ON BURNSIDE'S PROBLEM

BY
R. C. LYNDON

1. **Introduction.** Let B be the group on q generators defined by setting the pth power of every element, for some prime p, equal to the identity(1). A method, based on the free differential calculus of R. H. Fox, will be applied to study the quotients $Q_n = B_n/B_{n+1}$ of the lower central series of B, for $n \leqq p+2$(2). Our main results were obtained earlier by Philip Hall, using a different method(3).

To state these results, let $\psi(n)$ be the rank of the free abelian quotient F_n/F_{n+1}, where F is the free group on q generators. (Witt [11] has shown that $\psi(n) = n^{-1}\sum_{d|n} \mu(n/d)q^d$.) Then Q_n will be the direct product of a certain number $\kappa(n)$ of cyclic groups of order p, where $\kappa(n) \leqq \psi(n)$. We show that:

$$\text{(I)} \qquad \kappa(n) = \psi(n) \qquad \text{for } n < p;$$

$$\text{(II)} \qquad \kappa(p) = \psi(p) - \binom{p+q-1}{p} + q;$$

$$\text{(III)} \qquad \kappa(p+1) = \psi(p+1) - \binom{q}{2}\binom{p+q-2}{p-1} \qquad \text{for } p > 2;$$

$$\text{(IV)} \qquad \kappa(p+2) = \psi(p+2) - 3p + 1 \qquad \text{for } p > 3 \text{ and } q = 2.$$

2. **The Magnus series and Fox derivatives.** In this section we summarize, without proof, those known results that will be needed later.

Magnus has defined an isomorphic representation of a free group by power series. Let F be the free group on generators $x_1, \cdots, x_q$. Let Ω be the ring of all formal power series, with integer coefficients, in q noncommuting indeterminates denoted by $\Delta x_1, \cdots, \Delta x_q$. The Magnus representation $w \to 1 + \Delta w$

Presented to the Society, April 25, 1953; received by the editors May 18, 1953.

(1) For a general discussion of Burnside's problem, see Baer [1]. In addition to the papers mentioned in [1] we note a more recent paper of Magnus [10] that is in part parallel to the present investigation, and a paper of J. A. Green [5] in which he anticipates certain ideas of the present paper and establishes a remarkable theorem that supersedes similar results of ours.

(2) For the Fox calculus, see Fox [4]; for its application to the lower central quotients, see Chen-Fox-Lyndon [3]. The results cited in §2 are to be found in these papers and in the fundamental papers [9; 10] of Magnus and [11] of Witt. See also Hall [6]. *Added in proof:* These results are extended in a sequel to the present paper, to appear in Trans. Amer. Math. Soc.

(3) Hall, 1949, unpublished. Results I, II, III (at least for $q=2$), IV below. I am grateful for the opportunity to check my results against his (and to correct an error in my preliminary computation of $\kappa(p+2)$).

may be characterized as the unique multiplicative extension, F into Ω, of the correspondence $x_k \to 1+\Delta x_k$.

We write $w \to 1+\Delta w = 1+\omega_1+\omega_2+\cdots$ where ω_n is the sum of all terms of total degree n in the Δx_k. It is known that $\omega_1=\omega_2=\cdots=\omega_{n-1}=0$ if and only if w lies in the lower central group F_n. In this case ω_n is a Lie element in the Δx_k, of degree n, and it is known that the correspondence $w \to \omega_n$ defines an isomorphism of the abelian quotient F_n/F_{n+1} onto the module of all Lie elements of degree n contained in Ω. If $p\zeta$ is a Lie element, where p is an integer, then ζ is a Lie element.

The coefficients in the Magnus series are given by the Fox calculus. Let Γ be the group ring of F, with integer coefficients. For each generator x_k define $\partial/\partial x_k$ from F into Γ by the conditions

$$\frac{\partial x_j}{\partial x_k} = \delta_{jk}, \qquad \frac{\partial(uv)}{\partial x_k} = \frac{\partial u}{\partial x_k} + u\,\frac{\partial v}{\partial x_k}\cdot$$

By extending $\partial/\partial x_k$ linearly to a derivation from Γ into Γ, one defines the iterated derivatives $\partial^n/\partial x_{c_1}\cdots\partial x_{c_n}$. The coefficient sum $D_{c_1\ldots c_n}(w)$ of $\partial^n w/\partial x_{c_1}\cdots\partial x_{c_n}$ is then the coefficient of $\Delta x_{c_1}\cdots\Delta x_{c_n}$ in Δw:

$$\Delta w = \sum D_c(w)\cdot\Delta x_{c_1}\cdots\Delta x_{c_n},$$

summed over all nonempty finite sequences $c=c_1\cdots c_n$ of integers $c_k = 1, 2, \cdots, q$.

Let C_n be the set of all sequences c of length n, and define S_n to be the subset of those "standard" c that have the property of preceding lexicographically all of their own proper terminal segments $c_kc_{k+1}\cdots c_n$, $1<k\leqq n$. The operators D_c for c in C_n define homomorphisms of F_n/F_{n+1} into the additive group Z of integers, and the D_c for c in S_n form a basis for the group of all homomorphisms of F_n/F_{n+1} into Z. The operators D_c are homogeneous in the sense that $D_c(w)=0$ for w in F_n unless for each k the degree of w (as a commutator form) in x_k is equal to the number of occurrences of the symbol k in the sequence c.

The operators D_c, applied to the general element of F, are not independent, but are subject to certain "shuffle relations." Define a shuffle of two sequences a and b to be a pair of order-preserving one-to-one mappings embedding them as subsequences in a new sequence c; we require that c be precisely the union of the two subsequences, but not that they be disjoint. In these terms one has, for all w in F, the relations

$$D_a(w)\cdot D_b(w) = \sum D_c(w),$$

summed over all shuffles of a and b. All relations involving only a finite number of the operators D_c are consequences of these. In particular, by means of these relations it is possible to express the general operator D_c as a polynomial with rational coefficients in the D_c for c in S_n.

