that $g(\gamma(2i)) = (g\gamma)(2i)$ is in $\mathcal{F}$. By (2.3),

$$g\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) \Longrightarrow \mathrm{Im}((g\gamma)(2i)) = \frac{2}{4c^2+d^2} \geq \frac{\sqrt{3}}{2},$$

so $c = 0$ (since $2/4 = 1/2 < \sqrt{3}/2$). Then $ad = 1$, so $a = d = \pm 1$ and $(g\gamma)(2i) = (a(2i)+b)/d = 2i \pm b$. For $\mathrm{Re}((g\gamma)(2i))$ to be in $[-1/2, 1/2]$ forces $b = 0$, so $g\gamma = \pm I_2$. Thus $\gamma = \pm g^{-1}$. Since $-I_2 = S^2 \in G$, we get $\gamma \in G$. This finishes the proof of Theorem 1.1.

The region $\mathcal{F}$ above is called a *fundamental domain* for the action of $\mathrm{SL}_2(\mathbf{Z})$ on $\mathfrak{h}$. It is analogous to $[0,1]$ as a fundamental domain for the translation action of $\mathbf{Z}$ on $\mathbf{R}$: each point in the space ($\mathfrak{h}$ or $\mathbf{R}$) has a point of its orbit (by $\mathrm{SL}_2(\mathbf{Z})$ or $\mathbf{Z}$) in the fundamental domain ($\mathcal{F}$ or $[0,1]$) and all points in the fundamental domain lying in the same orbit are on the boundary. In Appendix A we use $\mathcal{F}$ to compute the stabilizer of each point in $\mathfrak{h}$.

Below is a decomposition of $\mathfrak{h}$ into translates $\gamma(\mathcal{F})$ as γ runs over $\mathrm{SL}_2(\mathbf{Z})$, with $\gamma = I_2$ corresponding to $\mathcal{F}$. The page https://roywilliams.github.io/play/js/sl2z/ animates $\mathrm{SL}_2(\mathbf{Z})$-orbits on this figure.



Different translates overlap only along boundary curves, and as we get closer to the x-axis $\mathfrak{h}$ is filled by infinitely many more of these translates. The fundamental domain and its translates are called *ideal triangles* since they are each bounded by three sides and have two endpoints in $\mathfrak{h}$ but one "endpoint" not in $\mathfrak{h}$: the third endpoint is either a rational number on the x-axis or (for the regions $T^n(\mathcal{F})$ with $n \in \mathbf{Z}$) is $i\infty$.

The description of $\mathcal{F}$ in (2.4) uses Euclidean geometry (the absolute value measures Euclidean distances in $\mathfrak{h}$). Using the hyperbolic metric d_H on $\mathfrak{h}$ (see Appendix B), the action of $\mathrm{SL}_2(\mathbf{Z})$ and more generally $\mathrm{SL}_2(\mathbf{R})$ by linear fractional transformations defines isometries for the hyperbolic metric and we can give another description of $\mathcal{F}$ using d_H:

$$\mathcal{F} = \{\tau \in \mathfrak{h} : d_H(\tau, 2i) \leq d_H(\tau, \gamma(2i)) \text{ for all } \gamma \in \mathrm{SL}_2(\mathbf{Z})\}.$$

That is, $\mathcal{F}$ is the points of $\mathfrak{h}$ whose distance (as measured by the hyperbolic metric) to $2i$ is minimal compared to the distance to all points in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of $2i$. The boundary of $\mathcal{F}$ is the points equidistant (for the hyperbolic metric) between $2i$ and one of its nearest $\mathrm{SL}_2(\mathbf{Z})$ translates $T(2i) = 2i+1$, $T^{-1}(2i) = 2i-1$, or $S(2i) = i/2$.[1] Part of what makes this geometric description of $\mathcal{F}$, called a *Dirichlet polygon*, attractive is that it also works for discrete groups actings by isometries on Euclidean spaces. For example, when $\mathbf{Z}$ acts on $\mathbf{R}$ by integer translations, for each $a \in \mathbf{R}$ the numbers whose distance to $a+\mathbf{Z} = \{a+n : n \in \mathbf{Z}\}$ is minimized at a form $[a-1/2, a+1/2]$ and this is a fundamental domain for $\mathbf{Z}$ acting on $\mathbf{R}$.

Example 2.1. We will carry out the algebraic proof of Theorem 1.1 to express $A = \left(\begin{smallmatrix} 17 & 29 \\ 7 & 12 \end{smallmatrix}\right)$ in terms of S and T.

Since $17 = 7 \cdot 2 + 3$, we want to subtract $7 \cdot 2$ from 17:

$$T^{-2}A = \begin{pmatrix} 3 & 5 \\ 7 & 12 \end{pmatrix}.$$

Now we want to switch the roles of 3 and 7. Multiply by S:

$$ST^{-2}A = \begin{pmatrix} -7 & -12 \\ 3 & 5 \end{pmatrix}.$$

Dividing -7 by 3, we have $-7 = 3 \cdot (-3) + 2$, so we want to add $3 \cdot 3$ to -7. Multiply by T^3:

$$T^3ST^{-2}A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}.$$

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

$$ST^3ST^{-2}A = \begin{pmatrix} -3 & -5 \\ 2 & 3 \end{pmatrix}.$$

Since $-3 = 2(-2) + 1$, we compute

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

Mutliply by S:

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

[1] We can replace $2i$ by yi for $y > 1$ and the same description of $\mathcal{F}$ works.

Since $-2 = 1(-2) + 0$, multiply by T^2:

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

Multiply by S:

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

Solving for A,

$$\begin{pmatrix} 17 & 29 \\ 7 & 12 \end{pmatrix} = A = T^2S^{-1}T^{-3}S^{-1}T^{-2}S^{-1}T^{-2}S^{-1}(S^2T) = T^2ST^{-3}ST^{-2}ST^{-2}ST \tag{2.5}$$

since $S^{-1} = -S$.

Remark 2.2. Readers familiar with continued fractions will like to know that multiplication by the matrices S and T is closely related to continued fractions for rational numbers, with the caveat that the continued fraction algorithm should use nearest integers from above rather than from below. To illustrate, the matrix $\left(\begin{smallmatrix} 17 & 29 \\ 7 & 12 \end{smallmatrix}\right)$ is in $\mathrm{SL}_2(\mathbf{Z})$, and to obtain an expression for it in terms of S and T, we look at the ratio in the first column, $17/7$:

$$\frac{17}{7} = 3 - \frac{4}{7} = 3 - \frac{1}{7/4} = 3 - \frac{1}{2 - 1/4}.$$

Using the entries 3, 2, and 4 as exponents for T,

$$T^3ST^2ST^4S = \begin{pmatrix} 17 & -5 \\ 7 & -2 \end{pmatrix},$$

whose first column is what we are after. To get the right second column, we solve $\left(\begin{smallmatrix} 17 & 29 \\ 7 & 12 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 17 & -5 \\ 7 & -2 \end{smallmatrix}\right)M$ for M, which is $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right) = T^2$, so

$$\begin{pmatrix} 17 & 29 \\ 7 & 12 \end{pmatrix} = \begin{pmatrix} 17 & -5 \\ 7 & -2 \end{pmatrix} T^2 = T^3ST^2ST^4ST^2.$$

This is a different expression for $\left(\begin{smallmatrix} 17 & 29 \\ 7 & 12 \end{smallmatrix}\right)$ than the one we found in (2.5).

Corollary 2.3. *The group* $\mathrm{SL}_2(\mathbf{Z})$ *is generated by two matrices of finite order.*

Proof. We have $\mathrm{SL}_2(\mathbf{Z}) = \langle S, T\rangle = \langle S, ST\rangle$, where $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has order 4 and $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ has order 6. (As a transformation on $\mathfrak{h}$, ST has order 3 since $(ST)^3 = -I_2$, which acts trivially on $\mathfrak{h}$.) □

Corollary 2.4. *Every homomorphism* $\mathrm{SL}_2(\mathbf{Z}) \to \mathbf{C}^\times$ *has image in the* 12*th roots of unity.*

Proof. By the previous corollary, $\mathrm{SL}_2(\mathbf{Z})$ is generated by an element S of order 4 and an element ST of order 6. Therefore a homomorphism $\mathrm{SL}_2(\mathbf{Z}) \to \mathbf{C}^\times$ has image in the subgroup generated by μ_4 and μ_6, which is μ_{12}. □

Example 2.5. To show Corollary 2.4 is not an empty result, here is an example of a homomorphism $\chi\colon \mathrm{SL}_2(\mathbf{Z}) \to \mathbf{C}^\times$ whose image is all the 12th roots of unity:

$$\chi\begin{pmatrix} a & b \\ c & d \end{pmatrix} = e^{\frac{2\pi i}{12}((1-c^2)(bd+3(c-1)d+c+3)+c(a+d-3))}.$$

For instance, $\chi(S) = -i$ and $\chi(T) = e^{2\pi i/12} = -i(\frac{-1+\sqrt{3}i}{2})$. We are pulling χ out of nowhere; it is not obvious it is a homomorphism! It occurs naturally in the theory of modular forms:

for $\tau \in \mathfrak{h}$, the function $\Delta(\tau) = e^{2\pi i\tau}\prod_{n\geq 1}(1-e^{2\pi in\tau})^{24}$ satisfies $\Delta(\gamma\tau) = (c\tau+d)^{12}\Delta(\tau)$ for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z})$ and its 12th root $f(\tau) = e^{2\pi i\tau/12}\prod_{n\geq 1}(1-e^{2\pi in\tau})^2$ satisfies $f(\gamma\tau) = \chi(\gamma)(c\tau+d)f(\tau)$ for all $\gamma \in \mathrm{SL}_2(\mathbf{Z})$: χ is a multiplying factor here.

