The proof is from the book "Groups and characters" by Larry C. Grove.

The simplicity of PSL(n,q)

If V is n-dimensional vector space over a field F then GL(V) denotes the general linear group of V. Choosing a basis for V provides an isomorphism of GL(V) with the group GL(n, F) of all invertible $n \times n$ matrices over F.

The determinant is a homomorphism from GL(V) onto the multiplicative group F^* ; its kernel is the subgroup SL(V), the special linear group.

The center of GL(V) is $Z(GL(V)) = \{a1 : a \in F^*\}$ and the center of SL(V) is $Z(SL(V)) = \{a1 : a \in F^*, a^n = 1\}.$

If F is finite, with q elements, the the matrix groups are denoted by GL(n,q) and SL(n,q).

If $0 \neq v \in V$ write [v] for the line $Fv = \{av : a \in F\}$ through the origin spanned by v, and call it a *projective point*. The set of all distinct projective points [v] is called the *projective space* of dimension n-1 based on V, and is denoted by $P_{n-1}(V)$, or P(V). There is a natural action of GL(V) on P(V), given by $\tau[v] = [\tau v]$ for all $\tau \in GL(V)$, $[v] \in P(V)$. The kernel of the action is Z(GL(V)), and likewise that the kernel of the action of SL(V) on P(V) is Z(SL(V)).

Define the projective special linear group on V to be

$$PSL(V) = SL(V)/Z(SL(V)).$$

This group acts faithfully on P(V).

Our goal is to show that except for PSL(2,2) and PSL(2,3) every PSL(V) is a simple group.

To accomplish this we'll use Iwasawa's theorem. But before we formulate and prove this theorem, the definition of primitive group action is needed.

Let a group G act on a set S. For $x \in G$ and $s \in S$ we will write s^x to denote the result of the action of x on s. With that convention we have $s^1 = s$ and $s^{xy} = (s^x)^y$. Define a block to be a subset $B \subseteq S$ such that $B \ne S$, |B| > 1, and if $x \in G$ then either $B^x = B$ or $B \cap B^x = \emptyset$. If G is transitive on S and has no blocks we say that G is primitive on S.

Theorem (Iwasawa) Suppose that G is faithful and primitive on S and $G^{(1)} = G$. Fix $s \in S$ and set $H = \operatorname{Stab}_G(s)$. Suppose there is a solvable subgroup $K \triangleleft H$ such that $G = \langle \bigcup \{K^x : x \in G\} \rangle$. Then G is simple.

Proof. Suppose that $1 \neq N \triangleleft G$. We will show that in this case N = G and therefore G has no proper normal subgroups.

Step 1. N is transitive on S.

Since G is faithful on S and $N \neq 1$ there is an N-orbit in S having more than one element, say $B = \operatorname{Orb}_N(a)$ and $|B| \geq 2$. If $x \in G$ then $B^x = a^{Nx} = a^{xN}$ (since $N \triangleleft G$). Thus $B^x = \operatorname{Orb}_N(a^x)$. Since S is partitioned into its N-orbits we have either $B^x = B$ or $B^x \cap B = \emptyset$. But G is primitive on S hence we must have B = S, or else B would be a block for G. Therefore $\operatorname{Orb}_N(a) = B = S$ and this means that S is transitive on S.

Step 2. $G = \operatorname{Stab}_G(s)N$ for all $s \in S$. In particular, G = HN.

If $x \in G$ then $s^x = s^y$ for some $y \in N$ since N is transitive on S. Thus $s^{xy^{-1}} = s$, so $xy^{-1} \in \operatorname{Stab}_G(s)$ and $x \in \operatorname{Stab}_G(s)y \subseteq \operatorname{Stab}_G(s)N$. Hence $G = \operatorname{Stab}_G(s)N$.

Step 3. G = KN.

