SIMPLICITY OF $\mathrm{PSL}_n(F)$

KEITH CONRAD

1. Introduction

For a field F and integer $n \geq 2$, the *projective special linear group* $\mathrm{PSL}_n(F)$ is the quotient group of $\mathrm{SL}_n(F)$ by its center: $\mathrm{PSL}_n(F) = \mathrm{SL}_n(F)/Z(\mathrm{SL}_n(F))$. In 1831, Galois claimed that $\mathrm{PSL}_2(\mathbf{F}_p)$ is a simple group for all primes $p > 3$, although he didn't give a proof. He had to exclude $p = 2$ and $p = 3$ since $\mathrm{PSL}_2(\mathbf{F}_2) \cong S_3$ and $\mathrm{PSL}_2(\mathbf{F}_3) \cong A_4$, and these groups are not simple. It turns out that $\mathrm{PSL}_n(F)$ is a simple group for all $n \geq 2$ and all fields F except when $n = 2$ and $F = \mathbf{F}_2$ and $\mathbf{F}_3$. The proof of this was developed over essentially 30 years, from 1870 to 1901:

- Jordan [4] for $n \geq 2$ and $F = \mathbf{F}_p$ except $(n,p) = (2,2)$ and (2,3).
- Moore [5] for $n = 2$ and F all finite fields of size greater than 3.
- Dickson for $n > 2$ and F finite [1], and for $n \geq 2$ and F infinite [2].

We will prove simplicity of $\mathrm{PSL}_n(F)$ using a criterion of Iwasawa [3] from 1941 that relates simple quotient groups and doubly transitive group actions. This criterion will be developed in Section 2, and applied to $\mathrm{PSL}_2(F)$ in Section 3 and $\mathrm{PSL}_n(F)$ for $n > 2$ in Section 4.

2. Doubly transitive actions and Iwasawa's criterion

An action of a group G on a set X is called *transitive* when, given two distinct x and y in X, there is a $g \in G$ such that $g(x) = y$. We call the action *doubly transitive* if each pair of distinct points in X can be carried to every other pair of distinct points in X by some element of G. That is, given two pairs (x_1, x_2) and (y_1, y_2) in $X \times X$, where $x_1 \neq x_2$ and $y_1 \neq y_2$, there is a $g \in G$ such that $g(x_1) = y_1$ and $g(x_2) = y_2$. Although the x_i's are distinct and the y_j's are distinct, we do allow an x_i to be a y_j. For instance, if x, x', x'' are three distinct elements of X then there is a $g \in G$ such that $g(x) = x$ and $g(x') = x''$. (Here $x_1 = y_1 = x$ and $x_2 = x'$, $y_2 = x''$.) Necessarily a doubly transitive action requires $|X| \geq 2$.

Example 2.1. The action of A_4 on $\{1, 2, 3, 4\}$ is doubly transitive.

Example 2.2. The action of D_4 on $\{1, 2, 3, 4\}$, as vertices of a square, is not doubly transitive since a pair of adjacent vertices can't be sent to a pair of nonadjacent vertices.

Example 2.3. For all fields F, the group $\mathrm{Aff}(F)$ acts on F by $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)x = ax + b$ and this action is doubly transitive.

Example 2.4. For all fields F, the group $\mathrm{GL}_2(F)$ acts on $F^2 - \{\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)\}$ by the usual way matrices act on vectors, but this action is not doubly transitive since linearly dependent vectors can't be sent to linearly independent vectors by a matrix.

Theorem 2.5. *If G acts doubly transitively on X then the stabilizer subgroup of each point in X is a maximal subgroup of G.*

A maximal subgroup is a proper subgroup contained in no other proper subgroup.

Proof. Pick $x \in X$ and let $H_x = \mathrm{Stab}_x$.

Step 1: For each $g \notin H_x$, $G = H_x \cup H_x g H_x$.

For $g' \in G$ such that $g' \notin H_x$, we will show $g' \in H_x g H_x$. Both gx and $g'x$ are not x, so by double transitivity with the pairs (x, gx) and $(x, g'x)$ there is some $g'' \in G$ such that $g''x = x$ and $g''(gx) = g'x$. The first equation implies $g'' \in H_x$, so let's write g'' as h. Then $h(gx) = g'x$, so $g' \in hgH_x \subset H_x g H_x$.

