FUNCTIONAL ANALYSIS AND OPERATOR THEORY
BANACH CENTER PUBLICATIONS, VOLUME 30
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 1994

CONJUGACY AND FACTORIZATION RESULTS ON MATRIX GROUPS

THOMAS J. LAFFEY

Department of Mathematics, University College Dublin
Belfield, Dublin 4, Ireland
E-mail: tlaffey@irlearn.bitnet

In this survey paper, we present (mainly without proof) a number of results on conjugacy and factorization in general linear groups over fields and commutative rings. We also present the additive analogue in matrix rings of some of these results. The first section deals with the question of expressing elements in the commutator subgroup of the general linear group over a field as (simple) commutators. In Section 2, the same kind of problem is discussed for the general linear group over a commutative ring. In Section 3, the analogous question for additive commutators is discussed. The case of integer matrices is given special emphasis as this is an area of current interest. In Section 4, factorizations of an element $A \in GL(n, F)$ (F a field) in which at least one of the factors preserves some form (e.g. is symmetric or skew-symmetric) is considered. An application to the size of abelian subgroups of finite p-groups is presented. In Section 5, a curious interplay between additive and multiplicative commutators in $M_n(F)$ (F a field) is identified for matrices of small size and a general factorization theorem for a polynomial using conjugates of its companion matrix is presented.

Notation. The notation is standard. In particular, $GL(n, R)$ denotes the group of invertible $n \times n$ matrices A such that A and A^{-1} have entries in the ring R and in the case R is commutative and has an identity, $SL(n, R)$ denotes the subgroup of those elements A of $GL(n, R)$ with $\det A = 1$. A matrix A is called *nonderogatory* (or cyclic) if its minimal polynomial equals its characteristic polynomial. Equivalently A is nonderogatory if the only matrices which commute with A are the polynomials in A; cf. [G-L-R, pp. 299–300]. An *involution* is an element

1991 *Mathematics Subject Classification*: 15A23, 15A33, 15A36, 15A63.
The paper is in final form and no version of it will be published elsewhere.

J in a group with $J^2 = I$. If V is a vector space equipped with a bilinear form g, a subspace U of V is *isotropic* if $g(u, v) = 0$ for all $u, v \in U$.

A^T denotes the transpose of the matrix A.

$[X, Y]$ denotes the multiplicative commutator $X^{-1}Y^{-1}XY$.

(X, Y) denotes the additive commutator $XY - YX$.

$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ denote as usual the sets of integers, rational numbers, real numbers and complex numbers, respectively.

1. Commutators over fields. Let F be a field and let $A \in SL(n, F)$. A famous theorem of R. C. Thompson [THO1] states that if $(n, |F|) \neq (2, 2)$, then A is a commutator $[X, Y]$ for $X, Y \in GL(n, F)$. Thompson's proof depends on an analysis of (a variant of) the rational canonical form and is difficult, particularly for small fields. A simpler proof for large fields was constructed by Grunenfelder, Paré and Radjavi [G-P-R], and independently Sourour [SOU1] and the author [LAF2] improved their argument to give the following very useful factorization theorem.

THEOREM 1. *Let F be a field and let $A \in GL(n, F)$ be nonscalar. Let $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ be any elements of F which satisfy the relation*

$$\det A = x_1 \ldots x_n y_1 \ldots y_n .$$

Then there exist elements T,L,$U \in GL(n, F)$ with L lower-triangular and having diagonal

$$\operatorname{diag}(L) = (x_1, \ldots, x_n)$$

and U upper-triangular and having diagonal

$$\operatorname{diag}(U) = (y_1, \ldots, y_n)$$

such that $T^{-1}AT = LU$.

To prove the theorem, we use induction on n. Since A is not scalar, there exists a vector $v \in F^n$, the space of column n-tuples over F, such that v and Av are linearly independent. Put $v_1 = v$, $v_2 = z_1^{-1}Av - v_1$ where $z_1 = x_1 y_1$ and extend to a basis $v_1, v_2, \ldots, v_n$ of F^n. Using this change of basis, we see that there exists an element $T_1 \in GL(n, F)$ such that

$$T_1^{-1}AT_1 = \begin{bmatrix} z_1 & z_1 & 0 & \cdots & 0 \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix} \quad \text{(for some } b_{ij} \in F\text{)}$$

$$= \begin{bmatrix} z_1 & 0 & \cdots & 0 \\ b_{21} & & & \\ \vdots & & I_{n-1} & \\ b_{n1} & & & \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & & & & \\ \vdots & & & A_2 & \\ 0 & & & & \end{bmatrix}$$

where $A_2 \in GL(n-1,F)$. If A_2 is not scalar, then using induction, we may assume that there exist $T_2, L_2, U_2 \in GL(n-1,F)$ with L_2 lower-triangular and $\operatorname{diag}(L_2) = (z_2, \ldots, z_n)$ and U_2 upper-triangular with $\operatorname{diag}(U_2) = (1, \ldots, 1)$ such that $T_2^{-1} A_2 T_2 = L_2 U_2$, where here $z_i = x_i y_i$ $(i = 2, \ldots, n)$. If A_2 is scalar, then one can show that a different choice of elements $v_2, \ldots, v_n$ will lead to a nonscalar A_2. It then follows that there exist $T_3, L_3, U_3 \in GL(n,F)$ with L_3 lower-triangular and $\operatorname{diag}(L_3) = (z_1, \ldots, z_n)$, and U_3 upper-triangular and $\operatorname{diag}(U_3) = (1, \ldots, 1)$ such that $T_3^{-1} A T_3 = L_3 U_3$. But $L_3 U_3 = LU$ where $L = L_3 D^{-1}$, $U = DU_3$ and $D = \operatorname{diag}(y_1, \ldots, y_n)$. This completes the proof.

Suppose $A \in SL(n,F)$ is not scalar and that F has at least $n+1$ elements. Then $x_1, \ldots, x_n$ can be taken to be distinct in Theorem 1. Also, since $\det A = 1$, we may take $y_i = x_i^{-1}$ for $i = 1, \ldots, n$. But then L is similar to the diagonal matrix $D = \operatorname{diag}(x_1, \ldots, x_n)$ and U is similar to D^{-1} and thus A is similar to a matrix of the form $DZ^{-1}D^{-1}Z$ and thus A is a commutator $X^{-1}Y^{-1}XY$. If A is scalar, one uses instead the fact that if P is the permutation matrix corresponding to the n-cycle $(1\,2\,3 \ldots n)$ and $w \in F$ with $w^n = 1$, then wP is similar to P^{-1}, so

$$wI = P^{-1} S^{-1} P S$$

for some $S \in GL(n,F)$.

No simplification of Thompson's argument for small fields appears to be available. One can use character theory (Honda's Theorem [HON]) since the group $GL(n,F)$ is finite in this case, but this approach is not simpler than his.

The problem of writing $A \in SL(n,F)$ as a commutator $[X,Y]$ with $X, Y \in SL(n,F)$ is also an interesting one, particularly in the scalar case, and it has been resolved by Thompson [THO2].

Theorem 1 has a number of interesting consequences. It is well known that a matrix $A \in GL(n,F)$ is similar to its inverse if and only if A is the product of two involutions. If $A = RS$ with $R^2 = I$, $S^2 = I$, then $R^{-1} A R = A^{-1}$. The proof of the converse is more difficult and has been the subject of several papers. All rely on the rational canonical form or some variant thereof. See Djoković [DJO], Wonenberger [WON2] for example.

