COMMUTATORS IN THE SPECIAL AND GENERAL LINEAR GROUPS

Thesis by
Robert Charles Thompson

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1960

ACKNOWLEDGEMENTS

The author wishes to acknowledge his large debt to Dr. O. Taussky-Todd for suggesting the problem studied in this thesis, for encouragement and advice, and for many hours devoted to reading drafts of this thesis. He also wishes to acknowledge a useful suggestion made by Dr. H. Zassenhaus during the course of the research presented in the thesis. He wishes to thank the National Research Council (of Canada) for financial support (1958-60) and the California Institute of Technology for the use of its facilities and for financial support (1957-60). Finally, he wishes to thank Miss E. Gibson for the expert typing of the thesis.

ABSTRACT

Let GL(n, K) denote the multiplicative group of all non-singular nxn matrices with coefficients in a field K; SL(n, K) the subgroup of GL(n, K) consisting of all matrices with determinant unity; C(n, K) the centre of SL(n, K); PSL(n, K) the factor group SL(n, K)/C(n, K); I_n the nxn identity matrix; $GF(p^n)$ the finite field with p^n elements. We determine when every element of SL(n, K) is a commutator of SL(n, K) or of GL(n, K). Theorem 1. Let A $\in$ SL(n, K). Then it follows that A is a commutator $BCB^{-1}C^{-1}$ of SL(n, K) unless: (i) n = 2 and K = GF(2); (ii) n = 2 and K = GF(3); or (iii) K has characteristic zero and $A = aI_n$ where a is a primitive n^{th} root of unity in K and $n \equiv 2 \pmod 4$. In case (i), SL(2, GF(2)) properly contains its commutator subgroup. In case (ii), SL(2, GF(3)) properly contains its commutator subgroup. Furthermore, every element of SL(2, GF(3)) is a commutator of GL(2, GF(3)). In case (iii), aI_n is always a commutator of GL(n,K). Moreover, aI_n is a commutator of SL(n, K) when, and only when, the equation $-1 = x^2 + y^2$ has a solution x, y $\in$ K. Hence: Theorem 2. Whenever PSL(n, K) is simple, every element of PSL(n, K) is a commutator of PSL(n, K). Theorem 1 simplifies and extends results due to K. Shoda (Jap. J. Math., 13 (1936), p. 361-365; J. Math. Soc. of Japan, 3 (1951), p. 78-81). Theorem 2 supports the suggestion made by O. Ore (Proc. Amer. Math. Soc., 2 (1951), p. 307-314) that in a finite simple group, every element is a commutator.

1. INTRODUCTION

Let K be a commutative field, A an n row square matrix (briefly, an $n \times n$ matrix) with coefficients in K, I_n the $n \times n$ identity matrix, $|A|$ the determinant of A. Let $GL(n, K)$ denote the multiplicative group of all $n \times n$ non-singular matrices with coefficients in K and let $SL(n, K)$ denote the subgroup of $GL(n, K)$ consisting of all matrices in $GL(n, K)$ with determinant one. It is known [1] that the centre $C(n, K)$ of $SL(n, K)$ consists of all scalar matrices with determinant unity. Let $PSL(n, K) = SL(n, K)/C(n, K)$. Finally let $GF(p^n)$ denote the finite field with p^n elements.

The following theorem has been known a long time; a proof may be constructed from material contained in [2], [3], and [4].

> THEOREM. (1). $SL(n, K)$ is its own commutator group, except when $n = 2$ and $K = GF(2)$ or when $n = 2$ and $K = GF(3)$. The commutator subgroup of $GL(2, GF(3))$ is $SL(2, GF(3))$.
>
> (2). $PSL(n, K)$ is a simple group except when $n = 2$ and $K = GF(2)$ or when $n = 2$ and $K = GF(3)$. Furthermore, $PSL(2, GF(2)) \cong SL(2, GF(2)) \cong GL(2, GF(2)) \cong S_3$ and $PSL(2, GF(3)) \cong A_4$, where S_3 is the symmetric group on three letters and A_4 is the alternating group on four letters.

In recent years a number of authors have studied the following problem: Given a set (usually a multiplicative group) of non-singular matrices with coefficients in a field, when is a matrix A in this set a commutator $BCB^{-1}C^{-1}$ or a product of commutators of matrices in the set? Clearly, a necessary condition is that $|A| = 1$. The first results in the converse direction are due to Shoda. In two papers, Shoda studied the problem of representing a matrix $A \in SL(n, K)$ as products of commutators of elements of $GL(n, K)$:

$$A = \prod_{i=1}^{m} (B_i C_i B_i^{-1} C_i^{-1}); \quad B_i, C_i \in GL(n, K).$$

In his first paper [5], Shoda showed that if K is algebraically closed, then only one commutator is required $(m = 1)$ and if K is real closed, then two commutators suffice $(m \leqq 2)$. Generalizing this result in his second paper [6], Shoda showed that if K has infinitely many elements, then not more than N commutators $(m \leqq N)$ are required, where N is the largest of the degrees over K of the characteristic values of A.

Other recent investigations of commutators of matrices have been made by Toyama, Taussky, and Fan. Toyama [7] proved that each element of the unitary unimodular group over the field C of complex numbers, the unitary symplectic group over C, or the proper orthogonal group of degree larger than two over the real number field, is a commutator in its respective group. Taussky [8] proved that if $X, Y \in GL(n, C)$, then matrices U, V exist in $GL(n, C)$ such that $X = UVYU^{-1}V^{-1}$ if, and only if, $|X| = |Y|$. Fan [9] reproved part of Toyama's results and also studied the problem of representing normal or Hermitian matrices with complex coefficients as commutators of normal or Hermitian matrices, respectively. Fan also extended Taussky's result by showing that if $x, y \in G$, an arbitrary group, then elements u, v exist in G such that $x = uvyu^{-1}v^{-1}$ if, and only if, xy^{-1} is a commutator of G.

An investigation of commutators in PSL(2, GF(p)) has been made by Villari [10]. Villari showed that if $p > 3$, every element of PSL(2, GF(p)) is a commutator of PSL(2, GF(p)). He also remarked that this result is false if $p = 2$ or 3.

The analogous commutator problem for permutation groups has been studied by Ore [11] and by Itô [12]. Ore and Itô proved simultaneously (but independently) that each element of A_n, the alternating group on n letters, is a commutator of A_n whenever $n \geq 5$. Since A_n is known to be a simple group whenever $n \geq 5$, Ore suggested that it may be true that every element of a finite simple group is a commutator.

It is the purpose of this thesis to determine when every element of SL(n, K) is a commutator of SL(n, K) or of GL(n, K). The theorem quoted above shows that for most integers n and most fields K the conjecture that every element of SL(n, K) is a commutator of SL(n, K) or of GL(n, K) is not an unreasonable one. Our results, which will be significant improvements of Shoda's results, will enable us to prove Ore's conjecture for those members of the class of groups PSL(n, K) which are simple and will, at the same time, reprove Villari's results.

2. RESULTS AND METHODS

Our main result is Theorem 1.

THEOREM 1. Let $A \in SL(n, K)$. Then, apart from the exceptional cases noted below, A is a commutator of $SL(n, K)$. The exceptional cases are:

(1) $n = 2$ and $K = GF(2)$. Here $SL(2, GF(2))$ properly contains its commutator subgroup.

(2) $n = 2$ and $K = GF(3)$. Here $A \in SL(2, GF(3))$ implies that A is a commutator of $GL(2, GF(3))$. Furthermore, $SL(2, GF(3))$ properly contains its commutator subgroup.

(3) K has characteristic zero and $A = aI_n$, where $n \equiv 2 \pmod 4$ and a is a primitive n^{th} root of unity in K. Here aI_{4m+2} is always a commutator of $GL(4m + 2, K)$. Moreover, the necessary and sufficient condition that aI_{4m+2} be a commutator of $SL(4m + 2, K)$ is that the equation $-1 = x^2 + y^2$ have a solution $x, y \in K$.

We proceed to deduce a number of corollaries of Theorem 1.

COROLLARY 1. Except in cases 1 and 2 (and 3 if -1 is not a sum of two squares within K), then $A \in SL(n, K)$ implies that A is a commutator of arbitrarily high weight in $SL(n, K)$.

PROOF. If $A = BCB^{-1}C^{-1}$, we simply have to reapply Theorem 1 to the matrices B and C and iterate, noting that neither B nor C

can be a scalar matrix if A is not the identity.

COROLLARY 2. Let $X, Y \in GL(n, K)$. Then, except when $n = 2$ and $K = GF(2)$, matrices C and D exist in $GL(n, K)$ such that $X = CDYC^{-1}D^{-1}$ if, and only if, $|X| = |Y|$.

PROOF. This is just the previously mentioned result of Taussky. In the same way, it is easy to determine when $X = CDYC^{-1}D^{-1}$ for $X, Y, C, D \in SL(n, K)$.

Since $S_3 \cong SL(2, GF(2)) = GL(2, GF(2))$ and since S_3 properly contains its derived group, the exceptional case 1 of Theorem 1 is genuine. Since $PSL(n, K)$ is a homomorphic image of $SL(n, K)$, and since A_4 properly contains its commutator subgroup, then from $PSL(2, GF(3)) \cong A_4$ we immediately see that the exceptional case 2 is also genuine. In the homomorphism from $SL(n, K)$ onto $PSL(n, K)$, the scalar matrices map onto the identity, which clearly is a commutator of $PSL(n, K)$. Since the exceptional cases 1 and 2 of Theorem 1 correspond just to those cases in which $PSL(n, K)$ is not simple, we have established Ore's conjecture for the simple groups belonging to the class $PSL(n, K)$ of groups. We state this as Theorem 2.

THEOREM 2. Whenever $PSL(n, K)$ is simple, then every element of $PSL(n, K)$ is a commutator of $PSL(n, K)$.

We now sketch our methods of proof. Given A with $|A| = 1$, we perform a similarity transformation which throws A into a rational canonical form. We then construct a triangular matrix D with coefficients in K such that $|D| = 1$. The elementary divisors of

D depend on its diagonal elements and the structure of its non-zero triangle. Our construction of this non-zero triangle of D will make the elementary divisors independent of the particular values of certain of the off-diagonal elements. We shall attempt to choose these off-diagonal elements such that AD and D have the same elementary divisors. This choice will involve solving a set of linear equations. If a solution can be found, the existence of a matrix S such that $AD = SDS^{-1}$ will be guaranteed. We shall show that we can satisfy $|S| = 1$. From this follows $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, K)$.

Our general methods will break down when the field K has five or fewer elements. Hence we shall have to give special arguments when the field K is GF(5), GF(4), GF(3) or GF(2).

For the matrix theory used in this thesis, we refer the reader to any standard text on matrix theory: for example, [13], Chapter 7.

3. NOTATION AND DEFINITIONS

In this section we shall describe some of the notation to be used. Part of the notation has already been described in the introduction.

If $p(\lambda) = \lambda^n + a_n \lambda^{n-1} + \cdots + a_1$ is a polynomial with coefficients in K, by $C(p(\lambda))$ we denote the companion matrix of $p(\lambda)$:

$$C(p(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & & & & \\ . & & & . & & & \\ . & & & & . & & \\ . & & & & & . & 0 \\ 0 & . & . & . & . & 0 & 1 \\ -a_1 & . & . & . & . & . & -a_n \end{pmatrix}, \quad n \geq 2;$$

$C(p(\lambda)) = (-a_1), \quad n = 1.$

Let $E_{\alpha,\beta}$ denote an nxn matrix with a one in row α and column β, and zeros elsewhere. The dimensions of $E_{\alpha,\beta}$ will always be clear from context. Let $S_{\alpha,\beta}(u) = I_n + uE_{\alpha,\beta}$. If $\alpha \neq \beta$, $S_{\alpha,\beta}(u)AS^{-1}_{\alpha,\beta}(u)$ is a matrix obtained from A by adding the u^{th} multiple of row β to row α, then adding the $(-u)^{th}$ multiple of column α to column β in the resulting matrix. Because of the associative law, $S_{\alpha,\beta}(u)AS^{-1}_{\alpha,\beta}(u)$ may also be obtained by performing first the column operation, then the row operation in the resulting matrix. Since such similarity transformations will occasionally occur, we give them a special name.

DEFINITION 1. An elementary similarity transformation of a matrix A is a similarity transformation of A which replaces A with $S_{\alpha,\beta}(u)AS^{-1}_{\alpha,\beta}(u)$ where $u \in K$ and $\alpha \neq \beta$.

We shall often specify α, β and u in an elementary similarity transformation by giving the row operation to be performed, or by giving the column operation.

If $a \in K$ and n is a positive integer, by $J_n(a)$ we denote the Jordan matrix of a dimension n:

$$J_n(a) = \begin{pmatrix} a & 1 & & & & \\ & a & 1 & & 0 & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & a & 1 \\ & 0 & & & & a \end{pmatrix}, \quad n \geq 2;$$

$J_1(a) = (a)$.

If A is an nxn matrix and B is an mxm matrix, by $A \dotplus B$ we denote the $(m + n) \times (m + n)$ matrix

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

The following matrix will be encountered so frequently that we give it a special name.

DEFINITION 2. The following nxn matrix D will be called a standard matrix.

$$D = \begin{pmatrix} d_1 & d_2 & d_3 & \cdot & \cdot & & d_n \\ & J_{s_1}(c_1) & & & 0 & & \\ & & J_{s_2}(c_2) & & & & \\ & & & \cdot & & & \\ & 0 & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & J_{s_r}(c_r) \end{pmatrix},$$

$$1 + s_1 + s_2 + \cdots + s_r = n, \quad n \geq 2;$$

$$D = (d_1), \quad n = 1.$$

The standard matrix is defined by the following parameters: the integers n, r, $s_1, \ldots, s_r$; and the field elements $d_1, d_2, \ldots, d_n$, $c_1, c_2, \ldots, c_r$. When describing a standard matrix, we shall always indicate the values to be assigned to these parameters in terms of the notation used in Definition 2 above.

In a matrix, an asterisk * will indicate elements whose precise values do not matter.

4. PRELIMINARY LEMMAS

In this section we collect together the lemmas that will be required later.

LEMMA 1. Let D be the standard matrix of Definition 2, with coefficients in a field K. Then, if $d_1 \neq c_i$ for $i = 1, 2, \ldots, r$, the elementary divisors of D are $(\lambda - d_1)$, $(\lambda - c_1)^{s_1}, \ldots, (\lambda - c_r)^{s_r}$. (If $n = 1$, the elementary divisor of D is $(\lambda - d_1)$.)

PROOF. The result is clear if $n = 1$. If $n \geqq 2$ and if $x \in K$ is suitably chosen, the matrix $S_{1,i}(x) D S_{1,i}^{-1}(x)$ $(i \geqq 2)$ has the same structure as D, except that d_i is replaced with 0 and d_{i+1} is altered. By a sequence of similarity transformations of this type, with $i = 2, 3, \ldots, n$, we may bring D to $J_1(d_1) \dotplus J_{s_1}(c_1) \dotplus \cdots \dotplus J_{s_r}(c_r)$. The structure of this matrix exhibits the required elementary divisors.

LEMMA 2. For $n \geqq 2$, let

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

be a matrix with coefficients in K such that $f_{1,2} f_{2,3} \ldots f_{n-1,n} \neq 0$. Then a matrix S exists with coefficients in K such that $|S| = 1$ and $SFS^{-1} = G$, where

$$G = \begin{pmatrix} 0 & f_{1,2} & 0 & . & . & . & 0 \\ 0 & 0 & f_{2,3} & 0 & . & . & 0 \\ . & & & . & & & . \\ . & & & & . & & . \\ . & & & & & . & 0 \\ 0 & & . & . & . & 0 & f_{n-1,n} \\ y_1 & y_2 & y_3 & . & . & . & y_n \end{pmatrix}$$

and

$$\left.\begin{aligned} y_n &= x_n, \\ y_{n-1} &= x_{n-1} + \sum_{i=1}^{n-2} a_{n-1,i}\, x_i, \\ &\ldots \\ y_j &= x_j + \sum_{i=1}^{j-1} a_{j,i}\, x_i, \\ &\ldots \\ y_1 &= x_1. \end{aligned}\right\} \quad (1)$$

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

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

$S_{n-2,n-1}(a_{n-2}) \ldots S_{1,n-1}(a_1)\, F\, S^{-1}_{1,n-1}(a_1) \ldots S^{-1}_{n-2,n-1}(a_{n-2})$ has the same structure as F, except that the coefficients standing above $f_{n-1,n}$ in the last column are replaced with zeros and the $(n - 1)^{st}$ column is the $(n - 1)^{st}$ column of F plus linear combinations of columns $1, \ldots, n-2$ of F. Repeating this procedure with columns $n-1, \ldots, 3$ produces the matrix G. The assertion about the determinant of S is clear since $|S_{i,j}(u)| = 1$ whenever $i \neq j$.

