CHAPTER 8

Some Simple Linear Groups

The Jordan–Hölder theorem tells us that once we know extensions and simple groups, then we know all finite groups. There are several infinite families of finite simple groups (in addition to the cyclic groups of prime order and the large alternating groups), and our main concern in this chapter is the most "obvious" of these, the *projective unimodular groups*, which arise naturally from the group of all matrices of determinant 1 over a field K. Since these groups are finite only when the field K is finite, let us begin by examining the finite fields.

Finite Fields

Definition. If K is a field, a ***subfield*** of K is a subring k of K which contains the inverse of every nonzero element of k. A ***prime field*** is a field k having no proper subfields.

Theorem 8.1. *Every field K contains a unique prime subfield k, and either $k \cong \mathbb{Q}$ or $k \cong \mathbb{Z}_p$ for some prime p.*

Proof. If k is the intersection of all the subfields of K, then it is easy to check that k is the unique prime subfield of K. Define $\chi\colon \mathbb{Z} \to K$ by $\chi(n) = n1$, where 1 denotes the "one" in K. It is easily checked that χ is a ring homomorphism with $\operatorname{im} \chi \subset k$. Since K is a field, $\operatorname{im} \chi$ is a domain and $\ker \chi$ must be a prime ideal in $\mathbb{Z}$. Therefore, either $\ker \chi = 0$ or $\ker \chi = (p)$ for some prime p. In the first case, $\operatorname{im} \chi \cong \mathbb{Z}$ and k contains an isomorphic copy F of the fraction field of $\mathbb{Z}$, namely, $\mathbb{Q}$; as k is a prime field, $k = F \cong \mathbb{Q}$. In the second case, k

contains an isomorphic copy E of $\mathbb{Z}/\ker \chi = \mathbb{Z}_p$, which is a field; as k is a prime field, $k = E \cong \mathbb{Z}_p$. ∎

Definition. If K is a field with prime field k, then K has ***characteristic 0*** if $k \cong \mathbb{Q}$ and K has ***characteristic p*** if $k \cong \mathbb{Z}_p$.

Observe that if K has characteristic $p \geq 0$, then $pa = 0$ for all $a \in K$.

Corollary 8.2. *If K is a finite field, then $|K| = p^n$ for some prime p and some $n \geq 1$.*

Proof. If k is the prime field of K, then $k \ncong \mathbb{Q}$ because K is finite; therefore, $k \cong \mathbb{Z}_p$ for some prime p. We may view K as a vector space over $\mathbb{Z}_p$ (the "vectors" are the elements of K, the "scalars" are the elements of k, and the "scalar multiplication" $a\alpha$, for $a \in k$ and $\alpha \in K$, is just their product in K); if K has dimension n, then $|K| = p^n$. ∎

There exist infinite fields of prime characteristic; for example, the field of all rational functions over $\mathbb{Z}_p$ (i.e., the fraction field of $\mathbb{Z}_p[x]$) is such a field.

The existence and uniqueness of finite fields are proven in Appendix VI: for every prime p and every integer $n \geq 1$, there exists a field with p^n elements (Theorem VI.19); two finite fields are isomorphic if and only if they have the same number of elements (Theorem VI.20). Finite fields are called ***Galois fields*** after their discoverer; we thus denote the field with $q = p^n$ elements by $\mathrm{GF}(q)$ (another common notation for this field is $\mathbb{F}_q$), though we usually denote $\mathrm{GF}(p)$ by $\mathbb{Z}_p$.

Recall that if E is a field, k is a subfield, and $\pi \in E$, then $k(\pi)$, the subfield of E obtained by ***adjoining*** π to k, is the smallest subfield of E containing k and π; that is, $k(\pi)$ is the intersection of all the subfields of E containing k and π.

Definition. A ***primitive element*** of a finite field K is an element $\pi \in K$ with $K = k(\pi)$, where k is the prime field.

Lemma 8.3. *There exists a primitive element π of $\mathrm{GF}(p^n)$; moreover, π may be chosen to be a root of an irreducible polynomial $g(x) \in \mathbb{Z}_p[x]$ of degree n.*

Proof. Let $q = p^n$ *and let* $K = \mathrm{GF}(q)$. By Theorem 2.18(ii), the multiplicative group $K^\times$ is cyclic; clearly, any generator π of $K^\times$ is a primitive element of K (there can be primitive elements of K that are not generators of $K^\times$). By Lagrange's theorem, $\pi^{q-1} = 1$ (for $|K^\times| = q - 1$), and so π is a root of $f(x) = x^{q-1} - 1$. Factoring $f(x)$ into a product of irreducible polynomials in $k[x]$ (where $k \cong \mathbb{Z}_p$ is the prime field) provides an irreducible $g(x) \in k[x]$ having π as a root. If $g(x)$ has degree d, then $k(\pi)$ is a subfield of K with $[k(\pi) : k] = d$ (Theorem VI.21 in Appendix VI); therefore, $|k(\pi)| = p^d$. But $k(\pi) = K$ (because π is a primitive element) and so $d = n$. ∎

Theorem 8.4. *If p is a prime, then the group* $\mathrm{Aut}(\mathrm{GF}(p^n))$ *of all field automorphisms of* $\mathrm{GF}(p^n)$ *is cyclic of order n.*

Proof. Let k be the prime field of $K = \mathrm{GF}(p^n)$. If π is a primitive element of K, as in the lemma, then there is an irreducible polynomial $g(x) \in k[x]$ of degree n having π as a root. Since every $\varphi \in \mathrm{Aut}(K)$ must fix k pointwise (because $\varphi(1) = 1$), Lemma 5.1 shows that $\varphi(\pi)$ is also a root of $g(x)$. As $K = k(\pi)$, Lemma 5.2 shows that φ is completely determined by $\varphi(\pi)$. It follows that $|\mathrm{Aut}(K)| \leq n$, because $g(x)$, having degree n, has at most n roots. The map $\sigma\colon K \to K$, given by $\sigma(\alpha) = \alpha^p$, is an automorphism of K. If $1 \leq i < n$ and $\sigma^i = 1$, then $\alpha = \alpha^{p^i}$ for every $\alpha \in K$. In particular, $\pi^{p^i-1} = 1$, contradicting π having order $p^n - 1$ in $K^\times$. Therefore, $\langle\sigma\rangle \leq \mathrm{Aut}(K)$ is cyclic of order n, and so $\mathrm{Aut}(K) = \langle\sigma\rangle$. ∎

We remark that $\mathrm{Aut}(\mathrm{GF}(p^n))$ is the Galois group $\mathrm{Gal}(\mathrm{GF}(p^n)/\mathbb{Z}_p)$, for every $\varphi \in \mathrm{Aut}(\mathrm{GF}(p^n))$ fixes the prime field $\mathbb{Z}_p$ pointwise.

The General Linear Group

Groups of nonsingular matrices are as natural an object of study as groups of permutations: the latter consists of "automorphisms" of a set; the former consists of automorphisms of a vector space.

Definition. If V is an m-dimensional vector space over a field K, then the ***general linear group*** $\mathrm{GL}(V)$ is the group of all nonsingular linear transformations on V (with composite as operation).

