ISRAEL JOURNAL OF MATHEMATICS 82 (1993), 281–297

ON EXTENSIONS OF THE BAER–SUZUKI THEOREM

BY

ROBERT M. GURALNICK*

Department of Mathematics
University of Southern Calfornia
Los Angeles, CA 90089-1113, USA
email: guralnic@mtha.usc.edu

AND

GEOFFREY R. ROBINSON**

Department of Mathematics
University of Florida
Gainesville, FL 32611, USA
email: robinson@math.ufl.edu

For John Thompson

ABSTRACT

We find a necessary and sufficient condition for an element of prime order in a finite group to be in a normal p-subgroup. This generalizes the Baer–Suzuki Theorem. Our proof depends on a result about elements of prime order contained in a unique maximal subgroup containing a result of Wielandt. We discuss various consequences, linear and algebraic group versions of the result.

1. Introduction

The Baer–Suzuki Theorem (cf [G1, p. 105]) asserts that if X is a subgroup of a finite group G and every pair of conjugates of X generates a nilpotent subgroup, then the normal closure of X in G is also nilpotent. There are some very short

* Partially supported by NSF grant DMS-91011407.
** Partially supported by NSF grant DMS-9208667.
Received September 21, 1992 and in revised form November 25, 1992

proofs of this result (cf [AL]). Observe also, that it suffices by induction to prove the result for subgroups of prime order.

In this article, we show that a weaker condition actually suffices. The proof, however, depends upon the classification of finite simple groups. We first introduce some notation. If $x, g \in G$, let $x^g = g^{-1}xg$ and let $[x, g] = x^{-1}x^g = x^{-1}g^{-1}xg$. Let $O_p(G)$ denote the largest normal p-subgroup of a finite group G. Let $O_{p'}(G)$ denote the largest normal subgroup of G which is a p'-group. Our first main result is:

THEOREM A: *Let p be a prime. Let G be a finite group. Let $x \in G$ be an element of order p such that $[x, g]$ is an p-element for every $g \in G$. Then $x \in O_p(G)$.*

An easy consequence of the previous result is:

COROLLARY B: *Let p be prime. Let G be a finite group. Let X be a p-subgroup of G such that $[x, g]$ is a p-element for every $x \in X$ and $g \in G$. Then $X \leq O_p(G)$.*

The Baer–Suzuki theorem is an immediate consequence of Corollary B. We should note that if $p = 2$, then Theorem A is equivalent to the Baer–Suzuki Theorem. Thus we need only prove the result for p odd (although our proof does go through for $p = 2$ as well). The proof proceeds as follows — if G is a minimal counterexample, it follows easily by a result of Wielandt [W] that x is contained in a unique maximal subgroup M. We use results of Aschbacher [A2] and Seitz [S] to prove the following result of independent interest:

THEOREM C: *Let x be an element of order p in the finite group G. Assume that x is contained in a unique maximal subgroup M of G and that M contains no nontrivial normal subgroup of G. Then the Sylow p-subgroup of G is cyclic or $O_p(M) = 1$ and $G = A_{2p}$ with $p > 13$.*

If M contains a nontrivial normal subgroup, an easy induction argument completes the proof. If the Sylow p-subgroup of G is cyclic, we use a block theoretic argument to obtain a contradiction.

We also consider the opposite situation — when each $[x, g]$ is a p'-element. In this case, the result follows from the Z^*-theorem and its analogue for odd primes (the proof of which depends upon the classification of finite simple groups).

THEOREM D: *Let G be a finite group. Let $x \in G$ be an element of prime order p.*

(i) *If $[x, g]$ is a p'-element for every $g \in G$, then x is central modulo $O_{p'}(G)$.*

(ii) *If $r \neq p$ is prime and $[x,g]$ is an r-element for every $g \in G$, then x is central modulo $O_r(G)$.*

Theorems A and D yield:

COROLLARY E: *Let G be a finite group and p a prime. Let X be a subgroup of G such that $[x,g]$ is a p-element for every $x \in X$ and $g \in G$. Then $[X,G] \leq O_p(G)$.*

As Aschbacher [A1] has observed, there is a linear version of the Baer–Suzuki Theorem (in fact, the two versions follow from one another). We mention two such results here (Aschbacher's version was under the assumption that $\langle X, X^g \rangle$ is unipotent for all $g \in G$). The first follows immediately from Corollary E. The second depends upon results on algebraic groups.

COROLLARY F: *Let V be a finite dimensional vector space over a field k. Let $G \leq GL(V)$. Let $X \leq G$ such that $[x,g]$ is unipotent for all $x \in X, g \in G$.*
(i) *$[G,X]$ is a unipotent normal subgroup of G.*
(ii) *If X is triangular, then $\langle X^g | g \in G \rangle$ is triangular.*

COROLLARY G: *Let V be a finite dimensional vector space over a field k. Let $G \leq GL(V)$. Assume that either k has characteristic 0 or that G is connected. If $X \leq G$ and $\langle X, X^g \rangle$ is triangular for each $g \in G$, then $\langle X^g | g \in G \rangle$ is triangular.*

The paper is organized as follows. In section 2, we prove Theorem A when the Sylow p-subgroup is cyclic. In section 3, we prove Theorem C, complete the proof of Theorem A and prove Corollary B. In section 4, we prove Theorem D. In section 5, we discuss some consequences of the theorems — in particular, linear versions of the results.