LEMMA 3. The matrix G of Lemma 2 is similar to $H = C(\lambda^n - w_n \lambda^{n-1} - \ldots - w_1)$, where

$$\left.\begin{aligned} w_n &= y_n, \\ w_{n-1} &= f_{n-1,n} y_{n-1}, \\ &\ldots \\ w_i &= f_{i,i+1} \cdots f_{n-1,n} y_i, \\ &\ldots \\ w_1 &= f_{1,2} \cdots f_{n-1,n} y_1. \end{aligned}\right\} \quad (2)$$

PROOF. Let $S = (s_{i,j})$ be a diagonal matrix with $s_{n,n} = 1$ and $s_{i,i} = f_{i,i+1}^{-1} \cdots f_{n-1,n}^{-1}$; $i = 1, \ldots, n-1$. Then $SGS^{-1} = H$.

LEMMA 4. Let

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

be an nxn matrix with coefficients in K. Let D be the nxn standard matrix of Definition 2 with coefficients in K and $|A|\, c_1 \cdots c_r \neq 0$. 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 $d_1, c_1, \ldots, c_r$, it is possible to choose $d_2, \ldots, d_n \in K$ in such a manner that $q(\lambda)$ is the characteristic and minimum polynomial of AD. (When $n = 1$, $q(\lambda) = \lambda - d_1$ and the characteristic and minimum polynomial of AD is

$(\lambda - |A| d_1)$.)

PROOF. For $n = 1$, there are no $d_2, \ldots, d_n$ to be chosen and the result is clear. Hence suppose $n \geq 2$. Compute AD.

$$AD = \begin{pmatrix} 0 & J_{s_1}(c_1) & 0 & \cdot & \cdot & 0 \\ 0 & 0 & J_{s_2}(c_2) & 0 & \cdot & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & 0 & J_{s_r}(c_r) \\ x_1 & x_2 & x_3 & \cdot & \cdot & x_n \end{pmatrix}$$

where the first side diagonal of AD above the main diagonal contains $c_1, \ldots, c_1, c_2, \ldots, c_2, \ldots, c_r, \ldots, c_r$ as coefficients (c_i appears s_i times, $i = 1, 2, \ldots, r$) and

$$\left.\begin{aligned} x_1 &= (-1)^{n-1} |A| d_1, \\ x_2 &= (-1)^{n-1} |A| d_2 + (\text{a linear expression in } a_2), \\ &\ldots \\ x_i &= (-1)^{n-1} |A| d_i + (\text{a linear expression in } a_i, a_{i-1}), \\ &\ldots \\ x_n &= (-1)^{n-1} |A| d_n + (\text{a linear expression in } a_n, a_{n-1}). \end{aligned}\right\} \quad (3)$$

The coefficients in the indicated linear expression depend on $c_1, \ldots, c_r$ but not on $d_1, \ldots, d_n$. Invoking Lemmas 2 and 3, we find that AD is similar to $C(p(\lambda))$ where $p(\lambda) = \lambda^n - w_n \lambda^{n-1} - \ldots - w_1$ and $w_1, \ldots, w_n$ are related to $d_1, \ldots, d_n$ by equations 1, 2, and 3. Since the characteristic polynomial of a companion matrix is also the minimum polynomial, the result will follow if we can determine

$d_2, \ldots, d_n$ such that $p(\lambda) = q(\lambda)$. The constant term of $p(\lambda)$ is $(-1)^n |A| d_1 c_1^{s_1} \ldots c_r^{s_r}$. But this is the constant term of $q(\lambda)$. Next, set $w_i = -q_i$, $i = 2, \ldots, n$ and note that equations 2 may be solved for $y_2, \ldots, y_n$ in terms of $w_2, \ldots, w_n$. From equations 1 we may determine $x_2, \ldots, x_n$ in terms of $y_2, \ldots, y_n$. Finally, since $|A| \neq 0$, equations 3 determine $d_2, \ldots, d_n$ in terms of $x_2, \ldots, x_n$. With $d_2, \ldots, d_n$ determined in this way, $p(\lambda) = q(\lambda)$ and the proof is complete.

The result just obtained will be the cornerstone of the proof of Theorem 1. It will most often be used in the following form.

LEMMA 5. Under the hypothesis of Lemma 4, if $|A| d_1, c_1, \ldots, c_r, 0$ are all distinct elements of K, then $d_2, \ldots, d_n$ may be determined such that the elementary divisors of AD are $(\lambda - |A| d_1), (\lambda - c_1)^{s_1}, \ldots, (\lambda - c_r)^{s_r}$. (When $n = 1$, the elementary divisor of AD is $(\lambda - |A| d_1)$.)

PROOF. Let

$$q(\lambda) = (\lambda - |A| d_1)(\lambda - c_1)^{s_1} \ldots (\lambda - c_r)^{s_r}$$

and choose $d_2, \ldots, d_n$ such that the characteristic and minimum polynomial of AD is $q(\lambda)$. But then the elementary divisors of AD are obtained by decomposing $q(\lambda)$ into its relatively prime constituents. From this observation the result immediately follows.

LEMMA 6. Let A and B be two matrices with coefficents in K, such that $SAS^{-1} = B$ for some S. If A possesses an elementary divisor $\lambda - \alpha$ with $\alpha \in K$, then a matrix T exists with coefficients in K such that, if t is any non-zero element of K, $TAT^{-1} = B$ and $|T| = t$.

PROOF. Since $\lambda - \alpha$ is an elementary divisor of A (and hence of B), matrices S_1 and S_2 exist with coefficients in K such that $S_1AS_1^{-1} = S_2BS_2^{-1} = (\alpha) \dot{+} A_1$ where A_1 is some $(n-1)\times(n-1)$ matrix. Let $t_1 = t|S_2S_1^{-1}|$ and set $T_1 = (t_1) \dot{+} I_{n-1}$. Then $T_1S_1AS_1^{-1}T_1^{-1} = S_1AS_1^{-1} = S_2BS_2^{-1}$, hence $S_2^{-1}T_1S_1AS_1^{-1}T_1^{-1}S_2 = B$. Set $T = S_2^{-1}T_1S_1$.

The following Lemma will be used in the study of matrices over GF(3).

LEMMA 7. For $n \geq 3$, let

$$H = \begin{pmatrix} 0 & d & c_1 & & & & & \\ 0 & 0 & d & c_2 & & & * & \\ \cdot & \cdot & & \cdot & \cdot & & & \\ \cdot & \cdot & & & \cdot & \cdot & & \\ \cdot & \cdot & & & & \cdot & \cdot & \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & d & c_{n-2} \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & d \\ h_1 & h_2 & \cdot & \cdot & \cdot & h_{n-2} & h_{n-1} & h_n \end{pmatrix}$$

Then the coefficient of λ in $|\lambda I_n - H|$ is $-(h_2d^{n-2} + h_1d^{n-3}(c_1 + \cdots + c_{n-2}))$.

PROOF. This coefficient is $(-1)^{n-1}$ (the sum of the principal minors of H of order $n-1$). The minor obtained by deleting the first row and column of H has value $(-1)^{n-2}d^{n-2}h_2$. The minor obtained by deleting the last row and column vanishes. The minor obtained by deleting row i and column $i\ (2 \leq i \leq n-1)$ is

$$\begin{vmatrix}
0 & d & & & & & & & \\
 & & d & & & & & * & \\
 & & & \cdot & & & & & \\
 & & & & \cdot & & & & \\
 & & & & & d & & & \\
 & O & & & & & c_{i-1} & & \\
 & & & & & & & d & \\
 & & & & & & & & \cdot \\
 & & & & & & & & \cdot \\
h_1 & & & & & * & & & d
\end{vmatrix}$$

$$= (-1)^{n-2}h_1 d^{n-3} c_{i-1}.$$

5. THE PROOF OF THEOREM 1 WHEN K CONTAINS SIX OR MORE ELEMENTS

Throughout this section A will denote an nxn matrix with coefficients in K, $|A| = 1$. Since any factorization $A = BCB^{-1}C^{-1}$ is preserved under a similarity transformation, we may perform a similarity transformation and throw A into a rational canonical form. Thus we may suppose that $A = A_1 \dotplus \cdots \dotplus A_m$, where A_i is a $j_i x j_i$ companion matrix of a polynomial with coefficients in K, $i = 1, 2, \ldots, m$. By rearranging the A_i if necessary, 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 and $j_1, \ldots, j_m$. Part of the proof presented in this section will be valid when K contains fewer than six elements.

CASE 1. The 2x2 matrices.

If A is 2x2 and not scalar, then A is similar to the companion matrix of a single polynomial with coefficients in K. Choose $\rho \in K$ such that $\rho^2 \neq 1, 0$. (This is possible if K is not GF(2) or GF(3).) Let

$$D = \begin{pmatrix} \rho & d_2 \\ 0 & \rho^{-1} \end{pmatrix}$$

be a standard 2x2 matrix. By Lemma 1, the elementary divisors of D are $(\lambda - \rho), (\lambda - \rho^{-1})$ since $\rho \neq \rho^{-1}$. By Lemma 5, if d_2 is properly chosen, the elementary divisors of AD are also $(\lambda - \rho), (\lambda - \rho^{-1})$. Hence, by Lemma 6, a matrix S exists in SL(2, K) such that $AD = SDS^{-1}$. Hence $A = SDS^{-1}D^{-1}$ where $S, D \in SL(2, K)$.

We shall give the proof for the scalar 2x2 matrices under a later case.

CASE 2. $j_m \geq 3$.

In the sequel, whenever we list the elementary divisors of a matrix and include a term $(\lambda - \gamma)^w$ where $w = 0$, then $(\lambda - \gamma)^w$ is to be deleted from the list.

Let δ_1 be any non-zero element of K and define

$$\left.\begin{aligned} \delta_2 &= |A_1|\,\delta_1, \\ &\dots \\ \delta_i &= |A_{i-1}|\,\delta_{i-1}, \\ &\dots \\ \delta_m &= |A_{m-1}|\,\delta_{m-1}. \end{aligned}\right\} \quad (4)$$

Then, since $|A_1 \dots A_m| = 1$,

$$\delta_1 = |A_m|\,\delta_m. \quad (5)$$

For $i = 1, 2, \dots, m-1$ let γ_i be an element of K distinct from δ_i, δ_{i+1}, 0. Let γ_m be any element of K distinct from δ_m, δ_1, 0 and define γ_m''' by the condition

$$\delta_1 \delta_2 \dots \delta_m \gamma_1^{j_1 - 1} \gamma_2^{j_2 - 1} \dots \gamma_{m-1}^{j_{m-1} - 1} \gamma_m^{j_m - 2} \gamma_m''' = 1. \quad (6)$$

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

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

These conditions prohibit at most four non-zero values of x. Hence, if K has six or more elements, a suitable x always exists. Let

$$\gamma'_m = \gamma_m x,$$
$$\gamma''_m = \gamma'''_m x^{-1}.$$

Then

$$\delta_1 \delta_2 \cdots \delta_m \gamma_1^{j_1-1} \gamma_2^{j_2-1} \cdots \gamma_{m-1}^{j_{m-1}-1} \gamma_m^{j_m-3} \gamma'_m \gamma''_m = 1.$$

For $i = 1, 2, \ldots, m-1$ let D_i be a $j_i x j_i$ standard matrix with $d_1 = \delta_i$ and (if $j_i \geq 2$), $r = 1$, $s_1 = j_i - 1$, $c_1 = \gamma_i$, and with $d_2, \ldots, d_{j_i}$ so chosen that the elementary divisors of $A_i D_i$ are $(\lambda - \delta_{i+1})$, $(\lambda - \gamma_i)^{j_i-1}$. (Because of the way in which the δ_i and γ_i have been selected, we may use Lemma 5 to make this choice.) By Lemma 1, the elementary divisors of D_i are $(\lambda - \delta_i)$, $(\lambda - \gamma_i)^{j_i-1}$. We now construct a matrix D_m. There are five different possibilities depending on the values of γ_m, γ'_m, γ''_m.

CASE 2.1.

If γ_m, γ'_m, γ''_m are all distinct, let D_m be a $j_m x j_m$ standard matrix with $d_1 = \delta_m$. If $j_m = 3$, we take $r = 2$, $c_1 = \gamma'_m$, $s_1 = 1$, $c_2 = \gamma''_m$, $s_2 = 1$. If $j_m > 3$, we set $r = 3$, take c_1, s_1, c_2, s_2 as just indicated, and $c_3 = \gamma_m$, $s_3 = j_m - 3$. In either case, it follows from Lemma 1 (in accordance with a convention indicated above) that the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma'_m)$, $(\lambda - \gamma''_m)$, $(\lambda - \gamma_m)^{j_m-3}$. Select $d_2, \ldots, d_{j_m}$ such that the elementary divisors of $A_m D_m$ are $(\lambda - \delta_1)$, $(\lambda - \gamma'_m)$, $(\lambda - \gamma''_m)$, $(\lambda - \gamma_m)^{j_m-3}$. (Lemma 5 and equation 5.) Now let $D = D_1 \dotplus \cdots \dotplus D_m$. Then, because of our choice of γ'_m, γ''_m, $|D| = 1$. The elementary divisors of D are

$$(\lambda - \delta_1),\ (\lambda - \gamma_1)^{j_1 - 1},\ (\lambda - \delta_2),\ (\lambda - \gamma_2)^{j_2 - 1},$$

$$\ldots,\ (\lambda - \delta_{m-1}),\ (\lambda - \gamma_{m-1})^{j_{m-1} - 1},\ (\lambda - \delta_m), \tag{7}$$

$$(\lambda - \gamma_m'),\ (\lambda - \gamma_m''),\ (\lambda - \gamma_m)^{j_m - 3}$$

and the elementary divisors of AD are

$$(\lambda - \delta_2),\ (\lambda - \gamma_1)^{j_1 - 1},\ (\lambda - \delta_3),\ (\lambda - \gamma_2)^{j_2 - 1},$$

$$\ldots,\ (\lambda - \delta_m),\ (\lambda - \gamma_{m-1})^{j_{m-1} - 1},\ (\lambda - \delta_1), \tag{8}$$

$$(\lambda - \gamma_m'),\ (\lambda - \gamma_m''),\ (\lambda - \gamma_m)^{j_m - 3}.$$

Since the display 8 is the same as 7 except for a rearrangement, D and AD have the same elementary divisors. Thus, by Lemma 6, a matrix S exists with coefficients in K and determinant unity such that $AD = SDS^{-1}$. Hence $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, K)$ as required.

CASE 2.2.

If $\gamma_m = \gamma_m' \neq \gamma_m''$, let D_m be a $j_m x j_m$ standard matrix with $d_1 = \delta_m$, $r = 2$, $c_1 = \gamma_m''$, $s_1 = 1$, $c_2 = \gamma_m$, $s_2 = j_m - 2$. Then the elementary divisors of D_m are $(\lambda - \delta_m), (\lambda - \gamma_m''), (\lambda - \gamma_m)^{j_m - 2}$ (Lemma 1) and by Lemma 5 and equation 5 we may determine $d_2, \ldots, d_{j_m}$ such that the elementary divisors of $A_m D_m$ are $(\lambda - \delta_1),\ (\lambda - \gamma_m''),\ (\lambda - \gamma_m)^{j_m - 2}$. Let $D = D_1 \dotplus \cdots \dotplus D_m$. Then the elementary divisors of D are

$$(\lambda - \delta_1),\ (\lambda - \gamma_1)^{j_1 - 1},\ \dots,\ (\lambda - \delta_{m-1}),$$

$$(\lambda - \gamma_{m-1})^{j_{m-1} - 1},\ (\lambda - \delta_m),\ (\lambda - \gamma''_m), \tag{9}$$

$$(\lambda - \gamma_m)^{j_m - 2}.$$

Furthermore, the elementary divisors of AD are

$$(\lambda - \delta_2),\ (\lambda - \gamma_1)^{j_1 - 1},\ \dots,\ (\lambda - \delta_m),$$

$$(\lambda - \gamma_{m-1})^{j_{m-1} - 1},\ (\lambda - \delta_1),\ (\lambda - \gamma''_m), \tag{10}$$

$$(\lambda - \gamma_m)^{j_m - 2}.$$

As before, 10 is merely a rearrangement of 9, hence, by Lemma 6, $AD = SDS^{-1}$, $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, K)$.

CASE 2.3.

Here we assume $\gamma_m = \gamma''_m \neq \gamma'_m$. In this case the proof is the same as the proof in case 2.2 except that γ'_m and γ''_m (in this case) play the roles of γ''_m and γ'_m (in the previous case) respectively.

CASE 2.4.

