COMMUTATORS IN THE SPECIAL AND GENERAL LINEAR GROUPS[1]

BY

R. C. THOMPSON

1. **Introduction.** We shall use the following notation: K denotes an arbitrary (commutative) field and $GF(p^n)$ denotes the finite field with p^n elements; $GL(n, K)$ denotes the multiplicative group of all nonsingular $n\times n$ matrices with coefficients in K and $SL(n, K)$ denotes the subgroup of $GL(n, K)$ consisting of the matrices in $GL(n, K)$ with determinant unity; I_n denotes the identity matrix in $GL(n, K)$ and $|A|$ denotes the determinant of the matrix A.

If x is an element of the group G and if $x=yzy^{-1}z^{-1}$ where $y, z\in G$, then x is said to be a commutator of G. It is known [1] that $SL(n, K)$ is its own commutator group except in the two cases $n=2$, $K=GF(2)$ and $n=2$, $K=GF(3)$. Also, $SL(2, GF(3))$ is the commutator subgroup of $GL(2, GF(3))$. The following question has been discussed by a number of authors for various groups, chiefly the symmetric groups and some linear groups [2; 3; 4; 8; 9; 10; 11; 13]: if $x\in G'$, the commutator subgroup of the group G, how many commutators are needed to express x as a product of commutators of G? It is the purpose of this paper to solve this problem for the groups $SL(n, K)$ and $GL(n, K)$ (with certain omissions to be stated presently). It is already known [10] that if $A\in SL(n, K)$ where K has infinitely many elements, then

$$A = \prod_{i=1}^{N} (B_iC_iB_i^{-1}C_i^{-1})$$

where B_i, $C_i\in GL(n, K)$ and where N depends upon the splitting of the characteristic polynomial of A into factors irreducible over K and is, in general, larger than one. In this paper we shall prove the following two theorems.

THEOREM 1. *Let $pI_n\in SL(n, K)$. Then pI_n is always a commutator of $GL(n, K)$. Moreover, pI_n is a commutator of $SL(n, K)$ unless p is a primitive nth root of unity in K and $n\equiv 2 \pmod 4$. In this exceptional case pI_n can always be expressed as a product of two commutators of $SL(n, K)$ and can be expressed*

Presented to the Society, February 20, 1960 under the title *Commutators of matrices*. III; received by the editors December 8, 1960.

[1] This paper presents material contained in the author's dissertation at the California Institute of Technology and announced in abstract 565-16 in the Notices of the American Mathematical Society, vol. 7 (1960) p. 194. The author wishes to acknowledge his large debt to Dr. O. Taussky-Todd for suggesting the problem considered here and for encouragement and advice during the course of the research presented in this paper. The work of the author was supported in part by the National Research Council of Canada.

as a single commutator of $SL(n, K)$ when, and only when, the equation $-1=x^2+y^2$ possesses a solution $x, y\in K$. This condition is always satisfied when K has characteristic different from zero.

THEOREM 2. *Let $A\in SL(n, K)$. If A is not scalar and if K has at least four elements, then A is a commutator of $SL(n, K)$.*

Our proof of Theorem 2 for general fields K depends heavily on the number of elements in K and breaks down when K has five or fewer elements. We are able to prove Theorem 2 when K is $GF(5)$ or $GF(4)$ by combining our general methods with the special properties of the elements of these fields. Our general methods break down drastically when K is $GF(2)$ or $GF(3)$. We have been able to prove the following theorem.

THEOREM. *Every element of $SL(n, GF(2))$ is a commutator of $SL(n, GF(2))$ when $n>2$. Every element of $SL(n, GF(3))$ is a commutator of $SL(n, GF(3))$ when $n>2$. Every element of $SL(2, GF(3))$ is a commutator of $GL(2, GF(3))$.*

It is hoped to publish the proof of this theorem elsewhere[2]. We remark that it is known [1] that $SL(2, GF(2))$ and $SL(2, GF(3))$ properly contain their commutator subgroups.

In §2 we shall introduce some additional notation and the lemmas required in the proofs of Theorems 1 and 2. In §3 we shall prove Theorem 1 and in §4 we shall prove Theorem 2 when K is not $GF(5)$ or $GF(4)$. The proofs of Theorem 2 for these fields are given in §§5 and 6. Some concluding remarks are appended in §7.

2. Preliminary material.

If

$$p(\lambda) = \lambda^n + a_n\lambda^{n-1} + \cdots + a_1$$

is a polynomial with coefficients in K, then by $C(p(\lambda))$ we denote the companion matrix of $p(\lambda)$:

$$C(p(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ -a_1 & -a_2 & -a_3 & \cdots & -a_n \end{pmatrix}$$

when $n\geqq 2$, and $C(p(\lambda))=(-a_1)$ when $n=1$. If $a\in K$ then $J_n(a)$ denotes the Jordan matrix of dimension n:

[2] A proof of this theorem is to be found in the author's dissertation. The proof in the case $K=GF(3)$ is like the proofs presented in this paper but involves complicated matrix calculations. The proof in the case $K=GF(2)$ uses matrix computations quite different from those presented in this paper.

$$J_n(a) = \begin{pmatrix} a & 1 & & & & \\ & a & 1 & & 0 & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & \cdot \\ & 0 & & & a & 1 \\ & & & & & a \end{pmatrix}$$

when $n \geqq 2$ and $J_1(a) = (a)$. The following $n \times n$ matrix will be used so frequently that we give it the name of ***standard matrix***:

$$D = \begin{pmatrix} d_1 & d_2 & d_3 & \cdots & d_n \\ & J_{s(1)}(c_1) & & & \\ & & J_{s(2)}(c_2) & & 0 \\ & & & \cdot & \\ & & & & \cdot \\ & 0 & & & \cdot \\ & & & & J_{s(r)}(c_r) \end{pmatrix}$$

when $n \geqq 2$ and

$$s(1) + s(2) + \cdots + s(r) = n - 1; \tag{1}$$

and $D = (d_1)$ when $n = 1$. The $n \times n$ standard matrix D is defined by the field element d_1 when $n = 1$ and, when $n \geqq 2$, D is defined by the integers r, $s(1)$, $\cdots$, $s(r)$ satisfying (1) and by the field elements $d_1, d_2, \cdots, d_n, c_1, c_2, \cdots, c_r$. Standard matrices will always be described in terms of the notation just introduced.

If $A \in GL(n, K)$ and $B \in GL(m, K)$ then $A \dotplus B$ denotes the following element of $GL(m+n, K)$:

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

LEMMA 1. *Let D be the standard matrix introduced above. When $n \geqq 2$, if $d_1 \neq c_i$ for $i = 1, 2, \cdots, r$, then the elementary divisors of D are $(\lambda - d_1)$, $(\lambda - c_1)^{s(1)}, \cdots, (\lambda - c_r)^{s(r)}$. When $n = 1$, the only elementary divisor of D is $(\lambda - d_1)$.*

Proof. The result is clear when $n = 1$. When $n \geqq 2$, let $S_{i,j}(u) = I_n + uE_{i,j}$ where $E_{i,j}$ is an $n \times n$ matrix with one at the intersection of row i and column j, and zeros elsewhere. Then, for $2 \leqq i \leqq n$, the matrix $S_{1,i}(u_i) D S_{1,i}(u_i)^{-1}$ is again D except that d_i is replaced with zero if u_i is properly chosen, and d_{i+1}

is altered (when $i<n$). Hence, transforming by $S_{1,i}(u_i)$ for $i=2, 3, \cdots, n$, we find that D is similar to

$$d_1 \dotplus J_{e(1)}(c_1) \dotplus \cdots \dotplus J_{e(r)}(c_r).$$

This matrix exhibits the required elementary divisors.