If one chooses an ordered basis $\{e_1, \ldots, e_m\}$ of V, then each $T \in \mathrm{GL}(V)$ determines a matrix $A = [\alpha_{ij}]$, where $Te_j = \sum_i \alpha_{ij} e_i$ (the jth *column* of A consists of the coordinates of Te_j). The function $T \mapsto A$ is an isomorphism $\mathrm{GL}(V) \to \mathrm{GL}(m, K)$, where $\mathrm{GL}(m, K)$ is the multiplicative group of all $m \times m$ nonsingular matrices over K. When $K = \mathrm{GF}(q)$, we may write $\mathrm{GL}(m, q)$ instead of $\mathrm{GL}(m, K)$.

Theorem 8.5. $|\mathrm{GL}(m, q)| = (q^m - 1)(q^m - q)\ldots(q^m - q^{m-1})$.

Proof. Let V be an m-dimensional vector space over a field K, and let $\{e_1, \ldots, e_m\}$ be an ordered basis of V. If $\mathscr{B}$ denotes the family of all ordered bases of V, then there is a bijection $\mathrm{GL}(V) \to \mathscr{B}$: if T is nonsingular, then $\{Te_1, \ldots, Te_m\}$ is an ordered basis of V; if $\{v_1, \ldots, v_m\}$ is an ordered basis, then there exists a unique nonsingular T with $Te_i = v_i$ for all i.

Let $\{v_1, \ldots, v_m\}$ be an ordered basis of V. Since there are q^m vectors in V, there are $q^m - 1$ candidates for v_1 (the zero vector is not a candidate). Having

chosen v_1, the candidates for v_2 are those vectors in V not in $\langle v_1 \rangle$, the subspace spanned by v_1; there are thus $q^m - q$ candidates for v_2. More generally, having chosen an independent set $\{v_1, \ldots, v_i\}$, we may choose v_{i+1} to be any vector not in $\langle v_1, \ldots, v_i \rangle$, and so there are $q^m - q^i$ candidates for v_{i+1}. The result follows. ▩

Notation. If V is an m-dimensional vector space over $K = \mathrm{GF}(q)$, if t is a nonnegative integer, and if π is a primitive element of K, then

$$M(t) = \{A \in \mathrm{GL}(V) \colon \det A \text{ is a power of } \pi^t\}.$$

Lemma 8.6. *If $\Omega = |\mathrm{GL}(m, q)|$ and if t is a divisor of $q - 1$, then $M(t)$ is a normal subgroup of $\mathrm{GL}(m, q)$ of order Ω/t. Moreover, if $q - 1 = p_1 \ldots p_r$, where the p_i are (not necessarily distinct) primes, then the following normal series is the beginning of a composition series*:

$$\mathrm{GL}(m, q) = M(1) > M(p_1) > M(p_1 p_2) > \cdots > M(q-1) > 1.$$

Proof. Let $K = \mathrm{GF}(q)$. Use the correspondence theorem in the setting

$$\det \colon \mathrm{GL}(m, q) \to K^\times.$$

If t divides $q - 1 = |K^\times|$, then the cyclic subgroup $\langle \pi^t \rangle$ of $K^\times$ is normal ($K^\times$ is abelian), has order $(q-1)/t$, and has index t. Since $M(t)$ is the subgroup of $\mathrm{GL}(m, q)$ corresponding to $\langle \pi^t \rangle$, it is a normal subgroup of index t hence order Ω/t. Now $|M(p_1 \ldots p_i)/M(p_1 \ldots p_{i+1})| = (\Omega/p_1 \ldots p_i)/(\Omega/p_1 \ldots p_{i+1}) = p_{i+1}$; since the factor groups have prime order, they are simple. ▩

Definition. A matrix (or linear transformation) having determinant 1 is called ***unimodular***.

The subgroup $M(q-1)$ consists of all the unimodular matrices, for $\pi^{q-1} = 1$.

Definition. If V is an m-dimensional vector space over a field K, then the ***special linear group*** $\mathrm{SL}(V)$ is the subgroup of $\mathrm{GL}(V)$ consisting of all the unimodular transformations.

Choosing an ordered basis of V gives an isomorphism $\mathrm{SL}(V) \to \mathrm{SL}(m, K)$, the group of all unimodular matrices. If $K = \mathrm{GF}(q)$, we may denote $\mathrm{SL}(m, K)$ by $\mathrm{SL}(m, q)$.

The following elementary matrices are introduced to analyze the structure of $\mathrm{SL}(m, K)$.

Definition. Let λ be a nonzero element of a field K, and let $i \neq j$ be integers between 1 and m. An ***elementary transvection*** $B_{ij}(\lambda)$ is the $m \times m$ matrix differing from the identity matrix E in that it has λ as its ij entry. A ***transvection*** is

a matrix B that is similar to some $B_{ij}(\lambda)$; that is, B is a conjugate of some $B_{ij}(\lambda)$ in $\mathrm{GL}(m, K)$.

Every transvection is unimodular. Note that the inverse of an elementary transvection is another such: $B_{ij}(\lambda)^{-1} = B_{ij}(-\lambda)$; it follows that the inverse of any transvection is also a transvection.

If $A \in \mathrm{GL}(m, K)$, then $B_{ij}(\lambda)A$ is the matrix obtained from A by adding λ times its jth row to its ith row.

Lemma 8.7. *Let K be a field. If $A \in \mathrm{GL}(m, K)$ and $\det A = \mu$, then $A = UD(\mu)$, where U is a product of elementary transvections and $D = \mathrm{diag}\{1, 1, \ldots, 1, \mu\}$.*

Proof. We prove, by induction on $t \leq m - 1$, that A can be transformed, by a sequence of elementary operations which add a multiple of one row to another, into a matrix of the form

$$A_t = \begin{bmatrix} E_t & * \\ 0 & C \end{bmatrix},$$

where E_t is the $t \times t$ identity matrix.

For the base step, note that the first column of A is not zero (A is nonsingular). Adding some row to the second row if necessary, we may assume that $\alpha_{21} \neq 0$. Now add $\alpha_{21}^{-1}(1 - \alpha_{11})$ times row 2 to row 1 get entry 1 in the upper left corner. We may now make the other entries in column 1 equal corner. We may now make the other entries in column 1 equal to zero by adding suitable multiples of row 1 to the other rows, and so A has been transformed into A_1.

For the inductive step, we may assume that A has been transformed into a matrix A_t as displayed above. Note that C is nonsingular (for $\det A_t = \det C$). Assuming that C has at least two rows, we may further assume, as in the base step, that its upper left corner $\gamma_{t+1,t+1} = 1$ (this involves only the rows of C, hence does not disturb the top t rows of A_t). Adding on a suitable multiple of row $t + 1$ to the other rows of A_t yields a matrix A_{t+1}.

We may now assume that A has been transformed into

$$\begin{bmatrix} E_t & * \\ 0 & \mu \end{bmatrix}$$

where $\mu \in K$ and $\mu \neq 0$. Adding suitable multiples of the last row to the other rows cleans out the last column, leaving $D(\mu)$.

In terms of matrix multiplication, we have shown that there is a matrix P, which is a product of elementary transvections, with $PA = D(\mu)$. Therefore, $A = P^{-1}D(\mu)$; this completes the proof because the inverse of an elementary transvection is another such. ∎

Theorem 8.8.