Here we assume that $\gamma'_m = \gamma''_m \neq \gamma_m$. Let D_m be a $j_m x j_m$ standard matrix with $d_1 = \delta_m$. If $j_m = 3$, we set $r = 1$; $c_1 = \gamma'_m$; $s_1 = 2$. If $j_m > 3$, we set $r = 2$; c_1, s_1 as just indicated; and $c_2 = \gamma_m$; $s_2 = j_m - 3$. The elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma'_m)^2$, $(\lambda - \gamma_m)^{j_m - 3}$, and we may choose $d_2, \dots, d_{j_m}$ such that the elementary divisors of $A_m D_m$ are

$(\lambda - \delta_1)$, $(\lambda - \gamma'_m)^2$, $(\lambda - \gamma_m)^{j_m - 3}$. Let $D = D_1 \dot{+} \cdots \dot{+} D_m$ and complete the proof as before.

CASE 2.5.

If $\gamma_m = \gamma'_m = \gamma''_m$, let D_m be a $j_m x j_m$ standard matrix with $d_1 = \delta_m$, $r = 1$, $c_1 = \gamma_m$, $s_1 = j_m - 1$. Then the elementary divisors of D_m are $(\lambda - \delta_m)$, $(\lambda - \gamma_m)^{j_m - 1}$ and, if $d_2, \ldots, d_{j_m}$ are properly selected, the elementary divisors of $A_m D_m$ are $(\lambda - \delta_1)$, $(\lambda - \gamma_m)^{j_m - 1}$. Let $D = D_1 \dot{+} \cdots \dot{+} D_m$ and complete the proof as before.

The proof in case 2 is now complete.

CASE 3. $m \geq 2$, $j_m = j_{m-1} = 2$.

The proof in this case is similar to the proof in the previous case. Let δ_1 be any non-zero element of K and for $i = 2, 3, \ldots, m$ define δ_i by equations 4. For $i = 1, 2, \ldots, m - 1$ let γ_i be any element of K other than δ_i, δ_{i+1}, 0. Define γ'''_m by equation 6. Choose $x \in K$ such that

$$\gamma_{m-1} x \neq \delta_{m-1} \text{ or } \delta_m,$$

$$\gamma'''_m x^{-1} \neq \delta_m \text{ or } \delta_1.$$

Since K has six or more elements, a suitable x always exists. Let

$$\gamma'_{m-1} = \gamma_{m-1} x,$$

$$\gamma'_m = \gamma'''_m x^{-1}.$$

Construct the matrices $D_1, \ldots, D_{m-2}$ as in case 2. Let

$$D_{m-1} = \begin{pmatrix} \delta_{m-1} & d_2 \\ 0 & \gamma'_{m-1} \end{pmatrix},$$

$$D_m = \begin{pmatrix} \delta_m & d'_2 \\ 0 & \gamma'_m \end{pmatrix},$$

where d_2 is chosen such that the elementary divisors of $A_{m-1}D_{m-1}$ are $(\lambda - \delta_m)$, $(\lambda - \gamma'_{m-1})$, and d'_2 is chosen such that the elementary divisors of $A_m D_m$ are $(\lambda - \delta_1)$, $(\lambda - \gamma'_m)$. Set $D = D_1 \dot{+} \cdots \dot{+} D_{m-1} \dot{+} D_m$. Then $|D| = 1$ and, by the argument used in the previous case, $A = SDS^{-1}D^{-1}$ with $S, D \in SL(n, K)$.

CASE 4. $m \geqq 2$, $j_m = 2$, $j_{m-1} = \cdots = j_1 = 1$.

In this case, A is the direct sum of a diagonal matrix of order $n - 2$ and a matrix of order 2. Because of the fact that $C(p(\lambda)) \dot{+} C(q(\lambda))$ is similar to $C(p(\lambda)q(\lambda))$ if $p(\lambda)$ and $q(\lambda)$ are relatively prime, we may assume that $A = aI_{n-2} \dot{+} C(\lambda^2 - (a+b)\lambda + ab)$, where $a, b \in K$ and $a^{n-1}b = 1$. (Otherwise, after a similarity transformation of A, we could fall back on a matrix studied in cases 2 or 3.)

As far as possible, we shall use the technique of proof used in the previous cases. For non-zero $\delta \in K$ we define $c(\delta)$ as a function of δ by

$$a^{(n-1)(n-2)/2} \delta^{n-1} c(\delta) = 1. \tag{11}$$

We attempt to choose δ such that

$$c(\delta) \neq \delta, \tag{12}$$

$$c(\delta) \neq a^{n-2}\delta. \tag{13}$$

If K has infinitely many elements, then a suitable δ always exists since the equations $c(\delta) = \delta$, $c(\delta) = a^{n-2}\delta$ have only finitely many roots. Let

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

Then $|D| = 1$ and because of inequality 13, the elementary divisors of D are $(\lambda - \delta)$, $(\lambda - a\delta)$, ... , $(\lambda - a^{n-2}\delta)$, $(\lambda - c(\delta))$. Because of 12, we may choose d_2 such that the elementary divisors of

$$\begin{pmatrix} 0 & 1 \\ -ab & a+b \end{pmatrix} \begin{pmatrix} a^{n-2}\delta & d_2 \\ 0 & c(\delta) \end{pmatrix}$$

are $(\lambda - \delta)$, $(\lambda - c(\delta))$. The elementary divisors of AD are then $(\lambda - a\delta)$, ... , $(\lambda - a^{n-2}\delta)$, $(\lambda - \delta)$, $(\lambda - c(\delta))$. Hence $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, K)$.

If K is a finite field, it is not clear that a suitable δ exists in K. To handle this situation, it is necessary to give a more complicated argument. We divide our argument into three cases. Let p be the characteristic of K.

CASE 4.1. $b = a$.

Assume first that $a^2 \neq 1$. Take $\delta = 1$. Then if $c(1) = 1$ or if $c(1) = a^{n-2}$, we find $a^2 = 1$ (using $a^n = 1$). Hence A is a commutator of $SL(n, K)$ if $a^2 \neq 1$. If $a^2 = 1$, then $C(\lambda^2 - 2a\lambda + a^2) \in SL(2, K)$. By case 1, if K is not $GF(2)$ or

GF(3), $C(\lambda^2 - 2a\lambda + a^2)$ is a commutator of SL(2, K). Two cases now arise: $a = 1$ or $a = -1$. If $a = 1$, then (since (1) is a commutator of SL(1, K)), A is a direct sum of commutators, hence is a commutator of SL(n, K). If $a = -1$, then n is even. It is known [14] that integers x and y exist such that $x^2 + y^2 \equiv -1 \pmod{p}$. This means that elements x, y exist in GF(p) and hence in K (since K contains GF(p)) such that $x^2 + y^2 = -1$. Let

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}.$$

Then, $X, Y \in SL(2, K)$ and $-I_2 = XYX^{-1}Y^{-1}$. Hence A is again a direct sum of commutators of SL(2, K) and so is a commutator of SL(n, K).

CASE 4.2. $b \neq a$, n is odd.

Since $b \neq a$, A is similar to $A_1 = aI_{n-1} \dotplus (b)$. Let $D = (\delta_1) \dotplus (a\delta_1) \dotplus \cdots \dotplus (a^{n-1}\delta_1)$ where $\delta_1 = a^{-(n-1)/2}$. Then $|D| = 1$ and it easily follows that A_1 (and hence A) is a commutator of SL(n, K).

CASE 4.3. $b \neq a$, n is even.

Here we prove that an element δ exists in K such that 11, 12 and 13 hold*. If $\delta \in K$, we say δ is admissible if 11 and 12

---

* When n is even, the obvious device of passing to a diagonal matrix A_1 and attempting to find a diagonal matrix D such that A_1D is similar to D fails since fields exist within which $|D| = 1$ cannot be satisfied.

hold. We first show the existence of admissible elements in K by noting that if 1 is not admissible, then a must be. For if $c(1) = 1$, $c(a) = a$, we have

$$a^{(n-1)(n-2)/2} = 1 = a^{(n-1)(n-2)/2}\, a^n .$$

Hence $a^n = 1$ so that $a = b$ (since $a^{n-1}b = 1$). This contradiction establishes the existence of admissible elements in K. Suppose now that δ' is admissible. Four mutually exclusive possibilities exist:

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

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

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

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

If (iii) holds, then

$$a^{(n-1)(n-2)/2}(\delta')^n a^{n-2} = 1, \tag{14}$$

$$a^{(n-1)(n-2)/2} a^n (\delta')^n a^{n-2} = 1,$$

so that $a^n = 1$ and hence $b = a$. If (iv) holds, then we have 14 and

$$a^{(n-1)(n-2)/2}\, a^n\, (\delta')^n = 1.$$

Hence $a^n = a^{n-2}$, so that $a^2 = 1$. Since n is even and $a^{n-1}b = 1$, we deduce that $a = b$. Thus we must have (i) or (ii), which proves the existence of the desired element in K^*.

The proof for case 4 is now complete.

CASE 5. The scalar matrices.

We observe that if A is not scalar, we may, after a similarity transformation of A, study the representability of A as a commutator under one of the previous cases. Hence only scalar matrices remain to be considered. The argument about to be presented will not depend on the number of elements in the field K.

In this section a will denote a primitive n^{th} root of unity in K. Observe that n is determined by the roots of unity that exist in K and is in general not arbitrary. We shall first show that $aI_{mn} = aI_n \dotplus aI_n \dotplus \cdots \dotplus aI_n$ (m terms) is a commutator of $GL(mn, K)$ for every integer $m \geq 1$. Next, when n is odd, we shall establish that aI_{mn} is, in fact, a commutator of $SL(mn, K)$ for every integer $m \geq 1$. We shall then prove that when n is even, aI_{mn} is a commutator of $SL(mn, K)$ for every integer $m > 1$. Finally, we shall determine when aI_n (n even) is a commutator of $SL(n, K)$.

Let $D = (1) \dotplus (a) \dotplus (a^2) \dotplus \cdots \dotplus (a^{n-1})$. Then D and $aI_nD = aD$ have the same elementary divisors since $a^n = 1$. By Lemma 6, S exists in $SL(n, K)$ such that $aI_n = SDS^{-1}D^{-1}$. Since the direct sum of commutators is again a commutator, it follows that

---

* Since the multiplicative group $K - \{0\}$ of the field K is cyclic, it it possible to search for δ as a power ξ^x of the generating element ξ of $K - \{0\}$. In this context, 12 and 13 demand that x is not a solution of either of two congruences. When n is even, we may establish directly the existence of a suitable non-solution x. The above argument establishes this result in a slightly easier manner.

aI_{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 $aI_n = TDT^{-1}D^{-1}$.

When n is odd, $|D| = a^{n(n-1)/2} = 1$. Hence aI_{mn} is, in fact, a commutator of $SL(mn, K)$ for all $m \geq 1$.

When n is even, $|D| = -1$ since $a^{n/2} = -1$. We showed above that S exists in $SL(n, K)$ such that $aI_n = SDS^{-1}D^{-1}$. Applying this result to $(aI_n)^{-1}$, we deduce the existence of matrices U, V in $GL(n, K)$ such that $aI_n = UVU^{-1}V^{-1}$ and $|U| = -1$, $|V| = 1$. Now note that

$$aI_{2n} = (S \dotplus S)(D \dotplus D)(S \dotplus S)^{-1}(D \dotplus D)^{-1},$$

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

where $(S \dotplus S), (D \dotplus D) \in SL(2n, K)$ and $(S \dotplus U \dotplus T)$, $(D \dotplus V \dotplus D) \in SL(3n, K)$. By writing aI_{mn} as a direct sum of matrices aI_{2n} or aI_{3n}, we immediately see that aI_{mn} is a commutator of $SL(mn, K)$ whenever $m > 1$.

For the remainder of this section, we suppose that n is even, and, if possible, suppose that $aI_n = BCB^{-1}C^{-1}$ where $B, C \in SL(n, K)$. Then $aCB = BC$. It is well known that BC and CB have the same characteristic values. (Proof: $BC = B(CB)B^{-1}$.) Let α be a characteristic value of BC (in a suitable extension field of K, if necessary). Then α is a characteristic value of CB, so that $a\alpha$ is a characteristic value of aCB and hence of BC. Iterating this argument, we find that $\alpha, a\alpha, a^2\alpha, \ldots, a^{n-1}\alpha$ are all characteristic values of BC and, since a is a primitive n^{th} root of unity, are all the characteristic values of BC. Since $|BC| = 1$, $a^{n(n-1)/2}\alpha^n = 1$, or, since $a^{n/2} = -1$, $\alpha^n + 1 = 0$. This means that the characteristic polynomial of BC is $p(\lambda) = \lambda^n + 1$. Since BC is a non-derogatory matrix, a matrix S with coefficients in K exists such that $SBCS^{-1} = C(p(\lambda)) = Z$, say. Then $aI_n = SaI_nS^{-1} = ZYZ^{-1}Y^{-1}$ where $Y = SCS^{-1}$, $|Y| = |C|$. Thus, if aI_n is a commutator within $SL(n, K)$, then aI_n is the commutator of Z and another matrix Y. We shall now deduce the form of Y. Let $Y = (y^i_j)$ where the superscript indicates the row index. The equation $ZY = aYZ$ gives the following matrix equation.

$$\begin{pmatrix} y^2_1 & \cdots & y^2_n \\ \cdot & & \cdot \\ y^n_1 & \cdots & y^n_n \\ -y^1_1 & \cdots & -y^1_n \end{pmatrix} = \begin{pmatrix} -ay^1_n & ay^1_1 & \cdots & ay^1_{n-1} \\ -ay^2_n & ay^2_1 & \cdots & ay^2_{n-1} \\ \cdot & \cdot & \cdots & \cdot \\ -ay^n_n & ay^n_1 & \cdots & ay^n_{n-1} \end{pmatrix}.$$

Hence, for $2 \leq j \leq n$, we find

$$\begin{aligned} y^j_1 &= -ay^{j-1}_n, \\ y^{j-1}_n &= ay^{j-2}_{n-1}, \\ &\;\;\cdots \\ y^2_{n-j+3} &= ay^1_{n-j+2}, \end{aligned} \quad \Bigg\} \text{ absent when } j = 2$$

$$-y^1_{n-j+2} = ay^n_{n-j+1},$$

$$\begin{aligned} y^n_{n-j+1} &= ay^{n-1}_{n-j}, \\ &\;\;\cdots \\ y^{j+1}_2 &= ay^j_1. \end{aligned} \quad \Bigg\} \text{ absent when } j = n$$

Thus, beginning with the last of these equations, we find

$$\left.\begin{aligned} y_2^{j+1} &= a y_1^j, \\ &\cdots \\ y_{n-j}^{n-1} &= a^{n-1-j} y_1^j, \\ y_{n-j+1}^{n} &= a^{n-j} y_1^j, \end{aligned}\right\} \text{absent when } j = n$$

$$y_{n-j+2}^{1} = -a^{n-j+1} y_1^j,$$

$$\left.\begin{aligned} y_{n-j+3}^{2} &= -a^{n-j+2} y_1^j, \\ &\cdots \\ y_n^{j-1} &= -a^{n-1} y_1^j. \end{aligned}\right\} \text{absent when } j = 2$$

Similarly,

$$\begin{aligned} y_1^1 &= a y_n^n, \\ y_n^n &= a y_{n-1}^{n-1}, \\ &\cdots \\ y_2^2 &= a y_1^1, \end{aligned}$$

so that

$$y_{1+i}^{1+i} = a^i y_1^1; \quad i = 1, 2, \ldots, n-1.$$

Hence,

$Y =$

$$\begin{pmatrix}
y_1^1 & -ay_1^n & & & & & & -a^{n-j+1}y_1^j & & -a^{n-1}y_1^2 \\
\cdot & ay_1^1 & & & & & & \cdot & & \\
\cdot & & a^2y_1^1 & & & & & & \cdot & \\
\cdot & & & \cdot & & & & & & -a^{n-1}y_1^j \\
y_1^j & & & & \cdot & & & & & \\
 & ay_1^j & & & & \cdot & & & & \\
\cdot & & a^2y_1^j & & & & a^{n-j}y_1^1 & & & \\
\cdot & & & \cdot & & & & a^{n-j+1}y_1^1 & & \\
\cdot & & & & \cdot & & & \cdot & & \\
 & & & & & \cdot & & & \cdot & -a^{n-1}y_1^n \\
y_1^n & & & & & & a^{n-j}y_1^j & & & a^{n-1}y_1^1
\end{pmatrix} .$$

Conversely, for this Y, $aYZ = ZY$.