Lemma 2. *For $n \geqq 2$, let*

$$F = \begin{pmatrix} 0 & f_{1,2} & f_{1,3} \cdots f_{1,n} \\ 0 & 0 & f_{2,3} \cdots f_{2,n} \\ \cdot & \cdot & \cdot \quad \cdots \quad \cdot \\ 0 & 0 & 0 \quad \cdots f_{n-1,n} \\ x_1 & x_2 & x_3 \quad \cdots \quad x_n \end{pmatrix}$$

be a matrix with coefficients in K and $f_{i,i+1} \neq 0$ for $i=1, 2, \cdots, n-1$. Then S exists in $GL(n, K)$ such that $SFS^{-1} = C(p(\lambda))$ where

$$p(\lambda) = \lambda^n - w_n\lambda^{n-1} - \cdots - w_1 \tag{2}$$

and where

$$\begin{aligned} w_n &= x_n, \\ w_i &= f_{i,i+1} \cdots f_{n-1,n}\left(x_i + \sum_{j=1}^{i-1} a_{i,j}x_j\right) \qquad (i = n-1, \cdots, 2), \\ w_1 &= f_{1,2} \cdots f_{n-1,n}x_1. \end{aligned} \tag{3}$$

The coefficients $a_{i,j} \in K$.

Proof. If $a_i = -f_{n-1,n}^{-1}f_{i,n}$, the matrix

$$S_{n-2,n-1}(a_{n-2}) \cdots S_{1,n-1}(a_1)FS_{1,n-1}(a_1)^{-1} \cdots S_{n-2,n-1}(a_{n-2})^{-1}$$

has the same structure as F except that the coefficients standing above $f_{n-1,n}$ in the last column are now zeros, and column $n-1$ is now column $n-1$ of F plus linear combinations of columns $n-2, n-3, \cdots, 1$ of F. Repeating this procedure with columns $n-1, \cdots, 3$ produces a matrix F_1 similar to F in which the coefficients standing above $f_{i,i+1}$ in column $i+1$ are all zeros for $i=2, 3, \cdots, n-1$. Now let S_1 be a diagonal matrix with $s_{n,n}=1$ and $s_{i,i} = (f_{i,i+1} \cdots f_{n-1,n})^{-1}$ for $i=1, 2, \cdots, n-1$. Then $S_1F_1S_1^{-1}$ is the required companion matrix.

Lemma 3. *Let*

$$A = C(\lambda^n - a_n\lambda^{n-1} - \cdots - a_2\lambda - (-1)^{n-1}|A|)$$

be an element of $GL(n, K)$. Let $D \in GL(n, K)$ be the $n \times n$ standard matrix introduced above. When $n \geqq 2$, let

$$q(\lambda) = \lambda^n + q_n\lambda^{n-1} + \cdots + q_2\lambda + (-1)^n |A| d_1(c_1)^{s(1)} \cdots (c_r)^{s(r)}$$

be a polynomial with coefficients in K. Then, for fixed integers r, $s(1), \cdots, s(r)$ and fixed field elements $d_1, c_1, \cdots, c_r$, it is possible to choose $d_2, \cdots, d_n \in K$ in such a manner that $q(\lambda)$ is both the characteristic and minimum polynomial of AD. When $n=1$, let $q(\lambda) = \lambda - |A| d_1$. Then the characteristic and minimum polynomial of AD is again $q(\lambda)$.

Proof. When $n=1$, there are no d_i to be chosen and the result is obvious. Suppose that $n \geqq 2$ and compute AD. We find that AD is a matrix like the matrix F described in Lemma 2, where the elements on the side diagonal just above the main diagonal are $c_1, \cdots, c_1, c_2, \cdots, c_2, \cdots, c_r, \cdots, c_r$ (c_i appears $s(i)$ times for $i = 1, 2, \cdots, r$) and where

$$(4) \qquad \begin{aligned} x_1 &= (-1)^{n-1} |A| d_1, \\ x_i &= (-1)^{n-1} |A| d_i + \text{terms free of } d_2, \cdots, d_n \qquad (i = 2, \cdots, n). \end{aligned}$$

Invoking Lemma 2, we find $S \in GL(n, K)$ such that $S(AD)S^{-1} = C(p(\lambda))$ where $p(\lambda)$ is given by (2) and (3). Since the characteristic polynomial of a companion matrix is also the minimum polynomial, the lemma will be established if we can choose $d_2, \cdots, d_n \in K$ such that $p(\lambda) = q(\lambda)$. First note that $p(\lambda)$ and $q(\lambda)$ have the same constant term. Next, set $w_i = -q_i$ for $i = 2, \cdots, n$, and note that, since (3) is a triangular system and no $f_{i,i+1}$ vanishes, we may solve for $x_2, \cdots, x_n$ in terms of $w_2, \cdots, w_n$. But then, from (4), we may solve for $d_2, \cdots, d_n$ since $|A| \neq 0$. With $d_2, \cdots, d_n$ determined in this way, $p(\lambda) = q(\lambda)$ and the proof is complete.

For the remainder of this paper we make the convention that whenever we list the elementary divisors of a matrix and include a term $(\lambda - \gamma)^v$, then $(\lambda - \gamma)^v$ is to be deleted from the list whenever $v = 0$.

The next result is the cornerstone of the proof of Theorem 2.

LEMMA 4. *Suppose that $A \in GL(n, K)$ is the companion matrix of a polynomial. Suppose that a field element $d_1 \in K$ and polynomials $(\lambda - \gamma_1)^{v(1)}, \cdots, (\lambda - \gamma_t)^{v(t)}$ are given where $v(1), \cdots, v(t)$ are nonnegative integers such that*

$$v(1) + \cdots + v(t) = n - 1,$$

where $|A| d_1, \gamma_1, \cdots, \gamma_t$, 0 are distinct elements of K, and where $d_1 \neq \gamma_i$ for $i = 1, 2, \cdots, t$. Then a standard matrix $D \in GL(n, K)$ may be found such that:

(i) *the elementary divisors of D are*

$$(5) \qquad (\lambda - d_1), (\lambda - \gamma_1)^{v(1)}, \cdots, (\lambda - \gamma_t)^{v(t)};$$

(ii) *the elementary divisors of AD are*

$$(6) \qquad (\lambda - |A| d_1), (\lambda - \gamma_1)^{v(1)}, \cdots, (\lambda - \gamma_t)^{v(t)}.$$

Proof. The result is obvious when $n = 1$. When $n \geqq 2$, suppose $v(i_1), \cdots,$

$v(i_r)$ are the nonzero integers among $v(1), \cdots, v(t)$. Let $D \in GL(n, K)$ be a standard matrix whose parameters are the integers r, $s(1)=v(i_1), \cdots, s(r)=v(i_r)$, and the field elements d_1, $c_1=\gamma_{i_1}, \cdots, c_r=\gamma_{i_r}$, together with as yet undetermined field elements $d_2, \cdots, d_n$. By Lemma 1, for any choice of $d_2, \cdots, d_n$, the elementary divisors of D are given by (5). By Lemma 3, we may select $d_2, \cdots, d_n \in K$ such that the characteristic and minimum polynomial of AD is

$$q(\lambda) = (\lambda - |A| d_1)(\lambda - \gamma_1)^{v(1)} \cdots (\lambda - \gamma_t)^{v(t)}.$$

But then the elementary divisors of AD are obtained by decomposing $q(\lambda)$ into powers of (nonconstant) polynomials which are irreducible over K and relatively prime. From this observation it is clear that the elementary divisors of AD are given by (6).

The following lemma will be found useful when K is $GF(5)$ or $GF(4)$.