(i) $\mathrm{GL}(m, K)$ *is a semidirect product of* $\mathrm{SL}(m, K)$ *by* $K^\times$.
(ii) $\mathrm{SL}(m, K)$ *is generated by elementary transvections.*

Proof. (i) We know that $\mathrm{SL} \lhd \mathrm{GL}$ (because SL = ker det), and it is easy to see that $\Delta = \{D(\mu): \mu \in K^\times\}$ $(\cong K^\times)$ is a complement of SL.

(ii) By (i), each $A \in \mathrm{GL}$ has a unique factorization $A = UD(\mu)$, where $U \in \mathrm{SL}$, $D(\mu) \in \Delta$, and $\det A = \mu$. Therefore, A is unimodular if and only if $A = U$. The result now follows from Lemma 8.7. ∎

Notation. If V is an m-dimensional vector space over a field K, let $\mathrm{Z}(V)$ denote the subgroup of $\mathrm{GL}(V)$ consisting of all scalar transformations, and let $\mathrm{SZ}(V)$ be the subgroup of $\mathrm{Z}(V)$ consisting of all unimodular scalar transformations.

Let $\mathrm{Z}(m, K) \cong \mathrm{Z}(V)$ denote the subgroup of all $m \times m$ scalar matrices αE, and let $\mathrm{SZ}(m, K) \cong \mathrm{SZ}(V)$ denote the subgroup of all αE with $\alpha^m = 1$. If $K = \mathrm{GF}(q)$, we may also denote these subgroups by $\mathrm{Z}(m, q)$ and $\mathrm{SZ}(m, q)$, respectively.

Theorem 8.9.

(i) *The center of* $\mathrm{GL}(V)$ *is* $\mathrm{Z}(V)$.
(ii) *The center of* $\mathrm{SL}(m, K)$ *is* $\mathrm{SZ}(m, K)$.

Proof. (i) If $T \in \mathrm{GL}(V)$ is not a scalar transformation, then there is $v \in V$ with $\{v, Tv\}$ independent; extend this to a basis $\{v, Tv, u_3, \ldots, u_m\}$ of V. It is easy to see that $\{v, v + Tv, u_3, \ldots, u_m\}$ is also a basis of V, so that there is a (non-singular) linear transformation $S: V \to V$ with $Sv = v$, $S(Tv) = v + Tv$, and $Su_i = u_i$ for all $i \geq 3$. Now T and S do not commute, for $TS(v) = Tv$ while $ST(v) = v + TV$. Therefore, $T \notin \mathrm{Z}(\mathrm{GL}(V))$, and it follows that $\mathrm{Z}(\mathrm{GL}(V)) = \mathrm{Z}(V)$.

(ii) Assume now that $T \in \mathrm{SL}(V)$, that T is not scalar, and that S is the linear transformation constructed in (i). The matrix of S relative to the basis $\{v, Tv, u_3, \ldots, u_m\}$ is the elementary transvection $B_{12}(1)$, so that $\det(S) = 1$ and $S \in \mathrm{SL}(V)$. As in (i), $T \notin \mathrm{Z}(\mathrm{SL}(V))$; that is, if $T \in \mathrm{Z}(\mathrm{SL}(V))$, then $T = \alpha E$ for some $\alpha \in K$. Finally, $\det(\alpha E) = \alpha^m$, and so $\alpha^m = 1$, so that $\mathrm{SZ}(V) = \mathrm{Z}(\mathrm{SL}(V))$. ∎

Theorem 8.10. $|\mathrm{SZ}(m, q)| = d$, *where* $d = (m, q - 1)$.

Proof. Let $K = \mathrm{GF}(q)$. We first show, for all $\alpha \in K^\times$, that $\alpha^m = 1$ if and only if $\alpha^d = 1$. Since d divides m, $\alpha^d = 1$ implies $\alpha^m = 1$. Conversely, there are

integers a and b with $d = am + b(q-1)$. Thus

$$\alpha^d = \alpha^{am+b(q-1)} = \alpha^{ma}\alpha^{(q-1)b} = \alpha^{ma},$$

because $\alpha^{q-1} = 1$. Hence $\alpha^m = 1$ gives $1 = \alpha^{ma} = \alpha^d$.

It follows that $\mathrm{SZ}(m, q) \cong \{\alpha \in K^\times : \alpha^m = 1\} \cong \{\alpha \in K^\times : \alpha^d = 1\}$. Therefore, if π is a generator of $K^\times$, then $\mathrm{SZ}(m, q) \cong \langle \pi^{n/d} \rangle$ and hence $|\mathrm{SZ}(m, q)| = d$. ∎

Our preceding discussion allows us to lengthen the normal series in Lemma 8.6 as follows:

$$\mathrm{GL}(m, q) > M(p_1) > M(p_1 p_2) > \cdots > \mathrm{SL}(m, q) > \mathrm{SZ}(m, q) > 1.$$

The center $\mathrm{SZ}(m, q)$ is abelian and so its composition factors are no secret (they are cyclic groups of prime order, occurring with multiplicity, for all primes dividing $q - 1$). We now consider the last factor group in this series.

Definition. If V is an m-dimensional vector spaces over a field K, the ***projective unimodular group*** $\mathrm{PSL}(V)$ is the group $\mathrm{SL}(V)/\mathrm{SZ}(V)$.

A choice of ordered basis of V induces an isomorphism $\varphi: \mathrm{SL}(V) \xrightarrow{\sim} \mathrm{SL}(m, K)$ with $\varphi(\mathrm{SZ}(V)) = \mathrm{SZ}(m, K)$, so that $\mathrm{PSL}(V) \cong \mathrm{SL}(m, K)/\mathrm{SZ}(m, K)$. The latter group is denoted by $\mathrm{PSL}(m, K)$. When $K = \mathrm{GF}(q)$, we may denote $\mathrm{PSL}(m, K)$ by $\mathrm{PSL}(m, q)$.

We shall see, in Chapter 9, that these groups are intimately related to projective geometry, whence their name.

Theorem 8.11. *If* $d = (m, q-1)$, *then*

$$|\mathrm{PSL}(m, q)| = (q^m - 1)(q^m - q)\ldots(q^m - q^{m-1})/d.$$

Proof. Immediate from Theorems 8.5 and 8.10. ∎

EXERCISES

8.1. Let $H \lhd \mathrm{SL}(2, K)$, and let $A \in H$. Using the factorization $A = UD(\mu)$ (in the proof of Theorem 8.8), show that if A is similar to $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, then there is $\mu \in K^\times$ such that H contains $\begin{bmatrix} \alpha & \mu^{-1}\beta \\ \mu\gamma & \delta \end{bmatrix}$.

8.2. Let $B = B_{ij}(1) \in \mathrm{GL}(m, K) = G$. Prove that $C_G(B)$ consists of all those nonsingular matrices $A = [a_{ij}]$ whose ith column, aside from a_{ii}, and whose jth row, aside from a_{jj}, consist of all 0's.

8.3. Let $\Delta \leq \mathrm{GL}(m, K)$ be the subgroup of all nonsingular diagonal matrices.
(i) Show that Δ is an abelian *self-centralizing* subgroup; that is, if $A \in \mathrm{GL}(m, K)$ commutes with every $D \in \Delta$, then $A \in \Delta$.
(ii) Use part (i) to give another proof that $\mathrm{Z}(\mathrm{GL}(m, K)) = \mathrm{Z}(m, K)$ consists of the scalar matrices.