In order to simplify the notation in a computation that will be presently made, we set $y_1^i = y_i$, $i = 1, \ldots, n$. Let

$$Y_1 = \begin{pmatrix}
y_1 & -y_n & -y_{n-1} & \cdots & -y_2 \\
y_2 & y_1 & -y_n & \cdots & -y_3 \\
y_3 & y_2 & y_1 & \cdots & -y_4 \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
y_{n-1} & y_{n-2} & y_{n-3} & \cdots & -y_n \\
y_n & y_{n-1} & y_{n-2} & \cdots & y_1
\end{pmatrix} .$$

Then $|Y| = a^{n(n-1)/2}\,|Y_1|$ and hence (since $a^{n(n-1)/2} = -1$) the necessary and sufficient condition that aI_n be a commutator within $SL(n, K)$ is that field elements $y_1, \ldots, y_n$ exist such that $|Y_1| = -1$.

In order to investigate the values that $|Y_1|$ can assume, we require a known [15] formula for $|Y_1|$. For completeness, we include a derivation of this formula. Let ω be a primitive $(2n)^{th}$ root of unity in a suitable extension field of K, $\omega^2 = a$. Set $\omega_i = a^i\omega$, $i = 1, \ldots, n$. Let $v_i = (1, \omega_i, \omega_i^2, \ldots, \omega_i^{n-1})$. Then it is easy to see that $v_i Y_1 = (y_1 + \omega_i y_2 + \cdots + \omega_i^{n-1} y_n) v_i$, $i = 1, \ldots, n$. Thus v_i is a characteristic vector of Y_1 belonging to the characteristic value $y_1 + \omega_i y_2 + \cdots + \omega_i^{n-1} y_n$. Now, the nxn matrix with $v_1, \ldots, v_n$ as its rows is non-singular since it is a Vandermonde matrix and the ω_i are distinct. Consequently we have found all of the characteristic values of Y_1. Hence we obtain the known expression

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

Thus

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

$$= \prod_{i=1}^{n} \left(\sum_{j=1}^{n/2} a^{i(2j-2)} \omega^{2j-2} y_{2j-1} + \sum_{j=1}^{n/2} a^{i(2j-1)} \omega^{2j-1} y_{2j} \right)$$

$$= \prod_{i=1}^{n} \left(\sum_{j=1}^{n/2} a^{i(2j-2)} a^{j-1} y_{2j-1} + \sum_{j=1}^{n/2} a^{i(2j-1)} a^{j} \omega^{-1} y_{2j} \right)$$

$$= \prod_{i=1}^{n} \left(\sum_{j=1}^{n/2} a^{(j-1)(2i+1)} y_{2j-1} + \omega^{-1} \sum_{j=1}^{n/2} a^{j(2i+1)-i} y_{2j} \right).$$

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

$$|Y_1| = \prod_{i=1}^{n/2} \Big[\Big(\sum_{j=1}^{n/2} a^{(j-1)(2i+1)} y_{2j-1} + \omega^{-1} a^{-i} \sum_{j=1}^{n/2} a^{j(2i+1)} y_{2j} \Big)$$

$$\cdot \Big(\sum_{j=1}^{n/2} a^{(j-1)(2i+1)} y_{2j-1} - \omega^{-1} a^{-i} \sum_{j=1}^{n/2} a^{j(2i+1)} y_{2j} \Big) \Big]$$

$$= \prod_{i=1}^{n/2} \Big[\Big(\sum_{j=1}^{n/2} a^{(j-1)(2i+1)} y_{2j-1} \Big)^2 - a^{-1} a^{-2i} \Big(\sum_{j=1}^{n/2} a^{j(2i+1)} y_{2j} \Big)^2 \Big] .$$

Consider the following set of $n/2$ equations in $n/2$ unknowns:

$$\sum_{j=1}^{n/2} a^{(j-1)(2i+1)} y_{2j-1} = w_i, \quad i = 1, 2, \ldots, n/2, \tag{15}$$

where $w_1, \ldots, w_{n/2} \in K$. The matrix of coefficients of 15 is

$$\begin{pmatrix} 1 & a^3 & (a^3)^2 & \ldots & (a^3)^{(n/2-1)} \\ 1 & a^5 & (a^5)^2 & \ldots & (a^5)^{(n/2-1)} \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ 1 & a^{n+1} & (a^{n+1})^2 & \ldots & (a^{n+1})^{(n/2-1)} \end{pmatrix} .$$

This is a Vandermonde matrix and is non-singular since $a^3, a^5, \ldots, a^{n+1} = a$ are distinct. Hence, for any choice of $w_1, \ldots, w_{n/2}$ in K, $y_1, y_3, \ldots, y_{n-1}$ can be found in K such that 15 is satisfied. Similarly the set of $n/2$ equations in $n/2$ unknowns

$$\sum_{j=1}^{n/2} a^{j(2i+1)} y_{2j} = a^i w_i', \quad i = 1, 2, \ldots, n/2,$$

has a non-singular coefficient matrix so that a solution exists in K for every choice of $w_1', \ldots, w_{n/2}'$ in K.

Thus, in order to set $|Y_1| = -1$, it is necessary and sufficient that $w_1, \ldots, w_{n/2}, w'_1, \ldots, w'_{n/2}$ be found in K such that

$$-1 = \prod_{i=1}^{n/2} \left(w_i^2 - a^{-1}(w'_i)^2 \right). \tag{16}$$

If $n = 4m$, take $w_1 = a^m$, $w_2 = \cdots = w_{n/2} = 1$, $w'_1 = \cdots = w'_{n/2} = 0$. Then, since $a^{n/2} = -1$, equation 16 is satisfied.

If $n = 4m + 2$, then from $a^{2m+1} = -1$ we obtain $-a^{-1} = a^{2m}$. Set $w''_i = w'_i a^m$. Then

$$|Y_1| = \prod_{i=1}^{n/2} \left(w_i^2 + (w''_i)^2 \right). \tag{17}$$

Since the product of sums of two squares is again a sum of two squares [16], if elements in K exist such that $|Y_1| = -1$, then for certain elements W, W' of K we have

$$-1 = W^2 + (W')^2. \tag{18}$$

Conversely, if 18 has a solution in K, then if in 17 we take $w_1 = W$, $w''_1 = W'$, $w_2 = \cdots = w_{n/2} = 1$, $w''_2 = \cdots = w''_{n/2} = 0$, we find that $|Y_1| = -1$. Hence we have reached the following conclusion: If $n = 4m + 2$, then the necessary and sufficient condition that $aI_n = ZYZ^{-1}Y^{-1}$ where $Z, Y \in SL(n, K)$ is that equation 18 have a solution in K.

It is known that integers x and y always exist such that $x^2 + y^2 + 1 \equiv 0 \pmod p$ where p is a prime. This means that elements x and y always exist within GF(p) such that $x^2 + y^2 = -1$. Hence, since any field of characteristic p contains GF(p), 18 always has a solution if K has characteristic p.

If K has characteristic 0, then 18 is sometimes impossible. As an example we may take $n = 2$, $a = -1$, and K to be any formally real field (such as the field of rational numbers). In many other cases 18 possesses a solution. We list two such cases.

(1). If K contains the primitive $(2n)^{th}$ root of unity ω, then a solution of 18 is $W = \omega^{n/2}$, $W' = 0$.

(2). If for some divisor r of $2m + 1$ integers s and h exist such that $r(h + 1) = 2^s + 1$, then a solution of 18 can be found. For, using a technique due to Landau [17], we first note the following polynomial identity:

$$(1 + \lambda + \lambda^2 + \cdots + \lambda^{r-1})(1 + \lambda^r + (\lambda^r)^2 + \cdots + (\lambda^r)^h)$$

$$= 1 + \lambda + \lambda^2 + \cdots + \lambda^{rh+r-1}$$

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

Since $2r$ divides n, the field K contains the primitive root of unity $\rho = a^{n/2r}$ of order $2r$, so that $\rho^r + 1 = 0$ and hence

$$\rho^{r-1} - \rho^{r-2} + - \cdots - \rho + 1 = 0.$$

Using $-1 = \rho^r$, we obtain

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

Hence, if we take $\lambda = \rho^2$, 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 $8k + 3$ or $8k - 3$, 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 [18] that -1 is a sum of four squares in the field of the n^{th} roots of unity over the rationals, $n > 2$. Whether or not -1 is a sum of two squares in such fields remains to be determined.

6. THE PROOF OF THEOREM 1 WHEN $K = GF(5)$.

The field $GF(5)$ consists of the elements $0, 1, 2, 3, 4$. In order to prove Theorem 1 when $K = GF(5)$, we assume as before that $A = A_1 \dotplus \cdots \dotplus A_m$, where A_i is the $j_i x j_i$ companion matrix of a polynomial with coefficients in K, and $|A_1 \cdots A_m| = 1$. We make the additional assumption that $|A_{i_1} \cdots A_{i_k}| \neq 1$ if the subset $\{i_1, \ldots, i_k\}$ of $\{1, \ldots, m\}$ is proper. This additional assumption, which involves no loss of generality, serves to restrict the values that m can assume. We divide our discussion into a number of cases depending on the value of m.

If $m = 1$, the result is clear if $n = j_1 = 1$ and if $n = 2$ (by cases 1 and 5 of the proof in the preceding section.) If $n \geq 3$, choose $\rho \in K$ such that $\rho^2 \neq 1, 0$. Let D be a standard matrix with $d_1 = \rho$, $r = 2$, $s_1 = 1$, $s_2 = n - 2$, $c_1 = \rho^{-1}$, $c_2 = 1$. Then the elementary divisors of D are $(\lambda - \rho)$, $(\lambda - \rho^{-1})$, $(\lambda - 1)^{n-2}$. Choose $d_2, \ldots, d_n$ such that these are the elementary divisors of AD. Then, by Lemma 6, $A = SDS^{-1}D^{-1}$ with $S, D \in SL(n, K)$. This part of the proof also works if $K = GF(4)$.

If $m = 2$, then, after a rearrangement of the A_i if necessary, the two-tuple $(|A_1|, |A_2|)$ must be $(4, 4)$ or $(2, 3)$. In order to use the method of proof given previously, we select an element $\delta_1 \in GF(5)$, then choose $c_1^{(1)}, \ldots, c_{j_1-1}^{(1)}, c_1^{(2)}, \ldots, c_{j_2-1}^{(2)} \in GF(5)$ and distinct from δ_1, $\delta_2 = |A_1| \delta_1$ such that

$$\delta_1 \delta_2 c_1^{(1)} \cdots c_{j_1-1}^{(1)} c_1^{(2)} \cdots c_{j_2-1}^{(2)} = 1.$$

If we are able to do this, we construct matrices D_i $(i = 1, 2)$ in a manner analogous to the constructions in the previous section. Thus, if γ_1 and γ_2 are so chosen that $\gamma_1, \gamma_2, \delta_1, \delta_2$ are the four non-zero elements of $GF(5)$, we suppose e_1 of the elements $c_1^{(1)}, \ldots, c_{j_1-1}^{(1)}$ and e_2 of the elements $c_1^{(2)}, \ldots, c_{j_2-1}^{(2)}$ are γ_1 and the remaining elements are γ_2. For $i = 1, 2$ we let D_i be a $j_i x j_i$ standard matrix with $d_1 = \delta_i$ and elementary divisors $(\lambda - \delta_i), (\lambda - \gamma_1)^{e_i}, (\lambda - \gamma_2)^{j_i-1-e_i}$ such that the elementary divisors of A_iD_i are $(\lambda - \delta_{i+1}), (\lambda - \gamma_1)^{e_i}, (\lambda - \gamma_2)^{j_i-1-e_i}$ (where $\delta_3 = \delta_1$). Lemmas 1 and 5 guarantee the existence of D_i. Setting $D = D_1 \dotplus D_2$, we find $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(5))$ in the usual way.

If $(|A_1|, |A_2|) = (4, 4)$, then we wish the $c_k^{(i)}$ to be distinct from $\delta_1, 4\delta_1$. Thus we may suppose that exactly e of the $c_k^{(i)}$ are equal to $2\delta_1$ and the remaining $n - 2 - e$ are equal to $3\delta_1$. Thus it suffices to find an field element δ_1 and an integer e with $0 \leqq e \leqq n - 2$ such that

$$4\delta_1^n \, 2^e 3^{n-2-e} = 1,$$

or,

$$\delta_1^n \, 3^{n+2e} = 1.$$

Take $\delta_1 = 3$ and choose e (= 0 or 1) so that $2n + 2e \equiv 0 \pmod 4$.

If $(|A_1|, |A_2|) = (2, 3)$, we wish to select $n - 2$ elements $c_k^{(i)}$ different from δ_1 or $2\delta_1$. Thus, if e of the $c_k^{(i)}$ are $3\delta_1$ and $n - 2 - e$ are $4\delta_1$, it suffices to find a field element δ_1 and an integer e with $0 \leqq e \leqq n - 2$ such that

$$2\delta_1^n \; 3^e \; 4^{n-2-e} = 1,$$

or,

$$\delta_1^n \; 2^{1+2n+e} = 1.$$

If $n - 2 \geq 3$, take $\delta_1 = 1$ and e (= 1 or 3) such that $1 + 2n + e \equiv 0 \pmod 4$. If $n = 3$, take $\delta_1 = 1$, $e = 1$. If $n = 2$, the matrix $A = (2) \dotplus (3)$ and is similar to $C((\lambda - 2)(\lambda - 3))$ which has already been studied under the case $m = 1$. The case $n = 4$ requires special treatment.

If $j_1 = 1$, $j_2 = 3$, then $A = A_1 \dotplus A_2$ where $A_1 = (2)$, $A_2 = C(\lambda^3 - \alpha\lambda^2 - \beta\lambda - 3)$. Let

$$D_2 = \begin{pmatrix} 4 & d_2 & d_3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

where (Lemma 4) d_2 and d_3 are so chosen that the characteristic and minimum polynomial of A_2D_2 is $(\lambda - 4)(\lambda - 1)(\lambda - 2)$. Let $D = (1) \dotplus D_2$. Then the elementary divisors of D are $(\lambda - 1)$, $(\lambda - 4)$, $(\lambda - 2)$, $(\lambda - 2)$ and those of AD are $(\lambda - 2)$, $(\lambda - 4)$, $(\lambda - 1)$, $(\lambda - 2)$. Hence $A = SDS^{-1}D^{-1}$, where $S, D \in SL(4, GF(5))$. If $j_1 = 3$, $j_2 = 1$, the result follows immediately from the observation that the inverse of the matrix corresponding to this case is similar to $(2) \dotplus C(\lambda^3 - \alpha\lambda^2 - \beta\lambda - 3)$. If $j_1 = j_2 = 2$, then, if the characteristic polynomials of A_1 and A_2 are relatively prime, the result is immediate since A is similar to the companion matrix of a single polynomial, for which the result is already known (case $m = 1$). This will be the case if either of the characteristic polynomials is irreducible. (Equal characteristic polynomials are impossible since $|A_1| = 2$, $|A_2| = 3$.) 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 $r = 3/r$ are impossible in $GF(5)$, it follows that A is similar to a diagonal matrix. We may suppose this diagonal matrix to be $I_2 \dotplus (2) \dotplus (3)$ if $r = 1$; $(1) \dotplus (2) \dotplus (2) \dotplus (4)$ if $r = 2$; $(1) \dotplus (3) \dotplus (3) \dotplus (4)$ if $r = 3$; $(4) \dotplus (4) \dotplus (2) \dotplus (3)$ if $r = 4$. Since the matrices $(2) \dotplus (3)$ and $(4) \dotplus (4)$ have been discussed above, the proof is complete when $r = 1$ or 4. To complete the proof when $r = 2$, we note that it suffices to consider $(2) \dotplus C((\lambda - 2)(\lambda - 4))$ to which the discussion above applies. Similarly we complete the proof when $r = 3$.

The proof for the case $m = 2$ is now complete.