3. **Preliminary constructions.** $B=F/R$, where F is free on q generators, and R is generated by all pth powers of elements from F. Then $Q_n=B_n/B_{n+1}$ is a quotient group of F_n/F_{n+1}. Let V_n be the quotient of F_n/F_{n+1} by the pth powers of its own elements. Since F_n/F_{n+1} is free abelian of rank $\psi(n)$, V_n may be taken, in additive notation, as a vector space of dimension $\psi(n)$ over the field of integers modulo p. Since Q_n is abelian of exponent p, it may be identified with a quotient space of V_n:

$$Q_n = V_n/M_n.$$

The dimension of Q_n is $\kappa(n)=\psi(n)-\mu(n)$, where $\mu(n)$ is the dimension of M_n.

Given a set of elements r whose cosets span $F_n \cap R/F_{n+1} \cap R$, and a set of elements c of C_n that includes the set S_n, the matrix $\mathfrak{M}_n=[D_c(r)]$, with elements taken modulo p, is a relation matrix for $Q_n=V_n/M_n$. Hence $\mu(n)$ is the rank of $\mathfrak{M}_n$.

We are thus led to consider the Magnus series $1+\Delta w$ for $w=\prod u_i^p$ in R, and the behavior of its coefficients reduced modulo p. From the equation

$$1 + \Delta(u_1 \cdots u_m) = (1 + \Delta u_1) \cdots (1 + \Delta u_m),$$

for elements $u_1, \cdots, u_m$ in F, one has the "Leibniz rule"

PROPOSITION 3.1.

$$D_c(u_1 \cdots u_m) = \sum D_{c^1}(u_1) \cdots D_{c^m}(u_m),$$

summation over all "partitions" of the sequence $c=c_1 \cdots c_n$ into m segments c^k: $c=c^1 \cdots c^m$. In this context only we admit the possibility of empty sequences c^k, with the understanding that $D_{c^k}(u_k)=1$.

Let the terms in (3.1) be grouped according to the number r of nonempty segments in the corresponding partition of c. Setting all $u_k=u$ and collecting identical terms then gives

PROPOSITION 3.2. *If $c=c_1 \cdots c_n$ is of length n, then*

$$D_c(u^m) = \sum_{1 \leq r \leq m,n} \binom{m}{r} \sum D_{c^1}(u) \cdots D_{c^r}(u),$$

with summation now confined to partitions of c into nonempty parts: $c=c^1 \cdots c^r$.

PROPOSITION 3.3. *If c is of length n, and p is a prime, then*

$$D_c(u^p) \equiv 0 \pmod p \qquad \textit{for } n < p; \tag{3.31}$$

$$D_c(u^p) \equiv \sum_{c=c^1 \cdots c^p} \prod_{1 \leq k \leq p} D_{c^k}(u) \pmod p \qquad \textit{for } n \geq p. \tag{3.32}$$

COROLLARIES 3.4. *For c of length n and p prime:*
(3.41) *If u is in F_m and $pm>n$, then*

$$D_c(u^p) \equiv 0.$$

(3.42) *If* $u \equiv v$ (mod F_{n-p+2}), *then*

$$D_c(u^p) \equiv D_c(v^p).$$

(3.43) *If* $n < 2p$, *then*

$$D_c(u^p v^p) \equiv D_c(u^p) + D_c(v^p).$$

To prove (3.41), note that if $pm > n$ then every partition of c into p (nonempty) parts must contain some part c^k of length less than m; hence every term in (3.32) contains a factor $D_{c^k}(u) = 0$. To prove (3.42), note that in every partition of c into p (nonempty) parts, all parts must be of length less than $n-p+2$; hence each $D_{c^k}(u) = D_{c^k}(v)$. To prove (3.43), apply (3.1) to $D_c(u^p v^p)$ with $m=2$, and observe that by (3.31) every term containing a factor for c^k nonempty and of length less than p must vanish; hence only those terms corresponding to $c = c^1c^2$ with one part empty and the other equal to c remain.

If, in $\Delta w = \omega_1 + \omega_2 + \cdots$, all $\omega_k = 0$ for $k < n$, then w lies in F_n. What does it signify if all $\omega_k \equiv 0$ for $k < n$?

PROPOSITION 3.5. *For w in F_h, and $h \leqq k$, suppose that*

$$\Delta w \equiv \omega_k + \omega_{k+1} + \cdots;$$

then, provided that $2 \leqq h \leqq k < 2p$, there exists $w' = wr$ in F_k, where r is in R, such that

$$\Delta w' = \omega_k' + \omega_{k+1}' + \cdots,$$

with $\omega_k' \equiv \omega_k, \cdots, \omega_{2p-1}' \equiv \omega_{2p-1}$.

The case $h = k$ is trivial, while the general case follows by iteration of the case $k = h+1$. Since w is in F_h, ω_h is a Lie element; and $\omega_h \equiv 0$ implies that $\omega_h = -p\zeta$ where ζ is again a Lie element of degree h. Then ζ is the leading term of Δz for some z in F_h. Taking $r = z^p$, $w' = wr$ is in F_{h+1}, with $\omega_h' = 0$. And since, by (3.41), $D_c(r) \equiv 0$ for c of length $n < 2p$, $\Delta r \equiv \rho_{2p} + \rho_{2p+1} + \cdots$ and $\omega_n' \equiv \omega_n$ for $n < 2p$.

(REMARK: The same argument can be applied in the general situation $a \leqq h \leqq k \leqq ap$.)

A special application of the above is to the case of $w = (uv)^p u^{-p} v^{-p}$, for u in F and v in F_h, $h \leqq p$. Clearly w lies in $F_{h+1} \subset F_2$. By (3.43), $D_c(w) \equiv D_c((uv)^p) - D_c(u^p) - D_c(v^p)$ for $n < 2p$, hence for $n < h+p$. By (3.42), since $uv \equiv u$, $v \equiv 1$ (mod F_h), $D_c((uv)^p) \equiv D_c(u^p)$ and $D_c(v^p) \equiv 0$ for $h \geqq n-p+2$, hence for $n < h+p-1$. Therefore $D_c(w) \equiv 0$ for $n < h+p-1$, and $\Delta w \equiv \omega_{h+p-1} + \omega_{h+p} + \cdots$. Applying now (3.5) and noting that w in R implies $w' = wr$ is in R, one has

PROPOSITION 3.6. *Let* $w=(uv)^p u^{-p} v^{-p}$ *where* u *is in* F *and* v *in* F_h, $h \leqq p$. *Then* $\Delta w \equiv \omega_{h+p-1}+\omega_{h+p}+\cdots$ *and there exists* w' *in* R *such that* $\Delta w'=\omega'_{h+p-1}+\omega'_{h+p}+\cdots$ *where* $\omega'_{h+p-1}\equiv\omega_{h+p-1}$.