PSL(2, K)

In this section, we concentrate on the case $m = 2$ with the aim of proving that PSL(2, q) is simple whenever $q > 3$. We are going to see that elementary transvections play the same role here that 3-cycles play in the analysis of the alternating groups.

Definition. A field K is ***perfect*** if either it has characteristic 0 or it has prime characteristic p and every $\lambda \in K$ has a pth root in K.

If K has prime characteristic p, then the map $F\colon K \to K$, given by $\lambda \mapsto \lambda^p$, is an injective homomorphism. If K is finite, then F must be surjective; that is, every finite field is perfect. Clearly, every algebraically closed field K is perfect. An example of a nonperfect field is $K = \mathbb{Z}_p(x)$, the field of all rational functions with coefficients in $\mathbb{Z}_p$; the indeterminate x does not have a pth root in $\mathbb{Z}_p(x)$.

Lemma 8.12. *Let K be a field which either has characteristic $\neq 2$ or is perfect of characteristic 2. If a normal subgroup H of* SL(2, K) *contains an elementary transvection $B_{12}(\lambda)$ or $B_{21}(\lambda)$, then $H = \mathrm{SL}(2, K)$.*

Proof. Note first that if $B_{21}(\lambda) \in H$, then $UB_{21}(\mu)U^{-1} = B_{12}(-\mu)$, where

$$U = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

By Theorem 8.8(ii), it suffices to prove that H contains every elementary transvection. Conjugate $B_{12}(\lambda)$ by a unimodular matrix:

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = \begin{bmatrix} 1 - \lambda\alpha\gamma & \lambda\alpha^2 \\ -\lambda\gamma^2 & 1 + \lambda\alpha\gamma \end{bmatrix}.$$

In particular, if $\gamma = 0$, then $\alpha \neq 0$ and this conjugate is $B_{12}(\lambda\alpha^2)$. Since H is normal in SL, these conjugates lie in H. Define

$$\Gamma = \{0\} \cup \{\mu \in K\colon B_{12}(\mu) \in H\},$$

It is easy to see that Γ is a subgroup of the additive group K, and so it contains all elements of the form $\lambda(\alpha^2 - \beta^2)$, where $\alpha, \beta \in K$.

We claim that $\Gamma = K$, and this will complete the proof. If K has character-

istic $\neq 2$, then each $\mu \in K$ is a difference of squares:

$$\mu = [\tfrac{1}{2}(\mu + 1)]^2 - [\tfrac{1}{2}(\mu - 1)]^2.$$

For each $\mu \in K$, therefore, there are $\alpha, \beta \in K$ with $\lambda^{-1}\mu = \alpha^2 - \beta^2$, so that $\mu = \lambda(\alpha^2 - \beta^2) \in \Gamma$, and $\Gamma = K$. If K has characteristic 2 and is perfect, then every element in K has a square root in K. In particular, there is $\alpha \in K$ with $\lambda^{-1}\mu = \alpha^2$, and Γ contains $\lambda\alpha^2 = \mu$. ▩

The next theorem was proved by C. Jordan (1870) for q prime. In 1893, after F. Cole had discovered a simple group G of order 504, E.H. Moore recognized G as PSL(2, 8), and then proved the simplicity of PSL(2, q) for all prime powers $q > 3$.

Theorem 8.13 (Jordan–Moore). *The groups* PSL(2, q) *are simple if and only if* $q > 3$.

Proof. By Theorem 8.11,

$$|\mathrm{PSL}(2, q)| = \begin{cases} (q^2 - 1)(q^2 - q) & \text{if } q = 2^n, \\ (q^2 - 1)(q^2 - q)/2 & \text{if } q = p^n, \ p \text{ an odd prime.} \end{cases}$$

Therefore, PSL(2, 2) has order 6 and PSL(2, 3) has order 12, and there are no simple groups of these orders.

Assume now that $q \geq 4$. It suffices to prove that a normal subgroup H of SL(2, q) which contains a matrix not in SZ(2, q) must be all of SL(2, q).

Suppose that H contains a matrix

$$A = \begin{bmatrix} \alpha & 0 \\ \beta & \alpha^{-1} \end{bmatrix},$$

where $\alpha \neq \pm 1$; that is, $\alpha^2 \neq 1$. If $B = B_{21}(1)$, then H contains the commutator $BAB^{-1}A^{-1} = B_{21}(1 - \alpha^{-2})$, which is an elementary transvection. Therefore, $H = \mathrm{SL}(2, q)$, by Lemma 8.12.

To complete the proof, we need only display a matrix in H whose top row is $[\alpha \ 0]$, where $\alpha \neq \pm 1$. By hypothesis, there is a matrix M in H, not in SZ(2, q), and M is similar to either a diagonal matrix or a matrix of the form

$$\begin{bmatrix} 0 & -1 \\ 1 & \beta \end{bmatrix},$$

for the only rational canonical forms for a 2×2 matrix are: two 1×1 blocks (i.e., a diagonal matrix) or a 2×2 companion matrix (which has the above form because it is unimodular). In the first case, Exercise 8.1 shows that

$$C = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \in H;$$

since C is unimodular, $\alpha\beta = 1$; since M is not in SZ(m, q), $\alpha \neq \beta$. It follows

that $\alpha \neq \pm 1$, and C is the desired matrix. In the second case, Exercise 8.1 shows that H contains

$$D = \begin{bmatrix} 0 & -\mu^{-1} \\ \mu & \beta \end{bmatrix}.$$

If $T = \operatorname{diag}\{\alpha^{-1}, \alpha\}$, where α is to be chosen, then H contains the commutator

$$U = TDT^{-1}D^{-1} = \begin{bmatrix} \alpha^{-2} & 0 \\ \mu\beta(\alpha^2 - 1) & \alpha^2 \end{bmatrix}.$$

We are done if $\alpha^{-2} \neq \pm 1$; that is, if $\alpha^4 \neq 1$. If $q > 5$, then such an α exists, for a field contains at most four roots of $x^4 - 1$. If $q = 4$, then every $\mu \in K$ satisfies the equation $x^4 - x = 0$, so that $\alpha \neq 1$ implies $\alpha^4 \neq 1$.

Only the case $\mathrm{GF}(5) \cong \mathbb{Z}_5$ remains. Consider the factor β occurring in the lower left corner $\lambda = \mu\beta(\alpha^2 - 1)$ of U. If $\beta \neq 0$, choose $\alpha = [2] \in \mathbb{Z}_5$; then $\alpha^2 - 1 \neq 0$ and $U = B_{21}(\lambda)$. Hence H contains the elementary transvection $U^2 = B_{21}(-2\lambda)$ and we are done. If $\beta = 0$, then

$$D = \begin{bmatrix} 0 & -\mu^{-1} \\ \mu & 0 \end{bmatrix} \in H.$$

Therefore, the normal subgroup H contains

$$B_{12}(v)DB_{12}(-v) = \begin{bmatrix} \mu v & -\mu v^2 - \mu^{-1} \\ * & * \end{bmatrix}$$

for all $v \in \mathbb{Z}_5$. If $v = 2\mu^{-1}$, then the top row of this last matrix is $[2\ 0]$, and the theorem is proved. ∎

Corollary 8.14. *If K is an infinite field which either has characteristic $\neq 2$ or is perfect of characteristic 2, then* $\mathrm{PSL}(2, K)$ *is a simple group.*

Proof. The finiteness of K in the proof of the theorem was used only to satisfy the hypotheses of Lemma 8.12. ∎

Remark. In Theorem 9.48, we will prove that $\mathrm{PSL}(2, K)$ is simple for every infinite field.