It is difficult to see how these results can be improved much. Probably Theorem A is true if we only assume that x is a p-element and replace the condition that every commutator with x be a p-element by the condition that every commutator with x is either 1 or a p-singular element.

If $x \in S_n$ is a transposition, then for any $g \in S_n$, $[x,g]$ has order 1, 2 or 3. Thus, there is little hope of replacing p by a set of primes in Theorem A. Similarly, if $x \in A_n$ is a 3-cycle, then $[x,g]$ has order 1, 2, 3, or 5 for any $g \in S_n$ (and more generally if $x \in A_n$ is a p-cycle, then any commutator $[x,g]$ has order which is a product of primes all less than $2p$).

It is perhaps worth mentioning an easier result of a similar nature. Let G be a finite group and π is a set of primes. If every commutator in G is a π-element, then G' is a π-group. This follows immediately from the Focal Subgroup Theorem (cf [G1, p. 250]).

2. Cyclic Sylow Subgroups

We prove Theorem A when the Sylow p-subgroup of G is cyclic. Indeed, we will prove a stronger version. The proof is block theoretic. See [F, Chapter VII] for a summary of the results about blocks with cyclic defect groups.

THEOREM 2.1: *Let G be a finite group, x in G a p-element with p a prime. If*

(a) *G has cyclic Sylow p-subgroups, and*

(b) *$[x,g]=1$ or $[x,g]$ is p-singular for each $g \in G$,*
then $x \in O_p(G)$.

Proof: Note that if $p=2$ and G has a cyclic Sylow 2-subgroup, then G has a normal 2-complement. Then the hypothesis implies that $x \in Z(G)$ and the result follows. So assume p is odd. We also may assume that $x \neq 1$.

Let $H = C_G(x)$. Let $1 \in T$ be a right transversal to H in G. If χ is a complex irreducible character of G, then, since $C_x := \sum_{t\in T} x^t$ is represented (in any representation affording χ) by the scalar matrix of trace $[G:H]\chi(x)$, it follows that

$$(1) \qquad \sum_{t\in T} \chi(x^{-1}x^t) = [G:H]|\chi(x)|^2/\chi(1).$$

Let P be a Sylow p-subgroup of G containing x and $e = [N_G(P) : C_G(P)]$. Note that since P is cyclic, $e|(p-1)$. Set $p^a = |P|$.

Let B denote the principal p-block of G. Recall that since P is cyclic, all irreducible characters of B are constant on nonidentity p-sections (if y is a p-element of G, the p-section corresponding to y consists of all elements of G whose p-part is conjugate to y). There are e nonexceptional characters $1=\mu_1,\mu_2,\ldots,\mu_e$ in B such that for each i, there is a sign ϵ_i such that $\mu_i(y)=\epsilon_i$ for all p-singular elements y. There are $(p^a-1)/e$ exceptional characters in B which agree on p-regular elements.

If λ is a nontrivial linear character of P, let λ^* denote its orbit under $N_G(P)$. Let Λ denote a full set of representatives for the orbits of $N_G(P)$ of the nontrivial linear characters of P. For each λ^*, there is an exceptional character χ_{λ^*} with the property that there is a fixed sign ϵ_0 (independent of λ^*) such that for any

$y \in P^{\#} := P - \{1\}$, χ_{λ^*} takes on the constant value $\epsilon_0 \sum_{\beta\in\lambda^*} \beta(y)$ on the p-section of y.

Consider (1) when $\chi = \mu_i$ is a nontrivial nonexceptional character in B. Then $\mu_i(x) = \epsilon_i$. Similarly, since $x^{-1}x^t$ is p-singular for each $1 \neq t \in T$, $\chi(x^{-1}x^t) = \epsilon_i$. We thus obtain $[G : H]/(\mu_i(1)) = \mu_i(1) + ([G : H] - 1)\epsilon_i$, and so

$$[G : H]\epsilon_i(\epsilon_i - \mu_i(1))/(\mu_i(1)) = \mu_i(1) - \epsilon_i.$$

If $\mu_i(1) = \epsilon_i$, then $\epsilon_i = 1$ and P is contained in the kernel K of μ_i. By induction, $x \in O_p(K) \leq O_p(G)$, and the result follows. Otherwise, $-\epsilon_i[G : H] = \mu_i(1)$, forcing $\epsilon_i = -1$. So we may assume that $\epsilon_i = -1$ for $i = 2, \ldots, e$.

In particular, $\mu_i(1) = [G : H]$. Also, observe that if $e = 1$, then G has a normal p-complement (cf [G1, p. 252]) and the hypothesis implies that x is central and so in $O_p(G)$. So we assume that $e > 1$.

We provide two proofs at this point.

First, observe that $Z = Z(G)$ is a p'-group (since $e \neq 1$). If Z is nontrivial, then by induction $xZ \in O_p(G/Z)$. Hence $x \in F(G)$, as desired. Thus $Z = 1$. We next claim that $G = \langle x, x^g \rangle$ for some $g \in G$. Otherwise, by induction $\langle x \rangle$ is normalized by $\langle x, x^g \rangle$ for all $g \in G$. This implies that $x \in O_p(N) \leq O_p(G)$, where $N = \langle x^g | g \in G \rangle$.

Thus $H \cap H^g \leq Z = 1$. Since $G \neq HH^g$, it follows that $|G| > |H|^2$. On the other hand, we have shown above that there is an irreducible nonexceptional character μ in B with $\mu(1) = [G : H]$. Thus $[G : H]^2 \leq |G|$ and $|G| \leq |H|^2$. This contradiction completes the proof.