We now consider the case $m = 3$. Here $(|A_1|, |A_2|, |A_3|)$ is $(2, 2, 4)$ or $(3, 3, 4)$. By passing to A^{-1} if necessary, it suffices to consider the first of these possibilities. Let δ_1 $(\neq 0) \in GF(5)$ and let $\delta_2 = 2\delta_1$, $\delta_3 = 2\delta_2 = 4\delta_1$. We wish to select $c_1^{(i)}, \ldots, c_{j_i-1}^{(i)}$, $(i = 1, 2, 3) \in GF(5)$ such that for $i = 1, 2, 3$, we have $c_k^{(i)} \neq \delta_i$ or δ_{i+1} for $k = 1, 2, \ldots, j_i - 1$ (where $\delta_4 = \delta_1$). Hence, for $i = 1, 2, 3$, we wish to find integers e_i with $0 \leq e_i \leq j_i - 1$ such that

(i). $c_1^{(1)} = \cdots = c_{e_1}^{(1)} = 4\delta_1$, $c_{e_1+1}^{(1)} = \cdots = c_{j_1-1}^{(1)} = 3\delta_1$;

(ii). $c_1^{(2)} = \cdots = c_{e_2}^{(2)} = 3\delta_1$, $c_{e_2+1}^{(2)} = \cdots = c_{j_2-1}^{(2)} = \delta_1$;

(iii). $c_1^{(3)} = \cdots = c_{e_3}^{(3)} = 2\delta_1$, $c_{e_3+1}^{(3)} = \cdots = c_{j_3-1}^{(3)} = 3\delta_1$;

and

$$\delta_1 \delta_2 \delta_3 \prod_{i,k} c_k^{(i)} = 1.$$

If this can be accomplished, we construct matrices D_i $(i = 1, 2, 3)$ as in the previous cases. Thus D_i is a $j_i x j_i$ standard matrix with $d_1 = \delta_i$ and elementary divisors $(\lambda - \delta_i)$, $(\lambda - c_1^{(i)})^{e_i}$, $(\lambda - c_{e_i+1}^{(i)})^{j_i - 1 - e_i}$ such that the elementary divisors of $A_i D_i$ are $(\lambda - \delta_{i+1})$, $(\lambda - c_1^{(i)})^{e_i}$, $(\lambda - c_{e_i+1}^{(i)})^{j_i - 1 - e_i}$, where $\delta_4 = \delta_1$. (Lemmas 1 and 5.) If $D = D_1 \dotplus D_2 \dotplus D_3$, then $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(5))$ in the usual way.

Thus, it suffices to find a field element δ_1 and integers e_1, e_2, e_3 with $0 \leqq e_i \leqq j_i - 1$ such that

$$3\delta_1^n \, 4^{e_1} \, 3^{j_1 - 1 - e_1} \, 3^{e_2} \, 2^{e_3} \, 3^{j_3 - 1 - e_3} = 1,$$

or,

$$\delta_1^n \, 3^{3 + j_1 + j_3 + e_1 + e_2 + 2e_3} = 1. \qquad (19)$$

In the following table we give suitable values for δ_1, e_1, e_2, e_3 as functions of j_1, j_2, j_3. We may suppose the notation so chosen that $j_1 \geqq j_2$. Note that $e_1 + e_2$ can assume any of the integers $0, 1, \ldots, j_1 + j_2 - 2$, so that if $j_1 + j_2 \geqq 5$, we may find e_1 and e_2 such that $e_1 + e_2$ is congruent to any of 0, 1, 2, or 3 (mod 4). The right hand column of the table gives equations which verify equation 19 or equations from which the e_i can be computed so as to satisfy equation 19. All congruences are modulo 4. Those entries in the table which are not specified may assume arbitrary values (to the extent permitted by the other entries in the table).

Table.

j_1+j_2	j_1	j_2	j_3	δ_1	e_1	e_2	e_3	Equation
$\geqq 5$				1	$< j_1$	$< j_2$	0	$e_1+e_2+3+j_1+j_3 \equiv 0$
4	3 or 2	1 or 2	$\geqq 2$	1	$< j_1$	$< j_2$	$< j_3$	$e_1+e_2+2e_3+3+j_1+j_3 \equiv 0$
4	3 or 2	1 or 2	1	1	$< j_1$	$< j_2$	0	$e_1+e_2+j_1 \equiv 0$
3	2	1	$\geqq 2$	1	0, 1	0	0, 1	$e_1+2e_3+3+j_1+j_3 \equiv 0$
3	2	1	1					Exceptional case
2	1	1	even	1	0	0	$j_3/2$	$2e_3+3+j_1+j_3 = 4+4(j_3/2) \equiv 0$
2	1	1	odd	3	0	0	0	$n+3+j_1+j_3 = 8+4((j_3-1)/2) \equiv 0$

In the exceptional case in this table (in which $\delta_1^n = 1$ for each $\delta \neq 0$) we have $A = C(\lambda^2 - \alpha\lambda + 2) \dotplus (2) \dotplus (4)$. By a similarity transformation we may pass to $C(\lambda^2 - \alpha\lambda + 2) \dotplus C((\lambda - 2)(\lambda - 4))$ which has already been studied under $m = 2$.

Finally, we arrive at the last case of $m = 4$. Here $(|A_1|, |A_2|, |A_3|, |A_4|)$ is $(2, 2, 2, 2)$ or $(3, 3, 3, 3)$. By passing to A^{-1} if necessary it suffices to consider the first of these possibilities. Here, if $\delta_1 \in GF(5)$, we let $\delta_2 = 2\delta_1$, $\delta_3 = 2\delta_2 = 4\delta_1$, $\delta_4 = 2\delta_3 = 3\delta_1$. For $i = 1, 2, 3, 4$ we wish to find field elements $c_1^{(i)}, \ldots, c_{j_i-1}^{(i)}$ distinct from δ_i, δ_{i+1} (where $\delta_5 = \delta_1$) and hence we wish to find integers e_i with $0 \leqq e_i \leqq j_i - 1$ such that

(i) $c_1^{(1)} = \cdots = c_{e_1}^{(1)} = 3\delta_1$, $c_{e_1+1}^{(1)} = \cdots = c_{j_1-1}^{(1)} = 4\delta_1$;

(ii) $c_1^{(2)} = \cdots = c_{e_2}^{(2)} = 3\,\delta_1, \quad c_{e_2+1}^{(2)} = \cdots = c_{j_2-1}^{(2)} = \delta_1;$

(iii) $c_1^{(3)} = \cdots = c_{e_3}^{(3)} = 2\,\delta_1, \quad c_{e_3+1}^{(3)} = \cdots = c_{j_3-1}^{(3)} = \delta_1;$

(iv) $c_1^{(4)} = \cdots = c_{e_4}^{(4)} = 2\,\delta_1, \quad c_{e_4+1}^{(4)} = \cdots = c_{j_4-1}^{(4)} = 4\,\delta_1.$

For $i = 1, 2, 3, 4$ we construct a $j_i x j_i$ standard matrix D_i with $d_1 = \delta_i$ and elementary divisors $(\lambda - \delta_i)$, $(\lambda - c_1^{(i)})^{e_i}$, $(\lambda - c_{e_i+1}^{(i)})^{j_i-1-e_i}$ such that the elementary divisors of A_iD_i are $(\lambda - \delta_{i+1})$, $(\lambda - c_1^{(i)})^{e_i}$, $(\lambda - c_{e_i+1}^{(i)})^{j_i-1-e_i}$. (Lemmas 1 and 5.) If we set $D = D_1 \dotplus D_2 \dotplus D_3 \dotplus D_4$ we then find that $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(5))$ provided that

$$4\,\delta_1^n \prod_{i,k} c_k^{(i)} = 1.$$

Thus it suffices to find integers e_i and a field element δ_1 such that

$$4\,\delta_1^n\, 3^{e_1}\, 4^{j_1-1-e_1}\, 3^{e_2}\, 2^{e_3}\, 2^{e_4}\, 4^{j_4-1-e_4} = 1,$$

or,

$$\delta_1^n\, 3^{2(1+j_1+j_4)-e_1+e_2+3e_3+e_4} = 1.$$

If two of the j_i are greater than one, suppose $j_2 > 1$, $j_4 > 1$ and take $\delta_1 = 1$, $e_1 = e_3 = 0$, and e_2 and 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 $j_1 = j_2 = j_4 = 1$, $j_3 \geqq 2$. Let $\delta_1 = 3^k$, then choose $k = 0, 1, 2,$ or 3 and e_3, $0 \leqq e_3 \leqq j_3 - 1$, such that (using $n = j_1 + j_2 + j_3 + j_4$)

$$3^{k(3+j_3) + 3e_3 + 6} = 1,$$

or,

$$k(3 + j_3) + 3e_3 + 6 \equiv 0 \pmod{4}.$$

If $j_3 \equiv 0 \pmod{4}$, take $k = 2$, $e_3 = 0$. If $j_3 \equiv 2 \pmod{4}$, take $k = 2$, $e_3 = 0$. If $j_3 \equiv 3 \pmod{4}$ take $k = 1$, $e_3 = 0$. If $j_3 \equiv 1 \pmod{4}$ but $j_3 \neq 1$, take $k = 0$, $e_3 = 2$. If $j_1 = j_2 = j_3 = j_4 = 1$, then A is scalar and we may appeal to the results of case 5, section 5.

7. THE PROOF OF THEOREM 1 WHEN $K = GF(4)$

The field $GF(4)$ consists of the elements $0, 1, \theta, \theta + 1$ with $\theta^2 = \theta + 1$. As in the previous section, we assume that $A = A_1 \dotplus \cdots \dotplus A_m$ where A_i is the $j_i x j_i$ companion matrix of a polynomial over $GF(4)$, and $|A_{i_1} \cdots A_{i_k}| = 1$ (where $1 \leq i_1 < \cdots < i_k \leq m$) if, and only if, $k = m$. The possible m-tuples $(|A_1|, \ldots, |A_m|)$ with this condition are (1), $(\theta, \theta + 1)$, $(\theta + 1, \theta)$, (θ, θ, θ), $(\theta + 1, \theta + 1, \theta + 1)$. If $m = 1$, the proof in the previous section applies here also. If $m = 2$ and $(|A_1|, |A_2|) = (\theta, \theta + 1)$ let D_1 be a $j_1 x j_1$ standard matrix with $d_1 = \theta$ and elementary divisors $(\lambda - \theta), (\lambda - 1)^{j_1 - 1}$ such that the elementary divisors of $A_1 D_1$ are $(\lambda - \theta^2), (\lambda - 1)^{j_1 - 1}$. Let D_2 be a $j_2 x j_2$ standard matrix with $d_1 = \theta^2$ and elementary divisors $(\lambda - \theta^2), (\lambda - 1)^{j_2 - 1}$ such that the elementary divisors of $A_2 D_2$ are $(\lambda - \theta), (\lambda - 1)^{j_2 - 1}$. These constructions are possible by Lemmas 1 and 5. Set $D = D_1 \dotplus D_2$. Then $|D| = (\theta)^3 = 1$, and by the usual argument, $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(4))$. By appealing to the automorphism σ for which $\sigma(\theta) = \theta + 1$, the other $m = 2$ case automatically follows.

We now consider the case $m = 3$. Owing to the existence of σ, it is enough to assume that $(|A_1|, |A_2|, |A_3|)$ is (θ, θ, θ). First suppose that $j_1 \equiv j_3 \pmod 3$. Appealing as usual to Lemmas 1 and 5, we let D_1 be a $j_1 x j_1$ standard matrix with $d_1 = 1$ and elementary divisors $(\lambda - 1), (\lambda - \theta^2)^{j_1 - 1}$ such that the elementary divisors of $A_1 D_1$ are $(\lambda - \theta), (\lambda - \theta^2)^{j_1 - 1}$. Let D_2 be a $j_2 x j_2$ standard matrix with $d_1 = \theta$ and elementary divisors $(\lambda - \theta)$,

$(\lambda - 1)^{j_2 - 1}$ such that the elementary divisors of A_2D_2 are $(\lambda - \theta^2)$, $(\lambda - 1)^{j_2 - 1}$. Let D_3 be a $j_3 x j_3$ standard matrix 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 $|D| = 1$ since $2j_1 + j_3 \equiv 0 \pmod 3$. Hence $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(4))$.

Next, suppose that $j_i \not\equiv j_k \pmod 3$ for any pair i, k. Choose the notation so that $j_1 \equiv 2 \pmod 3$, $j_2 \equiv 0 \pmod 3$, $j_3 \equiv 1 \pmod 3$. Since $j_1 \geqq 1$, $j_2 \geqq 1$, it follows from these congruences that $j_1 \geqq 2$, $j_2 \geqq 2$. Let D_1 be a $j_1 x j_1$ standard matrix with $d_1 = 1$ and elementary divisors $(\lambda - 1)$, $(\lambda - \theta)$, $(\lambda - \theta)^{j_1 - 2}$ (see Lemma 1) such that the elementary divisors of A_1D_1 are $(\lambda - 1)$, $(\lambda - \theta^2)$, $(\lambda - \theta)^{j_1 - 2}$ (see Lemma 4). Invoking Lemmas 1 and 4 again we let D_2 be a $j_2 x j_2$ standard matrix with $d_1 = \theta$ and elementary divisors $(\lambda - \theta)$, $(\lambda - \theta^2)$, $(\lambda - \theta^2)^{j_2 - 2}$ such that $(\lambda - \theta)$, $(\lambda - 1)$, $(\lambda - \theta^2)^{j_2 - 2}$ are the elementary divisors of A_2D_2. By Lemmas 1 and 5, we may construct a $j_3 x j_3$ standard matrix D_3 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 $(\lambda - \theta)$, $(\lambda - \theta^2)^{j_3 - 1}$. We set $D = D_1 \dotplus D_2 \dotplus D_3$. Then $|D| = 1$ since $j_1 + 2j_2 + 2j_3 - 4$ is congruent to $0 \pmod 3$. Hence $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(4))$.

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

8. THE PROOF OF THEOREM 1 WHEN $K = GF(3)$

The field $GF(3)$ consists of the elements $-1, 1, 0$. Let $p_1(\lambda) = \lambda^2 + 1$, $p_2(\lambda) = \lambda^2 + \lambda + 1 = (\lambda - 1)^2$, $p_3(\lambda) = \lambda^2 - \lambda + 1 = (\lambda + 1)^2$. These are the only monic polynomials of degree two over $GF(3)$ with one as constant term. Let $C_i = C(p_i(\lambda))$; $i = 1, 2, 3$. We list here a set of five lemmas which we shall discuss below. Let $A \in SL(n, GF(3))$.

LEMMA 8. If A is the companion matrix of a polynomial, but not C_1, C_2, or C_3, then $A = CDC^{-1}D^{-1}$, where $C, D \in SL(n, GF(3))$.

LEMMA 9. If A is the companion matrix of a polynomial, then $A = CDC^{-1}D^{-1}$, where $C, D \in GL(n, GF(3))$ and $|C| = -|D| = 1$.

LEMMA 10. If $A = A_1 \dotplus A_2$ where A_i is the companion matrix of a power of an irreducible polynomial over $GF(3)$ and $|A_i| = -1$, $i = 1, 2$, then $A = CDC^{-1}D^{-1}$, where $C, D \in SL(n, GF(3))$.

LEMMA 11. Under the hypotheses of Lemma 10, $A = CDC^{-1}D^{-1}$, where $C, D \in GL(n, GF(3))$ and $|C| = -|D| = 1$.

LEMMA 12. $C_i \dotplus C_i = S_iD_iS_i^{-1}D_i^{-1}$, where $|S_i| = 1$, $|D_i| = -1$ and $S_i, D_i \in GL(n, GF(3))$; $i = 1, 2, 3$.

In the following discussion we shall show that the validity of these five lemmas is enough to establish Theorem 1 when $K = GF(3)$.

Let $A \in SL(n, GF(3))$. Throw A into rational canonical form and assume that A is the direct sum of companion matrices of powers of polynomials irreducible over $GF(3)$, so arranged that $A = \mathcal{A}_1 \dotplus \cdots \dotplus \mathcal{A}_m$, where either $|\mathcal{A}_i| = 1$ and $\mathcal{A}_i$ is the companion matrix of a power of an irreducible polynomial, or $\mathcal{A}_i = \mathcal{A}_i^{(1)} \dotplus \mathcal{A}_i^{(2)}$ where $|\mathcal{A}_i^{(1)}| = |\mathcal{A}_i^{(2)}| = -1$ and $\mathcal{A}_i^{(1)}$, $\mathcal{A}_i^{(2)}$ are companion matrices of powers of irreducible polynomials. To prove Theorem 1 when $n = 2$, we note that if A is not scalar, then A is similar to the companion matrix of a single polynomial so that Lemma 9 establishes the desired result. If A is scalar, then the discussion in section 5, case 5 provides the result. To prove Theorem 1 when $n \neq 2$, we note that the result is immediate when $n = 1$. If $n \geqq 3$, Theorem 1 immediately follows from Lemmas 8 and 10 if C_1, C_2, C_3 do not appear among the direct summands of A, since the direct sum of commutators is again a commutator. If C_1, C_2, C_3 are some of the direct summands of A, but an $\mathcal{A}_i$ exists which is not C_1, C_2, or C_3, then we obtain the claimed result by applying Lemmas 9 or 11 to the other direct summands of A, and Lemmas 8, 9, 10, or 11 to this particular $\mathcal{A}_i$. If C_1, C_2, C_3 are the only matrices which constitute the $\mathcal{A}_i$ then, if m is even, Lemma 9 suffices. If m is odd and two different C_i appear (for example, C_1 and C_2) then, since $C_1 \dotplus C_2$ is similar to $C(p_1(\lambda)p_2(\lambda))$, we may apply Lemma 9 to the C_i and to $C(p_1(\lambda)p_2(\lambda))$. Finally, if $A = C_i \dotplus \cdots \dotplus C_i$ for some i with m odd, an appeal to Lemmas 12 and 9 completes the proof.