Corollary 8.15. $\mathrm{SL}(2, 5)$ *is not solvable.*

Proof. Every quotient of a solvable group is solvable. ∎

We have exhibited an infinite family of simple groups. Are any of its members distinct from simple groups we already know? Using Theorem 8.11, we see that both $\mathrm{PSL}(2, 4)$ and $\mathrm{PSL}(2, 5)$ have order 60. By Exercise 4.37, all simple groups of order 60 are isomorphic:

$$\mathrm{PSL}(2, 4) \cong A_5 \cong \mathrm{PSL}(2, 5).$$

If $q = 7$, however, then we do get a new simple group, for $|\mathrm{PSL}(2, 7)| = 168$, which is neither prime nor $\frac{1}{2}n!$. If we take $q = 8$, we see that there is a simple group of order 504; if $q = 11$, we see a simple group of order 660. (It is known that the only other isomorphisms involving A_n's and PSLs, aside from those displayed above, are Exercise 8.12: $\mathrm{PSL}(2, 9) \cong A_6$ (these groups have order 360); Exercise 9.26: $\mathrm{PSL}(2, 7) \cong \mathrm{PSL}(3, 2)$ (these groups have order 168); Theorem 9.73: $\mathrm{PSL}(4, 2) \cong A_8$ (these groups have order 20, 160).)

EXERCISES

8.4. Show that the Sylow p-subgroups of $\mathrm{SL}(2, 5)$ are either cyclic (when p is odd) or quaternion (when $p = 2$). Conclude that $\mathrm{SL}(2, 5) \ncong S_5$.

8.5. What is the Sylow 2-subgroup of $\mathrm{SL}(2, 3)$?

8.6. (i) Show that $\mathrm{PSL}(2, 2) \cong S_3$.
(ii) Show that $\mathrm{SL}(2, 3) \ncong S_4$ but that $\mathrm{PSL}(2, 3) \cong A_4$.

8.7. What are the composition factors of $\mathrm{GL}(2, 7)$?

8.8. Show that if $H \lhd \mathrm{GL}(2, K)$, where K has more than three elements, then either $H \leq Z(\mathrm{GL}(2, K))$ or $\mathrm{SL}(2, K) \leq H$.

8.9. (i) What is the commutator subgroup of $\mathrm{GL}(2, 2)$?
(ii) What is the commutator subgroup of $\mathrm{GL}(2, 3)$?
(iii) If $q > 3$, prove that the commutator subgroup of $\mathrm{GL}(2, q)$ is $\mathrm{SL}(2, q)$.

8.10. Prove, for every field K, that all transvections are conjugate in $\mathrm{GL}(2, K)$.

8.11. Let A be a unimodular matrix. Show that A determines an involution in $\mathrm{PSL}(2, K)$ if and only if A has trace 0, and that A determines an element of order 3 in $\mathrm{PSL}(2, K)$ if and only if A has trace ± 1. (*Hint.* Use canonical forms.)

8.12. Prove that any two simple groups of order 360 are isomorphic, and conclude that $\mathrm{PSL}(2, 9) \cong A_6$. (*Hint.* Show that a Sylow 5-subgroup has six conjugates.)

PSL(m, K)

The simplicity of $\mathrm{PSL}(m, K)$ for all $m \geq 3$ and all fields K will be proved in this section. In 1870, C. Jordan proved this theorem for $K = \mathbb{Z}_p$, and L.E. Dickson extended the result to all finite fields K in 1897, four years after Moore had proved the result for $m = 2$. The proof we present, due to E. Artin, is much more elegant than matrix manipulations (though we prefer matrices when $m = 2$).

An $m \times m$ elementary transvection $B_{ij}(\lambda)$ represents a linear transformation T on an m-dimensional vector space V over K. There is an ordered basis $\{v_1, \ldots, v_m\}$ of V with $Tv_l = v_l$ for all $l \neq i$ and with $Tv_i = v_i + \lambda v_j$. Note that T fixes every vector in the $(m - 1)$-dimensional subspace H spanned by all $v_l \neq v_i$.

Definition. If V is an m-dimensional vector space over a field K, then a ***hyperplane*** H in V is a subspace of dimension $m - 1$.

The linear transformation T arising from an elementary transvection fixes the hyperplane H pointwise. If $w \in V$ and $w \notin H$, then $\langle w \rangle = \{\mu w: \mu \in K\}$ is a transversal of H in V: the vector space V, considered as an additive group, is the disjoint union of the cosets $H + \mu w$. Hence, every vector $v \in V$ has a unique expression of the form

$$v = \mu w + h, \qquad \mu \in K, \quad h \in H.$$

Lemma 8.16. *Let H be a hyperplane in V and let $T \in \mathrm{GL}(V)$ fix H pointwise. If $w \in V$ and $w \notin H$, then*

$$T(w) = \mu w + h_0$$

for some $\mu \in K$ and $h_0 \in H$. Moreover, given any $v \in V$,

$$T(v) = \mu w + h',$$

for some $h' \in H$.

Proof. We observed above that every vector in V has an expression of the form $\lambda w + h$. In particular, $T(w)$ has such an expression. If $v \in V$, then $v = \lambda w + h''$ for some $\lambda \in K$ and $h'' \in H$. Since T fixes H,

$$\begin{aligned} T(v) &= \lambda T(w) + h'' = \lambda(\mu w + h_0) + h'' \\ &= \mu(\lambda w + h'') + [(1 - \mu)h'' + \lambda h_0] \\ &= \mu v + h'. \end{aligned}$$

The scalar $\mu = \mu(T)$ in Lemma 8.16 is thus determined uniquely by any T fixing a hyperplane pointwise.

Definition. Let $T \in \mathrm{GL}(V)$ fix a hyperplane H pointwise, and let $\mu = \mu(T)$. If $\mu \neq 1$, then T is called a ***dilatation***; if $\mu = 1$ and if $T \neq 1_V$, then T is called a ***transvection***.

The next theorem and its corollary show that the transvections just defined are precisely those linear transformations arising from matrix transvections.

Theorem 8.17. *Let $T \in \mathrm{GL}(V)$ fix a hyperplane H pointwise, and let $\mu = \mu(T)$.*

(i) *If T is a dilatation, then T has a matrix $D(\mu) = \mathrm{diag}\{1, \ldots, 1, \mu\}$ (relative to a suitable basis of V).*

(ii) *If T is a transvection, then T has matrix $B_{12}(1)$ (relative to a suitable basis of V). Moreover, T has no eigenvectors outside of H in this case.*

Proof. Every nonzero vector in H is an eigenvector of T (with eigenvalue 1);

are there any others? Choose $w \in V$ with $w \notin H$; since T fixes H pointwise,

$$Tw = \mu w + h, \qquad \text{where} \quad h \in H.$$

If $v \in V$ and $v \notin H$, the lemma gives

$$Tv = \mu v + h',$$

where $h' = (1 - \mu)h'' + \lambda h_0 \in H$. If v is an eigenvector of T, then $Tv = \beta v$ for some $\beta \in K$. But $Tv = \beta v$ if and only if $\beta = \mu$ and $\lambda h = (\mu - 1)h''$: sufficiency is obvious; conversely, if $\beta v = \mu v + h'$, then $(\beta - \mu)v = h' \in \langle v \rangle \cap H = 0$.