If F does not have characteristic 2 and $\det A = 1$, and A is similar to its inverse, then we show [LAF2] that $A = J_1 J_2$ where J_1, J_2 are each similar to

$$J_0 = \operatorname{diag}(\underbrace{1, \ldots, 1}_{k}, \underbrace{-1, \ldots, -1}_{n-k})$$

where $k = [(n+1)/2]$.

The corresponding question for the other classical groups has also been considered. See Gow [GOW2], Wonenberger [WON1] for some typical results.

If $A \in GL(n,F)$ with $\det A = \pm 1$ we may apply Theorem 1 with $x_1 = \ldots = x_{n-1} = y_1 = \ldots = y_n = 1$, $x_n = \det A$. Then L is similar to its inverse and U is also similar to its inverse. Hence L, U are each a product of two involutions and hence we see that A is the product of four involutions. [If A is scalar, we may use

the permutation matrix of $(1\,2\,3\ldots n)$ and the fact that $(1\,2\,3\ldots n)$ is the product of two involutions to achieve a similar result.] This result was first proved using rational canonical forms by Gustafson, Halmos and Radjavi [G-H-R].

It is not the case that every $A \in GL(n, F)$ with $\det A = \pm 1$ is the product of three involutions. This can be seen (if F does not have characteristic 2) easily as follows. If $A = J_1 J_2 J_3$ with $J_i^2 = I$ $(i = 1, 2, 3)$, then $J_1 A = J_2 J_3$ is similar to its inverse. Suppose A has an eigenvalue z with $z^2 \neq \pm 1$ such that the corresponding eigenspace V has dimension greater than $3n/4$. Since J_1 is an involution, it has eigenspaces U_1, U_2 corresponding to ± 1 of dimension k, l with $k + l = n$. Now

$$\dim(V \cap U_1) = \dim(V + U) + \dim V - \dim U > -n + \frac{3n}{4} + k$$

and

$$\dim(V \cap U_2) > -n + \frac{3n}{4} + l\,,$$

so

$$\dim(V \cap U_1) + \dim(V \cap U_2) > \frac{n}{2}\,.$$

On $V \cap U_1$, AJ_1 has an eigenvalue z and on $V \cap U_2$, AJ_1 has an eigenvalue $-z$. But then $J_1 A$ must also have eigenvalues z^{-1} and $-z^{-1}$ with the corresponding multiplicities. But this forces $z = \pm z^{-1}$, contrary to hypothesis. Since the only condition on A is that $\det A = \pm 1$ it is easy to construct examples with such an eigenvalue z.

A detailed (but not completely decisive) discussion of products of three involutions has been provided by Liu [LIU]. See also Wu [WU2].

When F is the field of real numbers $\mathbb{R}$ and $A \in GL(n, \mathbb{R})$ is not scalar and $\det A > 0$, we can choose the elements $x_1, \ldots, x_n, y_1, \ldots, y_n$ in Theorem 1 to be all distinct and positive. But it is well known and easy to prove that if $B \in GL(n, \mathbb{R})$, then B is the product of two positive definite symmetric matrices if and only if B is diagonalizable with positive real eigenvalues. Hence for this choice of x_i, y_j, we find that L, U are each products of two positive definite matrices and since being a product of four positive definite symmetric matrices is invariant under similarity, we conclude that A is a product of four positive definite matrices. This argument is due to Sourour [SOU1] and is much easier than the original proof of this result by Ballantine [BAL].

If F contains a primitive nth root of unity, ω say, we take the numbers $x_j = y_j = \omega^j$ and then we deduce that $A = [X, Y]$ where X is periodic of order n. Thus X is similar to the permutation matrix P corresponding to the n-cycle $(1\,2\,3\ldots n)$. However, even if the field F does not contain a primitive nth root of unity, a factorization of $A \in SL(n, F)$ as $[X, Y]$ with X similar to P can often be achieved. Eleanor Meehan and the author [L-M3] have found a new approach to this type of factorization when n is odd. We present the principal steps here.

Let R be a commutative ring with identity and assume that the matrices occurring in this section have entries in R.

Let $n \geq 3$ be odd,

$$A = \begin{bmatrix} a_1 & y_1 & 0 & \cdots & 0 & 0 \\ 0 & a_2 & y_2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & a_{n-1} & y_{n-1} \\ 0 & 0 & 0 & \ldots & 0 & a_n \end{bmatrix}$$

and $P = (p_{ij})$ the permutation matrix corresponding to the n-cycle $(1\,2\,3\ldots n)$ (so $p_{i,i+1} = 1$ $(i = 1, \ldots, n-1)$, $p_{n1} = 1$, $p_{ij} = 0$ otherwise).

Let $B = AP^{-2}$. Let z be an indeterminate.

Using the Laplace expansion along row one, we obtain a recurrence relation for the determinant

$$\begin{aligned} \det(zI + B) = z^n + a_2 a_n z^2 \Delta(3, \ldots, n-2) - a_1 z \Delta(2, \ldots, n-1) \\ - a_n y_1 z \Delta(2, \ldots, n-2) + a_1 a_2 \ldots a_n \end{aligned}$$

where

$$\Delta(k+1, \ldots, k+l) = \begin{vmatrix} y_{k+1} & z & 0 & \ldots & 0 & 0 & 0 \\ a_{k+2} & y_{k+2} & z & \ldots & 0 & 0 & 0 \\ 0 & a_{k+3} & y_{k+3} & \ldots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & y_{k+l-2} & z & 0 \\ 0 & 0 & 0 & \ldots & a_{k+l-1} & y_{k+l-1} & z \\ 0 & 0 & 0 & \ldots & 0 & a_{k+l} & y_{k+l} \end{vmatrix}.$$

Using the Laplace expansions along the top row and along the bottom row, respectively, we obtain the following two recurrence relations:

$$\Delta(k+1, \ldots, k+l) = y_{k+1} \Delta(k+2, \ldots, k+l) - a_{k+2} z \Delta(k+3, \ldots, k+l)$$

and

$$(*) \quad \Delta(k+1, \ldots, k+l) = y_{k+l} \Delta(k+1, \ldots, k+l-1) - z a_{k+l} \Delta(k+1, \ldots, k+l-2)\,.$$

Hence

$$\begin{aligned} (1) \quad \det(zI + B) = & z^n + z[z(a_2 a_n \Delta(3, \ldots, n-2) + a_1 a_{n-1} \Delta(2, \ldots, n-3)) \\ & - (a_1 y_{n-1} + a_n y_1) \Delta(2, \ldots, n-2)] + a_1 \ldots a_n\,. \end{aligned}$$

Using $(*)$ again we may write (for $n \geq 5$)

$$\begin{aligned} (2) \quad & a_2 a_n \Delta(3, \ldots, n-2) + a_1 a_{n-1} \Delta(2, \ldots, n-3) \\ & = (a_2 a_n y_{n-2} + a_1 a_{n-1} y_2) \Delta(3, \ldots, n-3) \\ & \quad - z(a_1 a_{n-1} a_3 \Delta(4, \ldots, n-3) + a_2 a_n a_{n-2} \Delta(3, \ldots, n-4)) \end{aligned}$$

We now may use $(*)$ again on the terms

$$a_1 a_{n-1} a_3 \Delta(4, \ldots, n-3) + a_2 a_n a_{n-2} \Delta(3, \ldots, n-4)$$

and continue the process indefinitely.

