A Remark on Finite Groups Having a Split BN-pair

of Rank One with Characteristic Two

鈴木 通夫
Michio Suzuki

1. Introduction

A BN-pair of rank one in a group $G$ is a pair of subgroups $(B, N)$ of $G$ which satisfy the following two conditions:

(BN 1) The subgroup $H$ defined by

$$H = B \cap N$$

is a normal subgroup of index 2 in $N$;

(BN 2) The group $G$ is the union of $B$ and $BNB$.

In order to define a split BN-pair, we need to introduce further notations. By (BN 1), there is an element $t$ of $N$ such that

$$t^2 \in H \quad \text{and} \quad N = \langle H, t \rangle = H \langle t \rangle .$$

A BN-pair $(B, N)$ is said to be split if the following additional condition is satisfied:

(BN 3) There is a normal subgroup $U$ of $B$ such that $B$ is a split extension of $U$ by $H$ and such that we have

$$B \cap tUt^{-1} = \{1\} .$$

If the split BN-pair $(B, N)$ of a finite group $G$ satisfies a further condition:

(BN 4) The subgroup $U$ contains a Sylow 2-subgroup of $G$, then $G$ is called a group with a split BN-pair of rank one with characteristic two.
The class of finite groups having split BN-pairs of rank one was studied during the 1960's. The complete determination of the simple groups which belong to this class was achieved in Suzuki [7] for the characteristic two case and in Hering-Kantor-Seitz [3] and Shult [5] for the other cases, and this was the first step in the eventual classification of simple groups of finite order. (For more information, consult Suzuki [8] where a complete list of references can be found.)

In studying the structure of a finite group $G$ with a split BN-pair of rank one and characteristic two, one of the most important ideas is the concept of the associated prime number $\chi(G)$ for $G$. (See Suzuki [7], §10.) The number $\chi(G)$ is defined as the order of the product of two involutions which are uniquely determined (up to conjugation) by the properties of the group $G$. It is not at all obvious why this order should be a prime number. In [7], the proof of the fact that $\chi(G)$ is indeed a prime number depends, among other things, on the classification of the Zassenhaus groups of characteristic two (cf. Suzuki [6]) and is indirect.

The purpose of this paper is to prove, by a direct method, that the integer $\chi(G)$ is prime. In order to make this paper reasonably self-contained, we have added a few elementary discussions on the structure of $G$ and on the definition of $\chi(G)$. It is hoped that the method of this paper, or some ramification of it, might simplify the long argument of [7] which leads to the determination of the structure of $G$. 
2. **Preliminaries**  Let $G$ be a finite group having a split BN-pair of rank one with characteristic two. We will use the notation introduced in §1 throughout this paper. Thus, we have

$$H = B \cap N, \quad N = H \langle t \rangle, \quad \text{and} \quad B = U H = H U \triangleright U.$$  

It is clear that $BNB = BtB$ in (BN 2). So, we have

$$G = B \cup BtB.$$  

Therefore, as a permutation group on the cosets of $B$, $G$ is doubly transitive. The normal subgroup $U$ of $B$ in (BN 3) acts regularly on the cosets different from $B$. (Thus, the group $G$ is really an (L)-group as defined in §8 of [7].) The above representation of $G$ as a permutation group is quite useful. For example, $B$ is the only coset fixed by an arbitrary nonidentity element of $U$. This fact leads to the following proposition ([Suzuki [7], Lemma 10(ii)].

(A) If $u$ is any nonidentity element of $U$, then its centralizer $C_G(u)$ is contained in $B$.

In the condition (BN 3), the conjugate subgroup $tU t^{-1}$ does not depend on the particular choice of $t$ as long as we choose $t \in N - H$. By (BN 3) and (BN 4), the group $H$ is isomorphic to $B/U$ and, hence, has odd order. It follows that the element $t$ can be chosen to be an involution. We will henceforth assume that we have done so. Thus, we have $t^2 = 1$. Since $H \vartriangleleft N$, the element $t$ induces an automorphism of order 2 in the group $H$ of odd order. A simple counting argument proves the following lemma (Gorenstein–Herstein [2]).
(B) There are exactly \(| H : C_H(t) |\) elements of \(H\) that satisfy \(x^t = x^{-1}\). Any such element \(x\) can be written in the form 
\[ x = y^{-1}y^t \]
with \(y \in H\).

We have \(tH = Ht\). So,
\[ BtB = BtHU = BtU = UHtU. \]

(C) Every element \(x \in G - B\) can be expressed uniquely in the form 
\[ x = f g t h \quad (f, h \in U, g \in H). \]

The uniqueness of the expression comes from the condition that we have \(B \cap tU t^{-1} = \{1\}\). The above expression for \(x\) is called the canonical form of the element \(x\) of \(G - B\).