Corollary 2.6. *The group* $\mathrm{SL}_2(\mathbf{Z})$ *is generated by* $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ *and* $U = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$.

Proof. Both T and U are in $\mathrm{SL}_2(\mathbf{Z})$, so $\langle T, U\rangle \subset \mathrm{SL}_2(\mathbf{Z})$. Conversely, since $S = T^{-1}UT^{-1}$, $\langle T, U\rangle \supset \langle S, T\rangle = \mathrm{SL}_2(\mathbf{Z})$. □

Theorem 2.7. *Elements of finite order in* $\mathrm{SL}_2(\mathbf{Z})$ *have order* $1, 2, 3, 4,$ *or* 6.

Proof. The following examples show each of the indicated orders occur: I_2 has order 1, $-I_2$ has order 2. $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has order 4, $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ has order 6, and $(ST)^2 = \left(\begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has order 3.

Suppose $A \in \mathrm{SL}_2(\mathbf{Z})$ has finite order n, so $A^n - I_2 = O$. We want to show n is 1, 2, 3, 4, or 6. Since A is a 2×2 matrix with determinant 1, its characteristic polynomial is $X^2 - tX + 1$, where t is the trace of A. Therefore the Cayley–Hamilton theorem tells us $A^2 - tA + I_2 = O$. Since A is annihilated by both $X^n - 1$ and $X^2 - tX + 1$, it is annihilated by $\gcd(X^n - 1, X^2 - tX + 1)$. This gcd has a limited number of choices since the integer t is limited: t is the sum of the eigenvalues of A, which have to be roots of unity since A has finite order, so $|t| \leq 2$.

Case 1: $t = 2$. Since $X^n - 1$ has distinct roots and $X^2 - 2X + 1 = (X-1)^2$, we have $\gcd(X^n - 1, X^2 - 2X + 1)$ is $X - 1$. Thus $A - I_2 = O$, so $A = I_2$, which has order 1.

Case 2: $t = -2$. Since $X^n - 1$ has distinct roots and $X^2 + 2X + 1 = (X+1)^2$, we have $\gcd(X^n - 1, X^2 - 2X + 1) = X + 1$ if n is even and the gcd is 1 if n is odd. Since A is annihilated by the gcd, the gcd must be $X + 1$, so $A + I_2 = O$ and thus $A = -I_2$, so A has order 2.

Case 3: $t = 1$. Since $X^2 - X + 1$ is a factor of $X^3 + 1 = (X+1)(X^2 - X + 1)$, we have $A^3 = -I_2$, so $A^6 = I_2$. Since $A^2 - A + I_2 = O$ we can't have $A^2 = I_2$, so A has order 6.

Case 4: $t = -1$. Since $X^2 + X + 1$ is a factor of $X^3 - 1 = (X-1)(X^2 + X + 1)$, we have $A^3 = I_2$. Since $A^2 + A + I_2 = O$ we can't have $A = I_2$, so A has order 3.

Case 5: $t = 0$. In this case, $A^2 = -I_2$, so $A^4 = I_2$ and A has order 4. □

Remark 2.8. Obviously I_2 is the only matrix in $\mathrm{SL}_2(\mathbf{Z})$ of order 1. The proof above shows $-I_2$ is the only matrix in $\mathrm{SL}_2(\mathbf{Z})$ of order 2. In fact, $-I_2$ is the only matrix in $\mathrm{SL}_2(\mathbf{R})$ of order 2. (Many matrices in $\mathrm{GL}_2(\mathbf{Z})$ have order 2, such as $\left(\begin{smallmatrix} -1 & n \\ 0 & 1 \end{smallmatrix}\right)$.) Up to conjugation in $\mathrm{SL}_2(\mathbf{Z})$, a matrix of order 3 is conjugate to $\left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} -1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, a matrix of order 4 is conjugate to $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, and a matrix of order 6 is conjugate to $\left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right)$. A description of representatives for all the conjugacy classes in $\mathrm{SL}_2(\mathbf{Z})$ is at https://mathoverflow.net/questions/236151/.

3. Congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$

For an "arithmetically" defined group such as $\mathrm{SL}_2(\mathbf{Z})$ (a discrete group of integral matrices), its most important subgroups are those of finite index. The most basic way to find finite-index subgroups of $\mathrm{SL}_2(\mathbf{Z})$ is through the finite groups $\mathrm{SL}_2(\mathbf{Z}/(N))$. For each integer $N > 1$, the natural reduction map $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/(N))$ is a homomorphism with kernel

$$\Gamma(N) = \ker(\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/(N))) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod N \right\}.$$

Of course this subgroup is defined for $N = 1$ too, and $\Gamma(1) = \mathrm{SL}_2(\mathbf{Z})$. Each $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbf{Z})$, since $\mathrm{SL}_2(\mathbf{Z})/\Gamma(N)$ embeds into the finite group $\mathrm{SL}_2(\mathbf{Z}/(N))$, so each subgroup of $\mathrm{SL}_2(\mathbf{Z})$ containing some $\Gamma(N)$ has finite index.

Theorem 3.1. *The group* $\Gamma(2) = \{A \in \mathrm{SL}_2(\mathbf{Z}) : A \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod 2\}$ *is generated by the matrices* $-I_2$, T^2, *and* U^2, *where*

$$T^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad U^2 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Proof. All the matrices $-I_2$, T^2, and U^2 are in $\Gamma(2)$, so $\langle -I_2, T^2, U^2\rangle \subset \Gamma(2)$.

To get the reverse inclusion, we adapt the algebraic proof that $\mathrm{SL}_2(\mathbf{Z}) = \langle S, T\rangle$, except instead of the usual division theorem in $\mathbf{Z}$ we will use the modified division theorem in $\mathbf{Z}$: if $a, b \in \mathbf{Z}$ with $b \neq 0$ then $a = bq + r$ where $|r| \leq (1/2)|b|$ (perhaps $r < 0$).

Pick $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(2)$, so a and d are odd while b and c are even. If A has lower left entry 0 then $A = \pm\left(\begin{smallmatrix} 1 & m \\ 0 & 1 \end{smallmatrix}\right)$ for some $m \in \mathbf{Z}$. Since A is in $\Gamma(2)$, m must be even. Writing $m = 2k$, $A = \pm\left(\begin{smallmatrix} 1 & 2k \\ 0 & 1 \end{smallmatrix}\right) = \pm T^{2k} \in \langle -I_2, T^2\rangle$.

If the lower left entry of A is not 0 then we will multiply A by a suitable power of T^2 or U^2 on the left to reduce the value of $\max(|a|, |c|)$. Since a and c have opposite parity, $a \neq \pm c$, so $|a| \neq |c|$ and therefore $\max(|a|, |c|)$ is either $|a|$ or $|c|$ but not both.

If $|a| > |c|$ and $c \neq 0$, write $a = (2c)q + r$ where $|r| \leq (1/2)|2c| = |c|$. Then $T^{-2q}A = \left(\begin{smallmatrix} 1 & -2q \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} r & b-2qd \\ c & d \end{smallmatrix}\right)$, with $\max(|r|, |c|) = |c| < |a| = \max(|a|, |c|)$.

If $|a| < |c|$, then (since $a \neq 0$, as a is odd) write $c = (2a)q + r$ where $|r| \leq (1/2)|2a| = |a|$. Now $U^{-2q}A = \left(\begin{smallmatrix} 1 & 0 \\ -2q & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ r & d-2qb \end{smallmatrix}\right)$, with $\max(|a|, |r|) = |a| < |c| = \max(|a|, |c|)$.