At each stage, we get an expression of the form $U - Vz$. We then set the term $U = 0$ in all these equations. This leads to the following system of equations:

$$\begin{aligned} a_n y_1 + a_1 y_{n-1} &= 0\,, \\ a_n a_2 y_{n-2} + a_1 a_{n-1} y_2 &= 0\,, \\ a_n a_2 a_{n-2} y_3 + a_1 a_{n-1} a_3 y_{n-3} &= 0\,, \\ a_n a_2 a_{n-2} a_4 y_{n-4} + a_1 a_{n-1} a_3 a_{n-3} y_4 &= 0\,, \\ a_n a_2 a_{n-2} a_4 a_{n-4} y_5 + a_1 a_{n-1} a_3 a_{n-3} a_5 y_{n-5} &= 0\,, \\ &\vdots \end{aligned} \tag{3}$$

The lth equation is

$$\begin{aligned} &a_n a_2 a_{n-2} a_4 a_{n-4} \dots a_{n-2k} y_{2k+1} \\ &\qquad + a_{n-1} a_1 a_{n-3} a_3 \dots a_{2k-1} a_{n-(2k-1)} a_{2k+1} y_{n-(2k+1)} = 0 \end{aligned}$$

for $l = 2k + 1$ and

$$a_n a_2 a_{n-2} a_4 \cdots a_{2k} y_{n-2k} + a_1 a_{n-1} a_3 a_{n-3} \dots a_{2k-1} a_{n-(2k-1)} y_{2k} = 0$$

for $l = 2k$. The last term occurs for $l = (n-1)/2$.

If $y_1, \dots, y_{n-1}$ are chosen to satisfy the system, then we obtain $\det(zI + B) = z^n + a_1 \dots a_n$, for that choice of A.

Suppose $a_1, \dots, a_n$ are nonzero. Then we can solve the system for $y_1, \dots, y_{n-1}$.

In fact, we may take

$$\begin{aligned} y_1 &= -a_1 x_1, \quad & y_{n-1} &= a_n x_1, \\ y_2 &= -a_n a_2 x_2, \quad & y_{n-2} &= a_1 a_{n-1} x_2, \\ y_3 &= -a_1 a_3 a_{n-1} x_3, \quad & y_{n-3} &= a_n a_2 a_{n-2} x_3, \quad \text{etc.}, \end{aligned}$$

for any elements $x_1, \dots, x_{(n-1)/2}$ of R.

Thus there are $(n-1)/2$ "free" parameters and for each choice, the corresponding matrix B has characteristic polynomial $x^n - a_1 \dots a_n$.

A particular case of interest is the case where $a_1, \dots, a_n$ are equal. Then we can take

$$y_1 = x, \quad y_2 = -x, \quad y_3 = x, \quad y_4 = -x, \quad \dots, \quad y_{n-1} = -x$$

for any x in R.

Note that if F is a field and $X \in GL(n, F)$ is a nonderogatory matrix with its eigenvalues in F, then X is similar over F to a matrix A of the form above with the y_i nonzero and hence if n is odd, it follows that $X = YZ$ where Z is similar to P and Y has characteristic polynomial $z^n - \det X$. In particular, if $\det X = 1$, then Y is similar to P (and to P^{-1}). [This is clear if the characteristic of F does not divide n since Y has characteristic polynomial $x^n - 1$, while an argument based on the minors of $xI - Y$ (which we omit) yields the result in general.] Thus we have

THEOREM 2. *Let F be a field, n an odd integer and $A \in SL(n,F)$. Suppose A is nonderogatory and has its eigenvalues in F. Then $A = X^{-1}Y^{-1}XY$ with $X, Y \in GL(n,F)$ and X conjugate to the permutation matrix of the n-cycle $(1\,2\,3\ldots n)$.*

We conclude this section with an application of Theorem 1 to a factorization theorem of P. Y. Wu. Wu proved that if $A \in M_n(\mathbb{C})$ has determinant 0, then A is the product of two nilpotent matrices except when $n = 2$ and $A \neq A^2 = 0$. His proof relies on the Jordan normal form. Using Theorem 1, the result can be proved for all fields. If A is nilpotent, then A is similar to its Jordan form over F and Wu's arguments work. The other extreme case is where 0 is a simple eigenvalue of A. In this case, using a similarity and Theorem 1, we may write

$$A = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & L & \\ 0 & & \end{pmatrix} \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & U & \\ 0 & & \end{pmatrix}$$

with $L = (l_{ij})$ lower-triangular and $U = (u_{ij})$ upper-triangular.

But then

$$A = \begin{pmatrix} 0 & \ldots\ldots\ldots\ldots & 0 \\ l_{11} & 0\ \ldots\ldots\ldots & 0 \\ l_{21} & l_{22}\ \ 0\ \ \cdots & 0 \\ \vdots & \ddots & \vdots \\ l_{m1} & \ldots\ldots\ldots & l_{mm}\ \ 0 \end{pmatrix} \begin{pmatrix} 0 & u_{11}\ldots\ldots\ldots & u_{1m} \\ 0 & 0\ \ u_{22}\ \ \cdots & u_{2m} \\ \vdots & & \ddots\ \ \vdots \\ \vdots & & \ u_{mm} \\ \ldots\ldots\ldots\ldots & 0\ \ 0 & \end{pmatrix}$$

(where $m = n - 1$), proving the result.

In the general case, A is similar to $A_1 \oplus A_2$ where A_1 is nilpotent and A_2 is nonsingular. The result is obtained by combining Wu's argument for A_1 with the argument above for $0 \oplus A_2$. The details are omitted. Sourour [SOU2] has obtained another proof of this result.

2. Multiplicative commutators over rings. Let R be a commutative ring with identity and let $A \in SL(n,R)$. One can ask whether A can be written as a commutator $[X,Y]$ with $X, Y \in GL(n,R)$. That the answer is No in general even for "nice" rings can be immediately seen from the fact that $SL(2,\mathbb{Z})/[SL(2,\mathbb{Z}), SL(2,\mathbb{Z})]$ has order 12. Newman [NEW] considered the problem of determining whether every element of $SL(n,\mathbb{Z})$ can be expressed as a bounded (as a function of n) number of commutators and proved that for $n \geq 3$, the answer is Yes with a bound of the form $c \log n + d$ where c, d are explicitly given constants. He posed the problem of whether the number required is bounded. Dennis and Vaserstein [D-V] using very ingenious methods proved that for all sufficiently large n, every element $A \in SL(n,\mathbb{Z})$ can be expressed as the product of six commutators. The problem of extending Theorem 1 to rings has been considered by Vaserstein and Wheland [V-W]. They have proved that it holds with R replaced by a ring with Bass stable rank one. ($\mathbb{Z}$ has Bass stable

rank two.) In the case of $\mathbb{Z}$, they show that every $A \in SL(n,\mathbb{Z})$ for n large can be expressed as the product of six unipotent matrices.

A key result in the discussion of this is the following result of Carter and Keller [C-K1].

THEOREM 3. *Every $A \in SL(n,\mathbb{Z})$ $(n \geq 3)$ can be expressed as the product of $(3n^2 - n)/2 + 36$ elementary matrices.*

No such bound exists for $n=2$ as it is easy to show that if $a, b \in \mathbb{Z}$ with highest common factor $(a,b) = 1$ and $c, d \in \mathbb{Z}$ with $ad - bc = 1$, then expressing $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ as the product of elementary matrices is essentially equivalent to performing the Euclidean algorithm to calculate the highest common factor of a, b. The number of elementary matrices required is at least half the number of steps in the Euclidean algorithm for (a,b) and this can be arbitrarily large.