By the condition (BN4), the group \(U\) contains an involution. For any involution \(u\) of \(U\), the conjugate \(u t u^{-1}\) is in \(G - B\) (by (BN 3)). So, let 
\[ u t u^{-1} = f g t h \]
be its canonical form. Since \(u = u^{-1}\), we get
\[ f = h^{-1} \quad \text{and} \quad g^t = g^{-1}. \]

By (B), we can write \(g = k^{-1}k^t\) for some \(k \in H\). Then, for the involution \(s = k^t u k^{-t}\), we have
\[ t s t = r^{-1} t r \]
where \(r = k h k^{-1} \in U\). Thus, we have proved the following proposition ([7], Lemma 16).
(D) Let \( u \) be any involution of \( U \). Then, there is a conjugate \( s \) of \( u \) such that

\[ \text{tst} = r^{-1}tr \]

for some \( r \in U \).

An identity of the above form is called a structure identity for \( G \) ([7], p.522). An important property of the structure identity is the following.

(E) Let \( s \) be an involution such that the pair \((s, t)\) satisfies the above structure identity for \( G \). If \((s_1, t_1)\) is a pair of involutions such that

\[ s_1 \in U, \quad t_1 \in N, \quad \text{and} \quad t_1s_1t_1 = r_1^{-1}t_1r_1 \]

for some \( r_1 \in U \), then there is an element \( k \) of \( H \) such that

\[ t_1 = t^k, \quad s_1 = s^k, \quad \text{and} \quad r_1 = r^k. \]

**Proof** Since \( \langle t \rangle \) and \( \langle t_1 \rangle \) are \( S_2 \)-subgroups of \( N \), they are conjugate in \( N \). So, there is an element \( k \) of \( H \) such that \( t_1 = t^k \). We replace the original structure identity for \( G \) by its conjugate and we assume that \( t_1 = t \). Then, for \( u = tr_1^{-1}rt^{-1} \), we have \( u^{-1}s_1u = s \). This implies that the element \( s_1 \) fixes the coset \( uB \). Hence, we get \( u \in B \). On the other hand, the definition of \( u \) shows that \( u \) is an element of \( tUt^{-1} \). So, it follows from (BN 3) that \( u = 1 \). Thus, we have \( s_1 = s \) and \( r_1 = r \). \( \square \)

In fact, we have proved the stronger property that we have \( s_1 = s^k \) (and \( r_1 = r^k \)) whenever \( t_1 = t^k \). Thus, for a fixed involution \( t \) of \( N \), there is a unique involution \( s \) of \( U \) which satisfies \( \text{tst} = r^{-1}sr \).
for some \( r \in U \).

From now on, let \((s, t)\) be the pair of involutions which satisfies the structure identity for \(G\) given in (D).

(F) If \(s_1\) is any involution of \(U\), then \(s_1\) is conjugate to \(s\) by an element of \(H\); i.e. there is an element \(k\) of \(H\) such that \(s_1 = s^k\).

Proof By (D), some conjugate of \(s_1\) satisfies the structure identity. So, we have \(s_1 = s^k\) for some \(k \in H\) by (E).

(G) We have \(C_H(s) = C_H(t)\).

Proof Proposition (E) implies that \(C_H(t) \subseteq C_H(s)\). Conversely, if \(k \in C_H(s)\), then we have \(tkst = tskt\). Thus,

\[
k^tr^{-1}tr = r^{-1}trk^t = r^{-1}ktk^{-1}tk.
\]

The uniqueness of the canonical form implies that we have \(k^t = k\). So, \(C_H(s) \subseteq C_H(t)\).

(H) If \(k\) is a nonidentity element of \(H\) such that \(k^t = k^{-1}\), then we have \(C_G(k) \subseteq H\).

Proof Suppose \(k^t = k^{-1}\) and \(ku = uk\) for some element \(u \in G - B\). Let \(u = fgh\) be the canonical form of the element \(u\). Then, we have

\[kfgth = fghk.\]

The canonical form of the left side is \(kg^{-1}kgth\), while that of the right side is \(fgk^t\). The uniqueness of the canonical form implies that

\[kg = gk^t = gk^{-1}, \text{ or } g^{-1}kg = k^{-1}.\]

Since \(g\) and \(k\) are elements of the group \(H\) which has odd order, we
must have $k = 1$. Thus, if $k^t = k^{-1} \neq 1$, then $C_G(k) \subseteq B$.

Therefore,

$$C_G(k) = C_G(k^{-1}) \subseteq B^t.$$  

Hence, we have $C_G(k) \subseteq B \cap B^t = H$.