4. **The quotient Q_n for $n<p$.** The dimension $\mu(n)$ of M_n is the rank of the matrix $\mathfrak{M}_n=[D_c(r)]$ with columns indexed by c in C_n, rows by r in $F_n \cap R$, and elements taken modulo p. Define $\mathfrak{N}_n=[D_c(r)]$ in the same way, but with rows for all $r=u^p$ in R. Every r in R can be written as $r=\prod u_t^{p\lambda_t}$, whence by (3.43), provided $n<2p$, $D_c(r)\equiv \sum \lambda_t D_c(u_t^p)$. It follows that the rows of $\mathfrak{M}_n$ are certain linear combinations of the rows of $\mathfrak{N}_n$.

For $n<p$, all $D_c(u_t^p)\equiv 0$ by (3.31), whence $\mathfrak{N}_n$, and so $\mathfrak{M}_n$, is a 0-matrix. Thus

THEOREM I. $\mu(n)=0$ *for* $n<p$.

5. **The quotient Q_p.** If c is of length p, it follows by (3.42) that $D_c(u^p)$, modulo p, depends upon u only modulo F_2, hence only upon the $D_k(u)=\alpha_k$ modulo p, for $k=1, 2, \cdots, q$. Therefore we may write $[u]=[\alpha_1, \cdots, \alpha_q]$ for the row of $\mathfrak{N}_p$ with elements $D_c(u^p)$.

LEMMA 5.1. *The linear combination* $L=\sum\lambda_t[u(t)]=\sum\lambda_t[\alpha(t)_1, \cdots, \alpha(t)_q]$ *belongs to the row space of* $\mathfrak{M}_p$ *if and only if*

$$\sum \lambda_t \alpha(t)_k \equiv 0 \qquad \textit{for } k = 1, 2, \cdots, q. \tag{5.1}$$

To prove this, first remark that L belongs to (the row space of) $\mathfrak{M}_p$ if and only if there exists some $r=\prod u(t)^{p\lambda_t}$ (order of factors immaterial) in $R\cap F_p$ for which $[u(t)]=[\alpha(t)_1, \cdots, \alpha(t)_q]$. If such r exists, a fortiori

$$r \equiv \prod_t \prod_k x_k^{\alpha(t)_k p\lambda_t} \equiv \left[\prod_k x_k^{\Sigma\lambda_t\alpha(t)_k}\right]^p \equiv 1 \pmod{F_2},$$

and, since F/F_2 is torsion-free, $\sum\lambda_t\alpha(t)_k=0$ for all k. For the converse, any given solution of (5.1) modulo p corresponds to a solution of the equations $\sum\lambda_t\alpha(t)_k=0$ in rational integers. Set $u(t)=\prod x_k^{\alpha(t)_k}$ and $w=\prod u(t)^{p\lambda_t}$. Then the $D_c(w)$ for c in C_p yield the entries in the row L. But w is in $R\cap F_2$, whence, by (3.43) and (3.41), $\Delta w\equiv\omega_p+\omega_{p+1}+\cdots$. By (3.5) there exists w' in $R\cap F_p$ with $\Delta w'=\omega'_p+\omega'_{p+1}+\cdots$ where $\omega'_p\equiv\omega_p$. Thus $D_c(w')\equiv D_c(w)$, and L is the row of $\mathfrak{M}_p$ indexed by w' in $R\cap F_p$.

Next consider the columns of $\mathfrak{N}_p$. For $c=c_1\cdots c_p$ of length p, (3.32) yields $D_c(u^p)\equiv D_{c_1}(u)\cdots D_{c_p}(u)=\alpha_1^{h_1}\cdots\alpha_q^{h_q}$ where $h_1, \cdots, h_q$ are the frequencies of the symbols $1, \cdots, q$ in the sequence c. Write $\phi_c(u)=\alpha_1^{h_1}\cdots\alpha_q^{h_q}$, and, for $L=\sum\lambda_t[u(t)]$, write $\phi_c(L)=\sum\lambda_t\phi_c(u(t))$. The column space of $\mathfrak{N}_p$, hence of $\mathfrak{M}_p$, is thus spanned by columns given by the ϕ_c for all distinct $(h)=(h_1, \cdots, h_q)$ belonging to some c in S_p. Now S_p contains none of the q sequences consisting of p repetitions of the same symbol; while for any other solution of the conditions $\sum h_k=p$, $0\leqq h_k\leqq p$, the sequence c

obtained by arranging the prescribed number of symbols $1, \cdots, q$ in nondescending order belongs to S_p. The number of distinct ϕ_c is therefore

$$\binom{p+q-1}{p} - q.$$

That the ϕ_c, clearly independent over $\mathcal{N}_p$, are independent over $\mathcal{M}_p$ follows from homogeneity considerations (§6). Or, directly, if any combination $\sum \nu_c \phi_c$ vanished on all the rows

$$[\alpha_1, \cdots, \alpha_k + 1, \cdots, \alpha_q] - [\alpha_1, \cdots, \alpha_k, \cdots, \alpha_q] - [0, \cdots, 1, \cdots, 0]$$

of $\mathcal{M}_p$, it would have to be independent of $\alpha_1, \cdots, \alpha_q$, whence all the $\nu_c \equiv 0$.

THEOREM II.

$$\mu(p) = \binom{p+q-1}{p} - q.$$

REMARK. For $p=2$, this gives $\kappa(2) = \psi(2) - \mu(2) = 0$, hence $Q_2 = 1$; in fact, $B_2 = 1$[4]. Since it follows that, for $p=2$, $Q_n = 1$ for all $n \geqq 2$, we may henceforth assume that $p > 2$.