The case $n = 3$ of Theorem 2 is the key one; a simple induction argument works for $n > 3$, once the case $n = 3$ has been done. In the case $n = 3$, most of the calculation arises in showing that

$$\begin{pmatrix} x & y & 0 \\ z & \omega & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad x, y, z, \omega \in \mathbb{Z},$$

with $x\omega - yz = 1$, can be expressed as the product of a bounded number K of elementary matrices. Carter and Keller [C-K1] show in an ingenious way that $K = 44$ will do. Earlier, using the Riemann hypothesis, van der Kallen had proved that such a K exists, but the Carter–Keller proof does not require any unproved hypotheses.

The results for matrices over $\mathbb{Z}$ do not extend easily to other rings or even to Euclidean domains. Van der Kallen [KAL] shows that in $SL(3, \mathbb{C}[z])$ no such result holds. The fact that the elementary matrices generate $SL(n,\mathbb{Z})$ for $n \geq 3$ and boundedness results on the length of products required have applications in K-theory. Such results do not hold for all principal ideal rings. See Grayson [GRA], Lenstra [LEN].

The situation is better over rings of algebraic integers. Assuming an extended Riemann hypothesis (for a class of L-functions) Cooke and Weinberger [C-W] proved that if $A \in SL(2,R)$ where R is the ring of algebraic integers in a finite extension of $\mathbb{Q}$ and the group of units $U(R)$ is infinite, then A can be written as the product of nine elementary matrices. It then follows that for every $n \geq 2$, there exists a function $f(n)$ such that every $A \in SL(n,R)$ is the product of $f(n)$ elementary matrices. Carter and Keller [C-K2] using class-field theory succeeded in proving a boundedness result (the nine in the Cooke–Weinberger theorem is replaced by a bound depending on the discriminant of the maximal order) without having to assume any unproved hypotheses. Note that from the Dirichlet unit theorem, if R is the ring of algebraic integers in a finite extension $\mathbb{Q}(\alpha)$ of $\mathbb{Q}$, then $U(R)$ is infinite if $[\mathbb{Q}(\alpha) : \mathbb{Q}] \neq 2$. Length questions for expressing elements

of $GL(n,R)$ and $SL(n,R)$ in terms of other generators have been considered particularly in the context of geometry. See, for example, Eller's survey [ELL].

The related problem of identifying sets of generators for the group $SL(n,\mathbb{Z})$ has also been considered by several authors. For $n \geq 3$, every subgroup of $SL(n,\mathbb{Z})$ of finite index contains a congruence subgroup of level m $(= \{A \in SL(n,\mathbb{Z}) \mid A \equiv I \bmod m\})$ for some $m \geq 1$ [B-M-S]. This result greatly restricts the normal structure of $SL(n,\mathbb{Z})$. Trott [TRO] has proved that if J is the upper Jordan block corresponding to $(x-1)^n$ and $B = I_n + E_{n1}$ where E_{n1} is the matrix with its $(n,1)$ entry equal to 1 and all other entries 0, then J, B generate $SL(n,\mathbb{Z})$. It is well known that J, J^T generate $M_n(\mathbb{Z})$ as a ring. Recently Gow and Tamburini [G-T] have proved the very interesting result that for $n \neq 4$, J, J^T generate $SL(n,\mathbb{Z})$ as a group.

The conjugacy problem in $GL(n,\mathbb{Z})$ has been the subject of much recent work.

For $A \in M_n(\mathbb{Z})$, let $\mathrm{orb}(A) = \{B \in M_n(\mathbb{Z}) \mid \text{there exists } Q \in M_n(\mathbb{Q}) \text{ with } Q^{-1}AQ = B\}$.

Then $\mathrm{orb}(A)$ is the union of $GL(n,\mathbb{Z})$-orbits. The Latimer–MacDuffee theorem states that if the characteristic polynomial of A is irreducible in $\mathbb{Z}[x]$, the number of $GL(n,\mathbb{Z})$ orbits in $\mathrm{orb}(A)$ equals the class number of $\mathbb{Z}[\theta]$ where θ is an eigenvalue of A in $\mathbb{C}$. One can deduce from this that $\mathrm{orb}(A)$ is the union of finitely many $GL(n,\mathbb{Z})$-orbits if and only if A is diagonalizable over $\mathbb{C}$. We have shown that if $A \in M_n(\mathbb{Z})$, then $\mathrm{orb}(A)$ consists of one $GL(n,\mathbb{Z})$-orbit if and only if the minimal polynomial $m(x)$ of A has the factorization

$$m(x) = p_1(x) \dots p_r(x)$$

where $r \geq 1$ and $p_1(x), \dots, p_r(x)$ are distinct irreducible polynomials such that

(i) $\mathrm{resultant}(p_i, p_j) = 1$ for $1 \leq i \neq j \leq r$ and

(ii) $\mathbb{Z}[\theta_i]$ has class number one where θ_i is a root of the equation $p_i(x) = 0$ $(i = 1, \dots, r)$.

In particular, if A has its eigenvalues in $\mathbb{Z}$, then $\mathrm{orb}(A)$ consists of one $GL(n,\mathbb{Z})$-orbit if and only if

$$(A - aI)(A - (a+1)I) = 0$$

for some $a \in \mathbb{Z}$. The special case where A is an idempotent is well known.

No satisfactory canonical form is known for representing a conjugacy class in $GL(n,\mathbb{Z})$. For $A \in GL(n,\mathbb{Z})$ with irreducible characteristic polynomial the Latimer–MacDuffee theorem [L-MAC], [TAU1] shows that there are only finitely many $GL(n,\mathbb{Z})$-conjugacy classes in $\mathrm{orb}(A)$. An interesting attempt to find a "companion-matrix-like" representative in each $GL(n,\mathbb{Z})$-class was made by Ochoa [OCH] and while the results do not appear to hold in the generality which Ochoa suggests (see Rehm [REH] for discussion and a derivation of some of Ochoa's results), they are important. The Ochoa matrix representations differ from companion matrices in that the last *two* rows have several nonzero entries. It may be worth pointing out here that even the first step in transforming a matrix

A to its companion (or rational canonical) form can fail over $\mathbb{Z}$, since no nonzero vector v need exist with the property that $\{v, Av\}$ can be included in an integral basis. For example, if $A \equiv aI \bmod p$ for some $a \in \mathbb{Z}$, and some prime p with $(p, n) = 1$, it is impossible for A to be integrally similar to a matrix with a zero on the diagonal, since if so, reading mod p yields a nonscalar matrix similar to aI over $GF(p)$. Deciding whether two elements $A, B \in M_2(\mathbb{Z})$ are integrally similar can be done using the continued fraction algorithm or the fact that $PSL(2, \mathbb{Z})$ is isomorphic to the free product $C(2) * C(3)$ of the cyclic groups of order 2, 3, respectively and that the conjugacy problem is algorithmically solvable in the free product. See Campbell and Trouy [C-T]. A discussion of the conjugacy problem for $GL(n, \mathbb{Z})$ $(n \geq 2)$ and some other arithmetic groups can be found in Grunewald [GRU]. See also Gustafson [GUS].