Applying these two alternating steps, for some $g \in \langle T^2, U^2\rangle$ the lower left entry of gA is 0, so by the argument above $gA \in \langle -I_2, T^2\rangle$. Thus $A = g^{-1} \cdot gA \in \langle -I_2, T^2, U^2\rangle$. □

Theorem 3.2. *The natural map* $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/(N))$ *is onto.*

Proof. Pick $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z}/(N))$. There is $b' \equiv b \bmod N$ such that a and b' are relatively prime. The proof that b' exists can be given using the Chinese remainder theorem, and is omitted. (In practice it doesn't take long to find b' by checking $b, b+N, b+2N, \ldots$.) Since $(a, b') = 1$, there are x and y in $\mathbf{Z}$ such that $ax - b'y = 1$. Using this x and y, set

$$c' = c + y(1 - (ad - b'c)), \quad d' = d + x(1 - (ad - b'c)).$$

The matrix $\left(\begin{smallmatrix} a & b' \\ c' & d' \end{smallmatrix}\right)$ is in $\mathrm{SL}_2(\mathbf{Z})$ by a direct check and is congruent to $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \bmod N$. □

Example 3.3. Let $A = \left(\begin{smallmatrix} 18 & 14 \\ 4 & 2 \end{smallmatrix}\right)$, so $\det A = -20 \equiv 1 \bmod 21$. We will find a matrix in $\mathrm{SL}_2(\mathbf{Z})$ that reduces to A in $\mathrm{SL}_2(\mathbf{Z}/(21))$.

The top two entries, 18 and 14, are not relatively prime, but if we change 14 to $14+21 = 35$ then they are relatively prime. A solution to $18x - 35y = 1$ is $x = 2$ and $y = 1$, giving

$$\begin{pmatrix} 18 & 14 \\ 4 & 2 \end{pmatrix} \equiv \begin{pmatrix} 18 & 35 \\ 109 & 212 \end{pmatrix} \bmod 21$$

and the second matrix is in $\mathrm{SL}_2(\mathbf{Z})$.

The corresponding reduction homomorphism $\mathrm{GL}_2(\mathbf{Z}) \to \mathrm{GL}_2(\mathbf{Z}/(N))$ is usually *not* onto. The matrices in $\mathrm{GL}_2(\mathbf{Z})$ have determinant ± 1 and $(\mathbf{Z}/(N))^\times$ has units $u \not\equiv \pm 1 \bmod N$ once $N > 6$, so $\left(\begin{smallmatrix} u & 0 \\ 0 & 1 \end{smallmatrix}\right)$ in $\mathrm{GL}_2(\mathbf{Z}/(N))$ can't be the reduction of a matrix in $\mathrm{GL}_2(\mathbf{Z})$ since the determinants won't match mod N.

Corollary 3.4. *For all integers* $N > 1$, $\mathrm{SL}_2(\mathbf{Z})/\Gamma(N) \cong \mathrm{SL}_2(\mathbf{Z}/(N))$.

Proof. The reduction map $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/(N))$ is onto, with kernel $\Gamma(N)$. □

Corollary 3.5. *The finite group* $\mathrm{SL}_2(\mathbf{Z}/(N))$ *is generated by* 2 *elements of order* N.

Proof. Since $\mathrm{SL}_2(\mathbf{Z})$ is generated by $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $U = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ (Corollary 2.6), reducing modulo N shows $\mathrm{SL}_2(\mathbf{Z}/(N))$ is generated by the reductions of T and U, which each have order N. □

Corollary 3.6. *In* $\mathrm{SL}_2(\mathbf{Z})$*, the subgroup* $\langle S, T^2\rangle$ *has index* 3.

Proof. We start by showing $\Gamma(2) \subset \langle S, T^2\rangle$. By Theorem 3.1, it is enough to show the three generators $-I_2, T^2$, and U^2 of $\Gamma(2)$ are in $\langle S, T^2\rangle$: $-I_2 = S^2$, $T^2 = T^2$, and $U^2 = ST^{-2}S^{-1}$.

To compute the index of $\langle S, T^2\rangle$ in $\mathrm{SL}_2(\mathbf{Z})$, it is equivalent to work modulo $\Gamma(2)$ and compute the index of the subgroup generated by S and T^2 in $\mathrm{SL}_2(\mathbf{Z})/\Gamma(2) \cong \mathrm{SL}_2(\mathbf{Z}/(2))$. Since $T^2 \in \Gamma(2)$, $S \notin \Gamma(2)$, and $S^2 = -I_2 \in \Gamma(2)$, the group $\langle S, T^2\rangle/\Gamma(2)$ has order 2, hence its index in $\mathrm{SL}_2(\mathbf{Z}/(2))$ is $6/2 = 3$. □

If we replace $\langle S, T^2\rangle$ with $\langle S, T^m\rangle$ for $m > 2$ then there is no analogue of Corollary 3.6: $\langle S, T^m\rangle$ does not have finite index in $\mathrm{SL}_2(\mathbf{Z})$ for $m > 2$! A proof of this, shown to me by V. Pasol, is based on the action of $\mathrm{SL}_2(\mathbf{Z})$ on the primitive vectors (relatively prime coordinates) in $\mathbf{Z}^2$. This action of $\mathrm{SL}_2(\mathbf{Z})$ has one orbit, so if $\langle S, T^m\rangle$ has finite index in $\mathrm{SL}_2(\mathbf{Z})$ then the action of $\langle S, T^m\rangle$ on primitive vectors in $\mathbf{Z}^2$ would have finitely many orbits (the number of orbits would be at most its index in $\mathrm{SL}_2(\mathbf{Z})$), but it turns out there are infinitely many $\langle S, T^m\rangle$-orbits if $m > 2$, so $\langle S, T^m\rangle$ must have infinite index in $\mathrm{SL}_2(\mathbf{Z})$.

A subgroup of $\mathrm{SL}_2(\mathbf{Z})$ that contains some $\Gamma(N)$ is called a *congruence subgroup*. The meaning of the terminology is that such a subgroup can be described by a finite set of congruence conditions (namely being congruent modulo N to a set of representatives for a subgroup of $\mathrm{SL}_2(\mathbf{Z}/(N))$).

Example 3.7. The proof of Corollary 3.6 shows $\langle S, T^2\rangle$ is a congruence subgroup since $\Gamma(2) \subset \langle S, T^2\rangle$. The image of $\langle S, T^2\rangle$ in $\mathrm{SL}_2(\mathbf{Z})/\Gamma(2) \cong \mathrm{SL}_2(\mathbf{Z}/(2))$ is $\{\overline{I}_2, \overline{S}\}$, so we can describe $\langle S, T^2\rangle$ by congruence conditions modulo 2:

$$\langle S, T^2\rangle = \left\{ A \in \mathrm{SL}_2(\mathbf{Z}) : A \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bmod 2 \right\}.$$

Theorem 3.8. *The commutator subgroup* $\mathrm{SL}_2(\mathbf{Z})'$ *is a congruence subgroup with index* 12.

Proof. Since $\mathrm{SL}_2(\mathbf{Z})$ is generated by S with order 4 and ST with order 6, where $S^2 = (ST)^3 = -I_2$, the abelianization $\mathrm{SL}_2(\mathbf{Z})/\mathrm{SL}_2(\mathbf{Z})'$ is generated by $g = \overline{S}$ and $h = \overline{ST}$ where $g^4 = 1$, $h^6 = 1$, and $g^2 = h^3$. Because of commutativity, every element of $\mathrm{SL}_2(\mathbf{Z})/\mathrm{SL}_2(\mathbf{Z})'$ has the form $g^i h^j$ where $0 \le i \le 3$ and $0 \le j \le 5$. From $g^2 = h^3$, we can restrict i further to $0 \le i \le 1$. The number of such different $g^i h^j$ is at most 12, so $[\mathrm{SL}_2(\mathbf{Z}) : \mathrm{SL}_2(\mathbf{Z})'] \le 12$.

Next we will show in two ways that $\mathrm{SL}_2(\mathbf{Z})$ has a cyclic quotient group of order 12. That implies $[\mathrm{SL}_2(\mathbf{Z}) : \mathrm{SL}_2(\mathbf{Z})']$ is at least 12, so the index is 12. From the construction of the quotient group, we will see that $\Gamma(12) \subset \mathrm{SL}_2(\mathbf{Z})'$.

Method 1. If the reader is willing to believe the incredible homomorphism χ in Example 2.5 exists, then $\mathrm{SL}_2(\mathbf{Z})/\ker\chi \cong \mu_{12}$ is abelian of order 12, so $\mathrm{SL}_2(\mathbf{Z})' = \ker\chi$ by our index bounds. Since $\Gamma(12) \subset \ker\chi$ by a direct computation, $\mathrm{SL}_2(\mathbf{Z})'$ is a congruence subgroup.

Method 2. By Corollary 3.4, the natural reduction map $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{SL}_2(\mathbf{Z}/(N))$ is surjective for all N. By the Chinese remainder theorem,

$$\mathrm{SL}_2(\mathbf{Z}/(12)) \cong \mathrm{SL}_2(\mathbf{Z}/(3)) \times \mathrm{SL}_2(\mathbf{Z}/(4)),$$

and combining this with reduction mod 12 gives us a surjective group homomorphism

$$\mathrm{SL}_2(\mathbf{Z}) \twoheadrightarrow \mathrm{SL}_2(\mathbf{Z}/(3)) \times \mathrm{SL}_2(\mathbf{Z}/(4)) \tag{3.1}$$

with kernel $\Gamma(12)$. We will show $\mathrm{SL}_2(\mathbf{Z}/(3))$ has a quotient group of order 3 (necessarily cyclic) and $\mathrm{SL}_2(\mathbf{Z}/(4))$ has a cyclic quotient group of order 4. Combining this with (3.1) gives us a surjective group homomorphism

$$\mathrm{SL}_2(\mathbf{Z}) \twoheadrightarrow \mathrm{SL}_2(\mathbf{Z}/(3)) \times \mathrm{SL}_2(\mathbf{Z}/(4)) \twoheadrightarrow \mathbf{Z}/(3) \times \mathbf{Z}/(4)$$

with a cyclic target group of order 12 and $\Gamma(12)$ is contained in the kernel.