Lemma 5. *Let $A \in GL(n, K)$ be the companion matrix of a polynomial, where $n \geqq 2$. Suppose that a field element $d_1 \in K$ and polynomials $(\lambda-\gamma_1)$, $(\lambda-\gamma_2)^{v(2)}$, $\cdots$, $(\lambda-\gamma_t)^{v(t)}$ are given, where $v(2), \cdots, v(t)$ are nonnegative integers such that*

$$v(2) + \cdots + v(t) = n - 2,$$

where d_1, $|A|\gamma_1, \gamma_2, \cdots, \gamma_t$, 0 are distinct elements of K, and where $d_1 \neq \gamma_1$. Then a standard matrix $D \in GL(n, K)$ may be found such that:

(i) *the elementary divisors of D are*

$$(\lambda - d_1), (\lambda - \gamma_1), (\lambda - \gamma_2)^{v(2)}, \cdots, (\lambda - \gamma_t)^{v(t)}; \tag{7}$$

(ii) *the elementary divisors of AD are*

$$(\lambda - d_1), (\lambda - |A| \gamma_1), (\lambda - \gamma_2)^{v(2)}, \cdots, (\lambda - \gamma_t)^{v(t)}. \tag{8}$$

Proof. Let $v(i_2), \cdots, v(i_r)$ be the nonzero integers among $v(2), \cdots, v(t)$. (We set $r=1$ if all $v(i)$ are zero.) Let $D \in GL(n, K)$ be a standard matrix where the parameters are the integers r, $s(1)=1$, $s(2)=v(i_2), \cdots, s(r)=v(i_r)$, and the field elements d_1, $c_1=\gamma_1$, $c_2=\gamma_{i_2}, \cdots, c_r=\gamma_{i_r}$, together with as yet undetermined field elements $d_2, \cdots, d_n$. Then Lemma 1 asserts that the elementary divisors of D are given by (7). Set

$$q(\lambda) = (\lambda - d_1)(\lambda - |A| \gamma_1)(\lambda - \gamma_2)^{v(2)} \cdots (\lambda - \gamma_t)^{v(t)}.$$

Complete the proof as in Lemma 4.

Lemma 6. *Let A, $B \in GL(n, K)$, where $B = TAT^{-1}$ for some $T \in GL(n, K)$. If A (and hence B) possesses an elementary divisor $\lambda-\alpha$ where $\alpha \in K$, then a matrix $S \in GL(n, K)$ exists such that, if s is any nonzero element of K, then $B = SAS^{-1}$ and $|S| = s$.*

Proof. Matrices T_1 and T_2 exist in $GL(n, K)$ such that $T_1AT_1^{-1}=T_2BT_2^{-1} = (\alpha)\dotplus A_1$, where $A_1\in GL(n-1, K)$. Set $s_1=s|T_2T_1^{-1}|$, and set $S_1=(s_1)\dotplus I_{n-1}$. Then

$$S_1T_1AT_1^{-1}S_1^{-1} = T_1AT_1^{-1} = T_2BT_2^{-1},$$

so that

$$T_2^{-1}S_1T_1AT_1^{-1}S_1^{-1}T_2 = B.$$

Set $S=T_2^{-1}S_1T_1$.

3. **The proof of Theorem** 1. In this section b will denote a primitive nth root of unity in K. Note that n is determined by the roots of unity that exist in K and is, in general, not arbitrary. We shall first show that bI_{mn} is a commutator of $GL(mn, K)$ for every integer $m\geqq 1$. We shall then show, when n is odd, that bI_{mn} is a commutator of $SL(mn, K)$ for all integers $m\geqq 1$, and, when n is even, that bI_{mn} is a commutator of $SL(mn, K)$ for all integers $m>1$. Next we shall determine when bI_n (n even) is a commutator of $SL(n, K)$. Finally, we shall show that bI_n is always a product of two commutators of $SL(n, K)$.

Let $D=(1)\dotplus(b)\dotplus(b^2)\dotplus\cdots\dotplus(b^{n-1})$. Then D and $bI_nD=bD$ have the same elementary divisors. By Lemma 6, S exists in $SL(n, K)$ such that $bD=SDS^{-1}$. Hence $bI_n=SDS^{-1}D^{-1}$. Since the direct sum of commutators is again a commutator, it follows that bI_{mn} is always a commutator of $GL(mn, K)$. For use later, we appeal again to Lemma 6 to find $T\in GL(n, K)$ such that $|T|=-1$ and $bI_n=TDT^{-1}D^{-1}$.

When n is odd, $|D|=(b^n)^{(n-1)/2}=1$. Hence bI_{mn} is always a commutator of $SL(mn, K)$ when n is odd and $m\geqq 1$.

When n is even, $|D|=(b^{n/2})^{n-1}=-1$ since $b^{n/2}=-1$ as b is a primitive nth root of unity in K. Applying results obtained above to $(bI_n)^{-1}$, we find matrices U, $V\in GL(n, K)$ such that $|U|=-1$, $|V|=1$, and $bI_n=UVU^{-1}V^{-1}$.

Now note that

$$bI_{2n} = (S\dotplus S)(D\dotplus D)(S\dotplus S)^{-1}(D\dotplus D)^{-1};$$

$$bI_{3n} = (S\dotplus U\dotplus T)(D\dotplus V\dotplus D)(S\dotplus U\dotplus T)^{-1}(D\dotplus V\dotplus D)^{-1};$$

where $(S\dotplus S)$, $(D\dotplus D)\in SL(2n, K)$ and $(S\dotplus U\dotplus T)$, $(D\dotplus V\dotplus D)\in SL(3n, K)$. Since, if $m>1$, either the integer m is even or $m-3$ is even, it follows that bI_{mn} is always a direct sum of commutators of matrices with determinant one. Hence bI_{mn} is a commutator of $SL(mn, K)$ for all integers $m>1$ when n is even.

Suppose now, if possible, that $bI_n=BCB^{-1}C^{-1}$ where n is even and B, $C\in SL(n, K)$. It is well known that BC and CB have the same characteristic values. Let β be a characteristic value of BC (in a suitable extension field of K). Then β is a characteristic value of CB, hence $b\beta$ is a characteristic value of $bCB=BC$. Hence β, $b\beta$, $\cdots$, $b^{n-1}\beta$ are characteristic values of BC

and, indeed, are all the characteristic values as b is a primitive nth root of unity. Since $|BC|=1$, $b^{n(n-1)/2}\beta^n=1$. Since $b^{n(n-1)/2}=-1$, we obtain $\beta^n+1=0$, so that the characteristic and minimum polynomial of BC is $p(\lambda)=\lambda^n+1$. Hence S exists in $GL(n, K)$ such that $S(BC)S^{-1}=C(p(\lambda))=Z$, say. Then $bI_n=S(bI_n)S^{-1}=ZYZ^{-1}Y^{-1}$ where $Y=SCS^{-1}\in SL(n, K)$. Thus, if bI_n is a commutator within $SL(n, K)$, then bI_n is the commutator of Z and another matrix Y. Let $Y=(y_{i,j})$. Comparison of the coefficients of the equal matrices bYZ and ZY yields recurrence relations on the coefficients of Y. We find that all coefficients of Y are determined by the coefficients in the first column of Y. Setting $y_i=y_{i,1}$ for $i=1, 2, \cdots, n$, we obtain

$$Y=\begin{pmatrix} y_1 & -by_n & -b^2y_{n-1} & \cdots & -b^{n-1}y_2 \\ y_2 & by_1 & -b^2y_n & \cdots & -b^{n-1}y_3 \\ y_3 & by_2 & b^2y_1 & \cdots & -b^{n-1}y_4 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ y_n & by_{n-1} & b^2y_{n-2} & \cdots & b^{n-1}y_1 \end{pmatrix}.$$

Conversely, for this Y, $bYZ=ZY$. Let

$$Y_1=\begin{pmatrix} y_1 & -y_n & -y_{n-1} & \cdots & -y_2 \\ y_2 & y_1 & -y_n & \cdots & -y_3 \\ y_3 & y_2 & y_1 & \cdots & -y_4 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ y_n & y_{n-1} & y_{n-2} & \cdots & y_1 \end{pmatrix}.$$

Then $|Y|=b^{n(n-1)/2}|Y_1|=-|Y_1|$. Thus we have proved that the necessary and sufficient condition that bI_n be a commutator of $SL(n, K)$ is that $y_1, \cdots, y_n\in K$ can be found such that $|Y_1|=-1$.

