Simplicity of $PSL_n(F)$ for $n > 2$

Kavi Duvvoori

October 2015

A Outline

To show the simplicity of $PSL_n(F)$ (for $n > 3$), we will consider a class of linear maps called transvections, or shear mappings - linear maps that fix some hyperplane, and move points in some fixed direction parallel to that hyperplane an amount proportional to their distance from it. Transvections turn out to both generate $SL_n(F)$ and be transitive under conjugation. We will then be able to show that any normal subgroup not contained in the center of $SL_n(F)$ must contain some non-trivial transvection, and hence all of $SL_n(F)$. It follows from that that $SL_n(F)/Z(SL_n(F))$ is a simple group. We will take advantage of the fact that in sufficient dimensions a hyperplane has at least two linearly independent vectors, and this is why we require $n > 2$.

It might also be possible to see these facts in terms of matrix representations of these mappings, but the proof repeatedly considers transvections over varying bases, so it is much more convenient to consider these transvections geometrically.

B Transvections

i Definition

A linear transformation $T \in GL_n(F)$ is a transvection if it is of the form $Tx = x + \lambda(x)u$ for some non-zero functional λ and some vector $u \in Ker(\lambda)$. Recall that a non-zero functional, as a map from an n-dimensional vector space to a 1-dimensional vector space, has as its kernel an $(n-1)$-dimensional subspace of F^n, some hyperplane: define $H_\lambda = Ker(\lambda)$. If you like that sort of thing, these maps have a geometric interpretation: they are linear maps that fix a hyperplane, and map all other vectors in a fixed direction parallel to it, an amount proportional to their distance from the hyperplane.

Let $T_{\lambda,u}$, for $u \in Ker(\lambda)$ be the the transvection where $T_{\lambda,u}x \mapsto x+\lambda(x)+u$. Then, if $u, v \in H_\lambda$, $T_{\lambda,u} \circ T_{\lambda,v}(x) = x + \lambda(x)v + \lambda(x + \lambda(x)v)u = x + \lambda(x)v + \lambda(x)u = x + \lambda(x)(u + v) = T_{\lambda,u+v}(x)$, since $u, v \in Ker(\lambda)$. Also, this implies that $T^{-1}_{\lambda,u}x = x - \lambda(x)u = T_{\lambda,-u}x$.

C Proposition 1: Transvections Generate $SL_n(F)$

First, we will sketch a proof for the fact from linear algebra that elementary matrices generate $SL_n(F)$: the elementary matrix $E_{ij}(c)$ with $i \neq j$ is the matrix consisting of 1s on the diagonal, a c in the ijth position, and 0s everywhere else. Note that multiplication by E_{ij} is equivalent to the row operation of adding c times the jth row to the ith row.

Now, suppose $S \in SL_n(F)$. Then, some row has non-zero first entry. Adding a suitable multiple to the first row, we can make the first entry of the first row equal to 1. If the first row is the only non-zero row, it can be added to the second row, and after a suitable multiple of the second row (which now has non-zero first entry), subtracted from the first row. Subtracting suitable multiples of the first row from all the other rows, we can then make all the other rows' first entry 0. Similarly, some row now has non-zero second entry. Adding a suitable multiple, we can make the second entry in the second row 1. Given this, we can make the second entry of all the other rows 0, by subtracting suitable multiples of the second. Note that this keeps the first entries fixed. As long as there is some lower row, it is possible to make the appropriate entry 1, even if the matrix is diagonal.

Continuing this process, we can attain a diagonal matrix. Since $det(S) = 1$, the final entry must also be 1, and this matrix must be the identity. Thus, S can be constructed from the identity by a series of row operations, and equivalently, can be written as a product of elementary matrices.

Now, we just have to notice that any elementary matrix represents a transvection: given the elementary matrix $E_{ij}(c)$, let λ be the functional mapping all basis vectors other than e_j to 0, and e_j to c. Then, $E_{ij}(x) = x + \lambda(x)e_i = T_{\lambda,e_i}$

D Proposition 2: For $n \geq 3$, the non-identity transvections form a single conjugacy class of $SL_n(F)$

Let $T_{\lambda,u}$ be an arbitrary transvection and A an arbitrary element of $SL_n(F)$. Then, $ATA^{-1}(x) = AT(A^{-1}x) = A(A^{-1}x + \lambda(A^{-1}x)u) = x + \lambda(A^{-1}x)Au = T_{\lambda',u'}$ where $\lambda' = \lambda \circ A^{-1}$ and $u' = Au$. Since A is an automorphism, $\lambda \circ A^{-1}$ is a linear functional, and $Au \neq 0$. Thus, the non-identity transvections are closed under conjugation.