For a second proof, note that it follows from the block orthogonality relations that

$$\sum_{\chi\in B} \chi(1)\chi(x) = 0. \tag{2}$$

If $e = p - 1$ and $a = 1$, then by a slight abuse we may regard all characters of B as nonexceptional. Then (2) yields the contradiction $1 - (p - 1)[G : H] = 0$. We therefore assume that $1 < e < p^a - 1$.

Letting χ_{λ^*} denote a fixed exceptional character in B with $\lambda^* = \{\beta_1, \ldots, \beta_e\}$, we have $\chi_{\lambda^*} = \epsilon_0(\beta_1 + \ldots + \beta_e)$ on $P^{\#}$. It follows from (2) that

$$1 - (e - 1)[G : H] - \epsilon_0\chi_{\lambda^*}(1) = 0.$$

Since $e > 1$, $\epsilon_0 = -1$. Summing (1) over all the exceptional characters, we obtain (observing that the exceptional characters sum to take value $-\epsilon_0$ on each p-singular element of G):

$$[G:H]([(p^a - e)/\chi_{\lambda^*}(1)] - 1) = [(p^a - 1)/e]\chi_{\lambda^*}(1) - 1,$$

forcing $\chi_{\lambda^*}(1) < (p^a - e)$, a contradiction as $\chi_{\lambda^*}(1) \equiv -e(\bmod p^a)$ (for notice that $\chi_{\lambda^*} + \beta_1 + \cdots + \beta_e$ vanishes on $P^\#$ and so must be a multiple of the character of the regular representation of P).

3. Unique Maximal Subgroups

We first prove Theorem C. The proof is based on a result of Aschbacher [A1, Theorem 2] which in turn partially depends on a result of Seitz [S].

Proof of Theorem C: First consider the case that $p = 2$. Assume that x is an involution of G and is contained in a unique maximal subgroup M of G which contains no normal subgroup of G. Let A be a proper normal subgroup. Since A is not contained in M, $G = \langle A, x\rangle$. If $g \in G$, set $W_g = W = \langle xx^g\rangle$. If W is normal in G, then $G = \langle Y, x\rangle$ for any nontrivial subgroup Y of W. It follows that G is dihedral of order $2r$ for some prime r. If r is odd, then the Sylow 2-subgroup of G is cyclic. If $r = 2$, then M is normal in G. So we may assume that W is not normal in G for any $g \in G$. Then $x \in N_G(W)$ and so $N_G(W) \le M$. In particular, $x^g \in M$ and so M contains the normal closure of x.

So assume p is odd and a Sylow p-subgroup of G is not cyclic. Thus $x \in E$, an elementary abelian noncyclic p-subgroup. Hence M is the unique maximal subgroup containing E. It follows from [A2, Theorem 2] that one of the following holds:

(a) G is a group of Lie type of rank 1 in characteristic p and M is a Borel subgroup,

(b) $p = 5$, $G = \mathrm{Aut}(Sz(32))$ and M is the normalizer of a Sylow 5-subgroup, or

(c) $G = A_{2p}$ with M the normalizer of $A_p \times A_p$.

First consider (a). If $G = L_2(q)$ with $p|q$, then there is a unique (in $\mathrm{Aut}(G)$) class of elements of order p. It follows that x is contained in a Borel subgroup and also a subgroup isomorphic to $L_2(p)$. Thus $q = p$ and the result holds.

If $G = U_3(q)$, there are two classes of subgroups of order p. Of course, x is contained in a Borel subgroup. If x is a transvection, then $x \in H \cong SL_2(p)$

(acting reducibly). If x is not a transvection, then $x \in H \cong L_2(p)$ (acting irreducibly). Thus, x is not contained in a unique maximal subgroup.

If $G = {}^2G_2(3^{2m+1})$ (with $p = 3$), then it follows by [Wa] that every element of order 3 is conjugate to an element of ${}^2G_2(3) \cong \mathrm{Aut}(L_2(8))$. Thus $m = 0$. Since the Sylow 3-subgroup of $L_2(8)$ is cyclic, $G = \mathrm{Aut}(L_2(8))$. Since $x \notin G'$, it follows that x induces a field automorphism on $L_2(8)$ (there is a unique class of subgroups of order 3 in G not contained in G') and $x \in N_G(S)$ for some Sylow 7-subgroup of G. Thus x is contained in more than one maximal subgroup of G.

Now consider (b). Since $x \notin G'$, it follows that x induces a field automorphism on G'. Thus x normalizes a Borel subgroup of G' and hence is not contained in a unique maximal subgroup.

Finally consider (c). A p-cycle is obviously contained in several maximal subgroups. Thus x is a product of two disjoint p-cycles. Clearly x is contained in N, the normalizer of $A_p \times A_p$. Note that N is maximal, for if $N \leq X$, then X is primitive and contains a p-cycle, whence $X = G$. Since $O_p(N) \neq 1$, it only remains to prove that $p > 13$. Note that if $2p = q + 1$ with q a prime power, then $x \in L_2(q) \leq A_{2p}$. This is the case for $p = 3, 5, 7$ and 13. If $p = 11$, then $x \in M_{22}$.

This completes the proof of the theorem.