The group $\mathrm{SL}_2(\mathbf{Z}/(3))$ has order 24, and by an explicit calculation it has 8 elements with 2-power order. Therefore the group has a unique 2-Sylow subgroup, which must be normal, and $\mathrm{SL}_2(\mathbf{Z}/(3))/\{\text{2-Sylow}\}$ has order $24/8 = 3$. (The 2-Sylow subgroup is isomorphic to Q_8 and $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ is a complementary subgroup of order 3, so $\mathrm{SL}_2(\mathbf{Z}/(3)) \cong Q_8 \rtimes \mathbf{Z}/(3)$.)

The group $\mathrm{SL}_2(\mathbf{Z}/(4))$ has order 48. In this group, let $x = \left(\begin{smallmatrix} 3 & 2 \\ 0 & 3 \end{smallmatrix}\right)$ and $y = \left(\begin{smallmatrix} 3 & 0 \\ 2 & 3 \end{smallmatrix}\right)$. Then $xy = yx = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$. These three matrices have order 2, so the subgroup $H = \langle x, y \rangle = \{I_2, x, y, xy\}$ has order 4. The matrix $z = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z}/(4))$ has order 3 and normalizes H since $zxz^{-1} = y$, $zyz^{-1} = xy$, and $zxyz^{-1} = x$. So $N = \langle x, y, z \rangle = \langle z, H \rangle = \{z^i h : i \in \mathbf{Z}, h \in H\}$ is a subgroup of order 12 and index 4.

The subgroup N is normal in $\mathrm{SL}_2(\mathbf{Z}/(4))$. To prove that, it suffices to check $gNg^{-1} \subset N$ when g is $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ since these two matrices generate $\mathrm{SL}_2(\mathbf{Z}/(4))$ by the proof of Corollary 3.5. If $g = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ then $gxg^{-1} = x$, $gyg^{-1} = xy$, and $gzg^{-1} = \left(\begin{smallmatrix} 3 & 3 \\ 1 & 0 \end{smallmatrix}\right) = z^2 x$, while if $g = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ then $gxg^{-1} = xy$, $gyg^{-1} = y$, and $gzg^{-1} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) = z^2 y$.

The quotient group $\mathrm{SL}_2(\mathbf{Z}/(4))/N$ has order $48/12 = 4$. Let's show the quotient group is cyclic. By an explicit calculation, all 8 elements of $N - H$ have order 3, so H is the unique 2-Sylow subgroup of N. The subgroup $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ of $\mathrm{SL}_2(\mathbf{Z}/(4))$ is cyclic of order 4 and intersects H trivially, so it intersects N trivially. Thus $\mathrm{SL}_2(\mathbf{Z}/(4))/N \cong \langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle \cong \mathbf{Z}/(4)$. (Since $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ is a complementary subgroup to N in $\mathrm{SL}_2(\mathbf{Z}/(4))$, $\mathrm{SL}_2(\mathbf{Z}/(4)) \cong N \rtimes \mathbf{Z}/(4)$. There are 5 groups of order 12 up to isomorphism and the only one with more than one subgroup of order 3 is A_4, so $N \cong A_4$. Thus $\mathrm{SL}_2(\mathbf{Z}/(4)) \cong A_4 \rtimes \mathbf{Z}/(4)$.) □

Remark 3.9. The commutator subgroup $\mathrm{SL}_2(\mathbf{Z})'$ turns out to be generated by the two commutators $[S, T] = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 2 \end{smallmatrix}\right)$ and $[S, T^{-1}] = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$.

For $n \geq 2$, a subgroup of $\mathrm{SL}_n(\mathbf{Z})$ is called a congruence subgroup if for some $N \in \mathbf{Z}^+$ it contains the kernel of the natural reduction map $\mathrm{SL}_n(\mathbf{Z}) \to \mathrm{SL}_n(\mathbf{Z}/(N))$ (which is onto, by a longer proof than Theorem 3.2). As in the case $n = 2$, every congruence subgroup of $\mathrm{SL}_n(\mathbf{Z})$ has finite index. We will see in Section 4 that $\mathrm{SL}_2(\mathbf{Z})$ has finite-index subgroups that are *not* congruence subgroups. It is a theorem of Bass, Lazard, and Serre (1964) and Mennicke (1965) that for $n > 2$, all finite-index subgroups of $\mathrm{SL}_n(\mathbf{Z})$ are congruence subgroups.[2] So in this regard the first group $\mathrm{SL}_2(\mathbf{Z})$ in the series of groups $\mathrm{SL}_n(\mathbf{Z})$ is *misleading* as to the behavior of the groups for higher n. (Compare to: A_n is simple for $n \geq 5$, $\mathrm{PSL}_2(\mathbf{Z}/(p))$ is simple for prime $p \geq 5, \dots$)

Among finite-index subgroups in $\mathrm{SL}_2(\mathbf{Z})$, the congruence subgroups are particularly important in number theory because of the modular forms associated to them. The theta-function of a binary quadratic form and the L-function of an elliptic curve are both natural

[2] A more general theorem in this direction was proved by Bass, Milnor, and Serre (1967): for a number field K, with ring of integers $\mathcal{O}_K$, all finite-index subgroups of $\mathrm{SL}_n(\mathcal{O}_K)$ ($n \geq 3$) are congruence subgroups if and only if K has at least one real embedding.

sources of modular forms for congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$. All finite-index subgroups of $\mathrm{SL}_2(\mathbf{Z})$ are important in geometry since the orbit space of $\mathfrak{h}$ under such a group is (after adding a finite set of "missing points") a smooth projective curve over the complex numbers.

Most finite-index subgroups of $\mathrm{SL}_2(\mathbf{Z})$ are not congruence subgroups, in a quantifiable sense: among subgroups of index n in $\mathrm{SL}_2(\mathbf{Z})$, the proportion of congruence subgroups tends to 0 as $n \to \infty$.

4. Non-congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$

The existence of non-congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$ (subgroups of finite index not containing some $\Gamma(N)$) was first announced by Klein in 1879. The first examples in print appeared in 1887 by Fricke and Pick, independently. Their construction of the subgroups used generators to define them. We will describe a construction of such subgroups using kernels. It will be a nice application of the Jordan–Hölder theorem, as codified in the following lemma.

Lemma 4.1. *Let S be a finite simple group. If $G_1, \dots, G_m$ are nontrivial finite groups such that none have S as a composition factor then S is not a composition factor of $G_1 \times \cdots \times G_m$. In particular, S is not a quotient group of $G_1 \times \cdots \times G_m$.*

Proof. The direct product $G := G_1 \times \cdots \times G_m$ has a normal series

$$\{(e, \dots, e)\} \lhd G_1 \times \{e\} \times \cdots \times \{e\} \lhd G_1 \times G_2 \times \{e\} \times \cdots \times \{e\} \lhd \cdots \lhd G_1 \times G_2 \times \cdots \times G_m$$

whose factors are isomorphic to $G_1, \dots, G_m$. This normal series can be refined to a composition series, whose simple factors are the composition factors for the G_i's. By the Jordan-Hölder theorem, the factors in *every* composition series for G must be one of these simple factors, so a simple group S that is not a composition factor for the G_i's is not a composition factor for $G_1 \times \cdots \times G_m = G$.

If G has a quotient group isomorphic to S then it has a normal series $\{e\} \lhd N \lhd G$ with $G/N \cong S$. This normal series for G can be extended to a composition series of G with S as the top factor, so S is a composition factor of G, which is a contradiction. □

Theorem 4.2. *For $n \geq 6$, the alternating group A_n is not a quotient of $\mathrm{SL}_2(\mathbf{Z}/(N))$ for each $N \geq 2$.*

Proof. Write $N = p_1^{r_1} \cdots p_m^{r_m}$, so $\mathbf{Z}/(N) \cong \prod_{i=1}^m \mathbf{Z}/(p_i^{r_i})$ by the Chinese remainder theorem. Then

$$\mathrm{SL}_2(\mathbf{Z}/(N)) \cong \prod_{i=1}^m \mathrm{SL}_2(\mathbf{Z}/(p_i^{r_i})),$$

so by Lemma 4.1 it suffices to show A_n for $n \geq 6$ is not a composition factor of $\mathrm{SL}_2(\mathbf{Z}/(p^r))$ for each prime power p^r.