G = HN = NH (since $N \triangleleft G$), and $G = \langle \bigcup \{K^x : x \in NH\} \rangle$. $K \triangleleft H$, hence for $n \in N$ and $h \in H$ we have $K^{nh} = nhKh^{-1}n^{-1} = \left(K^h\right)^n = K^n$. Thus $G = \langle \bigcup \{K^n : n \in N\} \rangle$. Every $g \in G$ can be written in the form $g = k_1^{n_1} k_2^{n_2} \dots k_r^{n_r} = n_1 k_1 n_1^{-1} n_2 k_2 n_2^{-1} \dots n_r k_r n_r^{-1}$. Since $N \triangleleft G$ we have partial commutativity $kn = \bar{n}k$ hence g = kn for some $k \in K$ and $n \in N$. Thus G = KN.

Step 4. N = G.

Since K is solvable $K^{(m)} = 1$ for some m.

$$\begin{array}{lcl} (KN)^{(1)} & = & \left\langle \{k_1n_1k_2n_2n_1^{-1}k_1^{-1}n_2^{-1}k_2^{-1}: k_1, k_2 \in K, \ n_1, n_2 \in N\} \right\rangle \\ & = & \left\langle \{k_1k_2k_1^{-1}k_2^{-1}n: k_1, k_2 \in K, \ n \in N\} \right\rangle \leq K^{(1)}N. \end{array}$$

Using induction we obtain that $(KN)^{(l)} \leq K^{(l)}N$ for all l. Since $G = G^{(1)}$ we have $G = G^{(m)} = (KN)^{(m)} \leq K^{(m)}N = N$. Therefore N = G and this ends the proof.

In order to apply Iwasawa's theorem to $\operatorname{PSL}(V)$ we must show that $\operatorname{PSL}(V)$ is primitive on P(V), $\operatorname{PSL}(V)^{(1)} = \operatorname{PSL}(V)$, and there exists a normal solvable subgroup in $\operatorname{Stab}_{\operatorname{PSL}(V)}([v])$ whose conjugates in $\operatorname{PSL}(V)$ generate $\operatorname{PSL}(V)$. To achieve this goal we will introduce transvections and prove some results for them.

A hyperplane in V is any subset of codimension 1. If $1 \neq \tau \in GL(V)$ then τ is called a transvection if there is a hyperplane W such that $\tau|_W = 1_W$ and $\tau v - v \in W$ for all $v \in V$; W is called the fixed hyperplane of τ .

If τ is a transvection with fixed hyperplane W choose a basis for V consisting of first some $v_1 \in V \setminus W$ and then a basis $\{v_2, \ldots, v_n\}$ for W. It is clear from the matrix representing τ relative to this basis that $\det \tau = 1$, so $\tau \in \operatorname{SL}(V)$.

Proposition 1 The inverse of a transvection is a transvection. Suppose that V is a subspace of a space V_1 , that $v \in V_1 \setminus V$, and that τ is a transvection on V with fixed hyperplane W, then τ can be extended to a transvection τ_1 on V_1 whose fixed hyperplane W_1 contains v.

Proof. If τ is a transvection then $\tau w = w$ for all $w \in W$. Multiplying both sides of this equality by $\tau^{-1} \neq 1$ we obtain $w = \tau^{-1}w$ for all $w \in W$. Also we have $\tau v - v = w \in W$, hence $v - \tau^{-1}v = w$, and we obtain $\tau^{-1}v - v = -w \in W$. Thus τ^{-1} is a transvection.

Now prove the second part of the proposition. Let $\{v_1, v_2, \ldots, v_k\}$ be a basis for V such that $\{v_2, \ldots, v_k\}$ is a basis for W. We can choose a basis for V_1 consisting of the following vectors $\{v_1, v_2, \ldots, v_k, v_{k+1} = v, \ldots, v_n\}$. Consider a hyperplane W_1 spanned by $\{v_2, \ldots, v_n\}$ and extend τ on V_1 by setting $\tau_1|_{V} = \tau|_{V}$, $\tau_1v_i = v_i$ for all i > k. Then $\tau_1|_{W_1} = 1_{W_1}$ and $\tau_1v_1 - v_1 = \tau v_1 - v_1 \in W \subseteq W_1$. So τ_1 is a transvection with fixed hyperplane containing $v_{k+1} = v$.