We note that for $G = A_{2p}$ and x a product of two disjoint p-cycles, then x will often be contained in a unique maximal subgroup (containing $A_p \times A_p$). This will be the case unless there exists a nontrivial primitive group of degree $2p$ (note that for $p > 5$, it follows from the classification of finite simple groups that primitive groups of degree $2p$ are in fact 2-transitive).

Proof of Theorem A: Let G be a minimal counterexample. Let $X = \langle x \rangle$. Clearly, we may assume that $O_p(G) = 1$. We may also assume that X is not contained in any proper normal subgroup K of G. For if so, then by minimality $X \leq O_p(K) \leq O_p(G)$. In particular, it follows that $G = \langle x^g | g \in G \rangle$. If $X \leq H$ with H a proper subgroup of G, then by the minimality, $X \leq O_p(H)$. Thus X is subnormal in H. It follows by [W] that either X is subnormal in G or X is contained in a unique maximal subgroup M of G. If X is subnormal in G, then $X \leq K$ with K normal in G, a contradiction as above.

So we may assume that X is contained in a unique maximal subgroup M and moreover that $X \leq O_p(M)$. In particular, M is a p-local subgroup.

Let K be a nontrivial normal subgroup of G contained in M, P a Sylow p-subgroup of G containing x and $Q = P \cap K$. By the Frattini argument, $G =$

$KN_G(Q)$ and $x \in N_G(Q)$. If $N_G(Q)$ is proper in G, then, by the uniqueness of M, $N_G(Q) \leq M$. This implies $G = M$, a contradiction. Since $O_p(G) = 1$, this implies $Q = 1$ and K is a p'-group. Also, by minimality $KX/K \leq O_p(G/K)$. Thus as G is the normal closure of X, $G = KP$ with P a Sylow p-subgroup of G. Since $[X, K] = 1$, it follows that $G = K \times P$ and so $X \leq P = O_p(G)$, a contradiction. Thus M contains no nontrivial normal subgroup of G.

By Theorem 2.1, the Sylow p-subgroup of G is not cyclic. By Theorem C, $O_p(M) = 1$. This contradiction completes the proof.

Proof of Corollary B: We may assume that $O_p(G) = 1$. We claim that $X = 1$. If not, choose $x \in X$ of order p. Then x satisfies the hypotheses of Theorem A and so $x \in O_p(G) = 1$, a contradiction.

4. The Proof of Theorem D

We first sketch a proof of the odd analogue of the Z^* theorem. It is well known to many that it follows easily from the classification of finite simple groups. See also [Ar].

THEOREM 4.1: *Let G be a finite group and p a prime. If $x \in G$ has order p and is not central modulo $O_{p'}(G)$, then x commutes with some conjugate $x^g \neq x$.*

Proof: Let G be a minimal counterexample. We claim $U = O_{p'}(G) = 1$. If not, then by minimality, there exists a conjugate y of x such that $xU \neq yU$ and $[x, y] \in U$. By Sylow's theorem, applied to $\langle U, x, y\rangle$, it follows by replacing y by some conjugate (under U), we can assume that x, y are contained in a p-subgroup. Then $[x, y]$ is both a p-element and a p'-element. Hence $[x, y] = 1$ and G is not a counterexample.

So $O_{p'}(G) = 1$. Let $X = \langle x\rangle$. So X is not central in G. If X is contained in a normal subgroup K, then $X \leq Z(K)$. So we may take K abelian. Then any two conjugates of x commute, a contradiction. It follows that G is the normal closure of X.

If a Sylow p-subgroup of G is cyclic, it follows that for any nontrivial p-subgroup Y, $N_G(Y) = C_G(Y)$ (since we may assume that $X \leq Y$ by conjugation - then $N_G(Y) \leq N_G(X) = C_G(X)$). Then G contains a normal p-complement (cf [G1, p. 253]). Since $O_{p'}(G) = 1$, this forces G to be a p-group and X to be central.

So we may assume that a Sylow p-subgroup of G is not cyclic. Moreover, we may assume that X is central in any proper overgroup H with $O_{p'}(H) = 1$. In particular, X is central in any p-subgroup containing it and X centralizes $O_p(G)$. Since the same is true for any conjugate of X, it follows that $O_p(G) = Z(G)$. In particular, $O_{p'}(G/O_p(G)) = O_p(G/O_p(G)) = 1$ and X is nontrivial in $G/O_p(G)$. If $O_p(G) \neq 1$, then there exists a conjugate of x, $y \neq x$ such that $[x, y] \in O_p(G)$. It follows that x, y generate a p-subgroup and as we observed above, this implies that $[x, y] = 1$, a contradiction to the fact that G is a counterexample. Hence $O_p(G) = 1$. It follows that G contains a normal subgroup A which is a direct product of groups L_i each isomorphic to a fixed nonabelian simple group L.

Since $O_{p'}(G) = 1$, p divides the order of L. If x does not normalize each L_i, then x will normalize but not centralize a Sylow p-subgroup of A, a contradiction. Thus x and every conjugate of x normalizes each normal simple subgroup of A. Hence A is simple.

It follows by [Gr] that x cannot induce an outer automorphism on A. Also, x cannot centralize A (for this would imply that G, the normal closure of x centralizes A). Thus x induces an inner automorphism on A and so we may assume that $G = A$. We now apply [G2, 4.250] to conclude that $G \cong U_3(p)$; $G \cong Mc, Co_2$, or Co_3 with $p = 5$; $G = G_2(q), q \neq 3^n$ or J_2 with $p = 3$; or $G = J_4$ with $p = 11$. It is straightforward to verify that in each of these cases, $N_G(X) \neq C_G(X)$, whence there exists $g \in G$ with $x \neq x^g \in X$. This completes the proof.