If $d = (d_3, d_4, \ldots, d_n)$ define the nxn matrix $\mathfrak{D}_n(g_1, g_2, g_3, g_4, d)$ in the following way.

$$\mathfrak{D}_n(g_1, g_2, g_3, g_4, d) = \begin{pmatrix} g_1 & g_2 & d_3 & d_4 & \cdots & & d_n \\ & g_3 & g_4 & 0 & \cdots & & 0 \\ & & 1 & 1 & \cdot & & \cdot \\ & & & \cdot & \cdot & \cdot & \cdot \\ & O & & & \cdot & \cdot & 0 \\ & & & & & 1 & 1 \\ & & & & & & 1 \end{pmatrix}, \quad n \geqq 4;$$

$$\mathfrak{D}_3(g_1, g_2, g_3, g_4, d) = \begin{pmatrix} g_1 & g_2 & d_3 \\ 0 & g_3 & g_4 \\ 0 & 0 & 1 \end{pmatrix};$$

$$\mathfrak{D}_2(g_1, g_2, g_3, g_4, d) = \begin{pmatrix} g_1 & g_2 \\ 0 & g_3 \end{pmatrix};$$

$$\mathfrak{D}_1(g_1, g_2, g_3, g_4, d) = (g_1).$$

When $n = 2$ or 1, the letter d in $\mathfrak{D}_n(g_1, g_2, g_3, g_4, d)$ is superfluous and no meaning is to be attached to it. Also, $d = 0$ will mean $d_3 = \cdots = d_n = 0$.

The proofs of Lemmas 8 and 10 will be complicated by the fact that we shall be unable to satisfy the hypotheses of Lemma 6. We now note two facts that will be used to circumvent this difficulty.

For $n \geqq 2$, if

$$M_1 = \begin{pmatrix} m_{1,1} & 1 & 0 & 0 & \cdot & \cdot & 0 \\ m_{2,1} & m_{2,2} & 1 & 0 & & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \cdot & \cdot & 0 \\ \cdot & & & & & \cdot & 1 \\ m_{n,1} & m_{n,2} & \cdot & \cdot & \cdot & m_{n,n-1} & m_{n,n} \end{pmatrix}$$

is a matrix with coefficients in GF(3), then a matrix S exists in $SL(n, GF(3))$ such that $SM_1S^{-1} = C(p(\lambda))$ where $p(\lambda)$ is some polynomial. We use elementary similarity transformations to find S. For fixed k, by transforming M_1 with a sequence of elementary similarity transformations which add the $(-m_{k,j})$ multiple of column $k+1$ to column j for $j = 1, 2, \ldots, k$, we obtain a matrix with the same structure as M_1, with the same coefficients in rows $1, 2, \ldots, k-1$, and with $m_{k,1}, m_{k,2}, \ldots, m_{k,k}$ replaced with zeros. Applying this result for $k = 1$, then for $k = 2, 3, \ldots, n-1$, we obtain the companion matrix of some polynomial.

Suppose now that $M_2 = (m_{i,j})$ is an nxn $(n \geq 3)$ matrix with $m_{1,1} = m_{2,2} = -1$; $m_{i,i} = 1$ for $i = 3, 4, \ldots, n$; $m_{i,i+1} = 1$ for all $i \neq 2$; and $m_{i,j} = 0$ whenever $i > j$. Then a matrix S exists in $SL(n, GF(3))$ such that $SM_2S^{-1} = \mathfrak{D}_n(-1, 1, -1, 1, 0)$. To find S, we shall first show that the coefficients of M_2 may be assumed to satisfy $m_{3,j} = m_{4,j} = \cdots = m_{j-2,j} = 0$ for $j = 5, 6, \ldots, n$. The following reduction to this special case, which will be established by induction on the columns of M_2, is to be omitted when $n = 3$ or 4. Suppose that for some integer k with

$5 \leq k \leq n$, we have $m_{3,j} = m_{4,j} = \cdots = m_{j-2,j} = 0$ for $j = k+1, k+2, \ldots, n$. Initially, $k = n$ and this set of equations is empty. Let $S_1 = S_{k-2,k-1}(-m_{k-2,k})S_{k-3,k-1}(-m_{k-3,k})\cdots S_{3,k-1}(-m_{3,k})$. Then, if we change notation and let $S_1M_2S_1^{-1} = (m_{i,j})$, we find that $S_1M_2S_1^{-1}$ satisfies all the hypotheses imposed on M_2 and, in addition, $m_{3,k} = \cdots = m_{k-2,k} = 0$. Thus, after a similarity transformation by an element of $SL(n, GF(3))$, we may suppose that $m_{3,j} = \cdots = m_{j-2,j} = 0$ for $j = 5, 6, \ldots, n$. Then, for any $x, y \in GF(3)$ and any integer $i > 2$, $S_{2,i}(x)S_{1,i}(y)M_2S_{1,i}^{-1}(y)S_{2,i}^{-1}(x)$ differs from M_2 only in those coefficients with coordinates $(1, i)$, $(2, i)$, $(1, i+1)$, $(2, i+1)$. Choosing x and y properly, we may replace $m_{1,i}$ and $m_{2,i}$ with arbitrary elements of $GF(3)$. Making use of this fact for $i = 3, \ldots, n$ produces the desired result.

PROOF OF LEMMA 8. Since $|A| = 1$ and A is not C_1, C_2 or C_3, if A is $n \times n$, then $n = 1$ or $n \geq 3$. The result is clear if $n = 1$. Assume $n \geq 3$. Let $A = C(\lambda^n - a_n\lambda^{n-1} - \cdots - a_2\lambda - (-1)^{n-1})$. Let $D = \mathfrak{D}_n(-1, -1, -1, d_2, d)$ where d_2 is the root of

$$(-1)^n - a_2 + (-1)^{n-1}(-d_2 + n - 3) = -E_1 \tag{20}$$

and E_1 is the coefficient of λ in $(\lambda+1)^2(\lambda-1)^{n-2}$. We shall later choose $d_3, \ldots, d_n$.

Let $S_1 = (-1) \dot{+} I_{n-1}$. Then, as noted above, we can find S_2 in $SL(n, GF(3))$ such that $S_2S_1DS_1^{-1}S_2^{-1} = \mathfrak{D}_n(-1, 1, -1, 1, 0)$. Hence, as outlined above, we find S_3 in $SL(n, GF(3))$ such that $S_3S_2S_1DS_1^{-1}S_2^{-1}S_3^{-1}$ is the companion matrix of some polynomial. This polynomial must be $(\lambda+1)^2(\lambda-1)^{n-2}$.

Now, compute AD. We find that $S_1ADS_1^{-1}$ is a matrix like the matrix F in Lemma 2 with $f_{1,2} = \cdots = f_{n-1,n} = 1$, and

$$\left.\begin{aligned}
x_1 &= (-1)^{n-1}, \\
x_2 &= (-1)^n - a_2, \\
x_3 &= (-1)^{n-1} d_3 + a_2 d_2 + a_3, \\
x_4 &= (-1)^{n-1} d_4 + a_3 + a_4, \\
&\cdots \\
x_n &= (-1)^{n-1} d_n + a_{n-1} + a_n.
\end{aligned}\right\} \quad (21)$$

By Lemma 2, a matrix S_4 exists in SL(n, GF(3)) such that $S_4S_1ADS_1^{-1}S_4^{-1} = C(\lambda^n - y_n\lambda^{n-1} - \cdots - y_1)$ where

$$\left.\begin{aligned}
y_n &= x_n, \\
y_{n-1} &= x_{n-1} + (\text{a linear combination of } x_1, x_2, \ldots, x_{n-2}), \\
&\cdots \\
y_2 &= x_2 + x_1(-d_2 + n - 3), \quad (\text{by Lemma 7}) \\
y_1 &= x_1.
\end{aligned}\right\} \quad (22)$$

We determine $d_3, \ldots, d_n$ such that $\lambda^n - y_n\lambda^{n-1} - \cdots - y_1$ is $(\lambda + 1)^2(\lambda - 1)^{n-2}$. The constant terms of these two polynomials agree, and, because of 20, so do the coefficients of λ. From equations 21 and 22, we may determine $d_3, \ldots, d_n$ so that the coefficients of the other powers of λ also agree. Hence $S_4S_1ADS_1^{-1}S_4^{-1} = S_3S_2S_1DS_1^{-1}S_2^{-1}S_3^{-1}$, from which it follows that $A = SDS^{-1}D^{-1}$ where $S, D \in SL(n, GF(3))$, as required.

PROOF OF LEMMA 9. If $n > 1$, let D be a standard matrix with $d_1 = -1$, $r = 1$, $s_1 = n-1$, $c_1 = 1$. Then, by Lemmas 1 and 5, we may choose $d_2, \ldots, d_n$ such that D and AD both have $\lambda + 1$, $(\lambda - 1)^{n-1}$ as elementary divisors. Hence $A = SDS^{-1}D^{-1}$ where $|S| = 1$, $|D| = -1$.

PROOF OF LEMMA 10. In the proofs of Lemmas 10 and 11 we assume $A = A_1 \dotplus A_2$ where $A_1 = C(\lambda^{j_1} - a_{j_1}\lambda^{j_1 - 1} - \cdots - a_2\lambda - (-1)^{j_1})$ and $A_2 = C(\lambda^{j_2} - a'_{j_2}\lambda^{j_2 - 1} - \cdots - a'_2\lambda - (-1)^{j_2})$. We may assume that the characteristic polynomials of A_1 and A_2 are powers of the same irreducible polynomial since otherwise we may find a similarity transformation which carries A into the companion matrix of a single polynomial, for which the claimed result has already been obtained in Lemma 8.

If $j_2 = 1$, then $A_2 = C(\lambda + 1)$, so that A_1 is $C((\lambda + 1)^{j_1})$. Since $|A_1| = -1$, j_1 is odd. If also $j_1 = 1$, $A = -I_2$ and we may appeal to the results obtained in section 5, case 5. Hence, if $j_2 = 1$, we may assume $j_1 \geqq 3$. Deferring until later the case $j_1 = j_2 = 2$, every conceivable situation is covered by one of the following three cases: (a) $j_1 \geqq 3$, $j_2 \geqq 3$; (b) $j_1 \geqq 3$, $j_2 = 2$; (c) $j_1 \geqq 3$, $j_2 = 1$.

In case (a), select b_1, b_2 in $GF(3)$ such that $b_1 b_2 \neq 0$ and such that

$$(-1)^{j_2}(b_1 + b_2 + j_2 - 3) = -a'_2 - E_2 \tag{23}$$

where E_2 is the coefficient of λ in $(\lambda + 1)(\lambda - 1)^{j_2 - 1}$. In case (b) let $b_2 = 1$ and $b_1 = -a_2'$. Then $b_1 \neq 0$ since, if $b_1 = 0$, A_2 is the companion matrix of $\lambda^2 - 1$, which is not a power of an irreducible polynomial. In case (c) set $b_1 = b_2 = 1$.

Let d_2 be the root of

$$(-1)^{j_1+1} b_1 - a_2 + (-1)^{j_1} (- d_2 + j_1 - 3) = -E_3 \tag{24}$$

where E_3 is the coefficient of λ in $(\lambda + 1)(\lambda - 1)^{j_1 - 1}$. Now let

$$D = \left(\begin{array}{c|c} \mathcal{D}_{j_1}(-1, -b_1, -1, d_2, d) & M \\ \hline 0 & \mathcal{D}_{j_2}(1, b_1, 1, b_2, d') \end{array}\right)$$

where $M = (m_{i,j})$ is a $j_1 x j_2$ matrix with $m_{1,k} = d_{j_1+k}$ $(k = 1, 2, \ldots, j_2)$; $m_{j_1-1,1} = b_1 b_2 x$; $m_{j_1,1} = b_1 b_2$; and all other $m_{i,j} = 0$. We set $x = d_2$ if $j_1 = 3$ and $x = 1$ otherwise. Here $d = (d_3, \ldots, d_{j_1}), d_{j_1+1}, \ldots, d_n$, $d' = (d_3', \ldots, d_{j_2}')$ will be determined later. Let

$$S_1 = (-b_1) \dot{+} I_{j_1-1} \dot{+} (b_1 b_2) \dot{+} (b_2) \dot{+} I_{j_2-2} \text{ in case (a);}$$

$$S_1 = (-b_1) \dot{+} I_{n-3} \dot{+} (b_1) \dot{+} (1) \text{ in case (b);}$$

$$S_1 = (-1) \dot{+} I_{n-1} \text{ in case (c).}$$

Then we may use the remarks preceding the proof of Lemma 8 to find $S_2, S_3 \in SL(n, GF(3))$ such that $S_2 S_1 D S_1^{-1} S_2^{-1} = \mathcal{D}_n(-1, 1, -1, 1, 0)$, $S_3 S_2 S_1 D S_1^{-1} S_2^{-1} S_3^{-1} = C((\lambda + 1)^2 (\lambda - 1)^{n-2})$.

Now, compute AD. We write down AD only for the case

$j_1 > 3$, $j_2 > 3$. The details of the other cases are similar and will be omitted.

AD =

$$
\left(\begin{array}{cccccccc|ccccccc}
0 & -1 & d_2 & 0 & \cdot & \cdot & \cdot & 0 & & & & & & & \\
0 & 0 & 1 & 1 & 0 & & & \cdot & & & & & & & \\
\cdot & & 0 & 1 & 1 & \cdot & & \cdot & & & & & O & & \\
\cdot & & & \cdot & \cdot & \cdot & \cdot & & & & & & & & \\
 & & & & \cdot & \cdot & \cdot & 0 & 0 & & & & & & \\
 & & & & & \cdot & 1 & 1 & b_1 b_2 x & 0 & & & & & \\
0 & & \cdot & \cdot & \cdot & & 0 & 1 & b_1 b_2 & 0 & & & & & \\
\alpha_1 & \alpha_2 & \cdot & \cdot & \cdot & \cdot & \cdot & \alpha_{j_1} & \alpha_{j_1+1} & \alpha_{j_1+2} & \cdot & \cdot & \cdot & \cdot & \alpha_n \\
\hline
 & & & & & & & & 0 & 1 & b_2 & 0 & \cdot & \cdot & 0 \\
 & & & & & & & & 0 & 0 & 1 & 1 & & & \cdot \\
 & & & & & & & & \cdot & & & \cdot & \cdot & & \cdot \\
 & & & O & & & & & \cdot & & & & \cdot & \cdot & 0 \\
 & & & & & & & & & & & & & 1 & 1 \\
 & & & & & & & & 0 & & & & & 0 & 1 \\
 & & & & & & & & \beta_1 & \beta_2 & \beta_3 & \cdot & \cdot & \cdot & \beta_{j_2}
\end{array}\right)
$$

where

$$
\left.\begin{aligned}
\alpha_1 &= (-1)^{j_1+1}, \\
\alpha_2 &= (-1)^{j_1+1} b_1 - a_2 \\
\alpha_3 &= (-1)^{j_1} d_3 + a_2 d_2 + a_3, \\
&\ \ \cdots \\
\alpha_{j_1} &= (-1)^{j_1} d_{j_1} + a_{j_1-1} + a_{j_1},
\end{aligned}\right\} \quad (25)
$$

$$\left.\begin{aligned} \alpha_{j_1+1} &= (-1)^{j_1} d_{j_1+1} + (a_{j_1-1}x + a_{j_1})(b_1 b_2), \\ \alpha_{j_1+2} &= (-1)^{j_1} d_{j_1+2}, \\ &\ldots \\ \alpha_n &= (-1)^{j_1} d_n, \end{aligned}\right\} \quad (26)$$

$$\left.\begin{aligned} \beta_1 &= (-1)^{j_2}, \\ \beta_2 &= (-1)^{j_2} b_1 + a'_2, \\ \beta_3 &= (-1)^{j_2} d'_3 + a'_2 b_2 + a'_3, \\ \beta_4 &= (-1)^{j_2} d'_4 + a'_3 + a'_4, \\ &\ldots \\ \beta_{j_2} &= (-1)^{j_2} d'_{j_2} + a'_{j_2-1} + a'_{j_2}. \end{aligned}\right\} \quad (27)$$

Let $S_4 = S_{j_1, j_1+1}(b_1 b_2)$. Then $S_4 A D S_4^{-1}$ differs from AD only in the submatrix in the $j_1 x j_2$ block in the upper right corner. The first $j_1 - 1$ rows of this $j_1 x j_2$ submatrix are now zeros only, and the j_1 st row is (in case (a))

$$(\alpha_{j_1+1} - b_1 b_2 \alpha_{j_1},\ \alpha_{j_1+2} + b_1 b_2,\ \alpha_{j_1+3} + b_1,\ \alpha_{j_1+4},\ \ldots,\ \alpha_n).$$

Now let $S_5 = (-1) \dot{+} I_{n-1}$. Then, by methods used in the proof of Lemma 2, we may transform $S_5 S_4 A D S_4^{-1} S_5^{-1}$ with a sequence of elementary similarity transformations which add multiples of row α to row β for $j_1 > \alpha > \beta \geq 1$ only so as to bring the matrix in the upper left $j_1 x j_1$ block of $S_5 S_4 A D S_4^{-1} S_5^{-1}$ into the form of the companion matrix of a polynomial. This means that S_6 (the direct sum of a triangular $j_1 x j_1$ matrix and I_{j_2}) exists in SL(n, GF(3)) such that $S_6 S_5 S_4 A D S_4^{-1} S_5^{-1} S_6^{-1}$ is the same as $S_5 S_4 A D S_4^{-1} S_5^{-1}$ except that the $j_1 x j_1$ block in the upper left corner is now $C(\lambda^{j_1} - \alpha'_{j_1} \lambda^{j_1 - 1} - \cdots - \alpha'_2 \lambda - \alpha'_1)$ where (by Lemmas 2 and 7)

$$\left.\begin{aligned}
\alpha'_{j_1} &= \alpha_{j_1}, \\
\alpha'_{j_1 - 1} &= \alpha_{j_1 - 1} + \sum_{k=1}^{j_1 - 2} a_{j_1 - 1, k} \alpha_k, \\
&\cdots \\
\alpha'_3 &= \alpha_3 + \sum_{k=1}^{2} a_{3,k} \alpha_k, \\
\alpha'_2 &= \alpha_2 - \alpha_1 (- d_2 + j_1 - 3) \quad \text{(by Lemma 7)}, \\
\alpha'_1 &= - \alpha_1 .
\end{aligned}\right\} \quad (28)$$

The coefficients $a_{i,j}$ are independent of the d_i, d'_i. We wish $\lambda^{j_1} - \alpha'_{j_1} \lambda^{j_1 - 1} - \cdots - \alpha'_2 \lambda - \alpha'_1 = (\lambda + 1)(\lambda - 1)^{j_1 - 1}$. The constant terms of these two polynomials agree. Because of equation 24, the coefficients of λ also agree. From equations 28 and 25, it is now possible to determine $d_3, \ldots, d_{j_1}$ such that the remaining coefficients agree.