3. Additive commutators. Let F be a field and let $A \in M_n(F)$ with $\operatorname{tr} A = 0$. Albert and Muckenhoupt [A-M] proved that A is an additive commutator $PQ - QP$ for some $P, Q \in M_n(F)$. This is the (much easier) additive analogue for the Lie algebra $sl_n(F)$ of Thompson's result for the group $SL(n, F)$ discussed in Section 1. Suppose $A \in M_n(F)$ is not scalar. Choose a vector v with v, Av linearly independent and set $v_1 = v$, $v_2 = Av$ and extend to a basis $v_1, v_2, \ldots, v_n$ of F^n. Using this basis we see that A is similar to a matrix with its $(1, 1)$ entry 0. Using induction one obtains the following result.

THEOREM 4. *Let F be a field and $A \in M_n(F)$ nonscalar with $\operatorname{tr} A = 0$. Then A is similar to an element $B = (b_{ij}) \in M_n(F)$ with $b_{11} = \ldots = b_{nn} = 0$.*

Several improvements of this result are known for large fields. For example, if F is the complex field $\mathbb{C}$, Gaines [GAI] and Fillmore [F] proved that the zero diagonal form can be achieved using a unitary similarity. Choi, Laurie and Radjavi [C-L-R] proved that if $F = \mathbb{C}$, $\operatorname{tr} A = 0$ and rank $A \geq 2$, then A is similar to a matrix B as above with zero diagonal and all off-diagonal entries nonzero. [Related to this is a result of Gaines [GAI] establishing that a nonscalar matrix A over an infinite field F is similar to one with all its entries nonzero.] West and the present author [L-W] have proved that if F has at least seven elements and $A \in M_n(F)$ is nonscalar, has trace zero and rank at least two, then A is similar over F to a matrix $B = (b_{ij})$ with the diagonal of B equal to zero, and all the entries $b_{i,i+1}$, $b_{i+1,i}$ $(i = 1, \ldots, n-1)$ nonzero. Writing B_L, respectively B_U, for the matrices obtained from B by replacing the entries b_{ij} $(j > i)$, respectively b_{ij} $(i > j)$, by 0, we see that

$$B = B_L + B_U = B_L - (-B_U)$$

is the sum and difference of similar nilpotent matrices.

Suppose R is a ring and $A \in M_n(R)$ has trace 0. It is a natural question to ask whether A must be expressible in the form $PQ - QP$ with $P, Q \in M_n(R)$. Taking x, y, z to be commuting indeterminates over a field F, it is not hard to show that $A = \left(\begin{smallmatrix} x & y \\ z & -x \end{smallmatrix}\right)$ cannot be expressed in this way over the polynomial ring

$F[x, y, z]$. The problem is, however, of great interest over $\mathbb{Z}$ and certain principal ideal domains. See [LIS1].

If R is a Euclidean ring and $A \in M_2(R)$ has trace zero, then one can deduce that A can be expressed as a commutator $PQ - QP$ with $P, Q \in M_2(R)$ from the fact that every element in $\mathbb{Z}^3$ can be expressed as a vector cross-product $u \times v$ with $u, v \in \mathbb{Z}^3$. This latter fact follows from the unimodular row lemma. It appears to have been initially observed by Hermite and had been studied by Lissner ([LIS1], [LIS2]), and his work has been further extended by Towber [TOW]. (I am grateful to Irving Kaplansky for this reference.) It was independently discovered by Vaserstein [VAS]. The problem for $n > 2$ requires a different approach and Vaserstein [VAS] poses the problem of whether every $A \in M_n(\mathbb{Z})$ $(n \geq 3)$ with trace zero can be expressed as a commutator.

This has now been answered affirmatively by the author and his student Robert Reams. The method used is as follows: Using the corresponding result for fields, one shows that if $A \in M_n(\mathbb{Z})$ with $\operatorname{tr} A = 0$, there exists a positive integer k such that $kA = PQ - QP$ for some $P, Q \in M_n(\mathbb{Z})$. One chooses such a representation with smallest possible k and if $k > 1$, one chooses a prime divisor p of k. If P or Q is nonderogatory when regarded as an element of $M_n(GF(p))$, one shows that such a representation exists with k replaced by k/p giving a contradiction and thus forcing $k = 1$ as required. If neither P nor Q is nonderogatory in $M_n(GF(p))$, then by studying the centralizer of Q, one attempts to replace P by a matrix P_0 with $P_0 \bmod p$ nonderogatory. When this strategy fails, an analysis of the situation where $P \bmod p$ has minimal polynomial of degree $n - 1$ is used to complete the (lengthy) proof. See [L-R] for details.

4. Factorizations preserving forms. Let F be a field and let $A \in M_n(F)$. It is well known (and the result is sometimes attributed to Frobenius) that A can be written as the product ST of two symmetric matrices with one, S say, invertible. The standard proof uses the rational canonical form and the fact that the result holds for companion matrices. In fact, if

$$f(x) = x^n - b_{n-1}x^{n-1} - \ldots - b_0$$

and the corresponding companion matrix

$$C(f) = \begin{bmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 0 & 1 \\ b_0 & b_1 & \ldots & \ldots & b_{n-2} & b_{n-1} \end{bmatrix}$$

one defines

$$V(f) = \begin{bmatrix} -b_1 & -b_2 & \cdots & -b_{n-1} & 1 \\ -b_2 & -b_3 & \iddots & 1 & 0 \\ -b_3 & \iddots & \iddots & 0 & 0 \\ \vdots & \iddots & \iddots & \vdots & \vdots \\ -b_{n-1} & \iddots & & & \vdots \\ 1 & 0 & \cdots & & 0 \end{bmatrix}$$

and observes that $V(f)C(f)$ is symmetric so $C(f) = (V(f))^{-1}(V(f)C(f))$ is the product of two symmetric matrices. This ingenious trick appears to have been first discovered by Williamson in the 1930s and has been rediscovered by several authors since. Because of the large number of zeros in $V(f)$, the argument enables one to get information on the signature of S when a matrix A is expressed as ST with S, T symmetric and S invertible. See [LAF3] for details.

Since $A = ST$, with S, T symmetric and S invertible, $S^{-1}AS = TS = A^T$, so A is similar to its transpose via a symmetric matrix. If A is nonderogatory and we also have $L^{-1}AL = A^T$, then LS^{-1} commutes with A, so $L = g(A)S$ for some polynomial $g(x) \in F[x]$. But $AS = SA^T$ implies that $g(A)S = S^T g(A^T) = L^T$, so L is symmetric. Hence every matrix which transforms A to A^T in this case is symmetric. A novel proof of this result was discovered by Taussky and Zassenhaus [T-Z] who proved by dimension arguments that for A nonderogatory, every matrix solution X of the linear system $AX - XA^T = 0$ is symmetric.

In discussing many similarity results, symmetric matrices and involutions appear to play dual roles. In Section 1, we have discussed factorization involving involutions. A new factorization combining symmetric matrices and involutions was obtained by Gow [GOW1]. He proved that if F does not have characteristic 2 and $A \in GL(n, F)$, then $A = SJ$ with S symmetric and J an involution. This result is best thought of in terms of congruence. Two elements $H, K \in M_n(F)$ are congruent if there exists $P \in GL(n, F)$ with $K = P^T HP$. The study of congruence over general fields is much more difficult than the study of similarity and no satisfactory canonical form is known. See Ballantine and Yip [B-Y], Riehm [RIE] and Waterhouse [WAT] for a number of results in this area. Gow's result can be stated in the equivalent form: if $A \in GL(n, F)$, then A is congruent to A^T and in fact $A^T = P^T AP$ for some involution P. If the field F is algebraically closed (and of characteristic not two), then from $A = SJ$ with S symmetric and $J^2 = I$, we deduce that A is congruent to $S^{1/2}JS^{-1/2}$ where we denote by $S^{1/2}$ a symmetric square root of S, so A is congruent to an involution. More generally, if S is congruent to the identity matrix, say $S = R^T R$, then $R^T AR = R^{-1}JR$ is an involution. This situation occurs not only if the field F is algebraically closed but also if $F = \mathbb{R}$, the field of real numbers, and S is positive definite, or if F is finite and $\det S$ is a square.