Proof of Theorem D: (i) We need to show that x is central in $G/O_{p'}(G)$. If not, then, by the previous result, x commutes with $x^g \neq x$. Then $[x, g] = x^{-1}x^g$ is a nontrivial p-element. This contradicts the hypothesis.

(ii) We may assume that $O_r(G) = 1$. By (i), x is central modulo $A = O_{p'}(G)$. Let Q be a Sylow q-subgroup of A for a prime $q \neq r$. By the Frattini argument, x normalizes some conjugate of Q and by hypothesis, x centralizes this conjugate. Similarly, x normalizes some Sylow r-subgroup R of A. Thus $A = RC_A(x)$ and so $[x, A] = [x, R]$ is a normal r-subgroup of A. Thus $[x, A] \leq O_r(G) = 1$. Since $B = \langle A, x \rangle$ is normal in G and $x \in Z(B)$, it follows that $[x, g]$ is a p-element for all $g \in G$. By hypothesis, it is also an r-element. Thus $x \in Z(G)$, as desired.

Proof of Corollary E: Let G be a minimal counterexample. We may assume that $O_p(G) = 1$. We need to show X is central in G. It suffices to assume that X is an r-subgroup for some prime r. If $r = p$, the result follows from Corollary B. If $r \neq p$, it follows from Theorem D(ii) that every element x of order r in X

is central.

In particular, $Y = X \cap Z(G) \neq 1$. Since $O_p(G/Y) = 1$, it follows by induction that X/Y is central in G/Y. Hence, if $g \in G$ and $x \in X$, $[g,x] \in Y$, a p'-group and $[g,x]$ is a p-element. Thus X is central in G as required.

5. Linear Groups

Let k be a field of characteristic r. A subset of matrices is said to be triangular if it is conjugate (in $GL_n(k)$) to a subset of the set of upper triangular matrices. It is fairly easy to translate between results about finite groups and linear groups. A result about linear groups will apply to finite groups since we can always embed the finite group in a linear group of arbitrary characteristic. Thus, for example, Aschbacher's linear version of the Baer–Suzuki Theorem [A1] certainly implies the classical Baer–Suzuki Theorem. Since finitely generated linear groups are residually finite (and even more), it is not too difficult to obtain consequences for linear groups from corresponding results about finite groups.

The mechanism for translating our results about finite groups to linear groups is provided by the next result.

LEMMA 5.1: *Let G be a finitely generated subgroup of $GL_n(k)$. Let r be the characteristic of k. Let $x_1, \ldots, x_m \in G^\#$ and $y_1, \ldots, y_d$ be elements of G which are not unipotent. There exists a homomorphism $\rho : G \to GL_n(F)$ with F a finite field with the following properties:*

(a) *For each i, j, $\rho(x_i) \neq 1$ and $\rho(y_j)$ is not unipotent.*
(b) *If $r > 0$, then F also has characteristic r. If $r = 0$, then F can be chosen to have arbitrarily large characteristic.*
(c) *ρ maps unipotent elements to unipotent elements.*

Proof: Let R be the finitely generated subring of k (containing 1) generated by all the matrix entries of the generators (and their inverses). Note that $G \leq GL_n(R)$. Let a_i be some nonzero entry of $I - x_i$ and let b_j be a nonzero entry of $(I - y_j)^n$. Let $S = R[a_i^{-1}, b_j^{-1}, 1 \leq i \leq m, 1 \leq j \leq d]$. If $r = 0$, replace S by $S[1/c!]$ for any fixed positive integer c. Let J be a maximal ideal of S and set $F = S/J$. Since F is a field which is finitely generated as a ring, it is a finite field. Note that F has characteristic r if r is positive and has characteristic $s > c$ if $r = 0$.

Consider the natural ring homomorphism $\rho : M_n(S) \to M_n(F)$. Then ρ induces a group homomorphism from G into $GL_n(F)$. Since $\rho(a_i) \neq 0 \neq \rho(b_j)$, $\rho(x_i) \neq 1$ and $\rho(y_j)$ is not unipotent.

Obviously, ρ maps unipotent elements to unipotent elements.

As an illustration, we prove the following well known result. We give a proof which is a bit different than the usual one, but is in the spirit of this section.

LEMMA 5.2: *Let G be a subgroup of $GL_n(k)$. The following are equivalent:*

(a) *G is triangular.*

(b) *G' consists of unipotent elements and if $g \in G$, then all eigenvalues of g are in k.*

(c) *Every commutator of G is unipotent and all eigenvalues of elements of G are in k.*

Proof: Clearly (a) implies (b) and (b) implies (c). We prove (c) implies (a). Since we are assuming all eigenvalues are in the field, there is no loss in assuming that k is algebraically closed. There is no loss of generality in assuming that G is finitely generated (if every finitely generated subgroup of G fixes a line, so does G).

If $H \leq GL_n(F)$, F a finite field of characteristic r with every commutator an r-element, then as we observed in the introduction, H' is an r-subgroup. In particular, H' is nilpotent (of class at most n). Thus, the same holds for G by the previous Lemma.