Proposition 2 If u and v are linearly independent in V then there is a transvection τ with $\tau u = v$.

Proof. Choose a hyperplane W in V with $u-v\in W$ but $u\notin W$, and define τ by means of $\tau|_W=1_W,\ \tau u=v.$ If $x\in V$ write x=au+w, where $a\in F$ and $w\in W.$ Then $\tau x-x=av+w-au-w=a(v-u)\in W,$ so τ is a transvection.

Proposition 3 Suppose that W_1 and W_2 are two distinct hyperplanes in V and that $v \in V \setminus (W_1 \cup W_2)$. Then there is a transvection τ with $\tau W_1 = W_2$ and $\tau v = v$.

Proof. Note that $W_1+W_2=V$, so $\dim W_1\cap W_2=\dim W_1+\dim W_2-\dim(W_1+W_2)=n-2$ and $W=W_1\cap W_2+Fv$ is another hyperplane. Write v=x+y, with $x\in W_1$ and $y\in W_2$. Then $x\notin W_2$, so $W_1=W_1\cap W_2+Fx$, and likewise $W_2=W_1\cap W_2+Fy$. Thus $V=W_1\cap W_2+Fx+Fy$. It follows that $x\notin W$, or else y=v-x is also in W and hence $V\subseteq W$, a contradiction. Define τ via $\tau|_W=1_W$ and $\tau x=-y$. If $z\in V$ write z=ax+w, where $a\in F$ and $w\in W$. Then $\tau z-z=-ay+w-ax-w=-a(y+x)=-av\in W$. So τ is a transvection, $\tau v=v$ since $v\in W$, and $\tau W_1=\tau(W_1\cap W_2+Fx)=W_1\cap W_2-Fy=W_2$.

Theorem 1 The set of transvections generates SL(V).

Proof. Fix $\rho \in \operatorname{SL}(V)$, then choose a hyperplane W in V and choose $v \in V \setminus W$. If v and ρv are linearly independent then by Proposition 2 there is a transvection τ_1 with $\tau_1 \rho v = v$. If v and ρv are linearly dependent then first choose a transvection τ_0 so that v and $\tau_0 \rho v$ are linearly independent, then a transvection τ_1' so that $\tau_1' \tau_0 \rho v = v$, and set $\tau_1 = \tau_1' \tau_0$. Thus in either case we have $\tau_1 \rho v = v$ and τ_1 is a product of transvections. Note that $v \notin \tau_1 \rho W$ (since $v \in V \setminus W$ and $\tau_1 \rho v = v$). If $\tau_1 \rho W = W$ set $\tau_2 = 1_V$. If $\tau_1 \rho W \neq W$ apply Proposition 3 to get a transvection τ_2 with $\tau_2 \tau_1 \rho W = W$ and $\tau_2 v = v$. Set $\sigma = \tau_2 \tau_1 \rho$. Since $\sigma \in \operatorname{SL}(V)$, $\sigma W = W$, and $\sigma v = v$ it follows that $\sigma|_W \in \operatorname{SL}(W)$. Now use induction on $n = \dim V$. If n = 2 then $\sigma|_W = 1_W$ (since $\operatorname{SL}(W) = 1_W$ when $\dim W = 1$), so $\sigma = 1$ and $\rho = \tau_1^{-1} \tau_2^{-1}$. Proposition 1 implies that ρ is a product of transvections. If n > 2 then by induction $\sigma|_W$ is a product of transvections on W, each of which extends to a transvection fixing v on V by Proposition 1. Thus σ is the product of the extended transvections, and $\rho = \tau_1^{-1} \tau_2^{-1} \sigma$, a product of transvections.

Proposition 4 If τ_1 and τ_2 are transvections on V then they are conjugate in GL(V). If n > 2 they are conjugate in SL(V).