We now examine the $j_2 x j_2$ matrix in the lower right corner of $S_6 S_5 S_4 A D S_4^{-1} S_5^{-1} S_6^{-1}$. By the proof of Lemma 2, we may, by a sequence of elementary similarity transformations of

$S_6S_5S_4ADS_4^{-1}S_5^{-1}S_6^{-1}$, throw the lower right $j_2 x j_2$ block of $S_6S_5S_4ADS_4^{-1}S_5^{-1}S_6^{-1}$ into the companion matrix of a polynomial. In these elementary similarity transformations the column operations are always the addition of a multiple of column α to column β with $j_1 + 1 \leq \alpha < \beta < n$ only. (These transformations are not required in cases (b) and (c).) This means that S_7 exists in SL(n, GF(3)) such that the lower right block of $S_7S_6S_5S_4ADS_4^{-1}S_5^{-1}S_6^{-1}S_7^{-1}$ is $C(\lambda^{j_2} - \beta'_{j_2}\lambda^{j_2-1} - \dots - \beta'_2\lambda - \beta'_1)$ and the upper right $j_1 x j_2$ block consists of zeros except for row j_1 which is $(\alpha'_{j_1+1}, \dots, \alpha'_n)$ where

$$\left.\begin{aligned}
\beta'_{j_2} &= \beta_{j_2}, \\
\beta'_{j_2-1} &= \beta_{j_2-1} + \sum_{k=1}^{j_2-2} b_{j_2-1,k}\beta_k, \\
&\dots \\
\beta'_3 &= \beta_3 + \sum_{k=1}^{2} b_{3,k}\beta_k, \\
\beta'_2 &= \beta_2 + \beta_1(b_2 + j_2 - 3), \quad \text{(by Lemma 7)}, \\
\beta'_1 &= \beta_1,
\end{aligned}\right\} \quad (29)$$

and

$$\alpha'_{j_1+i} = \alpha_{j_1+i} + (\text{a linear combination of } \alpha_{j_1}, \dots, \alpha_{j_1+i-1}); \quad i = 1, 2, \dots, j_2. \quad (30)$$

The coefficients of the linear combinations in 30 and the $b_{1,j}$ in 29 are independent of the d_i, d'_i. Since $d_3, \dots, d_{j_1}$ are already determined, α_{j_1} is now a known quantity. Hence, from equations 30

$$D = \begin{pmatrix} -1 & a & 1 & 0 \\ 0 & -1 & 0 & a \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

Then $S, D \in SL(4, GF(3))$ and $A = SDS^{-1}D^{-1}$.

PROOF OF LEMMA 11. To prove Lemma 11, we use the technique used in the proof of Lemma 10. Since we shall be able to use Lemma 6 in this proof, the details here are somewhat simpler than the details of the proof of Lemma 10. If $j_1 = j_2 = 1$, then $A = -I_2$ and we appeal to the results of section 5, case 5. Otherwise, we assume that $j_1 \geqq 2$. We need consider only the following four cases: (a) $j_1 \geqq 3$, $j_2 \geqq 3$; (b) $j_1 \geqq 3$, $j_2 = 2$; (c) $j_1 \geqq 3$, $j_2 = 1$; (d) $j_1 = 2$, $j_2 = 2$. In case (a), choose b_1, b_2 to be non-zero elements of $GF(3)$ such that

$$(-1)^{j_2}(b_1 + b_2 + j_2 - 3) + a_2' = -E_4$$

where E_4 is the coefficient of λ in $(\lambda + 1)(\lambda - 1)^{j_2 - 1}$. In cases (b) and (d), let $b_2 = 1$ and $b_1 = -a_2'$; then $b_1 \neq 0$ for the same reason as in the proof of Lemma 10. In case (c), let $b_1 = b_2 = 1$. Let

$$D = \left(\begin{array}{c|c} \mathfrak{D}_{j_1}(-1, d_2, 1, 1, d) & M \\ \hline 0 & \mathfrak{D}_{j_2}(1, b_1, 1, b_2, d') \end{array} \right)$$

where the coefficients of the $j_1 x j_2$ matrix $M = (m_{i,j})$ satisfy $m_{1,k} = d_{j_1+k}$, $k = 1, 2, \ldots, j_2$; $m_{j_1-1,1} = 1$ (if $j_1 > 2$); $m_{j_1,1} = 1$;

and all other $m_{i,j} = 0$. Here $d = (d_3, \dots, d_{j_1}), d_{j_1+1}, \dots, d_n$, $d' = (d'_3, \dots, d'_{j_2})$ are to be determined later. By applying the technique used to transform the matrix M_2 in the remarks preceding the proof of Lemma 8, we may transform D into a Jordan canonical form and show that the elementary divisors of D are $(\lambda + 1)$, $(\lambda - 1)^{n-1}$. Compute AD. Let $(AD)_1$ be the matrix obtained by adding row $j_1 + 1$ of AD to row j_1, then subtracting column j_1 from column $j_1 + 1$ in the resulting matrix. Now, by a sequence of elementary similarity transformations of $(AD)_1$ in which the row operations are addition of multiples of row α to row β for $j_1 > \alpha > \beta \geq 1$ only, we obtain a new matrix $(AD)_2$ which is the same as $(AD)_1$ except that the $j_1 x j_1$ submatrix in the upper left corner is now the companion matrix of a polynomial. Just as before, we may choose $d_2, \dots, d_{j_1}$ such that this polynomial is $(\lambda - 1)^{j_1}$. Next, by a sequence of elementary similarity transformations of $(AD)_2$ in which the column operations are the addition of multiples of columns α to columns β for $j_1 + 1 \leq \alpha < \beta < n$ only, we obtain a matrix $(AD)_3$ in which the $j_2 x j_2$ submatrix in the lower right corner is the companion matrix of a polynomial, and the $j_1 x j_2$ submatrix in the upper right corner consists of zeros in all rows except for $(\alpha'_{j_1+1}, \dots, \alpha'_n)$ in row j_1, where $\alpha'_{j_1+1}, \dots, \alpha'_n$ are related to $\alpha_{j_1+1}, \dots, \alpha_n$ by a system of equations like 30. Owing to the choice of b_1 and b_2, we may select $d'_3, \dots, d'_{j_2}$ such that the matrix in the lower right corner is $C((\lambda + 1)^2(\lambda - 1)^{j_2-1})$. Also, we may choose $d_{j_1+1}, \dots, d_n$ such that $(\alpha'_{j_1+1}, \dots, \alpha'_n)$ is $(1, 0, 0, \dots, 0)$. It now follows that the characteristic polynomial of $(AD)_3$ is $(\lambda + 1)(\lambda - 1)^{n-1}$. Since $(AD)_3$ is known to be similar to the companion matrix of a polynomial, it follows that the elementary divisors of AD are $(\lambda + 1)$, $(\lambda - 1)^{n-1}$. Appealing to Lemma 6, we we easily see that $A = SDS^{-1}D^{-1}$ where $S, D \in GL(n, GF(3))$ and $|S| = 1$, $|D| = -1$.

PROOF OF LEMMA 12. Let C be any one of C_1, C_2, C_3. The proof of Lemma 9 (with an appeal to Lemma 6) shows that we may have $C = STS^{-1}T^{-1}$ with $|S| = |T| = -1$. Applying Lemma 9 to C^{-1} (which we may, since C^{-1} is non-derogatory if C is), we see that $C = UVU^{-1}V^{-1}$ with $|U| = -1$, $|V| = 1$. Then

$C \dotplus C = (S \dotplus U)(T \dotplus V)(S \dotplus U)^{-1}(T \dotplus V)^{-1}$ and $|S \dotplus U| = 1$, $|T \dotplus V| = -1$.

9. THE PROOF OF THEOREM 1 WHEN $K = GF(2)$

The field $GF(2)$ consists of the elements 0, 1. All matrices appearing in this section are assumed to have coefficients in $GF(2)$.

LEMMA 13. The $n \times n$ $(n \geq 3)$ matrix $M_n = J_2(1) \dot{+} J_{n-2}(1)$ is a commutator over $SL(n, GF(2))$.

PROOF. We first consider the case in which n is even, $n \geq 8$. Let $j = (n-2)/2$. Then $j + 1 \geq 4$ and $j + 4 \leq n$. Let R be the following $n \times n$ matrix. (The numbers at the side and top indicate the rows and columns.)

$R =$

$$
\begin{array}{cccccccccccccc}
1 & 2 & 3 & 4 & \cdots & j & j+1 & j+2 & j+3 & j+4 & j+5 & \cdots & n & \\
1 & 0 & 1 & & & & & & & & & & & 1 \\
 & 1 & 1 & & & & & & & & & & & 2 \\
 & & 1 & & & & & & & & & & & 3 \\
 & & & 1 & & & & & & 0 & & & & 4 \\
 & & & & \ddots & & & & & & & & & \vdots \\
 & & & & & 1 & & & & & & & & j \\
 & & & & & & 1 & 1 & & & & & & j+1 \\
 & & & & & & & 1 & 1 & 0 & 1 & \cdots & 1 & j+2 \\
 & & & & & & & & 1 & 1 & 1 & \cdots & 1 & j+3 \\
 & & & 0 & & & & & & 1 & 1 & \cdots & 1 & j+4 \\
 & & & & & & & & & & 1 & \cdots & 1 & j+5 \\
 & & & & & & & & & & & \ddots & \vdots & \vdots \\
 & & & & & & & & & & & & 1 & n
\end{array}
$$

The elementary divisors of R are $(\lambda + 1)^2$, $(\lambda + 1)$, ..., $(\lambda + 1)$, $(\lambda + 1)^{n-j}$. To see this, note that the minimum polynomial of the principal submatrix of R formed from the last $n - j$ rows and columns of R is $(\lambda + 1)^{n-j}$, and that the elementary divisors of the principal submatrix formed from the first three rows and columns of R are $(\lambda + 1)$, $(\lambda + 1)^2$. Let $S = (s_{i,j})$ be an nxn matrix with $s_{i,i} = 1$ for $i = 1, 2, \ldots, j + 1, j + 3, \ldots, n$; $s_{j+1,j+2} = 1$; $s_{j+3,j+2} = 1$; $s_{j+2,k} = 1$ for $k = j + 3, j + 4, \ldots, n$; and all other $s_{i,j} = 0$. Then $|S| = 1$ and it is easy to see that $S(M_n R) = (J_{j+2}(1) \dotplus J_2(1) \dotplus I_{n-j-4})S$; hence, the elementary divisors of $M_n R$ are $(\lambda + 1)^{j+2}$, $(\lambda + 1)^2$, $(\lambda + 1)$, ..., $(\lambda + 1)$. But $j + 2 = n - j$. Hence R and $M_n R$ have the same elementary divisors, so that $M_n = QRQ^{-1}R^{-1}$ for some $Q \in SL(n, GF(2))$.

Next we consider the case in which n is odd, $n \geq 7$. Let $j = (n - 1)/2$. Then $j + 1 \geq 4$, $j + 4 \leq n$. Let R_1 be the matrix at the top of the next page. The elementary divisors of R_1 are $(\lambda + 1)^2$, $(\lambda + 1)$, ..., $(\lambda + 1)$, $(\lambda + 1)^{n-j}$. Let $S_1 = (s_{i,j})$ where $s_{i,i} = 1$ for $i = 1, 2, \ldots, j + 2, j + 4, \ldots, n$; $s_{j+4,j+3} = s_{j+3,j+4} = 1$; $s_{j+1,k} = 1$ for $k = j + 2, j + 3, \ldots, n$; and all other $s_{i,j}$ are zero. Then $S_1(M_n R_1) = (J_{j+1}(1) \dotplus J_2(1) \dotplus I_{n-j-3})S_1$ and $|S_1| = 1$, so that the elementary divisors of $M_n R_1$ are $(\lambda + 1)^{j+1}$, $(\lambda + 1)^2$, $(\lambda + 1)$, ..., $(\lambda + 1)$. Since $j + 1 = n - j$, the result follows as before.

$$R_1 =$$

$$\begin{array}{c} \\ \\ \end{array}
\begin{array}{cccccccccccccc|l}
1 & 2 & 3 & 4 & \cdots & j-1 & j & j+1 & j+2 & j+3 & j+4 & j+5 & \cdots & n & \\
\hline
1 & 0 & 1 & & & & & & & & & & & & 1 \\
 & 1 & 1 & & & & & & & & & & & & 2 \\
 & & 1 & & & & & & & & & & & & 3 \\
 & & & 1 & & & & & \text{\Large 0} & & & & & & 4 \\
 & & & & \ddots & & & & & & & & & & \vdots \\
 & & & & & 1 & & & & & & & & & j-1 \\
 & & & & & & 1 & & & & & & & & j \\
 & & & & & & & 1 & 1 & 1 & 1 & 1 & \cdots & 1 & j+1 \\
 & & & \text{\Large 0} & & & & & 1 & 1 & 0 & 1 & \cdots & 1 & j+2 \\
 & & & & & & & & & 1 & 1 & 1 & \cdots & 1 & j+3 \\
 & & & & & & & & & & 1 & 1 & \cdots & 1 & j+4 \\
 & & & & & & & & & & & 1 & \cdots & 1 & j+5 \\
 & & & & & & & & & & & & \ddots & \vdots & \vdots \\
 & & & & & & & & & & & & & 1 & n
\end{array}$$

We now complete the proof by exhibiting M_n as a commutator for $n = 3, 4, 5, 6$. Let

$$U_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad V_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$U_4 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad V_4 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$U_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad V_5 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix};$$

$$U_6 = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad V_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

Then $M_n = U_n V_n U_n^{-1} V_n^{-1}$ for $n = 3, 4, 5, 6$. This completes the proof of Lemma 13.

Let

$$U = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

We now state Lemma 14, which can be verified by direct computation.

LEMMA 14. $J_2(1) \dot{+} J_2(1) \dot{+} J_2(1) = UVU^{-1}V^{-1}$.

LEMMA 15. For $n \geq 3$, let

$$A_n = \begin{pmatrix} 1 & a_2 & a_3 & \cdot & \cdot & \cdot & a_n \\ & & & & & & 1 \\ & & 0 & & & \cdot & \\ & & & & \cdot & & \\ & & & \cdot & & & \\ & & 1 & & & 0 & \\ 0 & 1 & 0 & & & & \end{pmatrix}.$$

Then, if $a_2 + a_n = 1$, the elementary divisors of A_n are $(\lambda + 1)^3$, $(\lambda + 1)$, together with $(\lambda + 1)^2$ repeated $m - 2$ times when $n = 2m$ is even, and $(\lambda + 1)^3$ together with $(\lambda + 1)^2$ repeated $m - 1$ times when $n = 2m + 1$ is odd.