Now we do a double induction on the derived length of G and on n. If G is abelian (or in particular $n = 1$), the result is clear. Since the derived length of G' is less than that of G, it follows that G' is triangular. Since G' is generated by unipotent elements, it is thus unipotent. By induction on n, we may assume that G acts irreducibly. However, the fixed points of G' are G-invariant and are nonzero. Thus $G' = 1$, G is abelian and the result follows.

We now prove Corollary F, which can be viewed as the linear version of Corollary E (note that Lemma 5.2 is essentially the case $X = G$ - of course, the proof of Lemma 5.2 does not depend upon the classification of finite simple groups while the next result does). For the convenience of the reader, we restate the result.

THEOREM 5.3: *Let $G \leq GL_n(k)$. Let $X \leq G$. If $[x, g]$ is unipotent for every $x \in X$ and $g \in G$, then*

(a) *$[G,X]$ is a normal unipotent subgroup of G, and*
(b) *If X is triangular, then $\langle X^g | g \in G\rangle$ is triangular.*

Proof: (a) $[G,X]$ is always normal in G. We may assume that G is finitely generated. Suppose $y \in [G,X]$ is not unipotent. Then, by Lemma 5.1, there exists a homomorphism $\rho : G \to GL_n(F)$ with F a finite field of characteristic s such that $[\rho(x), \rho(g)]$ is unipotent for all $x \in X, g \in G$ (and in particular is an s-element) but $\rho(y)$ not unipotent. By Corollary E, $[\rho(G), \rho(X)]$ is an s-subgroup and so $\rho(y)$ is unipotent, a contradiction.

(b) We may assume by induction on n that G acts irreducibly. In particular, G contains no normal unipotent subgroups. Thus, by (a), X is central. Since X is triangular, the result follows.

We should note that it is not true that if $\langle X, X^g\rangle$ is triangular for each $g \in G$, then $\langle X^g | g \in G\rangle$ is necessarily triangular. An easy example is obtained by letting k be a field of characteristic 3, $G = S_n \leq GL_n(k)$ with $n > 3$ and $x \in G$ a transposition. Then $\langle x, x^g\rangle$ is either elementary abelian or is isomorphic to S_3. It follows by Lemma 5.2 that $\langle x, x^g\rangle$ is triangular. Clearly, $G = \langle x^g | g \in G\rangle$ is not triangular.

Note that in the above example, it follows that $[g,x,x]$ is unipotent for all $g \in G$, but $[G,x]$ is not.

A similar example can be constructed for fields of any odd characteristic. Let k be a field of characteristic p for some odd prime p. Let D be the dihedral group of order $2p$ and embed $D \leq S_p \leq GL_p(k)$. Let $G = ED$, where E consists of the diagonal matrices of determinant 1 with order at most 2. Let $x \in D$ be an involution. Then it is straightforward to check that for any $g \in G$, either x commutes with x^g or xx^g has order p. In particular, $\langle x, x^g\rangle$ is triangular. On the other hand, $G = \langle x^g | g \in G\rangle$. Since G' is not a p-group, G is not triangular. Thus we have shown:

PROPOSITION 5.4: *Let k be a field of odd characteristic. There exists a positive integer n, a finite group $G \leq GL_n(k)$ and $x \in G$ such that $\langle x, x^g\rangle$ is triangular for each $g \in G$ but $\langle x^g | g \in G\rangle$ is not triangular.*

The next two results show that for connected groups or over fields of characteristic 0, we do have a triangular version of the Baer–Suzuki theorem. The proofs use the theory of algebraic groups (cf [B] or [H]) and do not depend upon the classification of finite simple groups.

THEOREM 5.5: *Let G be a connected subgroup (in the Zariski topology) of $GL_n(k)$ with k a field. Let $x \in G$ and $N = \langle x^g | g \in G \rangle$. Assume that all the eigenvalues of x are in k. The following are equivalent:*

(a) *N is triangular.*

(b) *$\langle x, x^g \rangle$ is triangular for each $g \in G$.*

(c) *$[x, x^g]$ is unipotent for all $g \in G$.*

(d) *$[g, x, x]$ is unipotent for all $g \in G$.*

(e) *$[x, g]$ is unipotent for all $g \in G$.*

(f) *$[G, x]$ is a unipotent normal subgroup of G.*

Proof: We may assume that k is algebraically closed. Clearly (a) implies (b) and (b) implies (c). Also, (f) implies (e) and (e) implies both (c) and (d).

We next show the equivalence of (a) and (f). It is obvious that (f) implies (a). We prove (a) implies (f) by induction on n. We can assume that G acts irreducibly. Since N is triangular, N has a nonzero eigenspace (where N acts via scalars). Clearly G permutes the finitely many nonzero eigenspaces of N. The stabilizer of one of these subspaces is a closed subgroup of G of finite index in G. Since G is connected, this implies N consists of scalars and the result follows.

Let $\bar{G}$ be the closure of G in $GL_n(k)$ (in the Zariski topology). We now prove that (c) implies (f). Consider the map $\phi : \bar{G} \to \bar{G}$ by $\phi(g) = [x, x^g]$. Since $\phi(G) \subset U$, where U is the set of unipotent elements of G, $\phi(\bar{G}) \subset \bar{U}$. Since the set of unipotent elements in $\bar{G}$ is closed, there is no loss of generality in assuming that $G = \bar{G}$ is a connected algebraic group.