6. **Homogeneity of M_n.** The elements of V_n, regarded as commutator forms in F_n/F_{n+1} reduced modulo p (or as Lie elements), have well-defined degrees in each of the generators $x_1, \cdots, x_q$. For each solution $(h) = (h_1, \cdots, h_q)$ of $\sum h_k = n$, $0 \leqq h_k < n$, define $V(h)$ to be the subspace of all elements that are homogeneous of degree h_k in x_k for each $k = 1, \cdots, q$. Clearly V is the direct sum of the $V(h)$.

Define $M(h) = M_n \cap V(h)$.

LEMMA 6.1. *For $n=p$, for $n=p+1$, and for $p=2$ and $n=p+2$, M_n is the direct sum of its subspaces $M(h)$.*

The case $n=p$ is in fact implicit in the proof of Theorem II, but also falls out of a more general argument. If $L(x_1, \cdots, x_q)$ is a homogeneous form in $V(h)$, then "linear" substitution gives $L(x_1^{e_1}, \cdots, x_q^{e_q}) \equiv e_1^{h_1} \cdots e_q^{h_q} \cdot L(x_1, \cdots, x_q)$. Since R is a characteristic ("word") subgroup of F, the subspace M_n is closed in V_n under substitution. It follows by standard reasoning that M_n has a basis of forms with the property that one of them will contain terms in different $V(h)$ and $V(h')$ only if $e_1^{h_1} \cdots e_q^{h_q} \equiv e_1^{h'_1} \cdots e_q^{h'_q} \pmod{p}$ for all $e_1, \cdots, e_q$. This requires that $h_k = 0$ if and only if $h'_k = 0$, and that, for each k, $h_k \equiv h'_k \pmod{p-1}$.

If $n=p$, there exist no distinct (h) and (h') so related, whence M_n has a basis of elements lying in the various $M(h)$, and therefore is a direct sum.

For $n=p+1$, the pairs of (h) and (h') of this sort are all of the type

(4) Elementary; see Burnside [2].

$(h)=(1,\ p,\ 0,\cdots,0)$, $(h')=(p,\ 1,\ 0,\cdots,0)$. For $n=p+2$, provided $q=2$, they are of type $(h)=(1,\ p+1)$ and $(h')=(p,\ 2)$. Now, for $(h)=(1,n-1,0,\cdots,0)$, $S(h)$ contains only $c=122\cdots 2$, and $V(h)$ is of dimension 1, with basis element $\xi_n=(x_1,x_2,\cdots,x_2)$ ($n-1$ symbols x_2). The proof of Theorem II shows that, for $n=p$, $M(h)$ has dimension 1, hence $M(h)=V(h)$, and ξ_p lies in $R\cap F_{p+1}$. Since $\xi_{n+1}=(\xi_n,x_2)$, it follows inductively that ξ_n lies in $R\cap F_{n+1}$ for all $n\geqq p$, that $M(h)$ has dimension 1, hence that $M(h)=V(h)$. In particular, this gives $M(1,\ p,\ 0,\cdots,0)=V(1,\ p,\ 0,\cdots,0)$ and $M(1,\ p+1)=V(1,\ p+1)$, whence $M_n\cap(V(h)+V(h'))=M(h)+M(h')$, direct sum, in the two cases under consideration.

For each (h), let $C(h)$ consist of all sequences c in C that contain exactly h_k symbols k, for $k=1,\cdots,q$; and define $S(h)=S\cap C(h)$. Let $\mathcal{N}(h)$, $\mathcal{M}(h)$ be the submatrices of $\mathcal{N}_n$, $\mathcal{M}_n$ consisting of those columns indexed by c in $C(h)$, and let $\mu(h)$ be the rank of $\mathcal{M}(h)$. From the homogeneity of the operators D_c, as applied to F_n/F_{n+1}, one deduces

LEMMA 6.2. *For $n=p$, for $n=p+1$, and for $q=2$ and $n=p+2$, one has $\mu(n)=\sum\mu(h)$.*

7. **The quotient Q_{p+1}.** If c is of length $p+1$, it follows by (3.42) that $D_c(u^p)$, modulo p, depends upon u only modulo F_3, and hence only upon the numbers, taken modulo p, $D_k(u)=\alpha_k$ and $D_{ij}(u)=\gamma_{ij}$ for $1\leqq i<j\leqq q$. Therefore we write $[u]=[\alpha_1,\cdots,\alpha_q;\gamma_{12},\cdots,\gamma_{q-1,q}]$ for the row of $\mathcal{N}_{p+1}$ whose entries are $D_c(u^p)$.

LEMMA 7.1. *The linear combination $L=\sum\lambda_t u(t)$ belongs to the row space of $\mathcal{M}_{p+1}$ if and only if $\eta(L)\equiv 0$ for every form $\eta(\alpha_1,\cdots,\alpha_q)$ homogeneous of total degree p in the α_k.*

If L corresponds to some $r=\prod u(t)^{p\lambda_t}$ in $R\cap F_{p+1}$, then, since r is in $R\cap F_p$, all $\sum\lambda_t\alpha(t)_k\equiv 0$ by (5.1). Since, in fact, r is in F_{p+1}, all $D_c(r)=0$ for c in C_p, whence $D_c(r)\equiv\sum\lambda_t\alpha(t)_1^{h_1}\cdots\alpha(t)_q^{h_q}\equiv 0$ for all solutions of $\sum h_k=p$, $0\leqq h_k<p$. In the excluded cases, where some $h_k=p$, with the remaining $h_i=0$, one has $\sum\lambda_t\alpha(t)_k^p\equiv\sum\lambda_t\alpha(t)_k\equiv 0$. Hence $\eta(L)\equiv 0$ for all η.