(I) The involution $s$ of $U$ lies in the center of $U$.

**Proof**  If $s$ is the unique involution of $U$, then clearly $s$ is contained in the center of $U$. If $U$ contains more than one involution, $U$ contains exactly $|H : C_H(s)|$ involutions by (F). It follows from (G) and (B) that there is a nonidentity element $k$ of $H$ satisfying $k^t = k^{-1}$. We can choose $k$ to be an element of prime order. The conjugation by such an element $k$ induces an automorphism of $U$ of prime order which is fixed point free. So, by a theorem of Thompson [9], $U$ is nilpotent. Thus, some involution belongs to the center of $U$. Then, by (F), all involutions of $U$ are in the center.

3. **Definition of $\chi(G)$ and the statement of the theorem**

Let $(s, t)$ be the pair of involutions which satisfies the structure identity for $G$. Let $\chi(G)$ be the order of the element $st$ which is the product of the involutions $s$ and $t$.

**Theorem**  The integer $\chi(G)$ is a prime number.

We will prove that for any positive integer $n < \chi(G)$, the $n$-th power $(st)^n$ of $st$ is conjugate to $st$. If this is proved, the theorem clearly follows.
4. **Proof of the Theorem**  We will prove that for any positive integer \( n < \chi(G) \), there is an element \( u_n \) of \( U \) such that
\[
(st)^n = u_n^{-1}(st)u_n.
\]

First, we remark that the element \( u_n \), if it exists at all, is the unique element of \( U \) which satisfies \((st)^n = u_n^{-1}(st)u_n\). This is seen by noting that the right side is, as written, the canonical form of \((st)^n\) and by recalling the uniqueness of that form.

In order to prove the existence of an element \( u_n \), we proceed by induction on \( n \). If \( n = 1 \), the statement is obvious. Consider the case when \( n = 2 \). We have the structure identity \( tst = r^{-1}tr \). Hence, we get
\[
stst = (st)^2 = s r^{-1}tr = r^{-1}(st)r
\]
because \( s \) is in the center \( Z(U) \) of \( U \) by (I). Thus, we have
\[
u_2 = r.
\]
Suppose that \( n = 2m \) is even. Then, we have
\[
u_{m}^{-1}(st)u_{m} = (st)^m
\]
by the inductive hypothesis. Taking the conjugate of the above equation by the element \( r \), we get
\[
r^{-1}u_{m}^{-1}(st)u_{m}r = r^{-1}(st)^{m}r = (r^{-1}(st)r)^{m} = (st)^{2m}.
\]
Thus, with \( u_{2m} = u_{m}r \), we have \((st)^{2m} = u_{2m}^{-1}(st)u_{2m}\).

Finally, assume that \( n = 2m + 1 \) is odd. By the inductive hypothesis, we have (with \( u = u_{2m} \))
\[
(st)^{2m} = u^{-1}(st)u.
\]
We can write \((st)^n = (st)^{2m}st = st(st)^{2m}\). So, we get
(1) \((st)^n = u^{-1}stu = stu^{-1}stu\).

The element \(s\) is an involution in \(Z(U)\), so the terms between the two \(t\)'s in the middle and last expressions of (1) are inverse of each other:
\[(us)^{-1} = s^{-1}u^{-1} = u^{-1}s.\]

Since \(n < \chi(G)\), we have \((st)^n \neq 1\). Thus, \(us \neq 1\) and \(t(us)t\) is an element of \(G - B\). Let
\[(2) \quad t(us)t = fgth\]
be the canonical form. Since we have
\[t(u^{-1}s)t = t(us)^{-1}t^{-1} = [t(us)t^{-1}]^{-1},\]
the equation (1) gives us
\[u^{-1}sfgh = sh^{-1}tg^{-1}f^{-1}u.\]

So, the uniqueness of the canonical form implies
\[u^{-1}sf = sh^{-1}, \quad g^{-1} = g^t, \quad \text{and} \quad h = f^{-1}u.\]

Thus, we have
\[(3) \quad (st)^n = sh^{-1}gth = h^{-1}sgth\]
where \(g \in H\) and \(g^t = g^{-1}\). The last equality follows from the fact that \(s \in Z(U)\).

We need to show that \(g = 1\). By (B), we can write \(g = f^{-1}f^t\).