Step 2: H_x is a maximal subgroup of G.

The group H_x is not all of G, since H_x fixes x while G carries x to each element of X and $|X| \geq 2$. Let K be a subgroup of G strictly containing H_x and pick $g \in K - H_x$. By step 1, $G = H_x \cup H_x g H_x$. Both H_x and $H_x g H_x$ are in K, so $G \subset K$. Thus $K = G$. $\square$

The converse of Theorem 2.5 is false. If H is a maximal subgroup of G then left multiplication of G on G/H has H as a stabilizer subgroup, but this action is not doubly transitive if G has odd order because a finite group with a doubly transitive action has even order.

Theorem 2.6. *Suppose G acts doubly transitively on a set X. Any normal subgroup $N \lhd G$ acts on X either trivially or transitively.*

Proof. Suppose N does not act trivially: $nx \neq x$ for some $x \in X$ and some $n \neq 1$ in N. Pick arbitrary y and y' in X with $y \neq y'$. By double transitivity, there is $g \in G$ such that $gx = y$ and $g(nx) = y'$. Then $y' = (gng^{-1})(gx) = (gng^{-1})(y)$ and $gng^{-1} \in N$, so N acts transitively on X. $\square$

Example 2.7. The action of A_4 on $\{1, 2, 3, 4\}$ is doubly transitive and the normal subgroup $\{(1), (12)(34), (13)(24), (14)(23)\} \lhd A_4$ acts transitively on $\{1, 2, 3, 4\}$.

Example 2.8. For a field F, let $\mathrm{Aff}(F)$ act on F by $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)x = ax + b$. This is doubly transitive and the normal subgroup $N = \{\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right) : b \in F\}$ acts transitively (by translations) on F.

Example 2.9. The action of D_4 on the 4 vertices of a square is not doubly transitive. Consistent with Theorem 2.6, the normal subgroup $\{1, r^2\}$ of D_4 acts on the vertices neither trivially nor transitively.

Here is the main group-theoretic result we will use to prove $\mathrm{PSL}_n(F)$ is simple.

Theorem 2.10 (Iwasawa)**.** *Let G act doubly transitively on a set X. Assume the following:*

1. *For some $x \in X$ the group Stab_x has an abelian normal subgroup whose conjugate subgroups generate G.*
2. *$[G, G] = G$.*

Then G/K is a simple group, where K is the kernel of the action of G on X.

The kernel of an action is the kernel of the homomorphism $G \to \mathrm{Sym}(X)$; it's those g that act like the identity on X.

Proof. To show G/K is simple we will show the only normal subgroups of G lying between K and G are K and G. Let $K \subset N \subset G$ with $N \lhd G$. Let $H = \mathrm{Stab}_x$, so H is a maximal subgroup of G (Theorem 2.5). Since N is normal, $NH = \{nh : n \in N, h \in H\}$ is a subgroup of G, and it contains H, so by maximality either $NH = H$ or $NH = G$. By Theorem 2.6, N acts trivially or transitively on X.

If $NH = H$ then $N \subset H$, so N fixes x. Therefore N does not act transitively on X, so N must act trivially on X, which implies $N \subset K$. Since $K \subset N$ by hypothesis, we have $N = K$.

Now suppose $NH = G$. Let U be the abelian normal subgroup of H in the hypothesis: its conjugate subgroups generate G. Since $U \lhd H$, $NU \lhd NH = G$. Then for $g \in G$, $gUg^{-1} \subset g(NU)g^{-1} = NU$, which shows NU contains all the conjugate subgroups of U. By hypothesis it follows that $NU = G$.

Thus $G/N = (NU)/N \cong U/(N \cap U)$. Since U is abelian, the isomorphism tells us that G/N is abelian, so $[G,G] \subset N$. Since $G = [G,G]$ by hypothesis, we have $N = G$. $\square$

Example 2.11. We can use Theorem 2.10 to show A_5 is a simple group. Its natural action on $\{1,2,3,4,5\}$ is doubly transitive. Let $x = 5$, so $\mathrm{Stab}_x \cong A_4$, which has the abelian normal subgroup

$$\{(1),(12)(34),(13)(24),(14)(23)\}.$$

The A_5-conjugates of this subgroup generate A_5 since the (2,2)-cycles in A_5 are all conjugate in A_5 and they generate A_5. The commutator subgroup $[A_5, A_5]$ contains every (2,2)-cycle: if a,b,c,d are distinct then

$$(ab)(cd) = (abc)(abd)(abc)^{-1}(abd)^{-1}.$$

Therefore $[A_5, A_5] = A_5$, so A_5 is simple.