Proof. For i=1,2 write W_i for the fixed hyperplane of τ_i , choose $x_i \in V \setminus W_i$, and set $w_i = \tau_i x_i - x_i \in W_i$. Choose bases $\{w_1, u_3, \ldots, u_n\}$ for W_1 and $\{w_2, v_3, \ldots, v_n\}$ for W_2 . For each $a \in F^*$ define $\sigma_a \in \operatorname{GL}(V)$ by setting $\sigma_a x_1 = x_2$, $\sigma_a w_1 = w_2$, $\sigma_a u_i = v_i$ for $3 \leq i \leq n-1$, and, if n > 2, $\sigma_a u_n = a v_n$. Then $\sigma_a \tau_1 \sigma_a^{-1}(x_2) = \sigma_a \tau_1(x_1) = \sigma_a(x_1 + w_1) = x_2 + w_2 = \tau_2(x_2)$, $\sigma_a \tau_1 \sigma_a^{-1}(w_2) = \sigma_a \tau_1(w_1) = \sigma_a w_1 = w_2 = \tau_2(w_2)$, and $\sigma_a \tau_1 \sigma_a^{-1}(v_i) = \sigma_a \tau_1(u_i) = \sigma_a u_i = v_i = \tau_2(v_i)$. So $\sigma_a \tau_1 \sigma_a^{-1}$ and τ_2 agree on the basis $\{x_2, w_2, v_3, \ldots, v_n\}$, and hence $\sigma_a \tau_1 \sigma_a^{-1} = \tau_2$. If n > 2 we may set $b = (\det \sigma_1)^{-1}$ and obtain $\sigma_b \in \operatorname{SL}(V)$.

Proposition 5 Suppose that dim V = 2, and let $\{v_1, v_2\}$ be any basis for V. Every transvection is conjugate in SL(V) to one whose matrix relative to $\{v_1, v_2\}$ is of the form $\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$, $a \in F^*$.

Proof. If τ is a transvection with fixed hyperplane W choose $v \in V \setminus W$ and set $w = \tau v - v \in W$. Relative to the basis $\{v,w\}$ the matrix representing τ is $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. If M is the matrix that represents τ relative to $\{v_1,v_2\}$ then there is a matrix B in $\mathrm{GL}(2,F)$ with $B^{-1}MB = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. If $\det B = a^{-1}$ we may set $A = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}$. Then $BA \in \mathrm{SL}(2,F)$, and

$$(BA)^{-1}M(BA) = A^{-1} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}.$$

Theorem 2 If $n \geq 3$ and G = SL(V) then $G^{(1)} = G$, and also if G = PSL(V) then $G^{(1)} = G$.

Proof. If we exhibit a transvection in $G^{(1)}$ then all will be done since by Proposition 4 all transvections are conjugate in G, hence $G^{(1)}$ will contain all transvections, and by Theorem 1 transvections generate SL(V).

Choose a basis $\{v_1, \ldots, v_n\}$ for V and define τ_1, τ_2 via

$$\tau_1: v_1 \mapsto v_1 - v_2, v_i \mapsto v_i \text{ if } 2 \le i \le n,$$

$$\tau_2: v_1 \mapsto v_1, v_2 \mapsto v_2 - v_3, v_i \mapsto v_i \text{ if } 3 \le i \le n.$$

Then

$$\tau_1 \tau_2 \tau_1^{-1} \tau_2^{-1} : v_1 \mapsto v_1 - v_3, v_i \mapsto v_i \text{ if } 2 \le i \le n,$$

so $\tau_1 \tau_2 \tau_1^{-1} \tau_2^{-1}$ is a transvection in $G^{(1)}$.

PSL(V) = SL(V)/Z(SL(V)), hence

$$PSL(V)^{(1)} = \frac{SL(V)^{(1)}Z(SL(V))}{Z(SL(V))} = \frac{SL(V)}{Z(SL(V))} = PSL(V)$$

.

Theorem 3 If n = 2, |F| > 3, and G = SL(V), then $G^{(1)} = G$.