Suppose the result is false and let G be a counterexample of minimal dimension. The minimality implies that G is an irreducible subgroup of $GL_n(k)$. In particular, G is reductive. Let B be a Borel subgroup of G containing x. Let $P \geq B$ be a maximal parabolic subgroup of G. If y is any conjugate of x contained in P, then by induction, $[y, P]$ is a normal unipotent subgroup of P. It follows that y normalizes B and so $y \in B$. Thus $W := \langle x^g : g \in G, x^g \in B \rangle \leq B$ is a normal subgroup of P. If W is normal in G, then $W = N$ is triangular, whence (a) and so (f) holds. Thus, we may assume, by the maximality of P, that $P = N_G(W)$. Hence, there is a unique maximal parabolic subgroup containing B and so G' has rank one. So $G' \cong (P)SL_2(k)$. By multiplying x by a central element, we may assume that $x \in G' = G$ and that if x is semisimple, $a := \mathrm{tr}(x) \neq \pm 2$. If x is not semisimple, then $\pm x$ is a transvection and $[x, x^g]$ is not unipotent for any $g \notin B$.

If x is semsimple, then we may assume that $x = \text{diag}(\alpha, \alpha^{-1})$. Then

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

is conjugate to x and $[x, y]$ is not unipotent. This contradiction completes the proof of this case.

Essentially the same proof (reducing to a rank one group and verifying the equation there) yields (d) implies (f). This completes the proof.

If $\langle x \rangle$ is connected, then x is in the connected component of G and the previous result applies. In particular, this is true if x is unipotent and k has characteristic 0. In fact, we can drop the connectedness hypothesis if we assume characteristic zero. The characteristic 0 assumption implies that all unipotent subgroups are connected. It also implies that the result holds for groups of dimension 0 (since the group will have no unipotent elements).

Theorem 5.6: *Let G be an algebraic group over an algebraically closed field k of characteristic zero. Let $R_u(G)$ be the unipotent radical of G and G^0 the connected component of G containing 1. Let $x \in G$ and $N = \langle x^g | g \in G \rangle$. The following are equivalent:*

(a) $[NG^0, N] \leq R_u(G)$.
(b) *If $N_g = \langle x, x^g \rangle$, then $[N_g, N_g]$ is unipotent for each $g \in G$.*
(c) *$[x, x^g]$ is unipotent for all $g \in G$.*

Proof: Clearly, (a) implies (b) and (b) implies (c). We prove that (c) implies (a). Let G be a counterexample first of minimal dimension and then with $[G : G^0]$ minimal. Then $R_u(G) = 1$. If $G^0 = 1$, then N is a normal abelian subgroup and so (a) holds.

We claim $G = G^0\langle x \rangle$. If not, then by minimality, $[x, G^0] \leq R_u(G) = 1$ and so x centralizes G^0. Thus N centralizes G^0. Let $C = C_G(G^0)$. Then C^0 is a torus. If $C^0 = 1$, then N is a finite normal subgroup of G, whence $[G^0N, N] = [N, N]$ contains no unipotent elements. Thus N is abelian and (a) holds. Otherwise, by passing to G/C^0, we see that $[N, N] \leq C^0$ which contains no unipotent elements and again N is abelian and (a) holds. This proves the claim.

Suppose $G^0 = T$ is a torus. Then $[x, x^t] = 1$ for all $t \in T$, whence $[x, [x, T]] = 1$. Since x induces an automorphism of finite order on T, it follows that $[x, T] \cap C_T(x)$ is finite. Thus $[x, T]$ is finite and connected, whence $[x, T] = 1$. Since G contains no unipotent elements, it follows that N is abelian and so (a) holds.

By a result of Steinberg [St, 7.2], we can choose a Borel subgroup B of G such that $x \in N_G(B)$. Let U be the unipotent radicial of B. Since $B \neq G^0$ (otherwise, G^0 is a torus), it follows by minimality that $[x, B] \leq U$. Let P be a parabolic subgroup of G^0 containing B and assume that x normalizes P and P is maximal with respect to this. Let $W = \langle x^P \rangle$. By the minimality of G, $[WP, W] \leq R_u(P) \leq B$. In particular, $W \leq N_G(B)$. The same is true for any conjugate of x contained in $N_G(B)$ (and observe that if $y = x^g \in N_G(B)$, then P^x and P^y are G^0−conjugate (since $G = G^0\langle x \rangle$) and so y normalizes P). Let $V = \langle x^g | g \in G, x^g \in N_G(B) \rangle$. It follows that $[VP, V] \leq R_u(P) \leq U$. In particular, P normalizes V. If G^0 normalizes V, then $V = N$ and $[N, N]$ is a unipotent group. Since $R_u(G) = 1$, $R_u(N) = 1$ and N^0 is a torus. In particular, N contains no unipotent elements. Thus $[NB, N] \leq U \cap N = 1$ and the same is true for any Borel subgroup. The result follows in this case.

By the maximality of P, it thus follows that $P = N_{G^0}(V)$ and so there is precisely one such parabolic subgroup (i.e. there is a unique maximal element among the parabolics containing B which are x-invariant). Set $H = [G^0, G^0]$. We claim H is a simple algebraic group. If not, then x must permute the simple factors of H transitively (by the uniqueness of the parabolic subgroup) and then x will not centralize B/U as we observed above. Again, by the uniqueness of the maximal x-invariant parabolic subgroup containing B, it follows that x must induce a graph automorphism of the Dynkin diagram of H and have only one orbit. This implies that either H has rank one or H is of type $A_2(k)$. Again, in the latter case, it follows that x will not centralize B/U, a contradiction. So $H \cong (P)SL_2(k)$ and therefore x must induce an inner automorphism. As in the proof of the previous result, this cannot occur. This completes the proof.