3. Simplicity of $\mathrm{PSL}_2(F)$

Let F be a field. The group $\mathrm{SL}_2(F)$ acts on $F^2 - \{\binom{0}{0}\}$, but this action is not doubly transitive since linearly dependent vectors can't be sent to linearly independent vectors by a matrix. (We saw this for $\mathrm{GL}_2(F)$ in Example 2.4, and the same argument applies for its subgroup $\mathrm{SL}_2(F)$.) Linearly dependent vectors in F^2 lie along the same line through the origin, so let's consider the action of $\mathrm{SL}_2(F)$ on the linear subspaces of F^2: let $A \in \mathrm{SL}_2(F)$ send the line $L = Fv$ to the line $A(L) = F(Av)$. (Equivalently, we let $\mathrm{SL}_2(F)$ act on $\mathbf{P}^1(F)$, the projective line over F.)

Theorem 3.1. *The action of* $\mathrm{SL}_2(F)$ *on the linear subspaces of* F^2 *is doubly transitive.*

Proof. An obvious pair of distinct linear subspaces in F^2 is $F\binom{1}{0}$ and $F\binom{0}{1}$. It suffices to show that, given two distinct linear subspaces Fv and Fw of F^2, there is an $A \in \mathrm{SL}_2(F)$ that sends $F\binom{1}{0}$ to Fv and $F\binom{0}{1}$ to Fw, because we can then use $F\binom{1}{0}$ and $F\binom{0}{1}$ as an intermediate step to send a pair of distinct linear subspaces to every other pair of distinct linear subspaces.

Let $v = \binom{a}{c}$ and $w = \binom{b}{d}$. Since $Fv \neq Fw$, the vectors v and w are linearly independent, so $D := ad - bc$ is nonzero. Let $A = \left(\begin{smallmatrix} a & b/D \\ c & d/D \end{smallmatrix}\right)$, which has determinant $a(d/D) - (b/D)c = D/D = 1$, so $A \in \mathrm{SL}_2(F)$. Since $A\binom{1}{0} = \binom{a}{c} = v$ and $A\binom{0}{1} = \binom{b/D}{d/D} = (1/D)w$, A sends $F\binom{1}{0}$ to Fv and $F\binom{0}{1}$ to $F(1/D)w = Fw$. $\square$

We will apply Iwasawa's criterion (Theorem 2.10) to $\mathrm{SL}_2(F)$ acting on the set of linear subspaces of F^2. This action is doubly transitive by Theorem 3.1. It remains to check

- the kernel K of this action is the center of $\mathrm{SL}_2(F)$, so $\mathrm{SL}_2(F)/K = \mathrm{PSL}_2(F)$,
- the stabilizer subgroup of $\binom{1}{0}$ contains an abelian normal subgroup whose conjugate subgroups generate $\mathrm{SL}_2(F)$,
- $[\mathrm{SL}_2(F), \mathrm{SL}_2(F)] = \mathrm{SL}_2(F)$.

It is only in the third part that we will require $|F| > 3$. (At *some* point we need to avoid $F = \mathbf{F}_2$ and $F = \mathbf{F}_3$, because $\mathrm{PSL}_2(\mathbf{F}_2)$ and $\mathrm{PSL}_2(\mathbf{F}_3)$ are not simple.)

Theorem 3.2. *The kernel of the action of $\mathrm{SL}_2(F)$ on the linear subspaces of F^2 is the center of $\mathrm{SL}_2(F)$.*

Proof. A matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(F)$ is in the kernel K of the action when it sends each linear subspace of F^2 back to itself. If the matrix preserves the lines $F\binom{1}{0}$ and $F\binom{0}{1}$ then $c = 0$ and $b = 0$, so $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & 0 \\ 0 & d \end{smallmatrix}\right)$. The determinant is 1, so $d = 1/a$. If $\left(\begin{smallmatrix} a & 0 \\ 0 & 1/a \end{smallmatrix}\right)$ preserves the line $F\binom{1}{1}$ then $a = 1/a$, so $a = \pm 1$. This means $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Conversely, the matrices $\pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ both act trivially on the linear subspaces of F^2, so $K = \{\pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}$.