In these cases the problem of congruence is reduced to considering congruence on the conjugacy classes of involutions. In fact, the arguments of Gow and the author [G-L] show that in Gow's factorization theorem, J may be chosen in the

conjugacy class of

$$\operatorname{diag}(\underbrace{1,\dots,1}_{k},\underbrace{-1,\dots,-1}_{n-k})$$

where $k = [(n+1)/2]$.

Gow and the author [G-L] have shown that an element $A \in GL(n,F)$ is the product of two skew-symmetric matrices in $GL(n,F)$ if and only if n is even and A is similar to a matrix of the form $B \oplus B$ with $B \in GL(n/2,F)$. The proof uses the theory of bilinear forms and in particular, the theory of symplectic forms. The key step is a proof of the following result.

THEOREM 5. *Let F be a field and let f, g be nonsingular symplectic (= alternating) forms on a $2n$-dimensional space V over F. Then V has a decomposition of the form $V_1 \oplus V_2$ where V_1, V_2 are common maximal isotropic (of dimension n) subspaces for both forms.*

Graham Higman proved in the late 1950s that if f, g is a pair of symplectic forms as in the theorem, then f, g have a common maximal isotropic subspace and Alperin [ALP] used the result to get bounds on the size of maximal abelian subgroups of finite p-groups. To illustrate the connection here, suppose P is a finite p-group with its commutator subgroup $P' \le Z(P)$ and elementary abelian of order p^k and that $P/Z(P)$ is elementary abelian of order p^n. Let $z_1,\dots,z_k$ be a basis for P'. Then we define $\overline{P} = P/Z(P)$, $\overline{u} = uZ(P)$ for $u \in P$, and

$$[\overline{x},\overline{y}] = [x,y] = \prod_{i=1}^{k} z_i^{f_i(x,y)}$$

where $f_i : \mathbb{Z}_p \times \mathbb{Z}_p \to \mathbb{Z}_p$ is a symplectic form ($i = 1,\dots,k$) on the space $\overline{P}$. Note that x, y commute if and only if $\overline{x},\overline{y}$ belong to a common isotropic subspace for all the forms f_i. The case $k = 1$ arises when P is an extraspecial p-group and we conclude that $n = 2m$ is even and that all maximal abelian subgroups of P have order p^{m+1} (cf. Huppert [HUP], III, (13.7)). In the case $k = 2$, Theorem 5 implies that P has maximal abelian subgroups A_1, A_2 of order p^{m+2} such that $P = A_1A_2$. To obtain a corresponding result for general k, it is necessary to know the maximal dimension of a subspace isotropic with respect to k symplectic forms on an n-dimensional vector space over a field F. For F algebraically closed, this problem has been solved recently by Buhler, Gupta and Harris [B-G-H], using the methods of algebraic geometry (in particular, the theory of Schubert varieties). Their result is

THEOREM 6. *Let F be an algebraically closed field and let V be an n-dimensional vector space over F and let $f_1,\dots,f_k$ ($k > 1$) be symplectic forms on V. Then V has a subspace U of dimension $[(2n+k)/(k+2)]$ on which all the f_i are isotropic. Furthermore, there is a set of k forms for which $[(2n+k)/(k+2)]$ is the maximum dimension of such a subspace.*

For general fields F, they prove that the maximum dimension of a common isotropic subspace is bounded above by the one obtaining in the algebraically closed field case and thus they are able to use Theorem 6 to construct p-groups with all their maximal abelian subgroups of relatively small order. In particular, they prove that there exists a finite p-group ($p > 2$) of order p^n with all its abelian subgroups of order at most p^d where $d = [\sqrt{8n+9} - 3]$. In the opposite direction, Burnside observed that if Q is a finite p-group and A is a maximal abelian normal subgroup of Q, then $|Q| \leq h|A|$ where h is the p-part of $|GL(k, F)|$, F being the field of p elements, and where $|A| = p^k$, so $|A|$ is at least $p^{\sqrt{2n}}$ where $|Q| = p^n$.

By analogy with the result that every matrix is the product of two symmetric matrices, it is proved in [LAF5] that if F is algebraically closed of characteristic different from two, $n > 2$ is even and $A \in GL(n, F)$, then A is the product of five skew-symmetric elements of $GL(n, F)$. The number "five" is best possible here but no simple characterizations are available of products of three skew-symmetric matrices (which is a property invariant under congruence) or of four skew-symmetric matrices (which is a property invariant under similarity).

Over the field of real numbers $\mathbb{R}$, every matrix A is the product of a positive semidefinite symmetric matrix S and an orthogonal matrix V (this is the well-known polar decomposition). Over the complex field $\mathbb{C}$, the corresponding result holds with "symmetric" replaced by "Hermitian" and "orthogonal" by "unitary". Choudhury and Horn [cC-H] considered the problem of determining whether a matrix $A \in M_n(F)$ (where F is an algebraically closed field of characteristic not equal to two) can be factored as $A = SV$ with S symmetric and V orthogonal (that is, $S^T = S$, $V^T = V^{-1}$). Clearly, if this occurs $AA^T = S^2$ is similar to $A^T A = V^{-1} S^2 V$. They prove the decomposition holds if $A \in GL(n, F)$ and in some other cases. The problem has recently been completely settled by Kaplansky [KAP2] who shows that the condition that AA' and $A'A$ be similar is also sufficient for the decomposition. His proof uses the theory of bilinear forms in a clever manner. The crucial result required is that if $A \in M_n(F)$ with AA' nilpotent, then AA' has a symmetric square root if (and only if) AA' is similar to $A'A$.

5. Relationship between multiplicative and additive commutators. In this section, we write (P, Q) for the additive commutator $PQ - QP$ of two matrices P, Q.

Suppose F is a field and $A \in M_2(F)$ with $\operatorname{tr} A = a$, $\det A = b$. The Cayley–Hamilton theorem states that

$$A^2 - aA + bI = 0\,,$$

so if $X \in M_2(F)$, we have

$$(A^2 - aA + bI, X) = 0\,.$$

This yields

$$(A, X)A + A(A, X) - a(A, X) = 0\,,$$

so

$$(A, X)A = ((\operatorname{tr} A)I - A)(A, X)\,.$$

The matrix $A_2 = (\operatorname{tr} A)I - A$ is similar to A and if (A, X) is nonsingular, we have $A_2 = (A, X)A(A, X)^{-1}$. Also $(x^2 - ax + b)I = (xI - A_1)(xI - A_2)$ where $A_1 = A$, and $(A, X)^2$ is a scalar matrix for all $X \in M_2(F)$. This relationship between factorizing the characteristic polynomial of A as a product of linear factors $xI - B$ with B conjugate to A and additive commutators of the form (A, Y) can be generalized to $n > 2$.