In Theorem 5.6, if we only assume that $G \leq GL_n(k)$ with k of characteristic zero, (b) and (c) are still equivalent (since we can replace G by its closure). We can now prove the triangular version of Baer–Suzuki (Corollary G).

THEOREM 5.7: *Let $G \leq GL_n(k)$ with k a field. Assume that either k has characteristic 0 or that G is connected. Let X be a triangular subgroup of G and set $N = \langle X^g | g \in G \rangle$. The following are equivalent:*

(a) *$\langle X, X^g \rangle$ is triangular for each $g \in G$.*

(b) *N is triangular.*

Proof: Clearly (b) implies (a). So assume (a) holds. We may assume that G

acts irreducibly and in particular contains no normal unipotent subgroup. If G is connected, it follows by Theorem 5.5 that X is central in G. So assume k has characteristic zero.

We can assume that k is algebraically closed. Let $\bar{G}$ be the closure of G. Let $x \in X$ and $W_x = \langle x^g | g \in \bar{G} \rangle$. Since G is dense in $\bar{G}$, the condition of 5.6(c) holds for every $x \in X$ and $g \in \bar{G}$. By Theorem 5.6, $[W_x, W_x]$ is contained in the unipotent radical of $\bar{G}$ and so W_x is abelian. Since G is irreducible, every normal subgroup acts completely reducibly. It follows that W_x is diagonal. Since $\langle X, X^g \rangle$ is triangular, $[x, y^g]$ is unipotent for any $x, y \in X, g \in G$. Since also, $[x, y^g] \in W_x$, it is semisimple. Thus $[X, X^g] = 1$ and so N is generated by commuting semisimple elements and so is diagonal as desired.

Our final result is a linear version of Theorem A. The proof of this result does depend upon the classification of finite simple groups. The characteristic restriction is probably not necessary.

THEOREM 5.8: *Let $G \leq GL_n(k)$ with k a field of characteristic r. Assume $x \in G$ is unipotent and that $[x, g]$ is unipotent for all $g \in G$. If x has order r (for example if $r \geq n$) or $r = 0$, then $\langle x^g | g \in G \rangle$ is a unipotent subgroup of G.*

Proof: As usual, we may assume that G is finitely generated. Suppose $N = \langle x^g | g \in G \rangle$ is not a unipotent subgroup. Let $y \in N$ with y not unipotent. By Lemma 5.1, there exists a finite field F and a homomorphism $\rho : G \to GL_n(F)$ such that $\rho(y)$ is not unipotent. Let s denote the characteristic of F. If r is positive, then $r = s$ and x and $\rho(x)$ have prime order r. If $r = 0$, then we can assume that $s > n$. Thus $\rho(x)$ has prime order s. The result now follows by Theorem A.

Note added in proof: The authors have just become aware of the interesting article [X] where (among other results) Theorem A is proved for any p-element. The proof in [X] is much different than the one given here. It depends upon the fact that the finite simple groups of Lie type have a block of defect zero for any odd prime and uses the classification of finite simple groups. The proof given here when the Sylow p-subgroup is cyclic does not depend on the classification.

References

[AL] J. Alperin and R. Lyons, *Conjugacy classes of p-elements*, J. Algebra **19** (1971), 536–537.

[Ar] O. D. Artemovich, *Isolated elements of prime order in finite groups*, Ukranian Math. J. **40** (1988), 397–400.

[A1] M. Aschbacher, *The 27-dimensional module for E_6, IV*, J. Algebra **131** (1990), 23–39.

[A2] M. Aschbacher, *Overgroups of Sylow subgroups in sporadic groups*, Memoirs of the Amer. Math. Soc. **60** (1986), No. 343.

[B] A. Borel, *Linear Algebraic Groups*, 2nd Ed., Springer-Verlag, New York, 1991.

[F] W. Feit, *The Representation Theory of Finite Groups*, North-Holland Publishing Company, Amsterdam, 1982.

[G1] D. Gorenstein, *Finite Groups*, Harper & Row, New York, 1968.

[G2] D. Gorenstein, *Finite Simple Groups — An Introduction to their Classification*, Plenum Press, New York, 1982.

[Gr] F. Gross, *Automorphisms which centralize a Sylow p-subgroup*, J. Algebra **77** (1982), 202–233.

[H] J. Humphreys, *Linear Algebric Groups*, Springer-Verlag, New York, 1975.

[S] G. Seitz, *Generation of finite groups of Lie type*, Trans. Amer. Math. Soc. **271** (1982), 351–407.

[St] R. Steinberg, *Endomorphisms of linear algebraic groups*, Mem. Amer. Math. Soc., **80**, 1968.

[Wa] H. N. Ward, *On Ree's series of simple groups*, Trans. Amer. Math. Soc. **121** (1966), 62–89.

[W] H. Wielandt, *Kriterien für Subnormalität in endlichen Gruppen*, Math Z. **138** (1974), 199–203.

[X] Wen-jun Xiao, *Glauberman's conjecture, Mazurov's problem and Peng's problem*, Science in China Series A **34** (1991), 1025–1031.