Proof. Choose a basis $\{v_1, v_2\}$ for V and choose $a \in F^*$, $a \neq \pm 1$. Define $\sigma \in SL(V)$ via $\sigma(v_1) = a^{-1}v_1$, $\sigma(v_2) = av_2$, and for each $b \in F^*$ define $\tau_b \in SL(V)$ via $\tau_b(v_1) = v_1 + bv_2$, $\tau_b(v_2) = v_2$. Then $\sigma\tau_b\sigma^{-1}\tau_b^{-1}$ is represented by the matrix

$$\left[\begin{array}{cc}a^{-1}&0\\0&a\end{array}\right]\left[\begin{array}{cc}1&0\\b&1\end{array}\right]\left[\begin{array}{cc}a&0\\0&a^{-1}\end{array}\right]\left[\begin{array}{cc}1&0\\-b&1\end{array}\right]=\left[\begin{array}{cc}1&0\\ba(a-a^{-1})&1\end{array}\right].$$

The theorem follows from Theorem 1 and Proposition 5, since $b \in F^*$ is arbitrary.

Proposition 6 If $n \geq 2$ then SL(V) is primitive on the projective space P(V).

Proof. Take $[v_1] \neq [v_2]$ and $[w_1] \neq [w_2]$ in P(V), so $\{v_1, v_2\}$ and $\{w_1, w_2\}$ are linearly independent sets in V. If n=2 they are bases. If $n\geq 3$ set $V_1=Fv_1+Fv_2$ and $V_2=Fw_1+Fw_2$. Then either $V_1=V_2\neq V$ or else $V_1\neq V_2$, in which case $V_1\cup V_2$ is not a subspace – in either case $V_1\cup V_2\neq V$. If $v_3\in V\setminus (V_1\cup V_2)$ then both $\{v_1,v_2,v_3\}$ and $\{w_1,w_2,w_3\}$ are linearly independent. The argument may be repeated to obtain two bases for V, $\{v_1,\ldots,v_n\}$ and $\{w_1,\ldots,w_n\}$. For any $b\in F^*$ define $\tau_b\in \mathrm{GL}(V)$ via $v_1\mapsto bw_1,\ v_2\mapsto w_2,\ v_i\mapsto v_i$ for $1\leq i\leq n$. If $1\leq i\leq n$ if $1\leq$

Suppose $B \subseteq P(V)$, with |B| > 1 and $B \neq P(V)$. Choose $[v_1], [v_2] \in B$ and $[w] \in P(V) \setminus B$. Then choose $\tau \in \operatorname{SL}(V)$ with $\tau([v_1]) = [v_1]$ and $\tau([v_2]) = [w]$. Then $[v_1] \in \tau(B) \cap B$ so $\tau(B) \cap B \neq \emptyset$, and $[w] \in \tau(B) \setminus B$, so $\tau(B) \neq B$. Thus B is not a block for $\operatorname{SL}(V)$, hence $\operatorname{SL}(V)$ has no blocks and it is primitive.

Proposition 7 If $0 \neq v \in V$ and $A = \operatorname{Stab}_{\operatorname{SL}(V)}([v])$ then A has an abelian normal subgroup B whose conjugates in $\operatorname{SL}(V)$ generate $\operatorname{SL}(V)$.