(i) If T is a dilatation, then $\mu - 1 \neq 0$ and $h'' = \lambda(\mu - 1)^{-1}h$. It follows that $v = w + (\mu - 1)^{-1}h$ is an eigenvector of T for the eigenvalue μ. If $\{v_1, \ldots, v_{m-1}\}$ is a basis of H, then adjoining v gives a basis of V, and the matrix of T relative to this basis is $D(\mu) = \operatorname{diag}\{1, \ldots, 1, \mu\}$.

(ii) If T is a transvection, then $\mu = 1$. Choose $w \notin H$ so that $Tw = w + h$, where $h \in H$ and $h \neq 0$. If $v \notin H$ is an eigenvector of T, then $\alpha v = Tv = v + h$ for some $\alpha \in K$; hence, $(\alpha - 1)v \in \langle v \rangle \cap H = 0$, so that $\alpha = 1$ and $Tv = v$. It follows that $T = 1_V$, contradicting the proviso in the definition of transvection excluding the identity. Therefore, T has no eigenvectors outside of H. If $\{h, h_3, \ldots, h_m\}$ is a basis of H, then adjoining w as the first vector gives an ordered basis of V, and the matrix of T relative to this basis is $B_{12}(1)$. ▩

Corollary 8.18. *All transvections in* GL(*m*, *K*) *are conjugate.*

Proof. Since transvections are, by definition, conjugates of elementary transvections, it suffices to prove that any two elementary transvections are conjugate to $B_{21}(1)$. Let V be an m-dimensional vector space over K with basis $\{v_1, \ldots, v_m\}$, and let T be the linear transformation with $Tv_1 = v_1 + v_2$ and $Tv_l = v_l$ for all $l \geq 2$. If $i \neq j$ and $\lambda \neq 0$, define a new ordered basis $\{u_1, \ldots, u_m\}$ of V as follows: put v_1 in position i, put $\lambda^{-1}v_2$ in position j, and fill the remaining $m - 2$ positions with $v_3, \ldots, v_m$ in this order (e.g., if $m = 5$, $i = 2$, and $j = 4$, then $\{u_1, \ldots, u_5\} = \{v_3, v_1, v_4, \lambda^{-1}v_2, v_5\}$). The matrix of T relative to this new ordered basis is easily seen to be $B_{ij}(\lambda)$. Therefore $B_{21}(1)$ and $B_{ij}(\lambda)$ are similar, for they represent the same linear transformation relative to different choices of ordered basis. ▩

If $T \in \mathrm{GL}(V)$ is a transvection fixing a hyperplane H and if $w \notin H$, then $Tw = w + h$ for some nonzero $h \in H$. If $v \in V$, then $v = \lambda w + h''$ for some $\lambda \in K$ and $h'' \in H$, and (*) in the proof of Lemma 8.16 gives $Tv = v + \lambda h$ (because $1 - \mu = 0$). The function $\varphi\colon V \to K$, defined by $\varphi(v) = \varphi(\lambda w + h) = \lambda$ is a K-linear transformation (i.e., it is a ***linear functional***) with kernel H. For each transvection T, there is thus a linear functional φ and a vector $h \in \ker \varphi$ with

$$Tv = v + \varphi(v)h \qquad \text{for all} \quad v \in V.$$

Notation. Given a nonzero linear functional φ on V and a nonzero vector

$h \in \ker \varphi$, define $\{\varphi, h\}: V \to V$ by

$$\{\varphi, h\}: v \mapsto v + \varphi(v)h.$$

It is clear that $\{\varphi, h\}$ is a transvection; moreover, for every transvection T, there exist $\varphi \neq 0$ and $h \neq 0$ with $T = \{\varphi, h\}$.

Lemma 8.19. *Let V be a vector space over K.*

(i) *If φ and ψ are linear functionals on V, and if $h, l \in V$ satisfy $\varphi(h) = \psi(h) = \varphi(l)$, then*

$$\{\varphi, h\} \circ \{\varphi, l\} = \{\varphi, h + l\} \qquad \text{and} \qquad \{\varphi, h\} \circ \{\psi, h\} = \{\varphi + \psi, h\}.$$

(ii) *For all $\alpha \in K^\times$,*

$$\{\alpha\varphi, h\} = \{\varphi, \alpha h\}.$$

(iii) *$\{\varphi, h\} = \{\psi, l\}$ if and only if there is a scalar $\alpha \in K^\times$ with*

$$\psi = \alpha\varphi \qquad \text{and} \qquad h = \alpha l.$$

(iv) *If $S \in \mathrm{GL}(V)$, then*

$$S\{\varphi, h\}S^{-1} = \{\varphi S^{-1}, Sh\}.$$

Proof. All are routine. For example, let us prove half of (iii). If $\{\varphi, h\} = \{\psi, l\}$, then $\varphi(v)h = \psi(v)l$ for all $v \in V$. Since $\varphi \neq 0$, there is $v \in V$ with $\varphi(v) \neq 0$, so that $h = \varphi(v)^{-1}\psi(v)l$; if $\alpha = \varphi(v)^{-1}\psi(v)$, then $h = \alpha l$. To see that $\psi(u) = \alpha\varphi(u)$ for all $u \in V$, note that $\varphi(u) = 0$ if and only if $\psi(u) = 0$ (because both $h, l \neq 0$). If $\psi(u)$ and $\varphi(u)$ are nonzero, then $h = \varphi(u)^{-1}\psi(u)l$ implies $\varphi(u)^{-1}\psi(u) = \varphi(v)^{-1}\psi(v) = \alpha$, and so $\psi = \alpha\varphi$. ∎

Theorem 8.20. *The commutator subgroup of $\mathrm{GL}(V)$ is $\mathrm{SL}(V)$ unless V is a two-dimensional vector space over $\mathbb{Z}_2$.*

Proof. Now $\det: \mathrm{GL} \to K^\times$ has kernel SL and $\mathrm{GL}/\mathrm{SL} \cong K^\times$; since $K^\times$ is abelian, $(\mathrm{GL})' \leq \mathrm{SL}$.

For the reverse inclusion, let $\nu: \mathrm{GL} \to \mathrm{GL}/(\mathrm{GL})'$ be the natural map. By Corollary 8.18, all transvections are conjugate in GL, and so $\nu(T) = \nu(T')$ for all transvections T and T'; let d denote their common value. Let $T = \{\varphi, h\}$ be a transvection. If we avoid the exceptional case in the statement, then H contains a nonzero vector l (not necessarily distinct from h) with $h + l \neq 0$. By the lemma, $\{\varphi, h\} \circ \{\varphi, l\} = \{\varphi, h + l\}$ (these are transvections because $l \neq 0$ and $h + l \neq 0$). Applying ν to this equation gives $d^2 = d$ in $\mathrm{GL}/(\mathrm{GL})'$, whence $d = 1$. Thus, every transvection $T \in \ker \nu = (\mathrm{GL})'$. But SL is generated by the transvections, by Theorem 8.8(ii), and so $\mathrm{SL} \leq (\mathrm{GL})'$. ∎

If V is a two-dimensional vector space over $\mathbb{Z}_2$, then $\mathrm{GL}(V)$ is a genuine

exception to the theorem. In this case,

$$\mathrm{GL}(V) = \mathrm{SL}(V) \cong \mathrm{SL}(2, 2) \cong \mathrm{PSL}(2, 2) \cong S_3,$$

and $(S_3)' = A_3$, a proper subgroup.