For the converse, given an L such that $\eta(L)\equiv 0$ for all η, proceeding in the same manner as for Lemma 5.1 we can use the given λ_t and $\alpha(t)_k$ to construct an element $r=\prod u(t)^{p\lambda_t}$ in $R\cap F_{p+1}$ giving rise to a row L' in $\mathcal{M}_{p+1}$ with the same numbers $\alpha(t)_k$ as L. Since this construction provides no control over the γ_{ij}, to prove that L belongs to $\mathcal{M}_{p+1}$ we must show that $\mathcal{M}_{p+1}$ contains all rows of the form

$$K=[\alpha_k;\gamma_{ij}]-[\alpha_k;\gamma'_{ij}].$$

For this, let $\gamma'_{ij}=\gamma_{ij}+\gamma''_{ij}$ and choose u and v such that $[u]=[\alpha_k;\gamma_{ij}]$ and $[v]=[0;\gamma''_{ij}]$: more precisely, $v=\prod_{i<j}(x_i,x_j)^{\gamma_{ij}''}$. Then u is in F and v

in F_2, whence, taking $w=(uv)^p u^{-p} v^{-p}$, by (3.6) with $h=2$ there exists w' in $R \cap F_{p+1}$ such that $D_c(w') \equiv D_c(w)$ for all c in C_{p+1}. Thus w' gives rise to a row $[uv]-[u]-[v]$ in $\mathfrak{M}_{p+1}$. Since v is in F_2, $D_c(v^p) \equiv 0$ for all c in C_{p+1}, by (3.41), and $[v]=0$. Therefore $[uv]-[u]=[\alpha_k;\ \gamma'_{ij}]-[\alpha_k;\ \gamma_{ij}]$ and K belongs to $\mathfrak{M}_{p+1}$, as required.

Next we shall examine the columns of $\mathfrak{N}(h)$ and $\mathfrak{M}(h)$, for fixed (h). For $c=c_1 \cdots c_{p+1}$, (3.32) gives

$$\begin{aligned} D_c(u^p) &\equiv \sum_{k=1}^{p} D_{c_1}(u) \cdots D_{c_{k-1}}(u) D_{c_k c_{k+1}}(u) D_{c_{k+2}}(u) \cdots D_{c_{p+1}}(u) \\ &\equiv A \sum D_{c_k c_{k+1}}(u) / \alpha_{c_k} \alpha_{c_{k+1}} \end{aligned}$$

where $A = \alpha_1^{h_1} \cdots \alpha_q^{h_q}$. For $i<j$, we defined $\gamma_{ij}=D_{ij}(u)$. The shuffle relations $D_i \cdot D_j = D_{ij}+D_{ji}$ $(i \neq j)$ and $D_i \cdot D_i = D_{ii}+D_i+D_{ii}$ give

$$D_{ji} = \alpha_i \alpha_j - \gamma_{ij}, \qquad D_{ii} = \alpha_i^2/2 - \alpha_i/2.$$

For greater symmetry, define, for $i<j$,

$$\theta_{ij} = \frac{\gamma_{ij}}{\alpha_i \alpha_j} - \frac{1}{2}, \qquad \theta_{ji} = -\theta_{ij}, \qquad \theta_{ii} = 0.$$

Then, for $i<j$,

$$\begin{aligned} D_{ij}(u) &= \gamma_{ij} = \alpha_i \alpha_j \theta_{ij} + \alpha_i \alpha_j / 2, \\ D_{ji}(u) &= \alpha_i \alpha_j - \gamma_{ij} = -\alpha_i \alpha_j \theta_{ij} + \alpha_i \alpha_j / 2 = \alpha_j \alpha_i \theta_{ji} + \alpha_j \alpha_i / 2, \\ D_{ii}(u) &= \alpha_i^2/2 - \alpha_i/2 = \alpha_i \alpha_i \theta_{ii} + \alpha_i \alpha_i / 2 - \alpha_i / 2. \end{aligned}$$

In this notation,

$$D_c(u^p) \equiv A \sum_{1 \leqq k \leqq p} \left(\theta_{c_k c_{k+1}} + \frac{1}{2} \right) + \eta(\alpha_1, \cdots, \alpha_q)$$

where η is a form of total degree p in the α_k, and by (7.1) may be neglected in investigating the columns of $\mathfrak{M}_{p+1}$. If, for $1 \leqq i, j \leqq q$, we let h_{ij} be the number of consecutive pairs $c_k c_{k+1} = ij$ in the sequence c, the entries in the column indexed with c are given by

$$\begin{aligned} \phi_c(\alpha_k;\ \gamma_{ij}) &\equiv A \sum_{i,j} h_{ij} \theta_{ij} + \frac{1}{2} pA \\ &\equiv A \sum h_{ij} \theta_{ij}. \end{aligned}$$

To find a basis for these columns, first observe that if $h_i \neq 0$, $h_j \neq 0$, then $C(h)$ will contain, for some k, $c_2, \cdots, c_{p-1}$, sequences $c = k c_2 \cdots c_{p-1} ij$ and $c' = j k c_2 \cdots c_{p-1} i$. Comparing the h_{st} and h'_{st} gives

$$\psi_{ijk} = \phi_c - \phi_{c'} \equiv A(\theta_{ij} - \theta_{jk}).$$

Using $\theta_{jk} = -\theta_{kj}$, and choosing k' from the $c_2, \cdots, c_{p-1}$,

$$\begin{aligned}\psi_{ij} &= \psi_{ijk} + \psi_{ijk'} - \psi_{kjk'} \\ &\equiv A(\theta_{ij} - \theta_{jk} + \theta_{ij} - \theta_{jk'} - \theta_{kj} + \theta_{jk'}) \\ &\equiv 2A\theta_{ij}.\end{aligned}$$

From this it follows that the columns given by the ψ_{ij}, for $i<j$, span the column space of $\mathcal{N}(h)$ and so that of $\mathcal{M}(h)$. We shall show that the ψ_{ij} give independent columns of $\mathcal{M}(h)$. For $s<t$, choose u_{st} with all $\alpha_k=1$, and with all $\gamma_{ij}=0$ except $\gamma_{st}=1$. Choose u_0 with all $\alpha_k=1$ and all $\gamma_{ij}=0$. Evidently $L_{st}=[u_{st}]-[u_0]$ belongs to the row space of $\mathcal{M}_{p+1}$, by Lemma 7.1. But $\psi_{st}(L_{st}) = +1$, while all other $\psi_{ij}(L_{st}) = -1$.