Now let ω be a primitive $(2n)$th root of unity in an extension field of K: $\omega^2=b$. Set $\omega_i=b^i\omega$ for $i=1, \cdots, n$. Then it is known [7, p. 445] that

$$|Y_1|=\prod_{i=1}^{n}\left(\sum_{j=1}^{n}\omega_i^{j-1}y_j\right).$$

Hence

$$\begin{aligned}|Y_1| &= \prod_{i=1}^{n}\left(\sum_{j=1}^{n/2}b^{i(2j-2)}\omega^{2j-2}y_{2j-1}+\sum_{j=1}^{n/2}b^{i(2j-1)}\omega^{2j-1}y_{2j}\right)\\ &= \prod_{i=1}^{n}\left(\sum_{j=1}^{n/2}b^{(j-1)(2i+1)}y_{2j-1}+\omega^{-1}\sum_{j=1}^{n/2}b^{j(2i+1)-i}y_{2j}\right).\end{aligned}$$

Now $b^{k(2(n/2+i)+1)}=b^{k(2i+1)}$ if k is an integer, and $b^{-(n/2+i)}=-b^i$. Hence

$$\begin{aligned} | Y_1 | &= \prod_{i=1}^{n/2} \left[\left(\sum_{j=1}^{n/2} b^{(j-1)(2i+1)} y_{2j-1} + \omega^{-1} \sum_{j=1}^{n/2} b^{j(2i+1)-i} y_{2j} \right) \right. \\ &\qquad \left. \cdot \left(\sum_{j=1}^{n/2} b^{(j-1)(2i+1)} y_{2j-1} - \omega^{-1} \sum_{j=1}^{n/2} b^{j(2i+1)-i} y_{2j} \right) \right] \\ &= \prod_{i=1}^{n/2} \left[\left(\sum_{j=1}^{n/2} b^{(j-1)(2i+1)} y_{2j-1} \right)^2 - b^{-1-2i} \left(\sum_{j=1}^{n/2} b^{j(2i+1)} y_{2j} \right)^2 \right]. \end{aligned}$$

Consider the following system of $n/2$ equations in $n/2$ unknowns.

$$\sum_{j=1}^{n/2} b^{(j-1)(2i+1)} y_{2j-1} = w_{2i-1} \qquad (i = 1, 2, \cdots, n/2). \tag{9}$$

The coefficient matrix of (9) is a Vandermonde matrix and is nonsingular since b is a primitive nth root of unity. Hence, for any choice of $w_1, w_3, \cdots, w_{n-1} \in K$, we may find $y_1, y_3, \cdots, y_{n-1} \in K$ to satisfy (9). Similarly the set of $n/2$ equations in $n/2$ unknowns

$$\sum_{j=1}^{n/2} b^{j(2i+1)} y_{2j} = b^i w_{2i} \qquad (i = 1, 2, \cdots, n/2),$$

has a solution $y_2, y_4, \cdots, y_n$ in K for every choice of $w_2, w_4, \cdots, w_n \in K$.

Thus, in order to set $| Y_1 | = -1$, it is both necessary and sufficient that $w_1, w_2, \cdots, w_n$ be found in K such that

$$-1 = \prod_{i=1}^{n/2} ((w_{2i-1})^2 - b^{-1}(w_{2i})^2). \tag{10}$$

If $n = 4m$, take $w_1 = b^m$, $w_{2i-1} = 1$ for all $i \neq 1$, $w_{2i} = 0$ for all i. Then since $b^{2m} = -1$, (10) is satisfied.

If $n = 4m+2$ then, from $b^{2m+1} = -1$, we obtain $-b^{-1} = b^{2m}$. Set $(w_{2i})' = b^m w_{2i}$. Then

$$| Y_1 | = \prod_{i=1}^{n/2} ((w_{2i-1})^2 + ((w_{2i})')^2). \tag{11}$$

Since sums of two squares are closed under multiplication [6, p. 126], if elements in K exist such that $| Y_1 | = -1$, then, for certain elements $W, W' \in K$ we have

$$-1 = W^2 + (W')^2 \tag{12}$$

Conversely, if (12) has a solution in K, then if in (11) we set $w_1 = W$, $(w_2)' = W'$, $w_{2i-1} = 1$ and $(w_{2i})' = 0$ for all $i \neq 1$, we find that $| Y_1 | = -1$.Hence we have demonstrated that if $n = 4m+2$, then the necessary and sufficient condition that $bI_n = BCB^{-1}C^{-1}$ where $B, C \in SL(n, K)$ is that (12) have a solution $W, W' \in K$.

It is known [6, p. 135] that integers x, y always exist such that $x^2+y^2+1 \equiv 0 \pmod{p}$ where p is a prime. This means that elements x, y always exist in $GF(p)$ and hence in any field of characteristic p such that $x^2+y^2=-1$. Hence (12) is always solvable for such fields.

If K has characteristic zero, then (12) is sometimes impossible; as an example take $n=2$, $b=-1$, and K to be any formally real field. In many other cases (12) possesses a solution. We list two such cases.

(i) If K contains the primitive $(2n)$th root of unity ω then a solution of (12) is $W=\omega^{n/2}$, $W'=0$.

(ii) If for some divisor $r>1$ of $2m+1$ integers s and h exist such that $r(h+1)=2^s+1$, then a solution of (12) can be found. For, using a technique due to Landau [5], we first note the following polynomial identity[3].

$$\begin{aligned}(1+\lambda+\lambda^2+\cdots+\lambda^{r-1})&(1+\lambda^r+\lambda^{2r}+\cdots+\lambda^{hr})\\ &= 1+\lambda+\lambda^2+\cdots+\lambda^{rh+r-1}\\ &= (1+\lambda)(1+\lambda^2)(1+\lambda^4)\cdots(1+\lambda^{2^{s-1}})+\lambda^{2^s}.\end{aligned}$$

Let $\rho=b^{n/2r}$. Then $\rho^2\neq 1$ so that ρ satisfies

$$(\rho^{2r}-1)/(\rho^2-1)=\rho^{2r-2}+\rho^{2r-4}+\cdots+\rho^2+1=0.$$

Hence, if we set $\lambda=\rho^2$, we find

$$-\rho^{2^{s+1}}=(1+\rho^2)(1+\rho^4)\cdots(1+\rho^{2^s})$$

from which we deduce that -1 is a sum of two squares.

If $2m+1$ has a prime divisor p of the form $8k\pm 3$, then we may take $r=p$. For

$$\begin{aligned}2^{(p-1)/2} &\equiv (2/p) \pmod{p}\\ &= (-1)^{(p^2-1)/8} = -1\end{aligned}$$

so that p divides $2^{(p-1)/2}+1$. Here $(2/p)$ denotes the Legendre symbol.

It is known [12] that -1 is a sum of four squares in the field of the nth roots of unity over the field of rational numbers, $n>2$. When -1 is a sum of two squares in such fields remains to be determined.

We now complete the proof of Theorem 1 by showing that bI_n is always a product of two commutators of $SL(n, K)$ when $n\equiv 2 \pmod 4$ and K has characteristic zero. First note that if

$$P=\begin{pmatrix}5 & 14\\ -4 & -11\end{pmatrix}, \qquad Q=\begin{pmatrix}2 & 4\\ 0 & 1/2\end{pmatrix},$$

then $-I_2=(PQP^{-1}Q^{-1})^2$. This completes the proof when $n=2$. When $n=4m+2>2$ note that b^{2m} is a primitive root of unity of order $n/(n, 2m)$

[3] Note that $n>2$ since $r>1$. By hypothesis the field K in this case contains roots of unity since it contains the primitive nth root of unity b.

$=(2m+1)/(2m+1, m)=2m+1$. Thus b^{2m} is a primitive root of unity of odd order and hence $b^{2m}I_n$ is a commutator of $SL(n, K)$. Furthermore, $-I_n$ is a direct sum of $n/2$ commutators of $SL(2, K)$, hence is a commutator of $SL(n, K)$. Now, $(b^{2m}I_n)(bI_n)=-I_n$ since $b^{2m+1}=-1$. Thus bI_n is a product of two commutators of $SL(n, K)$.