To write down a composition series for $\mathrm{SL}_2(\mathbf{Z}/(p^r))$, we start with the reduction map $\mathrm{SL}_2(\mathbf{Z}/(p^r)) \to \mathrm{SL}_2(\mathbf{Z}/(p))$, which is onto. Let K be its kernel, so we have the normal series

$$\{I_2 \bmod p^r\} \lhd K \lhd \mathrm{SL}_2(\mathbf{Z}/(p^r)),$$

whose factors (up to isomorphism) are K and $\mathrm{SL}_2(\mathbf{Z}/(p))$. Therefore the composition factors for $\mathrm{SL}_2(\mathbf{Z}/(p^r))$ are the composition factors for K and for $\mathrm{SL}_2(\mathbf{Z}/(p))$.

What are the composition factors for K? The group $K = \{A \in \mathrm{SL}_2(\mathbf{Z}/(p^r)) : A \equiv I_2 \bmod p\}$ is a p-group: if $A \equiv I_2 \bmod p$ then $A^{p^k} \equiv I_2 \bmod p^{k+1}$ for all $k \geq 0$ (by induction), so $A^{p^{r-1}} \equiv I_2 \bmod p^r$. Therefore all elements of K have p-power order, and a finite group

whose elements have p-power order is a p-group (Cauchy!), so K is a p-group. (The exact order of K can be computed, but that's not important for us.) The composition factors of a finite p-group, such as K, are all cyclic of order p.

We now turn to $\mathrm{SL}_2(\mathbf{Z}/(p))$. For $p \geq 5$, a composition series for $\mathrm{SL}_2(\mathbf{Z}/(p))$ is $\{I_2\} \lhd \{\pm I_2\} \lhd \mathrm{SL}_2(\mathbf{Z}/(p))$, since $\mathrm{PSL}_2(\mathbf{Z}/(p)) = \mathrm{SL}_2(\mathbf{Z}/(p))/\{\pm I_2\}$ is simple for $p \geq 5$. Thus the composition factors for $\mathrm{SL}_2(\mathbf{Z}/(p))$ when $p \geq 5$ are $\mathbf{Z}/(2)$ and $\mathrm{PSL}_2(\mathbf{Z}/(p))$. What about for $p < 5$? Since $\mathrm{SL}_2(\mathbf{Z}/(2)) = \mathrm{GL}_2(\mathbf{Z}/(2)) \cong S_3$ and $\mathrm{SL}_2(\mathbf{Z}/(3))/\{\pm I_2\} \cong A_4$, the composition factors of $\mathrm{SL}_2(\mathbf{Z}/(2))$ and $\mathrm{SL}_2(\mathbf{Z}/(3))$ are cyclic (of order 2 or 3).

Thus for all prime powers p^r, $\mathrm{SL}_2(\mathbf{Z}/(p^r))$ has only one nonabelian composition factor when $p \geq 5$, namely $\mathrm{PSL}_2(\mathbf{Z}/(p))$. If $p \leq 3$ then all composition factors of $\mathrm{SL}_2(\mathbf{Z}/(p^r))$ are cyclic. So if A_n for $n \geq 6$ were a composition factor of some $\mathrm{SL}_2(\mathbf{Z}/(p^r))$, A_n would have to be isomorphic to $\mathrm{PSL}_2(\mathbf{Z}/(p))$ for some prime $p \geq 5$. The problem with this is that an alternating group and a projective special linear group hardly ever have the same size. The group $\mathrm{PSL}_2(\mathbf{Z}/(p))$ has order $(p^2-1)p/2$, so we ask: when can $(p^2-1)p/2 = n!/2$, or equivalently

$$(p-1)p(p+1) = n!?$$

(The punctuation there is: factorial, question mark.) If $n < p$ then $n!$ is not divisible by p and we have a contradiction. If $n = p$ then dividing both sides by $(p-1)p$ gives $p+1 = (p-2)!$, whose only solution is $p = 5$ (and $n = 5$). If $n = p+1$ then dividing both sides by $(p-1)p(p+1)$ gives $1 = (p-2)!$ so $p = 3$ (but we need $p \geq 5$). If $n \geq p+2$ then there is too much remaining on the right side when we divide through by $(p-1)p(p+1)$. Since we only found a solution when $p = n = 5$ (and indeed $\mathrm{PSL}_2(\mathbf{Z}/(5)) \cong A_5$), for $n \geq 6$ the group A_n is not a quotient group of $\mathrm{SL}_2(\mathbf{Z}/(N))$ for all $N \geq 2$. □

The bound $n \geq 6$ in Theorem 4.2 is optimal: $A_5 \cong \mathrm{PSL}_2(\mathbf{Z}/(5))$, $A_4 \cong \mathrm{PSL}_2(\mathbf{Z}/(3))$, and A_3 is isomorphic to the quotient of $\mathrm{SL}_2(\mathbf{Z}/(3))$ by its normal 2-Sylow subgroup.

While Theorem 4.2 says most A_n's do not arise as the quotient of the finite groups $\mathrm{SL}_2(\mathbf{Z}/(N))$, we will show most A_n's do arise as the quotient of $\mathrm{SL}_2(\mathbf{Z})$.

Theorem 4.3. *For $n \geq 9$, A_n is a quotient of $\mathrm{SL}_2(\mathbf{Z})$.*

Proof. We will actually get A_n as a quotient group of $\mathrm{PSL}_2(\mathbf{Z}) = \mathrm{SL}_2(\mathbf{Z})/\{\pm I_2\}$, but that also makes it a quotient group of $\mathrm{SL}_2(\mathbf{Z})$ by composing with the natural reduction map $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{PSL}_2(\mathbf{Z})$.

There are two things that make this result hold: A_n (for $n \geq 9$) is generated by two elements of order 2 and 3, and $\mathrm{PSL}_2(\mathbf{Z})$ is also *freely* generated by two elements of order 2 and 3. We will explain, in order, what these mean.

In 1901, G. A. Miller proved that for $n \geq 9$, the group A_n is generated by an element of order 2 and an element of order 3. His proof gave generators whose construction depends on a choice of a prime between $n/2$ and n when $n \geq 12$, and for smaller n he left it as an exercise for the reader to find elements of order 2 and 3 generating A_n. In 1971, Dey and Wiegold (unaware of Miller's work) gave an explicit pair of generators of order 2 and 3 for A_n without needing an auxiliary prime.

To see the group $\mathrm{PSL}_2(\mathbf{Z})$ is generated by elements of order 2 and 3, we work with the cosets of S and ST. Set $x = \overline{S} = \overline{\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)}$ and $y = \overline{ST} = \overline{\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)}$. Then $x^2 = -I_2 = I_2$ and $y^3 = -I_2 = I_2$ in $\mathrm{PSL}_2(\mathbf{Z})$. Because S and ST generate $\mathrm{SL}_2(\mathbf{Z})$, every element of $\mathrm{PSL}_2(\mathbf{Z})$ can be written as a word in x and y. Taking into account that x has order 2 and y has

order 3, we can write each product of x's and y's in the "reduced" form

$$y^{i_0}xy^{i_1}x\cdots y^{i_{n-1}}xy^{i_n},$$

where the exponents i_j are regarded in $\mathbf{Z}/(3)$ and all these exponents are nonzero modulo 3 except perhaps i_0 and i_n. It turns out such a representation is unique; that's the meaning of saying x and y freely generate $\mathrm{PSL}_2(\mathbf{Z})$: there are no relations on x and y in the group except for those that are logical consequences of $x^2 = 1$ and $y^3 = 1$. (For a proof, see Appendix C.) Because of the unique expression of each element of $\mathrm{PSL}_2(\mathbf{Z})$ as a word in x and y, each assignment to x and y of elements of order 2 and 3 in another group uniquely extends to a homomorphism from $\mathrm{PSL}_2(\mathbf{Z})$ to that group. Therefore, choosing a generating pair of order 2 and 3 for A_n, and sending x and y to them, leads to a homomorphism from $\mathrm{PSL}_2(\mathbf{Z})$ *onto* A_n. □

Example 4.4. The group A_9 turns out to be generated by

$$(14)(29)(37)(56) \text{ and } (123)(456)(789),$$

so one surjective homomorphism from $\mathrm{SL}_2(\mathbf{Z})$ to A_9 is the composite $\mathrm{SL}_2(\mathbf{Z}) \to \mathrm{PSL}_2(\mathbf{Z}) \to A_9$ where the first map is reduction mod $\pm I_2$ and the second is determined by $\overline{S} \mapsto (14)(29)(37)(56)$ and $\overline{ST} \mapsto (123)(456)(789)$.

Remark 4.5. The group A_n is generated by elements of order dividing 2 and 3 for all $n \geq 3$ except for $n = 6$, 7, and 8. Since the behavior is uniform once $n \geq 9$, we stated Theorem 4.3 in the simpler way excluding small n.