We have seen that any two transvections are conjugate in GL. It is easy to see that

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

are not conjugate in SL(2, 3); indeed, these transvections are not conjugate in SL(2, K) for any field K in which -1 is not a square. The assumption $m \geq 3$ in the next result is thus essential.

Theorem 8.21. *If $m \geq 3$, then all transvections are conjugate in* SL(V).

Proof. Let $\{\varphi, h\}$ and $\{\psi, l\}$ be transvections, and let $H = \ker \varphi$ and $L = \ker \psi$ be the hyperplanes fixed by each. Choose $v, u \in V$ with $\varphi(v) = 1 = \psi(u)$ (hence $v \notin H$ and $u \notin L$). There are bases $\{h, h_2, \ldots, h_{m-1}\}$ and $\{l, l_2, \ldots, l_{m-1}\}$ of H and L, respectively, and adjoining v and u gives bases $\{v, h, h_2, \ldots, h_{m-1}\}$ and $\{u, l, l_2, \ldots, l_{m-1}\}$ of V. If $S \in \mathrm{GL}(V)$ takes the first of these ordered bases to the second, then

$$(*) \qquad S(v) = u, \qquad S(H) = L, \qquad \text{and} \qquad S(h) = l.$$

Let $\det S = d$; we now show that we can force S to have determinant 1. Since $m \geq 3$, the first basis of V constructed above contains at least one other vector (say, h_{m-1}) besides v and h. Redefine S so that $S(h_{m-1}) = d^{-1}l_{m-1}$. Relative to the basis $\{v, h, h_2, \ldots, h_{m-1}\}$, the matrix of the new transformation differs from the matrix of the original one in that its last column is multiplied by d^{-1}. The new S thus has determinant 1 as well as the other properties $(*)$ of S.

Now $S\{\varphi, h\}S^{-1} = \{\varphi S^{-1}, Sh\} = \{\varphi S^{-1}, l\}$, by Lemma 8.19(iv). Since φS^{-1} and ψ agree on the basis $\{u, l, l_2, \ldots, l_{m-1}\}$ of V, they are equal. Therefore $\{\varphi, h\}$ and $\{\psi, l\}$ are conjugate in SL, as desired. ∎

Notation. If H is a hyperplane in a vector space V, then

$$\mathscr{T}(H) = \{\text{all transvections fixing } H\} \cup \{1_V\}.$$

Lemma 8.22. *Let H be a hyperplane in an m-dimensional vector space V over K.*

(i) *There is a linear functional φ with $H = \ker \varphi$ so that*

$$\mathscr{T}(H) = \{\{\varphi, h\} : h \in H\} \cup \{1_V\}.$$

(ii) *$\mathscr{T}(H)$ is an (abelian) subgroup of* SL(V), *and $\mathscr{T}(H) \cong H$.*

(iii) *The centralizer $C_{\mathrm{SL}}(\mathscr{T}(H)) = \mathrm{SZ}(V)\mathscr{T}(H)$.*

Proof. (i) Observe that linear functionals φ and ψ have the same kernel if and only if there is a nonzero $\alpha \in K$ with $\psi = \alpha\varphi$. Clearly $\psi = \alpha\varphi$ implies $\ker \psi = \ker \varphi$. Conversely, if H is their common kernel, choose $w \in V$ with $w \notin H$. Now $\psi(w) = \alpha\varphi(w)$ for some $\alpha \in K^\times$. If $v \in V$, then $v = \lambda w + h$, for some $\lambda \in K$ and $h \in H$, and $\psi(v) = \lambda\psi(w) = \lambda\alpha\varphi(w) = \alpha\varphi(\lambda w + h) = \alpha\varphi(v)$.

If $\{\varphi, h\}$, $\{\psi, l\} \in \mathscr{T}(H)$, then Lemma 8.19(ii) gives $\{\psi, l\} = \{\alpha\varphi, l\} = \{\varphi, \alpha l\}$. Since $\{\varphi, h\}^{-1} = \{\varphi, -h\}$, Lemma 8.19(i) gives $\{\varphi, h\} \circ \{\psi, l\}^{-1} = \{\varphi, h - \alpha l\} \in \mathscr{T}(H)$. Therefore, $\mathscr{T}(H) \leq \mathrm{SL}(V)$.

(ii) Let φ be a linear functional with $H = \ker \varphi$. By (i), each $T \in \mathscr{T}(H)$ has the form $T = \{\varphi, h\}$ for some $h \in H$, and this form is unique, by Lemma 8.19(iii). It is now easy to see that the function $\mathscr{T}(H) \to H$, given by $\{\varphi, h\} \mapsto h$, is an isomorphism.

(iii) Since $\mathscr{T}(H)$ is abelian, $\mathrm{SZ}(V)\mathscr{T}(H) \leq C_{\mathrm{SL}}(\mathscr{T}(H))$. For the reverse inclusion, assume that $S \in \mathrm{SL}(V)$ commutes with every $\{\varphi, h\}$: for all $h \in H$, $S\{\varphi, h\}S^{-1} = \{\varphi, h\}$. By Lemma 8.19(iv), $S\{\varphi, h\}S^{-1} = \{\varphi S^{-1}, Sh\}$, and so Lemma 8.19(iii) gives $\alpha \in K^\times$ with

$$(**) \qquad \varphi S^{-1} = \alpha\varphi \qquad \text{and} \qquad Sh = \alpha^{-1}h.$$

Hence αS fixes H pointwise, so that αS is either a transvection or a dilatation. If αS is a transvection, then $\alpha S \in \mathscr{T}(H)$, and so $S = \alpha^{-1}(\alpha S) \in \mathrm{SZ}(V)\mathscr{T}(H)$. If αS is a dilatation, then it has an eigenvector w outside of H, and $\alpha Sw = \mu w$, where $1 \neq \mu = \det \alpha S = \alpha^m$ (for $\det S = 1$); hence, $Sw = \alpha^{m-1}w$. But $\varphi S^{-1}w = \varphi(\alpha^{-m+1}w) = \alpha^{-m+1}\varphi(w)$, so that $(**)$ give $\varphi(w) = \alpha^m\varphi(w)$. Since $\varphi(w) \neq 0$ (because $w \notin H$), we reach the contradiction $\alpha^m = 1$. ∎

Theorem 8.23 (Jordan–Dickson). *If $m \geq 3$ and V is an m-dimensional vector space over a field K, then the groups* $\mathrm{PSL}(V)$ *are simple.*

Proof. We show that if N is a normal subgroup of $\mathrm{SL}(V)$ containing some A not in $\mathrm{SZ}(V)$, then $N = \mathrm{SL}(V)$; by Theorem 8.17, it suffices to show that N contains a transvection.