PROOF. To compute the elementary divisors of A_n, we reduce the polynomial matrix $\lambda I_n + A_n$ to a diagonal form $(p_1(\lambda)) \dotplus (p_2(\lambda)) \dotplus \cdots \dotplus (p_n(\lambda))$ where $p_i(\lambda)$ divides $p_{i+1}(\lambda)$, $i = 1, 2, \ldots, n - 1$, by transformations of the following two types: (1) interchange of two rows (or two columns); (2) addition of a polynomial multiple of one row (column) to another row (column). The row and column transformations necessary differ slightly in the two cases n even and n odd.

If n is even, $n = 2m$, we begin with $\lambda I_n + A_n$ and add λ times row $2m + 1 - k$ to row $k + 1$, then add a_{k+1} times row $2m + 1 - k$ to row 1 in the resulting matrix, for $k = 1, 2, \ldots, m - 1$. Next, add λ times column k to column $2m + 2 - k$ for $k = 2, 3, \ldots, m$. Then add a_k times column 1 to column $2m + 2 - k$ for

$k = 2, 3, \ldots, m$. At this juncture the coefficient with coordinates $(1, n)$ is $a_2 + a_n = 1$. Now add the $\lambda^2 + 1 = (\lambda + 1)^2$ multiple of the first row to the second, then add $a_{m+1}(\lambda + 1)$ times row $m + 1$ to row 2 and $a_{2m+2-k} + a_k$ times row k to row 2 for $k = m, m - 1, \ldots, 3$. At this point, the last column consists of zeros except for a 1 in the first row. Add $\lambda + 1$ times the last column to the first, a_{m+1} times the last column to column $m + 1$, and $a_{2m+2-k} + a_k$ times the last column to column $2m + 2 - k$ for $k = 3, 4, \ldots, m$. We now have a matrix which can be transformed to a diagonal form of the required type by permutations of its rows and columns.

If n is odd, $n = 2m + 1$, we begin with $\lambda I_n + A_n$ and add λ times row $2m + 2 - k$ to row $k + 1$ and a_{k+1} times row $2m + 2 - k$ to row 1, for $k = 1, 2, \ldots, m$. Then add λ times column k to column $2m + 3 - k$ for $k = 2, 3, \ldots, m+1$. Next, add a_k times column 1 to column $2m + 3 - k$ for $k = 2, 3, \ldots, m + 1$. The coefficient with coordinates $(1, n)$ is now $a_2 + a_n = 1$. Add the $(\lambda^2 + 1)$ multiple of the first row to row 2, then add $a_{2m+3-k} + a_k$ times row k to row 2, for $k = m + 1, m, \ldots, 3$. At this point, the last column has a 1 in the first row and zeros in the other rows. Finally, add $(\lambda + 1)$ times the last column to the first column and $a_{2m+3-k} + a_k$ times the last column to column $2m + 3 - k$ for $k = 3, \ldots, m + 1$. We now have a matrix which can be transformed into a diagonal matrix of the required type by permutations of its rows and of its columns.

LEMMA 16. For $n \geq 3$, let

$$C_n = \begin{pmatrix} & 0 & & & & 1 \\ & & & & 1 & \\ & & & \cdot & & \\ & & \cdot & & 0 & \\ & 1 & & & & \\ 1 & c_2 & c_3 & \cdot & \cdot & c_n \end{pmatrix} .$$

(1). If $n = 2m$ and if

$$\left.\begin{aligned} c_{m+2} + c_{m-1} &= 0, \\ c_{m+3} + c_{m-2} &= 0, \\ &\cdots \\ c_{2m-1} + c_2 &= 0, \end{aligned}\right\} \quad \text{(absent if } m = 2\text{)}$$

$$c_{2m} = 0,$$

$$c_m = 1,$$

$$c_{m+1} = 1,$$

then the elementary divisors of C_n are $(\lambda + 1)^3$, $(\lambda + 1)$, together with $(\lambda + 1)^2$ repeated $m - 2$ times.

(2). If $n = 2m + 1$, and if

$$\left.\begin{aligned} c_{m+2} + c_m &= 0, \\ c_{m+3} + c_{m-1} &= 0, \\ &\cdots \\ c_{2m} + c_2 &= 0, \end{aligned}\right\} \quad \text{(absent if } m = 1\text{)}$$

$$c_{2m+1} = 0,$$

$$c_{m+1} = 1,$$

then the elementary divisors of C_n are $(\lambda + 1)^3$ and $(\lambda + 1)^2$ repeated $m - 1$ times.

PROOF. As in the proof of Lemma 15, we use row and column transformations of the two types previously indicated to transform $\lambda I_n + C_n$ into a diagonal form from which the elementary divisors may be read off.

(1). In $\lambda I_n + C_n$ add λ times column $2m + 1 - k$ to column k for $k = 1, 2, \ldots, m$. In the resulting matrix, add λ times row k to row $2m + 1 - k$ for $k = 1, 2, \ldots, m$. Next, add c_{2m+1-k} times row k to row $2m$ for $k = 2, 3, \ldots, m$. Since $c_m = c_{m+1} = 1$, the coefficient with coordinates (n, m) is now $\lambda + 1$. Now add $\lambda + 1$ times column m to column 1 and c_{2m+1-k} times column m to column k for $k = 2, 3, \ldots, m - 1$. Add c_{2m+1-k} times row $2m + 1 - k$ to row $m + 1$ for $k = 2, 3, \ldots, m - 1$. At this juncture, the last row consists of zeros only except for $\lambda + 1$ in column m. Finally, add $\lambda + 1$ times row $2m$ to row $m + 1$. The matrix we now have can be brought to diagonal form by permutations of its rows and of its columns.

(2). When n is odd, we begin by adding λ times column $2m + 2 - k$ of $\lambda I_n + C_n$ to column k for $k = 1, 2, \ldots, m$. In the resulting matrix, add λ times row k to row $2m + 2 - k$ for $k = 1, 2, \ldots, m$, then add c_{2m+2-k} times row k to row $2m + 1$ for $k = 2, 3, \ldots, m$. Next, using $c_{m+1} = 1$, add $\lambda^2 + 1$ times column $m + 1$ to column 1 and $(\lambda + 1)c_{2m+2-k}$ times column $m + 1$ to column k for $k = 2, 3, \ldots, m$. Now add c_{2m+2-k} times row $2m + 2 - k$ to row $m + 1$

for $k = 2, 3, \ldots, m$. At this juncture, the last row consists entirely of zeros, except for a 1 in column $m + 1$. Finally, add $(\lambda + 1)$ times row $2m + 1$ to row $m + 1$. The matrix we now have may be brought to diagonal form by permutations of its rows and of its columns. This completes the proof.

We now turn to the crucial lemma of this section.

LEMMA 17. A non-derogatory, non-singular $n \times n$ matrix M_n with coefficients in $GF(2)$ is a commutator over $GF(2)$ unless $n = 2$ and M_2 is similar to $C((\lambda + 1)^2)$.

PROOF. Let $M_n = C(\lambda^n + b_2\lambda^{n-1} + \cdots + b_n\lambda + 1)$. Let A_n be as in Lemma 15, where the a_i are to be determined later. Let $C_n = M_nA_n$. Then C_n is as described in Lemma 16, with $c_i = a_i + b_i$, $i = 2, 3, \ldots, n$.

Case (1). $n = 2m$. If

$$\begin{aligned}
a_2 + a_{2m} &= 1, \\
\left.\begin{aligned}
c_{m+2} + c_{m-1} &= 0, \\
c_{m+3} + c_{m-2} &= 0, \\
&\cdots \\
c_{2m-1} + c_2 &= 0,
\end{aligned}\right\} &\quad \text{(absent if } m = 2) \\
c_{2m} &= 0, \\
c_m &= 1, \\
c_{m+1} &= 1,
\end{aligned}$$

then, by Lemmas 15 and 16, A_n and C_n have the same elementary divisors, which would imply the result. These equations become

$$\left.\begin{aligned}
a_{m+2} + b_{m+2} + a_{m-1} + b_{m-1} &= 0,\\
a_{m+3} + b_{m+3} + a_{m-2} + b_{m-2} &= 0,\\
&\cdots\\
a_{2m-1} + b_{2m-1} + a_2 + b_2 &= 0,
\end{aligned}\right\}\text{(absent if } m = 2) \qquad (31)$$

$$\begin{aligned}
a_{2m} + b_{2m} &= 0,\\
a_m + b_m &= 1,\\
a_{m+1} + b_{m+1} &= 1,\\
a_2 + a_{2m} &= 1.
\end{aligned}$$

If $m \geq 3$, then $2, m, m+1, 2m$ are distinct integers. Then take $a_{2m} = b_{2m}$, $a_m = 1 + b_m$, $a_{m+1} = 1 + b_{m+1}$, $a_2 = 1 + a_{2m}$. Choose $a_3, \ldots, a_{m-1}$ at will. Then equations 31 determine $a_{m+2}, \ldots, a_{2m-1}$. This completes the proof of case 1 when $m \geq 3$.

The cases $n = 4$ and $n = 2$ need special attention. When $n = 4$ we have unknowns a_2, a_3, a_4 and equations

$$\begin{aligned}
a_2 + b_2 &= 1,\\
a_3 + b_3 &= 1,\\
a_4 + b_4 &= 0,\\
a_2 + a_4 &= 1.
\end{aligned}$$

These equations have a solution if, and only if, $b_2 + b_4 = 0$. Thus the cases in which $b_4 = 1 + b_2$ are not covered by this proof. Let

$$S_1 = \begin{pmatrix} 1 & 1 & 1 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad T_1 = \begin{pmatrix} 1 & 0 & 0 & 0\\ 1 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 1 & 1 & 1 & 1 \end{pmatrix};$$

$$S_2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}, T_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Then $C(\lambda^4 + \lambda^3 + 1) = S_1 T_1 S_1^{-1} T_1^{-1}$ and $C(\lambda^4 + \lambda^3 + \lambda^2 + 1)$ equals $S_2 T_2 S_2^{-1} T_2^{-1}$. Furthermore, $C(\lambda^4 + \lambda + 1)$ is similar in $SL(4, GF(2))$ to $C^{-1}(\lambda^4 + \lambda^3 + 1)$ and $C(\lambda^4 + \lambda^2 + \lambda + 1)$ is similar in $SL(4, GF(2))$ to $C^{-1}(\lambda^4 + \lambda^3 + \lambda^2 + 1)$.

When $n = 2$, we note that every element of $SL(2, GF(2))$ is similar within $SL(2, GF(2))$ to I_2, $C(\lambda^2 + \lambda + 1)$, or $C((\lambda + 1)^2)$. Let $S_3 = J_2(1)$, $T_3 = C((\lambda + 1)^2)$. Then $C(\lambda^2 + \lambda + 1) = S_3 T_3 S_3^{-1} T_3^{-1}$.

The proof of Lemma 17 for even n is now complete.

Case (2). $n = 2m + 1$. The proof here is similar to the foregoing proof. If

$$\left.\begin{aligned} a_{m+2} + b_{m+2} + a_m + b_m &= 0, \\ a_{m+3} + b_{m+3} + a_{m-1} + b_{m-1} &= 0, \\ &\cdots \\ a_{2m} + b_{2m} + a_2 + b_2 &= 0, \end{aligned}\right\} \text{(absent if } m = 1) \qquad (32)$$

$$\begin{aligned} a_{2m+1} + b_{2m+1} &= 0, \\ a_{m+1} + b_{m+1} &= 1, \\ a_2 + a_{2m+1} &= 1, \end{aligned}$$

then, by Lemmas 15 and 16, A_n and C_n have the same elementary divisors and the result follows. If $m \geq 2$, then $2, m + 1, 2m + 1$ are distinct integers. Then let $a_{2m+1} = b_{2m+1}$, $a_{m+1} = 1 + b_{m+1}$, $a_2 = 1 + a_{2m+1}$. Choose $a_3, \ldots, a_m$ at will and solve equations 32 for $a_{m+2}, \ldots, a_{2m}$. The proof of case 2 for $m \geq 2$ is now finished.

When $m = 1$, we have unknowns a_2 and a_3 and equations

$$a_3 = b_3,$$
$$a_2 = 1 + b_2,$$
$$a_2 = 1 + a_3.$$

A solution exists when, and only when, $b_2 + b_3 = 0$. To complete the proof, we let

$$S_4 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad T_4 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then $C(\lambda^3 + \lambda + 1) = S_4 T_4 S_4^{-1} T_4^{-1}$ and $C(\lambda^3 + \lambda^2 + 1)$ is similar within $SL(3, GF(2))$ to $C^{-1}(\lambda^3 + \lambda + 1)$. Lemma 17 is now completely established.

With these lemmas at our disposal, we are now in a position to finish the proof of Theorem 1.

If $A \in GL(n, GF(2))$, we may, after a similarity transformation, suppose that A is the direct sum of companion matrices of powers of polynomials irreducible over $GF(2)$,
$A = C(p_1^{e_1}(\lambda)) \dotplus \cdots \dotplus C(p_r^{e_r}(\lambda))$. If no $p_i^{e_i}(\lambda) = (\lambda + 1)^2$, then each $C(p_i^{e_i}(\lambda))$ is a commutator by Lemma 17, hence A is a commutator. If exactly s of the $p_i^{e_i}(\lambda)$ are $(\lambda + 1)^2$ where $s > 1$, then by Lemma 14 and Lemma 13 (for $n = 4$) the $(2s)x(2s)$ matrix

$$C((\lambda + 1)^2) \dotplus \cdots \dotplus C((\lambda + 1)^2)$$

is a commutator, from which it follows that A is a commutator. If exactly one of the $p_i^{e_i}(\lambda)$ is $(\lambda + 1)^2$, then if $n \neq 2$, we must have $r > 1$. Suppose $p_1^{e_1}(\lambda) = (\lambda + 1)^2$. If $p_2(\lambda) = \lambda + 1$, Lemma 13 states that $C(p_1^{e_1}(\lambda)) \dotplus C(p_2^{e_2}(\lambda))$ is a commutator, and hence so is A. If $p_2(\lambda)$ is not $\lambda + 1$, then, since $p_2(\lambda)$ is irreducible over GF(2), $p_2(\lambda)$ is prime to $\lambda + 1$. But then A is similar to

$$C((\lambda + 1)^2 p_2^{e_2}(\lambda)) \dotplus C(p_3^{e_3}(\lambda)) \dotplus \cdots \dotplus C(p_r^{e_r}(\lambda)),$$

so that Lemma 17 is directly applicable. The proof is now complete.

10. BIBLIOGRAPHY

1. L. E. Dickson, Linear groups, (1901), p. 79, Corollary 1.

2. K. Iwasawa, Über die Einfachheit der speziallen projectiven Gruppen, Proc. Imperial Academy Tokyo, vol. 17 (1941), p. 57-59.

3. L. E. Dickson, op. cit., p. 78, Theorem 100.

4. O. Litoff, On the commutator subgroup of the general linear group, Proc. Amer. Math. Soc., vol. 6 (1955), p. 466, Theorem 2.

5. K. Shoda, Einige Sätze über Matrizen, Japanese J. Math., vol. 13 (1936), p. 361-365.

6. K. Shoda, Über den Kommutator der Matrizen, J. Math. Soc. of Japan, vol. 3 (1951), p. 78-81.

7. H. Tôyama, On commutators of matrices, Kodai Math. Seminar Reports, Nos. 5 and 6 (Dec. 1949), p. 1-2.

8. O. Taussky, Generalized commutators of matrices and permutations of factors in a product of three matrices, Studies presented to R. von Mises, (1954), p. 67-68.

9. K. Fan, Some remarks on commutators of matrices, Archiv der Math., vol. 5 (1954), p. 102-107.

10. G. Villari, Sui commutatori del gruppo modulare, Bollettino delle Unione Matematica Italiana, vol. 13 (1958), p. 196-201.

11. O. Ore, Some remarks on commutators, Proc. Amer. Math. Soc., vol. 2 (1951), p. 307-314.

12. N. Itô, A theorem on the alternating group $A_n (n \geq 5)$, Mathematica Japonicae, vol. 2 (1951), p. 59-60.

13. R. Stoll, Linear algebra and matrix theory, (1952).

14. W. LeVeque, Topics in number theory, (1956), vol. 1, p. 135.

15. T. Muir and W. H. Metzler, A treatise on the theory of determinants, (1930), p. 445.

16. W. LeVeque, op. cit., p. 126.

17. E. Landau, Über die Darstellung definiter Funktionen durch Quadrate, Math. Annalen, vol. 62 (1906), p. 271-285.

18. C. L. Siegel, Darstellung total positiver Zahlen durch Quadrate, Math. Zeitschrift, vol. 11 (1921), p. 246-275.