Then, \(gt = f^{-1}tf\) and (3) implies (by cancelling one \(s\) from the left)
\[t(st)^{2m} = h^{-1}f^{-1}tfh.\]

The left side is also a conjugate of \(t:\)
\[t(st)^{2m} = (st)^{-m}t(st)^m\]
because \((st)^{-1} = ts\). Therefore, we get
\[(st)^{-m}t(st)^m = h^{-1}f^{-1}tfh.\]

This will give us the information that a certain element commutes with
the involution $t$. It is more convenient to replace the middle $t$ by

$$t = rtst^{-1}r^{-1},$$

which is obtained from the structure identity. We get

$$\text{(st)}^{-m}rtst^{-1}r^{-1}(\text{st})^m = h^{-1}f^{-1}rtst^{-1}r^{-1}fh.$$  

Set

$$\text{(st)}^{-m}rt = h^{-1}f^{-1}rt w.$$

Then, the equation (4) is equivalent to saying that

$$w \in C_G(s).$$

By (A), (I), and (G), we have

$$C_G(s) = C_B(s) = C_H(s)U = C_H(t)U.$$

So, we can write

$$w = kv \quad (k \in C_H(t), \ v \in U).$$

It follows from the inductive hypothesis that

$$(\text{st})^m = u_m^{-1}(\text{st})u_m.$$

Then, the defining equation (5) of $w$ gives us

$$u_m^{-1}tsu_mrt = h^{-1}f^{-1}rtkv.$$

We have shown that $u_m^r = u_{2m} = u$. Thus, we get

$$tsut = u_m h^{-1}f^{-1}rtkv.$$

The canonical form of this element is

$$\text{(6)} \quad tsut = u_m h^{-1}f^{-1}rt \cdot f^{-1}k \cdot tv$$

where $u_m h^{-1}f^{-1}rt \in U$, $f^{-1}k \in H$, and $v \in U$. Since $s \in Z(U)$, the left side of (6) coincides with the left side of (2). The uniqueness of the canonical form implies, in particular, that

$$\text{(7)} \quad g = f^{-1}k.$$
On the other hand, the element \( f \) was defined by \( g = f^{-1}f^t \). So, the equation (7) gives us
\[
f^t = k.
\]
But, \( k \in C_H(t) \) and hence \( f = k^t = k \). This proves that
\[
g = f^{-1}k = 1.
\]
Therefore, the equation (3) can now be written as
\[
(st)^n = h^{-1}(st)h.
\]
This completes the inductive proof of the proposition.

5. Remarks For each odd prime number \( p \), there is a group \( G \) with a split BN-pair of rank one and characteristic two such that \( \chi(G) = p \).

Let \( G \) be the linear group \( L(F_p) \) of linear transformations
\[
x' = ax + b
\]
where \( a \neq 0 \) and \( a, b \) are elements of the finite field of \( p \) elements.

This group \( G \) has a split BN-pair \( (B, N) \) of rank one and characteristic two where
\[
B = U = \{ x' = ax \mid (a \neq 0) \},
\]
\[
N = \langle t \rangle, \quad t : x' = 1 - x, \text{ and}
\]
\[
H = \{ 1 \}.
\]
Similar groups can be constructed over any finite near-fields of odd characteristic. See [7], §5.

Let \( G \) be, as before, a finite group having a split BN-pair of rank one with characteristic two, and let \( p = \chi(G) \). The proof of §4 shows that the subgroup \( U \) contains a cyclic group of order \( p - 1 \).

In fact, the set of elements \( u_1, u_2, \ldots, u_{p-1} \) forms a subgroup which
is isomorphic to the group of automorphisms of the cyclic group $<st>$ of order $p$. We have

$$u_1 = 1, \quad u_2 = r, \quad \text{and} \quad u_{p-1} = s.$$ 

If the group $U$ contains only one involution, then $G$ is essentially a linear group over a near-field. See [7], Theorem 1. So, the interesting case is when $G$ is simple and $U$ contains more than one involution. In this case, $U$ is nilpotent (cf. the proof of (I)). It can be proved by using character theory that the group $U$ is indeed a 2-group. Then, the associated prime number $p = \chi(G)$ is a Fermat prime because $p - 1$ is a power of 2.

If the group $U$ is abelian, it is not hard to show that $G$ is the special linear group $SL(2,F)$ over a finite field $F$ of characteristic two. If $U$ is nonabelian, the property (F) together with the solvability of the group $H$ of odd order (cf. Feit-Thompson [1]) imposes a strong restriction on the 2-group $U$. This class of 2-groups was investigated by G. Higman [4]. Among others, Higman proved that the exponent of $U$ is at most 4. Since $U$ must contain a cyclic group of order $p - 1$, we must have $\chi(G) = p = 3$ or $5$.

It still requires a long argument to get the final conclusion that $G$ is either the 3-dimensional unitary group of characteristic two or the Suzuki group depending on whether $\chi(G) = 3$ or $\chi(G) = 5$. But, the above brief discussion explains the role of Higman's theorem on the special class of 2-groups in the classification of simple groups having a split BN-pair of rank one.
References