If a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is in the center of $\mathrm{SL}_2(F)$ then it commutes with $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, which implies $a = d$ and $b = c$ (check!). Therefore $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$. Since this has determinant 1, $a^2 = 1$, so $a = \pm 1$. Conversely, $\pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ commutes with all matrices. $\square$

Let $x = F\binom{1}{0}$. Its stabilizer subgroup in $\mathrm{SL}_2(F)$ is

$$\begin{aligned}
\mathrm{Stab}_{F\binom{1}{0}} &= \left\{ A \in \mathrm{SL}_2(F) : A\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in F\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{SL}_2(F) \right\} \\
&= \left\{ \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} : a \in F^\times, b \in F \right\}.
\end{aligned}$$

This subgroup has a normal subgroup

$$U = \left\{ \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} : \lambda \in F \right\},$$

which is abelian since $\left(\begin{smallmatrix} 1 & \lambda \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & \mu \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & \lambda+\mu \\ 0 & 1 \end{smallmatrix}\right)$.

Theorem 3.3. *The subgroup U and its conjugates generate $\mathrm{SL}_2(F)$. More precisely, each matrix of the form $\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)$ is conjugate to a matrix of the form $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$, and every element of $\mathrm{SL}_2(F)$ is the product of at most 4 elements of the form $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)$.*

This is the analogue for $\mathrm{SL}_2(F)$ of the 3-cycles generating A_n.

Proof. The matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ is in $\mathrm{SL}_2(F)$ and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & \lambda \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1 & 0 \\ -\lambda & 1 \end{smallmatrix}\right)$, so $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ conjugates $U = \{\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)\}$ to the group of lower triangular matrices $\{\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)\}$.

Pick $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(F)$. To show it is a product of matrices of type $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)$, first suppose $b \neq 0$. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (d-1)/b & 1 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ (a-1)/b & 1 \end{pmatrix}.$$

If $c \neq 0$ then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & (a-1)/c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & (d-1)/c \\ 0 & 1 \end{pmatrix}.$$

If $b = 0$ and $c = 0$ then the matrix is $\left(\begin{smallmatrix} a & 0 \\ 0 & 1/a \end{smallmatrix}\right)$, and

$$\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ (1-a)/a & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/a \\ 0 & 1 \end{pmatrix}. \quad \square$$

So far F has been a general field. Now we reach a result that requires $|F| \geq 4$.

Theorem 3.4. *If $|F| \geq 4$ then $[\mathrm{SL}_2(F), \mathrm{SL}_2(F)] = \mathrm{SL}_2(F)$.*

Proof. We compute an explicit commutator:

$$\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix}^{-1} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & b(a^2-1) \\ 0 & 1 \end{pmatrix}.$$

Since $|F| \geq 4$, there is an $a \neq 0, 1$, or -1 in F, so $a^2 \neq 1$. Using this value of a and letting b run over F shows $[\mathrm{SL}_2(F), \mathrm{SL}_2(F)]$ contains U. Since the commutator subgroup is normal, it contains every subgroup conjugate to U, so $[\mathrm{SL}_2(F), \mathrm{SL}_2(F)] = \mathrm{SL}_2(F)$ by Theorem 3.3. $\square$

Theorem 3.4 is false when $|F| = 2$ or 3: $\mathrm{SL}_2(\mathbf{F}_2) = \mathrm{GL}_2(\mathbf{F}_2)$ is isomorphic to S_3 and $[S_3, S_3] = A_3$. In $\mathrm{SL}_2(\mathbf{F}_3)$ there is a unique 2-Sylow subgroup, so it is normal, and its index is 3, so the quotient by it is abelian. Therefore the commutator subgroup of $\mathrm{SL}_2(\mathbf{F}_3)$ lies inside the 2-Sylow subgroup (in fact, the commutator subgroup is the 2-Sylow subgroup).