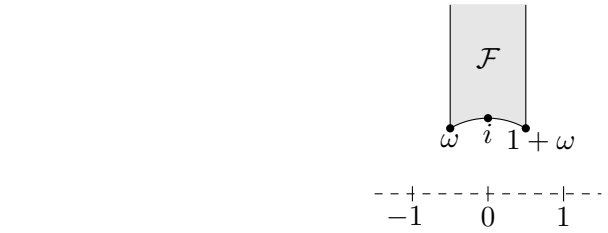
By Theorem 4.3, for all $n \geq 9$ there is a surjective homomorphism $\mathrm{SL}_2(\mathbf{Z}) \to A_n$. The (mysterious) kernel of such a homomorphism is a subgroup of $\mathrm{SL}_2(\mathbf{Z})$ with finite index. The kernel can't contain some $\Gamma(N)$, since otherwise A_n would be realizable as a quotient group of $\mathrm{SL}_2(\mathbf{Z}/(N))$, which is impossible by Theorem 4.2, so the kernel is a (finite-index) non-congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$. This description of the subgroup as a kernel does not provide an easily accessible set of generators for it, but it does provide a recipe for determining whether an individual matrix is in the subgroup. Here is the procedure. For $n \geq 9$, pick two elements x and y in A_n of respective orders 2 and 3 such that $A_n = \langle x, y\rangle$. For a matrix in $\mathrm{SL}_2(\mathbf{Z})$, write it (up to an overall sign) as a product of S and ST. Turn that word in S and ST into a word in x and y. The matrices whose corresponding word in x and y is trivial in A_n form a non-congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$.

Most of the nonabelian finite simple groups, not just the alternating groups A_n for $n \geq 9$, turn out to be generated by a pair of elements with order 2 and 3, and thus most nonabelian finite simple groups are quotient groups of $\mathrm{SL}_2(\mathbf{Z})$ by the same argument used for most A_n's. (Exceptions to this occur among some simple matrix groups defined in characteristics 2 and 3, such as the infinite family of Suzuki groups, whose orders are not divisible by 3.) A nonabelian finite simple group that is not isomorphic to $\mathrm{PSL}_2(\mathbf{Z}/(p))$ for $p \geq 5$ is not a quotient group of $\mathrm{SL}_2(\mathbf{Z}/(N))$ for $N \geq 2$ by the same Jordan-Hölder argument given before for alternating groups. So there is a tremendous number of ways to construct non-congruence subgroups of $\mathrm{SL}_2(\mathbf{Z})$, because most finite simple groups are quotients of $\mathrm{SL}_2(\mathbf{Z})$ but are not quotients of some $\mathrm{SL}_2(\mathbf{Z}/(N))$.

Amusingly, for large n (*e.g.*, $n \geq 28$), $\mathrm{SL}_n(\mathbf{Z})$ is generated by a pair of elements of order 2 and 3, so $\mathrm{SL}_n(\mathbf{Z})$ is a quotient group of $\mathrm{SL}_2(\mathbf{Z})$! (The group $\mathrm{SL}_3(\mathbf{Z})$ is known not to be generated by a pair of elements of order 2 and 3. I am not sure of the status of $4 \leq n \leq 27$.)

Appendix A. Stabilizers in $\mathrm{SL}_2(\mathbf{Z})$

For $z \in \mathfrak{h}$, let $\mathrm{Stab}_z = \{g \in \mathrm{SL}_2(\mathbf{Z}) : g(z) = z\}$ be its stabilizer subgroup in $\mathrm{SL}_2(\mathbf{Z})$. Since $\pm I_2$ both act trivially on $\mathfrak{h}$, they are in Stab_z. In this appendix we will compute Stab_z for all z, and it will turn out usually to be $\pm I_2$ but sometimes it is larger.



Example A.1. We show $\mathrm{Stab}_i = \langle S \rangle$, which is cyclic of order 4.[3]

For a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z})$ to fix i is equivalent to $ai + b = (ci + d)i = -c + di$, so $a = d$ and $b = -c$, or $d = a$ and $c = -b$. Then $1 = ad - bc = a^2 + b^2$. Since a and b are integers, $(a, b) = (\pm 1, 0)$ or $(0, \pm 1)$, so $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ is $\pm I_2$ or $\pm S$. Conversely, these four matrices all fix i, so $\mathrm{Stab}_i = \{\pm I_2, \pm S\} = \langle S \rangle$.

Example A.2. Let $\omega = e^{2\pi i/3} = (-1 + \sqrt{3}i)/2$, which is the nontrivial cube root of unity in $\mathfrak{h}$. Let's show $\mathrm{Stab}_\omega = \langle ST \rangle$, which is cyclic of order 6.

To have $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z})$ fix ω is equivalent to $a\omega + b = (c\omega + d)\omega = c\omega^2 + d\omega$. Since $\omega^2 = -1 - \omega$, $a\omega + b = (d - c)\omega - c$, so $b = -c$ and $a = d - c = d + b$. Thus $c = -b$ and $d = a - b$, so $1 = ad - bc = a(a - b) + b^2 = a^2 - ab + b^2$. Writing this as $1 = (a - b/2)^2 + (3/4)b^2$, the only possible values of b are 0, 1, and -1, and by taking cases we get $(a, b) = \pm(1, 0)$, $\pm(0, 1)$, or $\pm(1, 1)$, so $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ -b & a-b \end{smallmatrix}\right)$ has 6 possible values that turn out to be the powers of $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$. Since ST fixes ω (check!), its powers fix ω and thus $\mathrm{Stab}_\omega = \langle ST \rangle$.

Theorem A.3. *When $\mathrm{SL}_2(\mathbf{Z})$ acts on $\mathfrak{h}$ by linear fractional transformations, the stabilizer of a point $z \in \mathfrak{h}$ can be described as follows.*

1. *If z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i then $\mathrm{Stab}_z \cong \mathbf{Z}/(4)$.*
2. *If z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of ω then $\mathrm{Stab}_z \cong \mathbf{Z}/(6)$.*
3. *If z is not in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i or ω then $\mathrm{Stab}_z = \{\pm I_2\}$.*

Proof. Points in the same orbit of a group action have conjugate stabilizer subgroups, and conjugate subgroups are isomorphic, so the first two parts of the theorem follow from the calculations in Examples A.1 and A.2

It remains to show the third part: if $z \in \mathfrak{h}$ is not in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i or ω then the only $g \in \mathrm{SL}_2(\mathbf{Z})$ such that $g(z) = z$ are $\pm I_2$. We will prove the contrapositive: if Stab_z contains a matrix that is not $\pm I_2$ then z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i or ω.

Step 1: If $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)z = z$ and $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \neq \pm I_2$ then $c \neq 0$ and $d + a$ is 0, 1, or -1.

The condition $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)z = z$ is equivalent to $az + b = cz^2 + dz$, so

$$cz^2 + (d - a)z - b = 0. \tag{A.1}$$

We'll show $c \neq 0$ by contradiction. Suppose $c = 0$, so $b = (d - a)z$. From $1 = ad - bc = ad$, we get $a = d = \pm 1$ since a and d are integers. Thus $b = 0$, so $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right) = \pm I_2$.

[3] When $\mathrm{SL}_2(\mathbf{R})$ acts on $\mathfrak{h}$, the stabilizer subgroup of i is $\mathrm{SO}_2(\mathbf{R}) \subset \mathrm{SL}_2(\mathbf{R})$. See the appendix of https://kconrad.math.uconn.edu/blurbs/grouptheory/SL(2,R).pdf. All points of $\mathfrak{h}$ are in the same $\mathrm{SL}_2(\mathbf{R})$-orbit and thus their stabilizer subgroups in $\mathrm{SL}_2(\mathbf{R})$ are all conjugate to $\mathrm{SO}_2(\mathbf{R})$. Here we use $\mathrm{SL}_2(\mathbf{Z})$-orbits.

That contradicts $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \neq \pm I_2$ so $c \neq 0$. For (A.1) to have a root z in $\mathfrak{h}$, the discriminant $(d-a)^2 + 4bc$ is negative. Since $bc = ad - 1$,

$$(d-a)^2 + 4bc = d^2 - 2ad + a^2 + 4(ad-1) = d^2 + 2ad + a^2 - 4 = (d+a)^2 - 4. \tag{A.2}$$

Therefore $(d+a)^2 < 4$, so $|d+a| < 2$, which implies the integer $d+a$ is 1, 0, or -1.

Step 2: If $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z = z$ with $c \neq 0$ and $d + a = 0$ then z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i.

Since $c \neq 0$, by Step 1 we have $cz^2 + (d-a)z - b = 0$ and the quadratic polynomial on the left side has discriminant $(d+a)^2 - 4 = -4$ by (A.2), so by the quadratic formula

$$z = \frac{-(d-a) \pm \sqrt{-4}}{2c} = \frac{2a \pm 2i}{2c} = \frac{a \pm i}{c}.$$

Since $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z = \left(\begin{smallmatrix} -a & -b \\ -c & -d \end{smallmatrix}\right) z$, we can change signs on the matrix entries so that $c > 0$, and then $z = (a+i)/c$ since $z \in \mathfrak{h}$.