4. The proof of Theorem 2 when K contains at least six elements. Throughout this section A will denote an element of $SL(n, K)$. Since any factorization $A=BCB^{-1}C^{-1}$ is preserved under a similarity transformation by any element of $GL(n, K)$, it is enough to prove that some similarity transform of A is a commutator of $SL(n, K)$. Thus we may assume that

$$A = A_1 \dotplus \cdots \dotplus A_m$$

where A_i is the $j(i)\times j(i)$ companion matrix of a polynomial with coefficients in K, for $i=1, 2, \cdots, m$. Rearranging the A_i if necessary (a similarity transformation!), we may assume that $j(1)\leqq j(2)\leqq \cdots \leqq j(m)$. We divide the proof into a number of cases depending on the values of m, n, and $j(1), \cdots, j(m)$.

Case 1. $n=2$. If A is 2×2 and not scalar, then A is similar to the companion matrix of some polynomial. Hence we suppose A is such a companion matrix. Choose $\rho\in K$ such that $\rho^2\neq 1, 0$. (This is possible if K is not $GF(2)$ or $GF(3)$.) Invoking Lemma 4, we construct a standard matrix $D\in SL(2, K)$ with $d_1=\rho$ and elementary divisors $(\lambda-\rho)$, $(\lambda-\rho^{-1})$ such that the elementary divisors of AD are also $(\lambda-\rho)$, $(\lambda-\rho^{-1})$. By Lemma 6, we may find $S\in SL(2, K)$ such that $AD=SDS^{-1}$. Hence $A=SDS^{-1}D^{-1}$ where S, $D\in SL(2, K)$ as required.

Case 2. $j(m)\geqq 3$. Let δ_1 be any nonzero element of K and define $\delta_{i+1}=|A_i|\delta_i$ for $i=1, 2, \cdots, m$. Then, since $|A_1\cdots A_m|=1$, we have $\delta_{m+1}=\delta_1$. For $i=1, 2, \cdots, m$, let γ_i be any element of K distinct from δ_i, δ_{i+1}, 0. Define γ''' by

$$\left(\prod_{i=1}^{m}\delta_i\right)\left(\prod_{i=1}^{m-1}(\gamma_i)^{j(i)-1}\right)(\gamma_m)^{j(m)-2}\gamma''' = 1. \tag{13}$$

Choose $x\in K$ such that $x\neq 0$ and

$$\gamma_m x \neq \delta_m \text{ or } \delta_1, \qquad \gamma''' x^{-1} \neq \delta_m \text{ or } \delta_1. \tag{14}$$

The conditions (14) prohibit at most four nonzero elements of K so that, if K has at least six elements, a suitable x always exists. Let $\gamma'=\gamma_m x$, $\gamma''=\gamma''' x^{-1}$. Then

$$\left(\prod_{i=1}^{m}\delta_i\right)\left(\prod_{i=1}^{m-1}(\gamma_i)^{j(i)-1}\right)(\gamma_m)^{j(m)-3}\gamma'\gamma'' = 1. \tag{15}$$

For $i=1, 2, \cdots, m-1$, construct (by Lemma 4) a standard matrix

$D_i \in GL(j(i), K)$ with $d_1 = \delta_i$ and elementary divisors $(\lambda - \delta_i)$, $(\lambda - \gamma_i)^{j(i)-1}$ such that the elementary divisors of A_iD_i are $(\lambda - \delta_{i+1})$, $(\lambda - \gamma_i)^{j(i)-1}$. This is possible since $\gamma_i \neq \delta_i$, δ_{i+1}, 0. Again, by Lemma 4, construct a standard matrix $D_m \in GL(j(m), K)$ with $d_1 = \delta_m$ and:

(i) if γ_m, γ', γ'' are distinct elements of K, such that the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma_m)^{j(m)-3}$, $(\lambda - \gamma')$, $(\lambda - \gamma'')$ and such that the elementary divisors of A_mD_m are $(\lambda - \delta_1)$, $(\lambda - \gamma_m)^{j(m)-3}$, $(\lambda - \gamma')$, $(\lambda - \gamma'')$;

(ii) if $\gamma_m = \gamma' \neq \gamma''$, such that the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma_m)^{j(m)-2}$, $(\lambda - \gamma'')$ and such that the elementary divisors of A_mD_m are $(\lambda - \delta_1)$, $(\lambda - \gamma_m)^{j(m)-2}$, $(\lambda - \gamma'')$;

(iii) if $\gamma_m = \gamma'' \neq \gamma'$, such that the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma_m)^{j(m)-2}$, $(\lambda - \gamma')$ and such that the elementary divisors of A_mD_m are $(\lambda - \delta_1)$, $(\lambda - \gamma_m)^{j(m)-2}$, $(\lambda - \gamma')$;

(iv) if $\gamma' = \gamma'' \neq \gamma_m$, such that the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma')^2$, $(\lambda - \gamma_m)^{j(m)-3}$ and such that the elementary divisors of A_mD_m are $(\lambda - \delta_1)$, $(\lambda - \gamma')^2$, $(\lambda - \gamma_m)^{j(m)-3}$;

(v) if $\gamma_m = \gamma' = \gamma''$, such that the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma_m)^{j(m)-1}$ and such that the elementary divisors of A_mD_m are $(\lambda - \delta_1)$, $(\lambda - \gamma_m)^{j(m)-1}$.

Now set $D = D_1 \dotplus \cdots \dotplus D_m$. Then, because of (15), $D \in SL(n, K)$. Moreover, D and AD have the same elementary divisors. For example, in (i), the elementary divisors of D are

$$\begin{aligned}(\lambda - \delta_1), (\lambda - \gamma_1)^{j(1)-1}, \cdots, (\lambda - \delta_{m-1}), (\lambda - \gamma_{m-1})^{j(m-1)-1},\\ (\lambda - \delta_m), (\lambda - \gamma_m)^{j(m)-3}, (\lambda - \gamma'), (\lambda - \gamma''),\end{aligned} \tag{16}$$

and the elementary divisors of AD are

$$\begin{aligned}(\lambda - \delta_2), (\lambda - \gamma_1)^{j(1)-1}, \cdots, (\lambda - \delta_m), (\lambda - \gamma_{m-1})^{j(m-1)-1},\\ (\lambda - \delta_1), (\lambda - \gamma_m)^{j(m)-3}, (\lambda - \gamma'), (\lambda - \gamma'').\end{aligned} \tag{17}$$

Since (17) is merely a rearrangement of (16), D and AD have the same elementary divisors, including at least one linear elementary divisor. Hence, by Lemma 6, $AD = SDS^{-1}$ so that $A = SDS^{-1}D^{-1}$ where S, $D \in SL(n, K)$.

CASE 3. $m \geqq 2$; $j_m = j_{m-1} = 2$. For $i = 1, 2, \cdots, m$, let δ_i, γ_i be found as in Case 2. Define γ''' by (13) and γ', γ'' by

$$\begin{aligned}\gamma' &= \gamma_{m-1}x \neq \delta_{m-1} \text{ or } \delta_m,\\ \gamma'' &= \gamma'''x^{-1} \neq \delta_m \text{ or } \delta_1.\end{aligned}$$

Construct standard matrices $D_1, \cdots, D_{m-2}$ as before. Construct a standard matrix $D_{m-1} \in GL(2, K)$ with $d_1 = \delta_{m-1}$ and elementary divisors $(\lambda - \delta_{m-1})$, $(\lambda - \gamma')$ such that the elementary divisors of $A_{m-1}D_{m-1}$ are $(\lambda - \delta_m)$, $(\lambda - \gamma')$. Construct a standard matrix $D_m \in GL(2, K)$ with $d_1 = \delta_m$ and elementary divisors $(\lambda - \delta_m)$, $(\lambda - \gamma'')$ such that the elementary divisors of A_mD_m are $(\lambda - \delta_1)$,

$(\lambda-\gamma'')$. Set $D=D_1\dotplus\cdots\dotplus D_m$. Then $D\in SL(n, K)$ and D and AD have the same elementary divisors. Complete the proof as in Case 2.