Theorem 3.5. *If $|F| \geq 4$ then the group $\mathrm{PSL}_2(F)$ is simple.*

Proof. By the previous four theorems the action of $\mathrm{SL}_2(F)$ on the linear subspaces of F^2 satisfies the hypotheses of Iwasawa's theorem, and its kernel is the center of $\mathrm{SL}_2(F)$. $\square$

4. Simplicity of $\mathrm{PSL}_n(F)$ for $n > 2$

To prove $\mathrm{PSL}_n(F)$ is simple for all F when $n > 2$, we will study the action of $\mathrm{SL}_n(F)$ on the linear subspaces of F^n, which is the projective space $\mathbf{P}^{n-1}(F)$.

Theorem 4.1. *The action of $\mathrm{SL}_n(F)$ on $\mathbf{P}^{n-1}(F)$ is doubly transitive with kernel equal to the center of the group and the stabilizer of some point has an abelian normal subgroup.*

Proof. For nonzero v in F^n, write the linear subspace Fv as $[v]$. Pick $[v_1] \neq [v_2]$ and $[w_1] \neq [w_2]$ in $\mathbf{P}^{n-1}(F)$. We seek an $A \in \mathrm{SL}_n(F)$ such that $A[v_1] = [w_1]$ and $A[v_2] = [w_2]$.

Extend v_1, v_2 and w_1, w_2 to bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$ of F^n. Let $L \colon F^n \to F^n$ be the linear map where $Lv_i = w_i$ for all i, so $\det L \neq 0$ and on $\mathbf{P}^{n-1}(F)$ we have $L[v_i] = [w_i]$ for all i. In particular, $L[v_1] = [w_1]$ and $L[v_2] = [w_2]$. Alas, $\det L$ may not be 1. For $c \in F^\times$, let $L_c \colon F^n \to F^n$ be the linear map where $L_c v_i = w_i$ for $i \neq n$ and $L_c v_n = cw_n$, so $L = L_1$. Then L_c sends $[v_i]$ to $[w_i]$ for all i and $\det L_c = c \det L$, so $L_c \in \mathrm{SL}_n(F)$ for $c = 1/\det L$.

If $A \in \mathrm{SL}_n(F)$ is in the kernel of this action then $A[v] = [v]$ for all nonzero $v \in F^n$, so $Av = \lambda_v v$, where $\lambda_v \in F^\times$: every nonzero element of F^n is an eigenvector of A. The only matrices for which all vectors are eigenvectors are scalar diagonal matrices. To prove this, use the equation $Av = \lambda_v v$ when $v = e_i$, $v = e_j$, and $v = e_i + e_j$ for the standard basis $e_1, \ldots, e_n$ of F^n. The equation $A(e_i + e_j) = Ae_i + Ae_j$ implies $\lambda_{e_i+e_j} e_i + \lambda_{e_i+e_j} e_j = \lambda_{e_i} e_i + \lambda_{e_j} e_j$, so $\lambda_{e_i} = \lambda_{e_i+e_j} = \lambda_{e_j}$. Let λ be the common value of λ_{e_i} over all i, so $Av = \lambda v$ when v runs through the basis. By linearity, $Av = \lambda v$ for all $v \in F^n$, so A is a scalar diagonal matrix with determinant 1. It is left to the reader to check that the center of $\mathrm{SL}_n(F)$ is also the scalar diagonal matrices with determinant 1.

To show the stabilizer of some point in $\mathbf{P}^{n-1}(F)$ has an abelian normal subgroup, we look at the stabilizer H of the point

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbf{P}^{n-1}(F),$$

which is the group of $n \times n$ determinant 1 matrices

$$\begin{pmatrix} a & * \\ \mathbf{0} & M \end{pmatrix}$$

where $a \in F^\times$, $M \in \mathrm{GL}_{n-1}(F)$, and $*$ is a row vector of length $n-1$. For this to be in $\mathrm{SL}_n(F)$ means $a = 1/\det M$. The projection $H \to \mathrm{GL}_{n-1}(F)$ sending $\left(\begin{smallmatrix} a & * \\ \mathbf{0} & M \end{smallmatrix}\right)$ onto M has abelian kernel

$$U := \left\{ \begin{pmatrix} 1 & * \\ \mathbf{0} & I_{n-1} \end{pmatrix} \right\} \cong F^{n-1}. \tag{4.1}$$ $\square$

To conclude by Iwasawa's theorem that $\mathrm{PSL}_n(F)$ is simple, it remains to show

- the subgroups of $\mathrm{SL}_n(F)$ that are conjugate to U generate $\mathrm{SL}_n(F)$,
- $[\mathrm{SL}_n(F), \mathrm{SL}_n(F)] = \mathrm{SL}_n(F)$.