Since $\mathrm{SL}(V)$ is generated by transvections, there exists a transvection T which does not commute with A: the commutator $B = T^{-1}A^{-1}TA \neq 1$. Note that $N \lhd \mathrm{SL}$ gives $B \in N$. Thus

$$B = T^{-1}(A^{-1}TA) = T_1T_2,$$

where each T_i is a transvection. Now $T_i = \{\varphi_i, h_i\}$, where $h_i \in H_i = \ker \varphi_i$ for $i = 1, 2$; that is,

$$T_i(v) = v + \varphi_i(v)h_i \qquad \text{for all} \quad v \in V.$$

Let W be the subspace $\langle h_1, h_2\rangle \leq V$, so that $\dim W \leq 2$. Since $\dim V \geq 3$, there is a hyperplane L of V containing W. We claim that $B(L) \leq L$. If $l \in L$, then

$$\begin{aligned} B(l) &= T_1T_2(l) = T_2(l) + \varphi_1(T_2(l))h_1 \\ &= l + \varphi_2(l)h_2 + \varphi_1(T_2(l))h_1 \in L + W \leq L. \end{aligned}$$

We now claim that $H_1 \cap H_2 \neq 0$. This is surely true if $H_1 = H_2$. If $H_1 \neq H_2$, then $H_1 + H_2 = V$ (hyperplanes are maximal subspaces) and $\dim(H_1 + H_2) = m$. Since

$$\dim H_1 + \dim H_2 = \dim(H_1 + H_2) + \dim(H_1 \cap H_2),$$

we have $\dim(H_1 \cap H_2) = m - 2 \geq 1$.

If $z \in H_1 \cap H_2$ with $z \neq 0$, then

$$B(z) = T_1 T_2(z) = z.$$

We may assume that B is not a transvection (or we are done); therefore, $B \notin \mathscr{T}(L)$, which is wholly comprised of transvections. If $B = \alpha S$, where $S \in \mathscr{T}(L)$, then z is an eigenvector of S ($z = Bz = \alpha Sz$, and so $Sz = \alpha^{-1}z$). As eigenvectors of transvections lie in the fixed hyperplane, $z \in L$ and so $\alpha = 1$, giving the contradiction $S = B$. Therefore, $B \notin \mathrm{SZ}(V)\mathscr{T}(L) = C_{\mathrm{SL}}(\mathscr{T}(L))$, so there exists $U \in \mathscr{T}(L)$ not commuting with B:

$$C = UBU^{-1}B^{-1} \neq 1;$$

of course, $C = (UBU^{-1})B^{-1} \in N$. If $l \in L$, then

$$C(l) = UBU^{-1}B^{-1}(l) = UB(B^{-1}(l)) = l,$$

because $B^{-1}(l) \in L$ and $U^{-1} \in \mathscr{T}(L)$ fixes L. Therefore, the transformation C fixes the hyperplane L, and so C is either a transvection or a dilatation. But C is not a dilatation because $\det C = 1$. Therefore C is a transvection in N, and the proof is complete. ∎

We shall give different proofs of Theorems 8.13 and 8.22 in Chapter 9.

Observe that $|\mathrm{PSL}(3, 4)| = 20{,}160 = \frac{1}{2}8!$, so that PSL(3, 4) and A_8 are simple groups of the same order.

Theorem 8.24 (Schottenfels, 1900). PSL(3, 4) *and* A_8 *are nonisomorphic simple groups of the same order.*

Proof. The permutations (1 2)(3 4) and (1 2)(3 4)(5 6)(7 8) are even (hence lie in A_8), are involutions, and are not conjugate in A_8 (indeed, they are not even conjugate in S_8 for they have different cycle structures). We prove the theorem by showing that all involutions in PSL(3, 4) are conjugate.

A nonscalar matrix $A \in \mathrm{SL}(3, 4)$ corresponds to an involution in PSL(3, 4) if and only if A^2 is scalar, and A^2 is scalar if and only if $(\mathrm{PAP}^{-1})^2$ is scalar for every nonsingular matrix P. Thus A can be replaced by anything similar to it, and so we may assume that A is a rational canonical form. If A is a direct sum of 1×1 companion matrices, then $A = \mathrm{diag}\{\alpha, \beta, \gamma\}$. But A^2 scalar implies $\alpha^2 = \beta^2 = \gamma^2$; as GF(4) has characteristic 2, this gives $\alpha = \beta = \gamma$ and A is scalar, a contradiction. If A is a 3×3 companion matrix,

$$A = \begin{bmatrix} \alpha & 0 & 0 \\ 1 & \alpha & 0 \\ 0 & 1 & \alpha \end{bmatrix},$$

then A^2 has 1 as the entry in position (3, 1), and so A^2 is not scalar. We conclude that A is a direct sum of a 1×1 companion matrix and a 2×2 companion matrix:

$$A = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & \beta \\ 0 & 1 & \gamma \end{bmatrix}.$$

Now $\det A = 1 = \alpha\beta$ (remember that $-1 = 1$ here), so that $\beta = \alpha^{-1}$, and A^2 scalar forces $\gamma = 0$. Thus,

$$A = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & \alpha^{-1} \\ 0 & 1 & 0 \end{bmatrix}.$$

There are only three such matrices; if π is a primitive element of GF(4), they are

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} \pi & 0 & 0 \\ 0 & 0 & \pi^2 \\ 0 & 1 & 0 \end{bmatrix}; \quad C = \begin{bmatrix} \pi^2 & 0 & 0 \\ 0 & 0 & \pi \\ 0 & 1 & 0 \end{bmatrix}.$$

Note that $A^2 = E$, $B^2 = \pi^2 E$, and $C^2 = \pi E$. It follows that if $M \in \mathrm{SL}(2, 3)$ and $M^2 = E$ (a stronger condition, of course, than M^2 being scalar), then M is similar to A; that is, $M = PAP^{-1}$ for some $P \in \mathrm{GL}(3, 4)$. In particular, $\pi^2 B$ and πC are involutions, so there are $P, Q \in \mathrm{GL}(3, 4)$ with

$$PAP^{-1} = \pi^2 B \qquad \text{and} \qquad QAQ^{-1} = \pi C.$$

Since $[\mathrm{GL}(3, 4): \mathrm{SL}(3, 4)] = 3$ (for $\mathrm{GL}/\mathrm{SL} \cong \mathrm{GF}(4)^\times$) and since the matrix $\mathrm{diag}\{\pi, 1, 1\}$ of determinant $\pi \neq 1$ commutes with A, Exercise 3.7 allows us to assume that P and Q lie in SL(3, 4). It follows that A, B, and C become conjugate in PSL(3, 4), as desired. ∎

Theorem 8.24 can also be proved by showing that PSL(3, 4) contains no element of order 15, while A_8 does contain such an element, namely, (1 2 3)(4 5 6 7 8).

One can display infinitely many pairs of nonisomorphic simple groups having the same finite order, but the classification of the finite simple groups shows that there do not exist three nonisomorphic simple groups of the same order.

Classical Groups

At the end of the nineteenth century, the investigation of solutions of systems of differential equations led to complex Lie groups which are intimately related to simple Lie algebras of matrices over $\mathbb{C}$. There are analogues of these