The argument in the two previous cases has depended on the use of two "spare" diagonal positions in the triangular matrix D to control the value of $|D|$. If $n=2$, or if $n>2$ and $j(i)=1$ for $i=1, 2, \cdots, m-1$ and $j(m)=1$ or 2, these two spare diagonal positions do not exist. The case $n=2$ has already been discussed. When $n>2$, we first remark that $C(p(\lambda))\dotplus C(q(\lambda))$ is similar to $C(p(\lambda)q(\lambda))$ if the polynomials $p(\lambda)$ and $q(\lambda)$ are relatively prime. Thus, if A is diagonal but not scalar, then, after a similarity transformation of A by an element of $GL(n, K)$, we may pass to the case in which $j(i)=1$ for $i=1, 2, \cdots, m-1$ and $j(m)=2$. In this case we may make the further assumption that $A=fI_{n-2}\dotplus C((\lambda-f)(\lambda-g))$ where $f, g\in K$ since otherwise we may (by the remark above) transform A by an element of $GL(n, K)$ and so pass to one of the two cases $j(m)=3$ or $j(m-1)=j(m)=2$. Hence only Case 4 below remains to be considered.

CASE 4. $A=fI_{n-2}\dotplus C((\lambda-f)(\lambda-g))$; $f, g\in K$; $f^{n-1}g=1$. For nonzero $\delta\in K$ define $c(\delta)$ as a function of δ by

$$f^{(n-1)(n-2)/2}\delta^{n-1}c(\delta) = 1. \tag{18}$$

Imitating the proofs already used, we attempt to choose $\delta\in K$ such that

$$c(\delta) \neq \delta, \tag{19}$$

$$c(\delta) \neq f^{n-2}\delta. \tag{20}$$

If such a δ exists, define

$$D = (\delta) \dotplus (f\delta) \dotplus (f^2\delta) \dotplus \cdots \dotplus (f^{n-3}\delta) \dotplus \begin{pmatrix} f^{n-2}\delta & d_2 \\ 0 & c(\delta) \end{pmatrix}.$$

By (20), the elementary divisors of D are $(\lambda-\delta)$, $(\lambda-f\delta)$, $\cdots$, $(\lambda-f^{n-2}\delta)$, $(\lambda-c(\delta))$. By (19), we may choose $d_2\in K$ such that the elementary divisors of AD are $(\lambda-f\delta)$, $(\lambda-f^2\delta)$, $\cdots$, $(\lambda-f^{n-2}\delta)$, $(\lambda-\delta)$, $(\lambda-c(\delta))$. From this, the result follows as before. The desired element δ will always exist if K has infinitely many elements since the equations $c(\delta)=\delta$ and $c(\delta)=f^{n-2}\delta$ have only finitely many roots. Hence, to complete the proof, we may assume that K has characteristic $p\neq 0$.

CASE 4.1. $f=g$. Assume first that $f^2\neq 1$. Take $\delta=1$. Then if $c(1)=1$ or if $c(1)=f^{n-2}$, we find that $f^2=1$ (using $f^n=1$). Hence, if $f^2\neq 1$, the desired element in K is $\delta=1$. If $f^2=1$, then $C(\lambda^2-2f\lambda+f^2)\in SL(2, K)$ and is a commutator of $SL(2, K)$ by Case 1. The roots of $f^2=1$ in any field are $f=1$ and $f=-1$. If $f=1$, then A is the direct sum of a commutator of $SL(2, K)$ and several commutators of $SL(1, K)$. If $f=-1$ then, since $-I_2$ is a commutator of $SL(2, K)$ (by Theorem 1), A is the direct sum of commutators of $SL(2, K)$. This finishes Case 4.1.

CASE 4.2. $f\neq g$, n is even. We shall prove that we may take δ to be one of the elements $1, f, f^2$. An arbitrary nonzero $\delta\in K$ is said to be *admissible* if (18) and (19) hold. If $1, f$ are both not admissible, then $c(1)=1$ and $c(f)=f$ and hence we find that $f^n=1$. This contradicts the fact that $f^{n-1}g=1$ and $f\neq g$. Hence K contains admissible elements. If now δ' is admissible, then exactly one of the following four possibilities holds:

(i) $c(\delta')\neq f^{n-2}\delta'$;

(ii) $c(\delta')=f^{n-2}\delta'$; $f\delta'$ is admissible and $c(f\delta')\neq f^{n-2}(f\delta')$;

(iii) $c(\delta')=f^{n-2}\delta'$; $f\delta'$ is admissible and $c(f\delta')=f^{n-2}(f\delta')$;

(iv) $c(\delta')=f^{n-2}\delta'$; $f\delta'$ is not admissible.

If (iii) holds then we again deduce that $f^n=1$, contrary to hypothesis. If (iv) holds, we obtain $f^n=f^{n-2}$ so that $f^2=1$. Since n is even and $f^{n-1}g=1$, we find that $f=g$, again contrary to hypothesis. Thus we must have either (i) or (ii), which establishes the existence of a solution of (18), (19), and (20).

CASE 4.3. $f\neq g$, n is odd. Here we cannot show the existence of a suitable $\delta\in K$. However, A is similar over $GL(n, K)$ to $A_1=fI_{n-1}\dotplus(g)$. Let $D=(u)\dotplus(fu)\dotplus\cdots\dotplus(f^{n-1}u)$ where $u=f^{-(n-1)/2}$. Then $D\in SL(n, K)$ and it easily follows that A_1 and hence A is a commutator of $SL(n, K)$.

The proof of Theorem 2 is now complete if K is not $GF(5)$ or $GF(4)$.

5. **The case** $K=GF(5)$. The field $GF(5)$ consists of the elements 0, 1, 2, 3, 4. We shall prove that every matrix $A\in SL(n, GF(5))$ is a commutator of $SL(n, GF(5))$. As before, we may assume that $A=A_1\dotplus\cdots\dotplus A_m$ where A_i is the $j(i)\times j(i)$ companion matrix of a polynomial with coefficients in $GF(5)$. We make the additional assumption that $|A_{i_1}\cdots A_{i_k}|\neq 1$ if the nonempty subset $\{i_1, \cdots, i_k\}$ of $\{1, \cdots, m\}$ is proper. This assumption involves no loss of generality since a general A is similar within $GL(n, K)$ to a direct sum of such matrices. With this assumption and after a rearrangement of the A_i, if necessary, we find that $(|A_1|, \cdots, |A_m|)$ must be one of (1), (2, 3), (4, 4), (2, 2, 4), (3, 3, 4), (2, 2, 2, 2), or (3, 3, 3, 3). We divide our discussion into cases depending on the value of m.

If $m=1$, choose $\rho\in K$ such that $\rho^2\neq 1$, 0. By Lemma 4, construct a standard matrix $D\in SL(n, K)$ with $d_1=\rho$ such that the elementary divisors of D and AD are $(\lambda-\rho)$, $(\lambda-\rho^{-1})$, $(\lambda-1)^{n-2}$. (This for $n\geqq 2$; the case $n=1$ is obvious.) Then argue as previously. This part of the proof also works if K is $GF(4)$.

If $m\geqq 2$, we proceed as follows. If δ_1 is a nonzero element of $GF(5)$ then, for $i=1, 2, \cdots, m$, we define $\delta_{i+1}=|A_i|\delta_i$ and choose $\gamma_{i,1}$ and $\gamma_{i,2}$ such that δ_i, δ_{i+1}, $\gamma_{i,1}$, $\gamma_{i,2}$ are the four nonzero elements of $GF(5)$. Let $e(i)$ be an integer with $0\leqq e(i)\leqq j(i)-1$. Construct, by Lemma 4, a standard matrix $D_i\in GL(j(i), GF(5))$ with $d_1=\delta_i$ and elementary divisors $(\lambda-\delta_i)$, $(\lambda-\gamma_{i,1})^{e(i)}$, $(\lambda-\gamma_{i,2})^{j(i)-1-e(i)}$ such that the elementary divisors of A_iD_i are $(\lambda-\delta_{i+1})$, $(\lambda-\gamma_{i,1})^{e(i)}$, $(\lambda-\gamma_{i,2})^{j(i)-1-e(i)}$. If we set $D=D_1\dotplus\cdots\dotplus D_m$, then clearly $A=SDS^{-1}D^{-1}$ where $S\in SL(n, GF(5))$. In order that $D\in SL(n, GF(5))$ we have only to choose the field element δ_1 and the integers $e(i)$ such that

$$\prod_{i=1}^{m} (\delta_i(\gamma_{i,1})^{e(i)}(\gamma_{i,2})^{j(i)-1-e(i)}) = 1. \tag{21}$$

In each of the following cases, we attempt either to determine δ_1 and the $e(i)$ to satisfy (21) or to reduce the case to another in which the desired result is already established.