This will follow from a study of the elementary matrices $I_n + \lambda E_{ij}$ where $i \neq j$ and $\lambda \in F^\times$. An example of such a matrix when $n = 3$ is

$$I_3 + \lambda E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix $I_n + \lambda E_{ij}$ has 1's on the main diagonal and a λ in the (i,j) position. Therefore its determinant is 1, so such matrices are in $\mathrm{SL}_n(F)$. The most basic example of such an elementary matrix in U is

$$I_n + E_{12} = \begin{pmatrix} 1 & 1 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_{n-2} \end{pmatrix}. \tag{4.2}$$

Here are the two properties we will need about the elementary matrices $I_n + \lambda E_{ij}$:

(1) For $n > 2$, each $I_n + \lambda E_{ij}$ is conjugate in $\mathrm{SL}_n(F)$ to $I_n + E_{12}$.
(2) For $n > 2$, the matrices $I_n + \lambda E_{ij}$ generate $\mathrm{SL}_n(F)$.

These properties imply the conjugates of $I_n + E_{12}$ generate $\mathrm{SL}_n(F)$. Since $I_n + E_{12} \in U$, the subgroups of $\mathrm{SL}_n(F)$ that are conjugate to U generate $\mathrm{SL}_n(F)$, so the next theorem would complete the proof that $\mathrm{PSL}_n(F)$ is simple for $n > 2$.

Theorem 4.2. *For $n > 2$, $[\mathrm{SL}_n(F), \mathrm{SL}_n(F)] = \mathrm{SL}_n(F)$.*

Proof. We will show $I_n + E_{12}$ is a commutator in $\mathrm{SL}_n(F)$. Then, since the commutator subgroup is normal, the above two properties of elementary matrices imply that $[\mathrm{SL}_n(F), \mathrm{SL}_n(F)]$ contains every $I_n + \lambda E_{ij}$, and therefore $[\mathrm{SL}_n(F), \mathrm{SL}_n(F)] = \mathrm{SL}_n(F)$.

Set

$$g = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } h = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Then

$$ghg^{-1}h^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is $I_3 + E_{12}$. For $n \geq 4$, $I_n + E_{12}$ is the block matrix

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_{n-2} \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I_{n-3} \end{pmatrix} \\ &= \begin{pmatrix} g & O \\ O & I_{n-3} \end{pmatrix} \begin{pmatrix} h & O \\ O & I_{n-3} \end{pmatrix} \begin{pmatrix} g & O \\ O & I_{n-3} \end{pmatrix}^{-1} \begin{pmatrix} h & O \\ O & I_{n-3} \end{pmatrix}^{-1}. \end{aligned}$$ □

All that remains is to prove the two properties we listed of the elementary matrices, and this is handled by the next two theorems.

Theorem 4.3. *For $n > 2$, each $I_n + \lambda E_{ij}$ with $\lambda \in F^\times$ is conjugate in $\mathrm{SL}_n(F)$ to $I_n + E_{12}$.*

Proof. Let $T = I_n + \lambda E_{ij}$. For the standard basis $e_1, \ldots, e_n$ of F^n,

$$T(e_k) = \begin{cases} e_k, & \text{if } k \neq j, \\ \lambda e_i + e_j, & \text{if } k = j. \end{cases}$$

We want a basis $e_1', \ldots, e_n'$ of F^n in which the matrix representation of T is $I_n + E_{12}$, *i.e.*, $T(e_k') = e_k'$ for $k \neq 2$ and $T(e_2') = e_1' + e_2'$.