To show z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i, let z' be the point in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of z that's in $\mathcal{F}$. We'll show $z' = i$. Since $\mathrm{Stab}_{z'}$ is conjugate to Stab_z, $\mathrm{Stab}_{z'}$ contains a matrix $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ where the trace $d' + a'$ is 0, and necessarily $c' \neq 0$ by Step 1. Using the same calculations as in the previous paragraph starting from $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right) z' = z'$, we have $z' = (a'+i)/c'$ where without loss of generality we took $c' > 0$. Since $z' \in \mathcal{F}$, the condition $\mathrm{Im}\, z' \geq \sqrt{3}/2$ is the same as $c' \leq 2/\sqrt{3} \approx 1.15$, so $c' = 1$. Then $\mathrm{Re}(z') = a'/c' = a'$ is in $[-1/2, 1/2]$, so $a' = 0$ since a' is an integer. Thus $z' = i$.

Step 3: If $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z = z$ with $c \neq 0$ and $d + a = \pm 1$ then z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of ω.

Since $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z = \left(\begin{smallmatrix} -a & -b \\ -c & -d \end{smallmatrix}\right) z$, by changing the signs of all the matrix entries if necessary then we can suppose $d + a = 1$. By Step 1, $cz^2 + (d-a)z - b = 0$ where the quadratic polynomial on the left side has discriminant $(d+a)^2 - 4 = -3$ by (A.2), so

$$z = \frac{-(d-a) \pm \sqrt{-3}}{2c} = \frac{2a - 1 \pm \sqrt{3}i}{2c}.$$

If we have the $+$ sign then $z = (2a - 1 + \sqrt{3}i)/(2c) = (a + \omega)/c$. Then from $z \in \mathfrak{h}$ we get $c > 0$. If instead we have the $-$ sign then $z = (2a - 1 - \sqrt{3}i)/(2c) = (a - (1+\omega))/c = ((1-a) + \omega)/(-c)$, where the denominator $-c$ must be positive since $z \in \mathfrak{h}$.

As in Step 2, there is a number z' in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of z that lies in $\mathcal{F}$. Then $\mathrm{Stab}_{z'}$ contains a matrix $\left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ with trace 1 and (by Step 1) $c' \neq 0$. Calculations as in the previous paragraph show $z' = (A + \omega)/C$ for integers A and C such that $C > 0$. Since $z' \in \mathcal{F}$, $\mathrm{Im}\, z' \geq \sqrt{3}/2$ and $|\mathrm{Re}(z')| \leq 1/2$. The first inequality tells us $\sqrt{3}/(2C) \geq \sqrt{3}/2$, so $C \leq 1$. Thus $C = 1$, so the condition $\mathrm{Re}(z') \in [-1/2, 1/2]$ becomes $-1/2 \leq A - 1/2 \leq 1/2$, so $0 \leq A \leq 1$. Hence $z' = \omega$ or $z' = 1 + \omega = T(\omega)$, so z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of ω. □

Theorem A.4. *A number $z \in \mathfrak{h}$ is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i if and only if $z = (a+i)/c$ where a and c are integers such that $c > 0$ and $c \mid (a^2+1)$, and z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of ω if and only if $z = (a+\omega)/c$ where a and c are integers such that $c > 0$ and $c \mid (a^2 - a + 1)$.*

Proof. In Step 2 of the proof of Theorem A.3, where $d + a = 0$, we have $1 = ad - bc = a(-a) - bc = -a^2 - bc$, so $a^2 + 1 = c(-b)$. That suggests the following: if $z = (a+i)/c$ for some $a, c \in \mathbf{Z}$ such that $c > 0$ and $c \mid (a^2+1)$, then define the integer b by the condition $a^2 + 1 = c(-b)$ and define $d = -a$. Then $ad - bc = -a^2 + (a^2+1) = 1$ and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az+b}{cz+d} = \frac{a(a+i)/c + b}{c(a+i)/c + d} = \frac{a^2 + ai + bc}{ac + ci + dc} = \frac{a^2 + ai - 1 - a^2}{ci} = \frac{a+i}{c} = z.$$

So Stab_z contains the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z})$ with trace 0. By Step 2 of the proof of Theorem A.3, z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i.

Conversely, suppose z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i, so $z = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)i = (Ai+B)/(Ci+D)$ where $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z})$. Then

$$z = \frac{(B+Ai)(D-Ci)}{(D+Ci)(D-Ci)} = \frac{AC+BD+(AD-BC)i}{D^2+C^2} = \frac{AC+BD+i}{C^2+D^2}.$$

This is $(a+i)/c$ for $a = AC+BD$ and $c = C^2+D^2$, so $c > 0$. From $a+i = (B+Ai)(D-Ci)$, taking the norm of both sides in $\mathbf{Z}[i]$ shows $a^2+1 = (B^2+A^2)(D^2+C^2) = (B^2+A^2)c$, so $c \mid (a^2+1)$.

In Step 3 of the proof of Theorem A.3, if $d+a = 1$ then we have $1 = ad-bc = a(1-a)-bc = a-a^2-bc$, so $a^2-a+1 = c(-b)$. Therefore if $z = (a+\omega)/c$ for some $a, c \in \mathbf{Z}$ such that $c > 0$ and $c \mid (a^2-a+1)$, then define the integer b by the condition $a^2-a+1 = c(-b)$ and define $d = 1-a$. Then $ad-bc = a(1-a)+a^2-a+1 = 1$ and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{a(a+\omega)/c+b}{c(a+\omega)/c+d} = \frac{a^2+a\omega+bc}{ac+c\omega+dc} = \frac{a^2+a\omega-1+a-a^2}{c(1+\omega)} = \frac{a(1+\omega)-1}{c(1+\omega)},$$

which is $(a+\omega)/c$ since $-1/(1+\omega) = \omega$. We have shown Stab_z contains a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbf{Z})$ with trace 1. By Step 3 of the proof of Theorem A.3, z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of ω.

Now suppose, conversely, that z is in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of ω. Then $z = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)\omega = (A\omega+B)/(C\omega+D)$ for some $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z})$. Therefore

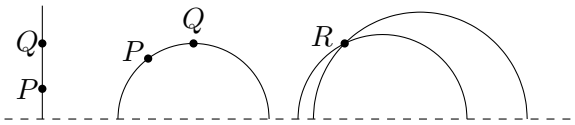
$$z = \frac{(B+A\omega)(D+C\overline{\omega})}{(D+C\omega)(D+C\overline{\omega})} = \frac{(BD+AC)+AD\omega+BC\overline{\omega}}{D^2-DC+C^2} = \frac{(AC-BC+BD)+\omega}{C^2-CD+D^2}.$$

This has the form $(a+\omega)/c$ where $a = AC-BC+BD$ and $c = C^2-CD+D^2$, so $c > 0$. From $a+\omega = (B+A\omega)(D+C\overline{\omega})$, taking norms of both sides in $\mathbf{Z}[\omega]$ gives us $a^2-a+1 = (B^2-BA+A^2)(D^2-DC+C^2) = (B^2-BA+A^2)c$, so $c \mid (a^2-a+1)$. □

Appendix B. The Hyperbolic Plane

The hyperbolic plane is the upper half-plane $\mathfrak{h}$ with a definition of lines (also called geodesics) and distances that differ from the usual meaning of these notions in the Euclidean plane $\mathbf{R}^2$.

Lines in $\mathfrak{h}$ are the vertical lines in $\mathfrak{h}$ or the semicircles in $\mathfrak{h}$ that meet the x-axis in a 90-degree angle (the x-axis is the diameter of the semicircle). In the picture below, if P and Q have the same x-coordinate then the line $\overline{PQ}$ through P and Q is the part of the usual Euclidean (vertical) line through P and Q that is in $\mathfrak{h}$. If P and Q do not have the same x-coordinate then $\overline{PQ}$ is the unique Euclidean semicircle through P and Q with diameter on the x-axis.



On the right side of the picture two lines drawn through a point R not on $\overline{PQ}$ don't intersect $\overline{PQ}$. This contradicts the parallel postulate of Euclidean geometry, which says a point not on a line L has exactly one line through it that doesn't meet L. In $\mathbf{R}^2$ the parallel postulate is true, but in $\mathfrak{h}$ it is not.

The hyperbolic distance between two points P and Q in $\mathfrak{h}$ is defined using integration along $\overline{PQ}$:

$$d_H(P,Q) = \int_P^Q \frac{\sqrt{(dx/dt)^2 + (dy/dt)^2}}{y(t)}\, dt,$$

where the integral is taken along the hyperbolic line in $\mathfrak{h}$ from P to Q using a smooth parametrization $(x(t), y(t))$ of the segment in $\overline{PQ}$ from P to Q.

Example B.1. To compute the distance between $y_0 i$ and $y_1 i$, parametrize the vertical line between them as $(x(t), y(t)) = (0, (1-t)y_0 + ty_1)$ for $0 \le t \le 1$. Then

$$d_H(y_0 i, y_1 i) = \int_0^1 \frac{\sqrt{0^2 + (y_1 - y_0)^2}}{(1-t)y_0 + ty_1}\, dt = |\log y_1 - \log y_0| = |\log(y_1/y_0)|.$$

For example, $d_H(yi, i) = |\log y|$ and the midpoint between $y_0 i$ and $y_1 i$ when $y_0 \neq y_1$ is $\sqrt{y_0 y_1}\, i$, which is (always) different from the Euclidean midpoint.