If $m=2$ and $(|A_1|, |A_2|)=(4, 4)$, we set $\gamma_{1,1}=\gamma_{2,1}=2\delta_1$ and $\gamma_{1,2}=\gamma_{2,2}=3\delta_1$. Then, if $e=e(1)+e(2)$, (21) becomes $\delta_1^n 3^{n+2e}=1$, where $0\leqq e\leqq n-2$. A solution is $\delta_1=3$ and $e=0$ or 1 such that $2n+2e\equiv 0 \pmod 4$.

If $m=2$ and $(|A_1|, |A_2|)=(2, 3)$ we set $\gamma_{1,1}=\gamma_{2,1}=3\delta_1$ and $\gamma_{1,2}=\gamma_{2,2}=4\delta_1$. Then, with $e=e(1)+e(2)$, (21) becomes $\delta_1^n 2^{1+2n+e}=1$. If $n\geqq 5$ or if $n=3$ take $\delta_1=1$ and $e=1$ or 3 such that $1+2n+e\equiv 0 \pmod 4$. If $n=2$ the matrix $A=(2)\dotplus(3)$ and is similar within $GL(2, GF(5))$ to $C((\lambda-2)(\lambda-3))$ which has already been treated under the case $m=1$. The case $n=4$ requires a more detailed examination.

If $n=4$ and $j(1)=1, j(2)=3$, then $A=(2)\dotplus A_2$ where A_2 is a companion matrix of a cubic polynomial and $|A_2|=3$. Using Lemma 5, construct a standard matrix $D_2\in SL(3, GF(5))$ with $d_1=4$ and elementary divisors $(\lambda-4)$, $(\lambda-2)$, $(\lambda-2)$ such that the elementary divisors of A_2D_2 are $(\lambda-4)$, $(\lambda-1)$, $(\lambda-2)$. Then if $D=(1)\dotplus D_2$ we find that $D\in SL(4, GF(5))$ and that D and AD have the same elementary divisors, completing the proof in this case. In the case $n=4, j(1)=3, j(2)=1$, we let $D_1\in SL(3, GF(5))$ be a standard matrix with $d_1=4$ and elementary divisors $(\lambda-4)$, $(\lambda-3)$, $(\lambda-3)$ such that the elementary divisors of A_1D_1 are $(\lambda-4)$, $(\lambda-1)$, $(\lambda-3)$. Set $D=D_1\dotplus(1)$. Thus, only the possibility $j(1)=j(2)=2$ remains. If the characteristic polynomials of A_1 and A_2 are relatively prime then A is similar within $GL(4, GF(5))$ to a companion matrix for which the result is already known (case $m=1$). This will hold if either characteristic polynomial is irreducible since equal characteristic polynomials are excluded by the fact that $|A_1|\neq|A_2|$. Thus, suppose that the characteristic roots of A_1 are r, $2/r$ and that the characteristic roots of A_2 are r, $3/r$, where $r\in GF(5)$. Since $r=2/r$ and $r=3/r$ cannot happen in $GF(5)$, our matrix A has linear elementary divisors and so is similar within $GL(4, GF(5))$ to: $I_2\dotplus C((\lambda-2)(\lambda-3))$ if $r=1$; $(2)\dotplus C((\lambda-1)(\lambda-2)(\lambda-4))$ if $r=2$; $(3)\dotplus C((\lambda-1)(\lambda-3)(\lambda-4))$ if $r=3$; $(4)\dotplus C((\lambda-4)(\lambda-2)(\lambda-3))$ if $r=4$. Thus we come to matrices already treated under previous cases.

The case $m=2$ is now finished.

When $m=3$, by passing to A^{-1} if necessary, we may assume that $(|A_1|, |A_2|, |A_3|)$ is $(2, 2, 4)$. Set $\gamma_{1,2}=\gamma_{2,1}=\gamma_{3,2}=3\delta_1$, $\gamma_{1,1}=4\delta_1$, $\gamma_{2,2}=\delta_1$, $\gamma_{3,1}=2\delta_1$. Then equation (21) becomes

$$\delta_1^n 3^{3+j(1)+j(3)+e(1)+e(2)+2e(3)} = 1. \tag{22}$$

We may assume without loss of generality that $j(1)\geqq j(2)$. In the following table we indicate values of δ_1, $e(1)$, $e(2)$, $e(3)$ (as functions of $j(1), j(2), j(3)$)

which satisfy (22). Those $j(i)$ which are not fully specified in the table may assume arbitrary values as the table permits. Those $e(i)$ which are not fully specified can always be chosen such that

$$3 + j(1) + j(3) + e(1) + e(2) + 2e(3) \equiv 0 \pmod 4.$$

In case (i) of the table note that, since $e(1)+e(2)$ may be any of the integers $0, 1, \cdots, j(1)+j(2)-2$, if $j(1)+j(2) \geqq 5$ we may find $e(1)$ and $e(2)$ such that $e(1)+e(2) \equiv 0$, 1, 2, or 3 (mod 4). In case (v) the matrix A is similar within $GL(n, GF(5))$ to some matrix $C(\lambda^2+u\lambda+2) \dotplus C((\lambda-2)(\lambda-4))$ which has already been discussed under $m=2$. In case (vi) note that

$$2e(3) + 3 + j(1) + j(3) = 4 + 4(j(3)/2) \equiv 0 \pmod 4.$$

In case (vii) note that (using $n=j(1)+j(2)+j(3)$)

$$n + 3 + j(1) + j(3) = 8 + 4((j(3) - 1)/2) \equiv 0 \pmod 4.$$

Case	$j(1)+j(2)$	$j(1)$	$j(2)$	$j(3)$	δ_1	$e(1)$	$e(2)$	$e(3)$
i	$\geqq 5$				1	$<j(1)$	$<j(2)$	0
ii	4	3 or 2	1 or 2	$\geqq 2$	1	$<j(1)$	$<j(2)$	$<j(3)$
iii	4	3 or 2	1 or 2	1	1	$<j(1)$	$<j(2)$	0
iv	3	2	1	$\geqq 2$	1	0, 1	0	0, 1
v	3	2	1	1				
vi	2	1	1	even	1	0	0	$j(3)/2$
vii	2	1	1	odd	3	0	0	0

Finally we arrive at the last case of $m=4$. Passing to A^{-1} if necessary, we assume that $(|A_1|, \cdots, |A_m|)$ is (2, 2, 2, 2). We set $\gamma_{1,1}=\gamma_{2,1}=3\delta_1$; $\gamma_{1,2}=\gamma_{4,2}=4\delta_1$; $\gamma_{2,2}=\gamma_{3,2}=\delta_1$; $\gamma_{3,1}=\gamma_{4,1}=2\delta_1$. Then (21) becomes

$$\delta_1^n 3^{2(1+j(1)+j(4))-e(1)+e(2)+3e(3)+e(4)} = 1.$$

If two of the $j(i)$ are greater than one, take $j(2)>1$, $j(4)>1$ and set $\delta_1=1$, $e(1)=e(3)=0$ and $e(2)$, $e(4)$ equal to 0 or 1 such that

$$e(2) + e(4) + 2(1 + j(1) + j(4)) \equiv 0 \pmod 4.$$

If only one $j(i)$ is not one, suppose $j(1)=j(2)=j(4)=1$ and $j(3)>1$. Set $\delta_1=3^k$ and choose $k=0$, 1, 2, or 3 and $e(3)$, $0 \leqq e(3) \leqq j(3)-1$, such that

$$k(3 + j(3)) + 3e(3) + 6 \equiv 0 \pmod 4.$$

Suitable values for k and $e(3)$ are: $k=2$ and $e(3)=0$ when $j(3)\equiv 0$ or $2 \pmod 4$; $k=1$ and $e(3)=0$ when $j(3)\equiv 3 \pmod 4$; $k=0$ and $e(3)=2$ when $j(3)\equiv 1 \pmod 4$ but $j(3)\neq 1$. Lastly, if all $j(i)$ are one, then A is scalar and we appeal to Theorem 1.