Define a basis $f_1, \ldots, f_n$ of F^n by $f_1 = \lambda e_i$, $f_2 = e_j$, and $f_3, \ldots, f_n$ is some ordering of the $n-2$ standard basis vectors of F^n besides e_i and e_j. Then

$$T(f_1) = \lambda T(e_i) = \lambda e_i = f_1, \quad T(f_2) = T(e_j) = \lambda e_i + e_j = f_1 + f_2, \quad T(f_k) = f_k \text{ for } k \geq 3,$$

so relative to the basis $f_1, \ldots, f_n$ the matrix representation of T is $I_n + E_{12}$. Therefore

$$T = A(I_n + E_{12})A^{-1},$$

where A is the matrix such that $A(e_k) = f_k$ for all k. (Check $T = A(I_n + E_{12})A^{-1}$ by checking both sides take the same values at $f_1, \ldots, f_n$.) There is no reason to expect $\det A = 1$, so the equation $T = A(I_n + E_{12})A^{-1}$ shows us T and $I_n + E_{12}$ are conjugate in $\mathrm{GL}_n(F)$, rather than in $\mathrm{SL}_n(F)$. With a small change we can get a conjugating matrix in $\mathrm{SL}_n(F)$, as follows. For all $c \in F^\times$ we have

$$T = A_c(I_n + E_{12})A_c^{-1},$$

where

$$A_c(e_k) = \begin{cases} f_k, & \text{if } k < n, \\ cf_n, & \text{if } k = n. \end{cases}$$

(Check both sides of the equation $T = A_c(I_n + E_{12})A_c^{-1}$ are equal at $f_1, \ldots, f_{n-1}, cf_n$, where we need $n > 2$ for both sides to be the same at f_2.) The columns of A_c are the same as the columns of A except for the nth column, where A_c is c times the nth column of A. Therefore $\det(A_c) = c \det A$, so if we use $c = 1/\det A$ then $A_c \in \mathrm{SL}_n(F)$. That proves T is conjugate to $I_n + E_{12}$ in $\mathrm{SL}_n(F)$. □

Example 4.4. Let

$$T = I_3 + \lambda E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix}.$$

Starting from the standard basis e_1, e_2, e_3 of F^3, introduce a new basis f_1, f_2, f_3 by $f_1 = \lambda e_2$, $f_2 = e_3$, and $f_3 = e_1$. Since $T(f_1) = f_1$, $T(f_2) = f_1 + f_2$, and $T(f_3) = f_3$, we have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ \lambda & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ \lambda & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{-1},$$

where the conjugating matrix

$$\begin{pmatrix} 0 & 0 & 1 \\ \lambda & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

has for its columns f_1, f_2, and f_3 in order. The determinant of this conjugating matrix is λ, so it is usually not in $\mathrm{SL}_3(F)$. If we insert a nonzero constant c into the third column then we get a more general conjugation relation between $I_3 + \lambda E_{23}$ and $I_3 + E_{12}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & c \\ \lambda & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & c \\ \lambda & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{-1}.$$

The conjugating matrix has determinant λc, so using $c = 1/\lambda$ makes the conjugating matrix have determinant 1, which exhibits an $\mathrm{SL}_3(F)$-conjugation between $I_3 + \lambda E_{23}$ and $I_3 + E_{12}$.

Theorem 4.5. *For $n \geq 2$, the matrices $I_n + \lambda E_{ij}$ with $i \neq j$ and $\lambda \in F^\times$ generate $\mathrm{SL}_n(F)$.*

Proof. This will be a sequence of tedious computations. By a matrix calculation,

$$E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell} = \begin{cases} E_{i\ell}, & \text{if } j = k, \\ O, & \text{if } j \neq k. \end{cases} \tag{4.3}$$

Therefore $(I_n + \lambda E_{ij})(I_n + \mu E_{ij}) = I_n + (\lambda + \mu)E_{ij}$, so $(I_n + \lambda E_{ij})^{-1} = 1 - \lambda E_{ij}$, so the theorem amounts to saying that every element of $\mathrm{SL}_n(F)$ is a product of matrices $I_n + \lambda E_{ij}$.