The action of $\mathrm{SL}_2(\mathbf{R})$ on $\mathfrak{h}$ by linear fractional transformations preserves hyperbolic distances: for each $A \in \mathrm{SL}_2(\mathbf{R})$, $d_H(A(P), A(Q)) = d_H(P,Q)$ for all P and Q in $\mathfrak{h}$. A function $\mathfrak{h} \to \mathfrak{h}$ that preserves distances is called an isometry, and $\mathrm{SL}_2(\mathbf{R})$ acting by linear fractional transformation is the group of all orientation-preserving isometries of the hyperbolic plane.[4] An example of an isometry of $\mathfrak{h}$ that reverses orientation is $\tau \mapsto -\overline{\tau}$, or equivalently $x + yi \mapsto -x + yi$, and every orientation-reversing isometry is this example composed with the action by a matrix in $\mathrm{SL}_2(\mathbf{R})$.

Appendix C. Generators and Relations for $\mathrm{PSL}_2(\mathbf{Z})$

By Corollary 2.3, $\mathrm{SL}_2(\mathbf{Z})$ is generated by $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, which have respective orders 4 and 6. Set $R = ST$, so every element of $\mathrm{SL}_2(\mathbf{Z})$ is a product of S's and R's. Since $S^2 = R^3 = -I_2$, every product of S's and R's can be brought to the form

$$(-I_2)^a R^{i_0} S R^{i_1} S \cdots R^{i_{n-1}} S R^{i_n},$$

where $a \in \mathbf{Z}/(2)$ and $i_j \not\equiv 0 \bmod 3$ for $0 < j < n$; that is, the outer R-powers R^{i_0} and R^{i_n} might be $\pm I_2$ but the inner R-powers are not. (If $n = 0$ this product is $(-I_2)^a R^{i_0}$.) We can't consider the exponents i_j to be in $\mathbf{Z}/(3)$ because R does not have order 3. However, if we pass to $\mathrm{PSL}_2(\mathbf{Z}) = \mathrm{SL}_2(\mathbf{Z})/\{\pm I_2\}$ then $x := \overline{S}$ has order 2, $y := \overline{R}$ has order 3 and every element of $\mathrm{PSL}_2(\mathbf{Z})$ has the form

$$y^{i_0} x y^{i_1} x \cdots y^{i_{n-1}} x y^{i_n}, \quad i_j \in \mathbf{Z}/(3),\ i_j \not\equiv 0 \bmod 3 \text{ for } 0 < j < n. \tag{C.1}$$

Note the condition on the exponents. It means the powers of y on the inside of the product are all nontrivial, but we do allow trivial y-powers for the outer terms. (Thus $x = y^0 x y^0$, for instance.)

Theorem C.1. *Each element of* $\mathrm{PSL}_2(\mathbf{Z})$ *can be written in the form* (C.1) *in exactly one way.*

Since x has order 2 and y has order 3 in $\mathrm{PSL}_2(\mathbf{Z})$, that (C.1) provides a unique representation for elements of $\mathrm{PSL}_2(\mathbf{Z})$ is described by saying $\mathrm{PSL}_2(\mathbf{Z})$ is a free product of $\mathbf{Z}/(2)$ and $\mathbf{Z}/(3)$.

[4] Strictly speaking, since A and $-A$ act in the same way, the group of orientation-preserving isometries is $\mathrm{SL}_2(\mathbf{R})/\{\pm I_2\}$.

Proof. Our argument is taken from [2, p. 12]. (There is a similar proof in [3, Prop. V.4.o].)

To start, suppose we can write the identity element of $\mathrm{PSL}_2(\mathbf{Z})$ in this way:

$$1 = y^{i_0}xy^{i_1}x\cdots y^{i_{n-1}}xy^{i_n},$$

If $n = 0$, so the product on the right is y^{i_0}, this representation works using $i_0 = 0$ and not for other i_0 in $\mathbf{Z}/(3)$. If $n = 1$, the right side is y^ixy^j for $i, j \in \mathbf{Z}/(3)$. A computation shows the only such product equal to the identity in $\mathrm{PSL}_2(\mathbf{Z})$ is that with $i, j \equiv 0 \bmod 3$. To show a representation of 1 as (C.1) is impossible for $n \geq 2$, assume there is such a representation and let n be minimal. Multiply both sides of the above equation on the left by y^{-i_0} and on the right by y^{i_0}:

$$1 = xy^{i_1}x\cdots y^{i_{n-1}}xy^{i_n+i_0}, \tag{C.2}$$

The inner exponents $i_1, \ldots, i_{n-1}$ are all nonzero modulo 3. We will show by contradiction that the last exponent is nonzero modulo 3 as well. If $i_n + i_0 \equiv 0 \bmod 3$ then we get

$$1 = xy^{i_1}x\cdots y^{i_{n-1}}x,$$

so multiplying both sides on the left and right by $x = x^{-1}$ gives

$$1 = y^{i_1}x\cdots y^{i_{n-1}}.$$

By the minimality of n, we must have $n - 1 = 0$, so $n = 1$. But $n \geq 2$. Therefore $i_n + i_0 \not\equiv 0 \bmod 3$. So in (C.2), we have written 1 as a product of xy's and xy^2's. Now let's looks at what xy and xy^2 actually are, as matrices (up to sign):

$$SR = S^2T = -T = -\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad SR^2 = -TST = -\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

If, in $\mathrm{PSL}_2(\mathbf{Z})$, we have a product of xy's and xy^2's equal to 1 then that means in $\mathrm{SL}_2(\mathbf{Z})$ there is a product of SR's and SR^2's equal to $\pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Since the matrices $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ have three positive entries and the other entry is 0, products of these matrices have entries that are also nonnegative, and in fact the sum of all the matrix entries will always increase under further multiplications. In particular, it is impossible for a product of finitely many copies of SR and SR^2 to equal $\pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, whose entries add up to ± 2. This completes the proof that in $\mathrm{PSL}_2(\mathbf{Z})$ the identity element can be written in the form (C.1) only in the trivial way: $n = 0$ and $i_0 = 0$.

Now consider a general equality

$$y^{i_0}xy^{i_1}x\cdots y^{i_{n-1}}xy^{i_n} = y^{i'_0}xy^{i'_1}x\cdots y^{i'_{m-1}}xy^{i'_m}.$$

where the inner exponents (not i_0, i_n, i'_0, or i'_m) are nonzero modulo 3. We want to show $m = n$ and $i_j = i'_j$ for all j. Bring the left side over to the right side using inversion:

$$\begin{aligned} 1 &= (y^{i_0}xy^{i_1}x\cdots y^{i_{n-1}}xy^{i_n})^{-1}(y^{i'_0}xy^{i'_1}x\cdots y^{i'_{m-1}}xy^{i'_m}) \\ &= y^{-i_n}x^{-1}y^{-i_{n-1}}\cdots x^{-1}y^{-i_1}x^{-1}y^{-i_0}y^{i'_0}xy^{i'_1}x\cdots y^{i'_{m-1}}xy^{i'_m} \\ &= y^{-i_n}xy^{-i_{n-1}}\cdots xy^{-i_1}xy^{i'_0-i_0}xy^{i'_1}x\cdots y^{i'_{m-1}}xy^{i'_m} \end{aligned}$$

The outer exponents $-i_n$ and i'_m are nonzero modulo 3. The inner exponents are the same as the inner exponents before, up to sign, except for $i'_0 - i_0$. So all inner exponents are nonzero except perhaps $i'_0 - i_0$. From what we know about representations of 1 as a product

of x's and y's, some inner exponent has to be 0. Therefore $i_0 = i'_0$ in $\mathbf{Z}/(3)$, which means $xy^{i'_0 - i_0}x = x^2 = 1$. So

$$1 = y^{-i_n}xy^{-i_{n-1}} \cdots xy^{i'_1 - i_1}x \cdots y^{i'_{m-1}}xy^{i'_m}.$$

Using induction on $\max(m, n)$, we obtain $m = n$ and $i_j = i'_j$ for all j. □

For another algebraic proof that $\mathrm{PSL}_2(\mathbf{Z})$ is generated by x and y with $x^2 = 1$, $y^3 = 1$, and no other relations, see [1]. (Warning: on the first page of [1], the definition of $\beta(z)$ should be $1 - 1/z$ and not $-1/z$.)

References

[1] R. C. Alperin, $\mathrm{PSL}_2(\mathbf{Z}) = \mathbf{Z}_2 * \mathbf{Z}_3$, Amer. Mathematical Monthly **100** (1993), 385–386.
[2] R. Rankin, *Modular Forms and Functions*, Cambridge Univ. Press, Cambridge, 1977.
[3] E. Schenkman, *Group Theory*, Van Nostrand, 1965.