Proof. Choose a hyperplane W with $v \notin W$, so V = W + Fv. If $\sigma \in A$ and $w \in W$ write $\sigma w = \sigma' w + a_w v$, with $\sigma' w \in W$ and $a_w \in F$. Let $x = aw_1 + bw_2$ then $\sigma(aw_1 + bw_2) = \sigma'(aw_1 + bw_2) + a_x v$ and $\sigma(aw_1 + bw_2) = a\sigma(w_1) + b\sigma(w_2) = a\sigma' w_1 + b\sigma' w_2 + (aa_{w_1} + ba_{w_2})v$. Therefore $\sigma'(aw_1 + bw_2) = a\sigma' w_1 + b\sigma' w_2$ since $v \notin W$. Also we have $\sigma v = \sigma' v + a_v v$, on the other hand $\sigma v = a_v v$ since $\sigma \in A = \operatorname{Stab}_{\operatorname{SL}(V)}([v])$, so $\sigma' v = 0$. $\sigma^{-1} w = (\sigma^{-1})' w + b_w v$, hence $w = \sigma^{-1}(\sigma w) = (\sigma^{-1})'(\sigma w) = (\sigma^{-1})'(\sigma' w + a_w v) = (\sigma^{-1})'\sigma' w$. Analogously $\sigma'(\sigma^{-1})' w = w$, therefore $\sigma' \in \operatorname{GL}(W)$. Consider a map $\varphi : \sigma \mapsto \sigma'$. $(\sigma_1 \sigma_2) w = (\sigma_1 \sigma_2)' w + a_w v$, $(\sigma_1 \sigma_2) w = \sigma_1(\sigma_2 w) = \sigma_1'(\sigma_2' w + b_w v) = \sigma_1'(\sigma_2' w) + a_{\sigma_2' w} v$, hence $(\sigma_1 \sigma_2)' = \sigma_1' \sigma_2'$, therefore φ is a homomorphism form A into $\operatorname{GL}(W)$. Thus $B = \ker \varphi$ is a normal subgroup of A.

Choose a basis $\{w_1,\ldots,w_{n-1}\}$ for W. For any $b\in F^*$ define $\tau_b\in A$ via $v_1\mapsto w_1+bv,\ w_i\mapsto w_i$ for $2\leq i\leq n-1$, and $v\mapsto v$. Then τ_b is a transvection (with hyperplane spanned by $\{w_2,\ldots,w_{n-1},v\}$), and if $w\in W$ then $\tau_bw=w+bv$, so $\tau_b'=1_W$ and $\tau_b\in B$. If n=2 the matrix representing τ_b relative to the basis $\{w_1,v\}$ is $\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$, and $b\in F^*$ is arbitrary. It follows from

Theorem 1 and Propositions 4 and 5 that the conjugates of B generate $\mathrm{SL}(V)$ for all $n \geq 2$. If $\sigma \in B$ then $\sigma' = 1_W$, so $\sigma w_i = w_i + a_i v$. Thus, if $\sigma_1, \sigma_2 \in B$, their representing matrices relative to the basis $\{w_1, \ldots, w_{n-1}, v\}$ have the partitioned form $\begin{bmatrix} I & 0 \\ u_1 & 1 \end{bmatrix}$ and $\begin{bmatrix} I & 0 \\ u_2 & 1 \end{bmatrix}$, where I is the $(n-1) \times (n-1)$ identity matrix and u_1, u_2 are $1 \times (n-1)$ matrices. Since

$$\left[\begin{array}{cc} I & 0 \\ u_1 & 1 \end{array}\right] \left[\begin{array}{cc} I & 0 \\ u_2 & 1 \end{array}\right] = \left[\begin{array}{cc} I & 0 \\ u_1 + u_2 & 1 \end{array}\right] = \left[\begin{array}{cc} I & 0 \\ u_2 & 1 \end{array}\right] \left[\begin{array}{cc} I & 0 \\ u_1 & 1 \end{array}\right]$$

we see that B is abelian.

Theorem 4 Except for PSL(2,2) and PSL(2,3), every PSL(V) is a simple group.

Proof. Choose $v \neq 0$ in V, and take $A = \operatorname{Stab}_{\operatorname{SL}(V)}([v])$ and $B \triangleleft A$ as in Proposition 7. Write $Z = Z(\operatorname{SL}(V))$ and set $H = A/Z = \operatorname{Stab}_{\operatorname{PSL}(V)}([v])$, $K = BZ/Z \triangleleft H$. We have that $\operatorname{PSL}(V)$ acts faithfully on P(V). It is also primitive on P(V) since by Proposition 6 $\operatorname{SL}(V)$ is primitive on P(V). $K \triangleleft H = \operatorname{Stab}_{\operatorname{PSL}(V)}([v])$ and K is abelian, hence it is solvable. The conjugates of K generate $\operatorname{PSL}(V)$. Then all the conditions of Iwasawa's Theorem are met, and $\operatorname{PSL}(V)$ is simple.