We already proved the theorem for $n = 2$ in Theorem 3.3, so we can take $n > 2$ and assume the theorem is proved for $\mathrm{SL}_{n-1}(F)$. Pick $A \in \mathrm{SL}_n(F)$. We will show that by multiplying A on the left or right by suitable elementary matrices $I_n + \lambda E_{ij}$ we can obtain a block matrix $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A' \end{smallmatrix}\right)$. Since this is in $\mathrm{SL}_n(F)$, taking its determinant shows $\det A' = 1$, so $A' \in \mathrm{SL}_{n-1}(F)$. By induction A' is a product of elementary matrices $I_{n-1} + \lambda E_{ij}$, so $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A' \end{smallmatrix}\right)$ would be a product of block matrices of the form $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & I_{n-1}+\lambda E_{ij} \end{smallmatrix}\right)$, which is $I_n + \lambda E_{i+1\ j+1}$. Therefore

$$(\text{product of some } I_n + \lambda E_{ij})A(\text{product of some } I_n + \lambda E_{ij}) = \text{product of some } I_n + \lambda E_{ij},$$

and we can solve for A to see that it is a product of matrices $I_n + \lambda E_{ij}$.

The effect of multiplying an $n \times n$ matrix A by $I_n + \lambda E_{ij}$ on the left or right is an elementary row or column operation:

$$(I_n + \lambda E_{ij})A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} + \lambda a_{j1} & \cdots & a_{in} + \lambda a_{jn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} i\text{th row} = i\text{th row of } A + \lambda(j\text{th row of } A)$$

and

$$A(I_n + \lambda E_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1j} + \lambda a_{1i} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} + \lambda a_{ni} & \cdots & a_{nn} \end{pmatrix}$$
$$j\text{th col.} = j\text{th col. of } A + \lambda(i\text{th col. of } A)$$

Looking along the first column of A, some entry is not 0 since $\det A \neq 0$. If some a_{k1} in A is not 0 where $k > 1$, then

$$\left(I_n + \frac{1 - a_{11}}{a_{k1}} E_{1k}\right) A = \begin{pmatrix} 1 & \cdots \\ \vdots & \ddots \end{pmatrix}. \tag{4.4}$$

If $a_{21}, \ldots, a_{n1}$ are all 0, then $a_{11} \neq 0$ and

$$\left(I_n + \frac{1}{a_{11}} E_{21}\right) A = \begin{pmatrix} a_{11} & \cdots \\ 1 & \cdots \\ \vdots & \ddots \end{pmatrix}.$$

Then by (4.4) with $k = 2$,

$$(I_n + (1 - a_{11})E_{12}) \left(I_n + \frac{1}{a_{11}} E_{21}\right) A = \begin{pmatrix} 1 & \cdots \\ \vdots & \ddots \end{pmatrix}.$$

Once we have a matrix with upper left entry 1, multiplying it on the left by $I_n + \lambda E_{i1}$ for $i \neq 1$ will add λ to the $(i, 1)$-entry, so with a suitable λ we can make the $(i, 1)$-entry of the matrix 0. Thus multiplication on the left by suitable matrices of the form $I_n + \lambda E_{ij}$ produces a block matrix $\left(\begin{smallmatrix} 1 & * \\ \mathbf{0} & B \end{smallmatrix}\right)$ whose first column is all 0's except for the upper left entry, which is 1. Multiplying this matrix on the right by $I_n + \lambda E_{1j}$ for $j \neq 1$ adds λ to the $(1, j)$-entry without changing column other than the jth column. With a suitable choice of λ we can make the $(1, j)$-entry equal to 0, and carrying this out for $j = 2, \ldots, n$ leads to a block matrix $\left(\begin{smallmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A' \end{smallmatrix}\right)$, which is what we need to conclude the proof by induction. □

References

[1] L. E. Dickson, "The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group," Ann. Math. **11** (1897), 161–183.

[2] L. E. Dickson, "Theory of linear groups in an arbitrary field," Trans. Amer. Math. Soc. **2** (1901), 363–394.

[3] K. Iwasawa, "Über die Einfachkeit der speziellen projection Gruppen," Proc. Imperial Acad. Tokyo **17** (1941), 57–59.

[4] C. Jordan, *Traité des Substitutions*, Gauthier-Villars, Paris, 1870.

[5] E. H. Moore, "A doubly-infinite system of simple groups," Bull. New York Math. Soc. **3** (1893), 73–78.