For $n = 3$ and $A \in M_3(F)$ with characteristic polynomial $x^3 - ax^2 + bx - c$, we have for all $X \in M_3(F)$,

$$(A^3 - aA^2 + bA - cI, X) = 0\,,$$

so

$$A^2C + ACA + CA^2 - a(AC + CA) + bC = 0$$

where $C = (A, X)$.

If C^{-1} exists, we write $D = C^{-1}AC$ and the last equation becomes

$$D^2 + DA + A^2 - aD - aA + bI = 0$$

and again taking the commutator with A, we obtain

$$D(D, A) + (D, A)D + (D, A)A - a(D, A) = 0\,.$$

Assuming (D, A) is invertible, we thus have

$$(D, A)^{-1}D(D, A) = aI - D - A$$

so $(\operatorname{tr} A)I - D - A$ is similar to D via an element of the form (A, Y).

A calculation essentially due to Wedderburn and given in Rowen [ROW] shows that $(D, A)^3$ is a scalar matrix. Using the Cayley–Hamilton theorem again, we can replace $(A, X)^{-1}$ and we deduce that if

$$P = ((A, X)^2A(A, X) - A(A, X)^3, A)$$

then P need not be scalar but P^3 is a scalar matrix for all $A, X \in M_3(F)$. So $M_3(F)$ has a "cube central" central polynomial.

Meehan and the author [L-M1] have proved the following factorization result for general n.

THEOREM 7. *Let F be an infinite field and let $f(x) \in F[x]$ be a monic polynomial of degree n. Let A_1 be the companion matrix of $f(x)$. Then there exists an $X \in M_n(F)$ for which the following sequences are defined* (*that is*, *the requisite matrices are invertible*):

$$D_1 = C_1 = (A_1, X), \qquad A_2 = C_1^{-1}A_1C_1,$$
$$C_2 = (A_2, X), \qquad D_2 = C_1 + C_2, \qquad A_3 = D_2^{-1}A_2D_2\,,$$

and in general

$$C_r = (A_r, X), \quad D_r = D_{r-1} + C_r, \quad A_{r+1} = D_r^{-1}A_rD_r \quad \textit{for } r = 1, 2, \ldots, n-1.$$

For any such X, we have the identity

$$f(x)I_n = (xI_n - A_1)\dots(xI_n - A_n)\,.$$

When F has characteristic zero, we can exhibit an example of such an X. We can take $X = D^{-1}JD$ where J is the lower-triangular Jordan block with characteristic polynomial $(x-1)^n$ and $D = \mathrm{diag}((n-1)!, (n-2)!, \dots, 4!, 3!, 2!, 1!, 1)$.

As a result, we can deduce several consequences on the existence of conjugates of a nonderogatory matrix A satisfying several identities. For example, by comparing the coefficients of x^{n-1} and x^0 we see that there exist conjugates $A_1, \dots, A_n$ of A_1 satisfying

$$(*) \qquad A_1 + \dots + A_n = (\mathrm{tr}\, A)I_n\,, \qquad A_1A_2\dots A_n = (\det A)I_n\,.$$

EXAMPLE. If $K = F(\theta)$ is a field extension of degree n and K is Galois over F and $f(x) = \mathrm{Irr}(\theta, F, x)$, the (monic) irreducible polynomial satisfied by θ over F, then all the roots of $f(x) = 0$ are of the form $g_i(\theta)$ for some $g(x) \in F[x]$, with $g_1(x) = x$.

So

$$f(x) = (x-\theta)(x-g_2(\theta))\dots(x-g_n(\theta))\,.$$

If A is the companion matrix of $f(x)$, then

$$f(x)I_n = (xI_n - A)(xI_n - g_2(A))\dots(xI_n - g_n(A))$$

and the factors $xI - g_2(A), \dots, xI - g_n(A)$ are uniquely determined up to order. In this example, the equations $(*)$ correspond to the evaluation of $\mathrm{trace}_{K/F}(\theta)$ and $\mathrm{Norm}_{K/F}(\theta)$.

In Theorem 7, the matrices $A_1, \dots, A_n$ cannot in general be taken to commute. A necessary and sufficient condition for the existence of a factorization with commuting $A_1, \dots, A_n$ (with A_1 the companion matrix of $f(x)$) over every field F has been found [L-M2]. In the (very) special case of an irreducible polynomial $f(x)$, a factorization of $f(x)I_n$ involving only commuting A_i exists if and only if $F[\theta]$ is the splitting field of $f(x)$, for θ any root of the equation $f(x) = 0$. In this case, the factorization is the one described in the Example. Of particular interest here is the case of finite fields F. In this case, every finite extension is Galois, so we have a factorization of every irreducible polynomial of degree n over F using commuting $n \times n$ matrices over F.

References

[ALB] A. A. Albert and B. Muckenhoupt, *On matrices of trace zero*, Michigan J. Math. 4 (1957), 1–3.

[ALP] J. L. Alperin, *Large Abelian subgroups of p-groups*, Trans. Amer. Math. Soc. 117 (1965), 10–20.

[BAL] C. S. Ballantine, *Products of positive definite matrices*, *III*, J. Algebra 10 (1968), 174–182; *IV*, Linear Algebra Appl. 3 (1970), 79–114.

[B-Y] C. S. Ballantine and E. L. Yip, *Congruence and conjunctivity of matrices*, Linear Algebra Appl. 32 (1980), 159–198.

[B-M-S] H. Bass, J. Milnor and J.-P. Serre, *Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$)*, IHES Publ. Math. 33 (1967), 59–137.

[B-G-H] J. Buhler, R. Gupta and J. Harris, *Isotropic subspaces for skewforms and maximal Abelian subgroups of p-groups*, J. Algebra 108 (1987), 269–279.

[C-T] J. T. Campbell and E. C. Trouy, *When are two elements of $GL(2,\mathbb{Z})$ similar*, Linear Algebra Appl. 157 (1991), 175–184.

[C-K1] D. Carter and G. Keller, *Elementary expressions for unimodular matrices*, Comm. Algebra 12 (1984), 379–389.

[C-K2] —, —, *Bounded elementary generation of $SL_n(\theta)$*, Amer. J. Math. 105 (1983), 673–687.

[C-L-R] M.-D. Choi, C. Laurie and H. Radjavi, *On commutators and invariant subspaces*, Linear and Multilinear Algebra 9 (1981), 329–340.

[C-H] D. Choudhury and R. A. Horn, *A complex orthogonal-symmetric analog of the polar decomposition*, SIAM J. Algebraic Discrete Methods 8 (1987), 218–225.

[C-W] G. Cooke and P. J. Weinberger, *On the construction of division chains in algebraic number rings, with applications to SL_2*, Comm. Algebra 3 (1975), 481–524.

[D-V] R. K. Dennis and L. N. Vaserstein, *On a question of M. Newman on the number of commutators*, J. Algebra 118 (1988), 150–161.

[DJO] D. Ž. Djoković, *Product of two involutions*, Arch. Math. (Basel) 18 (1967), 582–584.

[ELL] E. W. Ellers, *Classical Groups*, in: Generators and Relations in Groups and Geometries, NATO Adv. Sci. Inst. Ser. C, Kluwer, Dordrecht, 1991, 1–45.

[F] P. A. Fillmore, *On similarity and the diagonal of a matrix*, Amer. Math. Monthly 76 (1969), 167–169.

[GAI] F. J. Gaines, *Kato–Taussky–Wielandt commutator relations*, Linear Algebra Appl. 1 (1968), 127–138.