It follows that the rank of $\mathcal{M}_{p+1}$, $\mu(p+1) = \sum \mu(h)$, is the sum, over all (h), of the number of pairs $i<j$ for which $h_i \neq 0$, $h_j \neq 0$. Evidently, this is the sum over all $i<j$, of the number of (h) with $h_i \neq 0$, $h_j \neq 0$, which is evidently

$$\binom{q}{2}\binom{p+q-2}{p-1}.$$

THEOREM III.

$$\mu(p+1) = \binom{q}{2}\binom{p+q-2}{p-1}$$

for $p>2$.

REMARK. For $p=3$, this gives $\kappa(4) = \psi(4) - \mu(4) = 0$, hence $Q_4=1$; in fact, $B_4=1$ (5). Since it follows that, for $p=3$, all $Q_n=1$, $n \geqq 4$, we henceforth assume $p>3$.

8. **The quotient Q_{p+2} for $q=2$.** It is assumed henceforth that B is defined by two generators x_1, x_2, and that $p \geqq 5$. To avoid subscripts, we introduce the alternate notation $x=x_1$, $y=x_2$, $\alpha=\alpha_1=D_1(u)$, $\beta=\alpha_2=D_2(u)$, $\gamma=\gamma_{12} = D_{12}(u)$. If c is of length $p+2$, it follows by (3.42) that $D_c(u^p)$ modulo p depends upon u only through the numbers α, β, γ and $\sigma=D_{112}(u)$, $\tau=D_{122}(u)$. We write $[u]=[\alpha, \beta, \gamma, \sigma, \tau]$ for the row of $\mathcal{N}_{p+2}$ given by the $D_c(u^p)$.

LEMMA 8.1. *The combination* $L=\sum \lambda_t[u_t]$ *belongs to the row space of* $\mathcal{M}_{p+2}$ *if and only if*

(8.1) $$\eta(L) \equiv 0 \text{ for all forms } \eta(\alpha, \beta) \text{ of total degree } p,$$

(8.2) $$\sum \lambda_t \alpha_t^h \beta_t^k \equiv 2 \sum \lambda_t \alpha_t^{h-1} \beta_t^{k-1} \gamma_t$$

for all $1 \leqq h \leqq p$, $k=p+1-h$.

(5) See Burnside [2], Levi-van der Waerden [7].

Observing that, for $q=2$, the columns of $\mathfrak{M}_{p+1}$ are all given by polynomials

$$\psi_{12} = 2A\theta_{12} = 2\alpha^h\beta^k\left(\frac{\gamma}{\alpha\beta} - \frac{1}{2}\right),$$

the proof runs exactly parallel to that of Lemma 7.1.

Next we shall examine the columns of $\mathfrak{N}(h)$ and $\mathfrak{M}(h)$, for a fixed $(h)=(h, k)$, $0<h<p+2$, $k=p+2-h$. For the right member of (3.32), the partitions of $c=c_1 \cdots c_{p+2}$ into p parts are clearly of two kinds:

(i) one segment $c_ic_{i+1}c_{i+2}$, the rest c_j;

(ii) two segments c_ic_{i+1} and c_jc_{j+1}, the rest c_r. According as the c_i, c_j, etc., are 1 or 2, we classify these partitions in the obvious fashion into types

$$111, \cdots, 222, 11/11, \cdots, 22/22.$$

Define the integers $(111), \cdots, (22/22)$ to be the number of partitions of c falling into each of these types. Then, by (3.32),

$$\begin{aligned} D_c(u^p) \equiv{} & A \sum (ijk)D_{ijk}(u)/\alpha_i\alpha_j\alpha_k \\ & + A\sum (ij/rs)D_{ij}(u)D_{rs}(u)/\alpha_i\alpha_j\alpha_r\alpha_s \end{aligned}$$

with summation over all distinct partition types.

By means of the shuffle relations, the $D_{ijk}(u)$ and $D_{ij}(u)D_{rs}(u)$ are all expressible as polynomials in the α, β, γ, σ, τ. For example, from the shuffle relation $D_{12}\cdot D_1 = D_{121}+D_{112}+D_{12}+D_{112}$ we find that

$$\frac{D_{121}(u)}{\alpha^2\beta} = -2\frac{\sigma}{\alpha^2\beta} + 1\frac{\gamma}{\alpha\beta} - 1\frac{\gamma}{\alpha^2\beta}\cdot \tag{8.3}$$

Without entering into further details at this point, it follows that the $D_c(u^p)$ will all be given by polynomials, with certain coefficients $K_\sigma, \cdots, H'_\beta$ depending on c, of the general form

$$\begin{aligned} A\Big\{ & K_\sigma\frac{\sigma}{\alpha^2\beta} + K_\tau\frac{\tau}{\alpha\beta^2} + K_{\gamma\gamma}\frac{\gamma^2}{\alpha^2\beta^2} + K_\gamma\frac{\gamma}{\alpha\beta} + K_1 \\ & + H_\alpha\frac{\gamma}{\alpha^2\beta} + H_\beta\frac{\gamma}{\alpha\beta^2} + H'_\alpha\frac{1}{\alpha} + H'_\beta\frac{1}{\beta}\Big\} + \eta(\alpha, \beta), \end{aligned}$$

where η is a form of total degree p and may be ignored. Further, if $L = \sum\lambda_t[u_t]$ belongs to $\mathfrak{M}_{p+2}$, then by (8.1), since $(h-1)+k=p+1$, we have

$$\sum\lambda_t(H_\alpha\alpha_t^{h-2}\beta_t^{k-1}\gamma_t + H'_\alpha\alpha_t^{h-1}\beta_t^k) \equiv \sum\lambda_t(H_\alpha + 2H'_\alpha)\alpha_t^{h-2}\beta_t^{k-1}\gamma_t,$$

and it follows that, for the purpose of investigating $\mathfrak{M}_{p+2}$, we may describe $D_c(u^p)$ by the polynomial

$$\phi_c = A\left\{K_\sigma \frac{\sigma}{\alpha^2\beta} + K_\tau \frac{\tau}{\alpha\beta^2} + K_{\gamma\gamma}\frac{\gamma^2}{\alpha^2\beta^2} + K_\gamma \frac{\gamma}{\alpha\beta} + K_1 + K_\alpha \frac{\gamma}{\alpha^2\beta} + K_\beta \frac{\gamma}{\alpha\beta^2}\right\}, \tag{8.4}$$

where $K_\alpha = H_\alpha + 2H'_\alpha$ and $K_\beta = H_\beta + 2H'_\beta$.