6. **The case $K=GF(4)$.** This case is simpler than the previous case. The field $GF(4)$ consists of the elements 0, 1, θ, $\theta+1$ with $\theta^2=\theta+1$. As in the previous case, we assume $A=A_1\dotplus\cdots\dotplus A_m$ where A_i is a $j(i)\times j(i)$ companion matrix of a polynomial over $GF(4)$ and $|A_{i_1}\cdots A_{i_k}|\neq 1$ if the nonempty subset $\{i_1,\cdots,i_k\}$ of $\{1,\cdots,m\}$ is proper. With this supposition and after a reordering of the A_i, if necessary, we find that $(|A_1|,\cdots,|A_m|)$ is one of (1), $(\theta,\theta+1)$, (θ,θ,θ), $(\theta+1,\theta+1,\theta+1)$.

When $m=1$, the proof for the case $m=1$ in §5 works here also.

When $m=2$, construct by Lemma 4 a standard matrix $D_1\in GL(j(1),GF(4))$ with $d_1=\theta$ and elementary divisors $(\lambda-\theta)$, $(\lambda-1)^{j(1)-1}$ such that the elementary divisors of A_1D_1 are $(\lambda-\theta^2)$, $(\lambda-1)^{j(1)-1}$. Again, construct a standard matrix $D_2\in GL(j(2),GF(4))$ with $d_1=\theta^2$ and elementary divisors $(\lambda-\theta^2)$, $(\lambda-1)^{j(2)-1}$ such that the elementary divisors of A_2D_2 are $(\lambda-\theta)$, $(\lambda-1)^{j(2)-1}$. Set $D=D_1\dotplus D_2$. Then $D\in SL(n,GF(4))$ and D and AD have the same elementary divisors. The proof when $m=2$ is now completed in the usual way.

When $m=3$, we may assume that $(|A_1|,|A_2|,|A_3|)$ is (θ,θ,θ). First suppose that $j(1)$, $j(2)$, $j(3)$ are not distinct (mod 3), say $j(1)\equiv j(3) \pmod 3$. Construct a standard matrix $D_1\in GL(j(1),GF(4))$ with $d_1=1$ and elementary divisors $(\lambda-1)$, $(\lambda-\theta^2)^{j(1)-1}$ such that the elementary divisors of A_1D_1 are $(\lambda-\theta)$, $(\lambda-\theta^2)^{j(1)-1}$. Construct a standard matrix $D_2\in GL(j(2),GF(4))$ with $d_1=\theta$ and elementary divisors $(\lambda-\theta)$, $(\lambda-1)^{j(2)-1}$ such that the elementary divisors of A_2D_2 are $(\lambda-\theta^2)$, $(\lambda-1)^{j(2)-1}$. Construct a standard matrix $D_3\in GL(j(3),GF(4))$ with $d_1=\theta^2$ and elementary divisors $(\lambda-\theta^2)$, $(\lambda-\theta)^{j(3)-1}$ such that the elementary divisors of A_3D_3 are $(\lambda-1)$, $(\lambda-\theta)^{j(3)-1}$. Set $D=D_1\dotplus D_2\dotplus D_3$. Then $D\in SL(n,GF(4))$ and D and AD have the same elementary divisors, completing the proof.

Now suppose that $j(1)$, $j(2)$, $j(3)$ are distinct (mod 3), say $j(1)\equiv 2 \pmod 3$, $j(2)\equiv 0 \pmod 3$, $j(3)\equiv 1 \pmod 3$. Then $j(1)\geqq 2$ and $j(2)\geqq 3$. Construct, by Lemma 5, a standard matrix $D_1\in GL(j(1),GF(4))$ with $d_1=1$ and elementary divisors $(\lambda-1)$, $(\lambda-\theta)$, $(\lambda-\theta)^{j(1)-2}$ such that the elementary divisors of A_1D_1 are $(\lambda-1)$, $(\lambda-\theta^2)$, $(\lambda-\theta)^{j(1)-2}$. Construct, again by Lemma 5, a standard matrix $D_2\in GL(j(2),GF(4))$ with $d_1=\theta$ and elementary divisors $(\lambda-\theta)$, $(\lambda-\theta^2)$, $(\lambda-\theta^2)^{j(2)-2}$ such that the elementary divisors of A_2D_2 are $(\lambda-\theta)$, $(\lambda-1)$, $(\lambda-\theta^2)^{j(2)-2}$. Construct, by Lemma 4, a standard matrix $D_3\in GL(j(3),GF(4))$ with $d_1=1$ and elementary divisors $(\lambda-1)$, $(\lambda-\theta^2)^{j(3)-1}$ such that the elementary divisors of A_3D_3 are $(\lambda-\theta)$, $(\lambda-\theta^2)^{j(3)-1}$. Set $D=D_1\dotplus D_2\dotplus D_3$. Then $|D|=1$ and D and AD have the same elementary divisors.

The proof for $K=GF(4)$ is now complete.

7. **Concluding remarks.** Let $PSL(n, K)$ denote the factor group $SL(n, K)/C(n, K)$ where $C(n, K)$ is the centre of $SL(n, K)$ i.e. the scalar matrices in $SL(n, K)$. It is known [1] that $PSL(n, K)$ is a simple group except in the two cases $n=2$, $K=GF(2)$ and $n=2$, $K=GF(3)$. O. Ore [8] has raised the question of determining when every element of a finite simple group G can be expressed as a commutator of G. Our results enable us to answer this question for the class of (finite and infinite) simple groups $PSL(n, K)$ where K is not $GF(2)$ or $GF(3)$.

THEOREM 3. *If K is not $GF(2)$ or $GF(3)$, then every element of $PSL(n, K)$ is a commutator of $PSL(n, K)$.*

In fact, the results announced in §1 enable us to prove that whenever $PSL(n, K)$ is simple then every element of $PSL(n, K)$ is a commutator of $PSL(n, K)$. For $n=2$, $p>3$, and $K=GF(p)$, this result has already been established by Villari [13].

REFERENCES

1. L. E. Dickson, *Linear groups with an exposition of Galois field theory*, Dover, 1958.

2. K. Fan, *Some remarks on commutators*, Arch. Math. vol. 5 (1954) pp. 102–107.

3. M. Gotô, *A theorem on compact semi-simple groups*, J. Math. Soc. Japan vol. 1 (1949) pp. 270–272.

4. N. Itô, *A theorem on the alternating group $A_n(n \geqq 5)$*, Math. Japon. vol. 2 (1951) pp. 59–60.

5. E. Landau, *Über die Darstellung definiter Funktionen durch Quadrate*, Math. Ann. vol. 62 (1906) pp. 271–285.

6. W. LeVeque, *Topics in number theory*, vol. 1, Addison-Wesley, 1956.

7. T. Muir and W. H. Metzler, *A treatise on the theory of determinants*, Albany, New York, 1930.

8. O. Ore, *Some remarks on commutators*, Proc. Amer. Math. Soc. vol. 2 (1951) pp. 307–314.

9. K. Shoda, *Einige Sätze über Matrizen*, Jap. J. Math. vol. 13 (1936) pp. 361–365.

10. ———, *Über den Kommutator der Matrizen*, J. Math. Soc. Japan vol. 3 (1951) pp. 78–81.

11. H. Tôyama, *On commutators of matrices*, Kōdai Math. Seminar Reports (1949) pp. 1–2.

12. C. L. Siegel, *Darstellung total positiver Zahlen durch Quadrate*, Math. Z. vol. 11 (1921) pp. 246–275.

13. G. Villari, *Sui commutatori del gruppo modulare*, Boll. Un. Mat. Ital. vol. 13 (1958) pp. 196–201.

UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER, CANADA