[G-L-R] I. Gohberg, P. Lancaster and L. Rodman, *Invariant Subspaces of Matrices with Applications*, Wiley, New York, 1986.

[GOW1] R. Gow, *The equivalence of an invertible matrix to its transpose*, Linear and Multilinear Algebra 8 (1980), 329–336.

[GOW2] —, *Products of two involutions in classical groups of characteristic* 2, J. Algebra 71 (1981), 583–591.

[G-L] R. Gow and T. J. Laffey, *Pairs of alternating forms and products of two skew-symmetric matrices*, Linear Algebra Appl. 63 (1984), 119–132.

[G-T] R. Gow and C. Tamburini, *Generation of $SL(n,\mathbb{Z})$ by a Jordan unipotent matrix and its transpose*, to appear.

[GRA] D. R. Grayson, *SK_1 of an interesting principal ideal domain*, J. Pure Appl. Algebra 20 (1981), 157–163.

[G-P-R] L. Grunenfelder, L. Paré and H. Radjavi, *On a commutator theorem of R. C. Thompson*, Linear and Multilinear Algebra 16 (1984), 129–131.

[GRU] F. Grunewald, *Solution of the conjugacy problem in certain arithmetic groups*, in: Word Problems II, S. I. Adian, W. W. Boone and G. Higman (eds.), North-Holland, 1980, 101–139.

[GUS] W. Gustafson, *Modules and matrices*, Linear Algebra Appl. 157 (1991), 3–19.

[G-H-R] W. Gustafson, P. Halmos and H. Radjavi, *Products of involutions*, ibid. 13 (1976), 157–162.

[H-OM] A. J. Hahn and O. T. O'Meara, *The Classical Groups and K-theory*, Grundlehren Math. Wiss. 291, Springer, New York, 1989.

[HON] K. Honda, *On commutators in finite groups*, Comment. Math. Univ. St. Paul. 2 (1953), 9–12.

[HUP] B. Huppert, *Endliche Gruppen I*, Springer, Berlin, 1967.

[KAL] W. van der Kallen, $SL(\mathbb{C}[x])$ *does not have bounded word length*, in: Proc. Algebraic K-Theory Conf., Lecture Notes in Math. 996, Springer, 1982, 356–361.

[KAP1] I. Kaplansky, *Linear Algebra and Geometry*, Allyn and Bacon, 1963.

[KAP2] —, *Algebraic polar decomposition*, SIAM J. Matrix Anal. Appl. 11 (1990), 213–217.

[LAF1] T. J. Laffey, *Algebras generated by two idempotents*, Linear Algebra Appl. 35 (1985), 45–53.

[LAF2] —, *Factorizations of matrices involving symmetric matrices and involutions*, in: Current Trends in Matrix Theory, North-Holland, 1987, 175–198.

[LAF3] —, *Matrix factorization with symmetry properties*, in: Applications of Matrix Theory, Clarendon Press, Oxford, 1989, 63–70.

[LAF4] —, *Factorizations of integer matrices as products of idempotents and nilpotents*, Linear Algebra Appl. 120 (1989), 81–94.

[LAF5] *Products of matrices*, in: Generators and Relations in Groups and Geometries, NATO Adv. Sci. Inst. Ser. C, Kluwer, Dordrecht, 1991, 95–123.

[L-M1] T. J. Laffey and E. Meehan, *An extension of a factorization theorem of Wedderburn to matrix rings*, Linear Algebra Appl. 172 (1992), 243–260.

[L-M2] —, —, *Factorization of polynomials using commuting matrices*, ibid., to appear.

[L-M3] —, —, *Factorization of polynomials involving unipotent Jordan blocks*, Appl. Math. Lett. 5 (1992), 29–33.

[L-R] T. J. Laffey and R. Reams, *Integral similarity and commutators of integral matrices*, Linear Algebra Appl., to appear.

[L-W] T. J. Laffey and T. T. West, *Polynomial commutators*, Bull. Irish Math. Soc., to appear.

[L-MAC] C. G. Latimer and C. C. MacDuffee, *A correspondence between classes of ideals and classes of matrices*, Ann. of Math. 34 (1933), 313–316.

[LEN] H. W. Lenstra, *Grothendieck groups of Abelian group rings*, J. Pure Appl. Algebra 20 (1981), 173–193.

[LIS1] D. Lissner, *Matrices over polynomial rings*, Trans. Amer. Math. Soc. 98 (1961), 285–305.

[LIS2] —, *Outer product rings*, ibid. 116 (1965), 526–535.

[LIU] K.-M. Liu, *Decompositions of matrices into three involutions*, Linear Algebra Appl. 111 (1989), 1–24.

[NEW] M. Newman, *Unimodular commutators*, Proc. Amer. Math. Soc. 101 (1987), 605–609.

[OCH] J. Ochoa, *Un modelo elementel para las clases de ideales de un anillo algebraico*, Rev. Real Acad. Cienc. Madrid 63 (1974), 711–806.

[REH] H. P. Rehm, *On Ochoa's special matrices in matrix classes*, Linear Algebra Appl. 17 (1977), 181–188.

[ROW] L. H. Rowen, *Polynomial Identities in Ring Theory*, Academic Press, New York, 1980.

[SOU1] A. R. Sourour, *A factorization theorem for matrices*, Linear and Multilinear Algebra 19 (1986), 141–147.

[SOU2] —, *Nilpotent factorization of matrices*, ibid. 31 (1992), 303–308.

[TAU1] O. Taussky, *On a theorem of Latimer and MacDuffee*, Canad. J. Math. 1 (1949), 300–302.

[TAU2] —, *Positive definite matrices and their role in the study of the characteristic roots of general matrices*, Adv. in Math. 2 (1967), 175–186.

[T-Z] O. Taussky and H. Zassenhaus, *On the similarity transformation between a matrix and its transpose*, Pacific J. Math. 9 (1959), 893–896.

[THO1] R. C. Thompson, *Commutators in the special and general linear groups*, Trans. Amer. Math. Soc. 101 (1961), 16–33.

[THO2] —, *Commutators of matrices with prescribed determinants*, Canad. J. Math. 20 (1968), 203–221.

[TOW] J. Towber, *Complete reducibility in exterior algebras over free modules*, J. Algebra 10 (1968), 299–309.

[TRO] S. M. Trott, *A pair of generators for the unimodular group*, Canad. Math. Bull. 5 (1962), 245–252.

[VAS] L. N. Vaserstein, *Noncommutative number theory*, *algebraic K-theory and algebraic number theory*, in: Contemp. Math. 83, Amer. Math. Soc., 1985, 445–449.

[V-W] L. N. Vaserstein and E. Wheland, *Factorization of invertible matrices over rings of stable rank one*, preprint, 1990.

[WAT] W. C. Waterhouse, *Pairs of quadratic forms*, Invent. Math. 37 (1976), 157–164.

[WIL] J. Williamson, *The equivalence of non-singular pencils of Hermitian matrices in an arbitrary field*, Amer. J. Math. 57 (1935), 475–490.

[WON1] M. J. Wonenberger, *A decomposition of orthogonal transformations*, Canad. Math. Bull. 7 (1964), 379–383.

[WON2] —, *Transformations which are products of two involutions*, J. Math. Mech. 16 (1966), 327–338.

[WU1] P. Y. Wu, *Products of nilpotent matrices*, Linear Algebra Appl. 96 (1987), 227–232.

[WU2] —, *The operator factorization theorems*, ibid. 117 (1989), 35–63.