Now, let $T = T_{\lambda,u}$ and $T' = T_{\lambda',u'}$ be any two distinct non-trivial transvections. Then there are some $z, z' \in F^n$ with $Tz = T'z' = 1$. Now, we can extend u to a basis for H_λ and u' to a basis for $H_{\lambda'}$. With z and z' respectively, these yield two different bases for f^n. Since $n \geq 3$, H_λ contains some vector v independent from u and H'_λ v' from u'. Then, there is some change of basis matrix A between these two bases, with $Au = u'$, $AH_\lambda = AH'_\lambda$ and $Az = z'$. Scaling appropriately, we can have $Av = cv$ so that $det(A) = 1$. Then, $ATA^{-1}x = x + \lambda(A^{-1}x)Au$. Since $Au = u'$, A fixes H, and $Az = z' \Rightarrow \lambda(A^{-1}x) = \lambda'(x)$, then $ATA^{-1}x = x + \lambda'(x)u' = T'x$. Thus, the non-trivial transvections are transitive.

E Corollary 1: For $n \geq 3$, if $\forall A \in SL_n(F), AGA^{-1} \subseteq G$ and G contains a transvection, then $SL_n(F) \subseteq G$

By proposition 2, conjugates of the given transvection yield all transvections, and by proposition 1 the transvections generate $SL_n(F)$, hence G contains $SL_n(F)$.

F Lemma 3: $Z(SL_n(F))$ consists of the subgroup of scalar multiplications

Note that for $M \in SL_n(F)$ to be contained in the center $Z(SL_n(F))$, it must commute with all matrices of the form 1_{ij}, i.e. matrices with a 1 in the ith row and jth column and 0s everywhere else. Hence, M must have a zero on all non-diagonal entries, and the same value along the the diagonal - it must represent multiplication by some scalar. Scalar multiplication commutes with any element of $GL_n(F)$, hence $Z(SL_n(F))$ consists of scalar multiplications.

G Lemma 4: For $n \geq 3$, , if $\forall A \in SL_n(F), AGA^{-1} \subseteq G$ and $G \not\subseteq Z(SL_n(F))$, then G contains a non-identity transvection

We just want to show G contains a transvection. Since G is not contained in the center, there is some $A \in G$ where A moves a line, i.e. $\exists u \in F^n$ where $Au = v$ and v is not a scalar multiple of u. Recall that $AT_{\lambda,u}A^{-1}x = x + \lambda(A^{-1}x)Au = x + \lambda(A^{-1}x)v \neq T_{\lambda,u}$ since v and u are linearly independent. Thus, $B = AT_{\lambda,u}A^{-1}T^{-1} \neq I$.

$Bx = ATA^{-1}T^{-1}x = ATA^{-1}(x - \lambda(x)u) = x - \lambda(x)u + \lambda'(x - \lambda(x)u)v = x - \lambda(x)u + \lambda'(x)$. Thus, $Bx - x$ is contained in the plane (u, v). Then, let H be some hyperplane containing u and v - then $BH \subset H$ so $BH = H$ and $Bx - x \in H$. Since $A^{-1} \in G$, $T_{\lambda,u}A^{-1}T_{\lambda,u}^{-1}$, as a conjugate of an $SL_n(F)$ invariant subgroup, is in G, so $B = A(T_{\lambda,u}A^{-1}T_{\lambda,u}^{-1}) \in G$.

i Case 1: All transvections on H commute with B.

Consider any $w \in H$. $BT_w x = Bx + \lambda(x)Bw$ and $T_w Bx = Bx + \lambda(Bx)w = Bx + \lambda(x)w$. Then, since $BT_w = T_w B$, $Bw = w$. Thus, B fixes H. Since $Bx - x \in H$, B is then a transvection.

ii Case 2: Some transvection T_w does not commute with B

Let $C = BT_wB^{-1}T_w^{-1}$. Then, since B and T_w do not commute, $C \neq I$. Now, since $BH = H = B^{-1}H$, we have that BT_wB^{-1} and T_w^{-1} both are transvections with hyperplane H, so their product, C, is also a transvection. Finally, C is the product of B and a conjugate of B^{-1}, which by $SL_n(F)$ invariance is in G, so $C \in G$.

H Corollary 2: For $n \geq 3$, , if $\forall A \in SL_n(F), AGA^{-1} \subseteq G$ and $G \not\subseteq Z(SL_n(F))$, then $SL_n(F) \subseteq G$

Immediate from Lemma 4 and Lemma 3.

I Proposition 3: For $n \geq 3, PSL_n(F)$ is simple

Let $G \triangleleft PSl_n(F)$ be some non-trivial normal subgroup. Then, if $\phi : SL_n(F) \to SL_n(F)/Z(SL_n(F)) \cong PSL_n(F)$ is the quotient map, $\phi^{-1}(G) = \overline{G}$ is a normal subgroup of $SL_n(F)$. Since G is non-trivial, $\overline{G} \not\subset Z(SL_n(F))$. Then, by Lemma 5, $SL_n(F) \subseteq \overline{G}$, so $G = PSL_n(F)$.

J Source:

Adapted from

Lang, S. (1993). Algebra 3rd-edition. Pg. 541-545