Although we shall have later to prove only a small part of this fact, it may be noted that routine calculation shows that the monomials $A\sigma/\alpha^2\beta, \cdots, A\gamma/\alpha\beta^2$ define linearly independent functions over the row space of $\mathfrak{M}_{p+2}$.

9. **Continuation.** We next examine how the coefficients K in (8.4) depend upon the numbers (111), $\cdots$, (22/22). From equation (8.3), for example, it appears that each partition of c of the type 121 contributes -2 to K_σ, $+1$ to K_γ, -1 to H_α (and thus to K_α), and nothing to the remaining coefficients. We tabulate the result of analogous computations for the other types of partitions in Table 1.

TABLE 1

	K_σ	K_τ	$K_{\gamma\gamma}$	K_γ	K_1	H_α	H'_α	H_β	H'_β	K_α	K_β
(111)					1/6		$-1/2$			-1	
(222)					1/6				$-1/2$		-1
(112)	1										
(121)	-2			1		-1				-1	
(211)	1			-1	1/2	1	$-1/2$				
(122)		1									
(212)		-2		1				-1			-1
(221)		1		-1	1/2			1	$-1/2$		
(11/11)					1/4		$-1/2$			-1	
(11/22)					1/4		$-1/4$		$-1/4$	$-1/2$	$-1/2$
(22/22)					1/4				$-1/2$		-1
(11/12)				1/2		$-1/2$				$-1/2$	
(11/21)				$-1/2$	1/2	1/2	$-1/2$			$-1/2$	
(22/12)				1/2				$-1/2$			$-1/2$
(22/21)				$-1/2$	1/2			1/2	$-1/2$		$-1/2$
(12/12)			1								
(12/21)			-1	1							
(21/21)			1	-2	1						
	K_σ	K_τ	$K_{\gamma\gamma}$	K_γ	K_1	H_α	H'_α	H_β	H'_β	K_α	K_β

The question now arises of what values of the partition numbers (111), $\cdots$, (22/22) correspond to elements c in $S(h)$. Since these numbers are not independent, we first express them in terms of independent parameters. Every sequence c in $S(h)$ contains h symbols 1 and $p+2-h$ symbols 2; moreover, c must begin with a 1 and end with a 2. We define

$d=0$ or 1 according as c begins with 11 or with 12,
$e=0$ or 1 according as c ends with 22 or with 12,
$a=$ the number of couples $c_ic_{i+1}=12$ in c.

Then all the partition numbers for c are expressible in terms of d, e, a, $b=(112)$, and $f=(122)$. The specific equations are listed in Table 2. We illustrate the method by evaluating $(11/21)$. First, $(11/21)=(11)(21)-(211)$, the number of pairs of segments 11 and 21, minus the number that overlap. Since every 1 begins a pair, $h=(11)+(12)$, and $(11)=h-a$. Since c begins with a 1 and ends with a 2, $(12)=(21)+1$, and $(21)=a-1$. Finally, (11) is equal to the number of triples 111 or 112, and also is equal to the number of triples 111 or 211, plus 1 if $d=0$; hence $(111)+(112)=(111)+(211)+(1-d)$, and $(211)=(112)+d-1=b+d-1$. Combining these gives $(11/21)=(h-a)(a-1)-(b+d-1)$.

TABLE 2. (All entries modulo p.)

$(111)=h-a-b$	
$(222)=2-h-a-f$	
$(112)=b$	$(122)=f$
$(121)=a-f-e$	$(212)=a-b-d$
$(211)=b-d-1$	$(221)=f+e-1$
$(11/11)=\frac{1}{2}[(h-a)^2-(h-a)]-(h-a-b)$	
$(22/22)=\frac{1}{2}[(2-h-a)^2-(2-h-a)]-(2-h-a-f)$	
$(11/22)=(h-a)(2-h-a)$	
$(11/12)=(h-a)a-b$	$(11/21)=(h-a)(a-1)-(b+d-1)$
$(22/12)=(2-h-a)a-f$	$(22/12)=(2-h-a)(a-1)-(f+e-1)$
$(12/12)=\frac{1}{2}(a^2-a)$	$(21/21)=\frac{1}{2}[(a-1)^2-(a-1)]$
$(12/21)=a(a-1)-(a-f-e)-(a-b-d)$	

The results listed in Tables 1 and 2 can now be combined to express the coefficients K in terms of the parameters h, d, e, a, b, c. Straightforward computation gives

$$
\begin{aligned}
K_\sigma &= 2g + 2e + d - 1,\\
K_\tau &= 2g + e + 2d - 1,\\
K_{\gamma\gamma} &= -g - e - d + 1,\\
K_\gamma &= -g - e/2 - d/2,\\
K_1 &= g/12 + 1/12,\\
K_\alpha &= K_\sigma/2, \qquad K_\beta = K_\tau/2,
\end{aligned}
\tag{9.1}
$$

where $g=-a+b+f$.

We are now in a position to determine what polynomials ϕ_c correspond to columns in the matrix $\mathfrak{M}_{p+2}$. For this purpose we may restrict attention to c in $S(h)$. The cases $(h)=(1, p+1)$ and $(h)=(p+1, 1)$, where $\mu(h)=1$, may be dismissed. Since $7\leqq p+2=h+k$, odd, by symmetry we may suppose that $h>k\geqq 2$ and that $h\geqq 4$. Then c must begin with 11, and we may henceforth suppose that $d=0$.

First, let $h>4$, and $k>2$. Then $S(h)$ contains the three sequences listed below, with g and e as shown:

$$\begin{array}{ll} & \\ c = 11 \cdots 122 \cdots 22 \\ c' = 11 \cdots 122 \cdots 212 \\ c'' = 11 \cdots 122 \cdots 2112 \end{array} \qquad \begin{array}{|ccc|cc} a & b & f & g & e \\ \hline 1 & 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 & 1 \\ 2 & 2 & 1 & 1 & 1 \end{array}$$

If the corresponding polynomials are ϕ, ϕ', ϕ'', evidently $\phi_1=\phi''-\phi-\phi'$ has coefficients corresponding to setting $g=e=d=0$ in (9.1); $\phi_2=\phi-\phi_1$ to retaining only the coefficient of g in (9.1); and $\phi_3=\phi'-\phi_1$ to retaining that of e. Explicitly, the first three coefficients of these polynomials are

$$\text{(9.2)} \qquad \begin{array}{c|ccc} & \phi_1 & \phi_2 & \phi_3 \\ \hline K_\sigma & -1 & +2 & +2 \\ K_\tau & -1 & +2 & +1 \\ K_{\gamma\gamma} & +1 & -1 & -1 \end{array}$$

If $h=4$, $k\geqq 3$, and a similar argument applies with c'' replaced by

$$c'' = 1122 \cdots 21212 \qquad \begin{cases} \begin{array}{ccc|cc} 3 & 1 & 1 & -1 & 1 \\ \hline 3 & 1 & 0 & -2 & 1 \end{array} & \begin{array}{l} (\text{for } k>3) \\ (\text{for } k=3). \end{array} \end{cases}$$

If $k=2$, then $h\geqq 5$, and one uses

$$\begin{array}{l} c = 11 \cdots 122 \\ c' = 11 \cdots 1212 \\ c'' = 11 \cdots 12112 \end{array} \qquad \begin{array}{|ccc|cc} 1 & 1 & 1 & -1 & 0 \\ 2 & 1 & 0 & -1 & 1 \\ 2 & 2 & 0 & 0 & 1 \end{array}$$

In all cases, the same ϕ_1, ϕ_2, ϕ_3 define columns spanning $\mathfrak{M}(h)$, and it remains to show that these columns are independent.

Define three rows $L=\sum\lambda_t[u_t]=\sum\lambda_t[\alpha_t, \beta_t, \gamma_t, \sigma_t, \tau_t]$ of $\mathcal{N}_{p+2}$ as follows:

$$\begin{aligned} L_1 &= [1, 1, 0, 1, 0] - [1, 1, 0, 0, 0], \\ L_2 &= [1, 1, 0, 0, 1] - [1, 1, 0, 0, 0], \\ L_3 &= [1, 1, 2, 0, 0] + [1, 1, 0, 0, 0] - 2[1, 1, 1, 0, 0]. \end{aligned}$$

It is easily seen, in accordance with Lemma 8.1, that these lie in the row space of $\mathfrak{M}_{p+2}$. Applying ϕ_c, as given by (8.4), to L_1, one sees that all terms not containing σ cancel, hence that $\phi_c(L_1)\sim K_\sigma$. Similarly, $\phi_c(L_2)\sim K_\tau$. To evaluate $\phi_c(L_3)$, define $\Omega_\nu=[1, 1, \nu, 0, 0]-\nu[1, 1, 1, 0, 0]$; then $\phi_c(\Omega_\nu)$ contains only terms in γ:

$$\phi_c(\Omega_\nu) \sim \nu K_\gamma + \nu^2 K_{\gamma\gamma} + \nu H_\alpha + \nu H_\beta.$$

Since $L_3=\Omega_2-2\Omega_1$, in $\phi_c(L_3)$ those terms that are linear in ν cancel out, leaving

$$\phi_c(L_3) \sim 2^2\cdot K_{\gamma\gamma} - 2\cdot 1^2\cdot K_{\gamma\gamma} = 2K_{\gamma\gamma}.$$

Applying ϕ_1, ϕ_2, ϕ_3 to L_1, L_2, L_3 yields essentially the matrix (9.2) as a submatrix of $\mathfrak{M}(h)$; and since this matrix is clearly nonsingular, $\mu(h)=3$.

Combining this result, for $h=2, \cdots, p$, with the values $\mu(1, p+1)=\mu(p+1, 1)=1$ gives $\mu(p+1)=3(p-1)+2=3p-1$.

THEOREM IV. *$\mu(p+2)=3p-1$ for $p>3$.*

REFERENCES

1. R. Baer, *The higher commutator subgroups of a group*, Bull. Amer. Math. Soc. vol. 50 (1944) pp. 143–160.
2. W. Burnside, *On an unsettled question in the theory of discontinuous groups*, Quart. J. Math. vol. 33 (1902) pp. 230–238.
3. K. T. Chen, R. H. Fox, and R. C. Lyndon, *On the quotient groups of the lower central series* (to appear).
4. R. H. Fox, *Free differential calculus* I, II, Ann. of Math. vol. 56 (1953) pp. 547–560; vol. 59 (1954) pp. 196–210.
5. J. A. Green, *On groups with odd prime-power exponent*, J. London Math. Soc. vol. 27 (1952) pp. 476–485.
6. P. Hall, *A contribution to the theory of groups of prime-power order*, Proc. London Math. Soc. (2) vol. 36 (1934) pp. 29–95.
7. F. Levi and B. L. van der Waerden, *Über eine besondere Klasse von Gruppen*, Abh. Math. Sem. Hansischen Univ. vol. 9 (1932) pp. 154–158.
8. W. Magnus, *Beziehungen zwischen Gruppen und Idealen in einem speziellen Ring*, Math. Ann. vol. 111 (1935) pp. 259–280.
9. ———, *Über beziehungen zwischen höheren Kommutatoren*, J. Reine Angew. Math. vol. 177 (1937) pp. 105–115.
10. ———, *A connection between the Baker-Hausdorff formula and a problem of Burnside*, Ann. of Math. vol. 52 (1950) pp. 111–126.
11. E. Witt, *Treue Darstellung Liescher Ringe*, J. Reine Angew. Math. vol. 177 (1937) pp. 152–160.

PRINCETON UNIVERSITY,
PRINCETON, N. J.