ON BURNSIDE'S PROBLEM

BY

R. C. LYNDON

1. Introduction. Let $B$ be the group on $q$ generators defined by setting the $p$th 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 \leq 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$^1$ 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 $\kappa(n) \leq \psi(n)$. We show that:

(I) $\kappa(n) = \psi(n)$ for $n < p$;

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

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

(IV) $\kappa(p + 2) = \psi(p + 2) - 3p + 1$ for $p > 3$ 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 $\Omega$ 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 $w \rightarrow 1 + \Delta w$...
may be characterized as the unique multiplicative extension, $F$ into $\Omega$, of
the correspondence $x_k \to 1 + \Delta x_k$.

We write $w \to 1 + \Delta w = 1 + \omega_1 + \omega_2 + \cdots$ where $\omega_n$ is the sum of all terms of
total degree $n$ in the $\Delta x_k$. 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 $\omega_n$ is a Lie element in
the $\Delta x_k$, 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 $\Omega$. If $\rho \zeta$ is a Lie element, where $\rho$ is an
integer, then $\zeta$ is a Lie element.

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

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

By extending $\partial/\partial x_k$ linearly to a derivation from $\Gamma$ into $\Gamma$, 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 $\Delta w$:

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

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 $c_k c_{k+1} \cdots c_n$, $1 < k \leq n$.
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 inde-
dependent, 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 em-
bedding 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 poly-
nomial with rational coefficients in the $D_c$ for $c$ in $S_n$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
3. Preliminary constructions. $B = F/R$, where $F$ is free on $q$ generators, and $R$ is generated by all $p^\text{th}$ 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 $p^\text{th}$ 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 $M_n = [D_c(r)]$, with elements taken modulo $p$, is a relation matrix for $Q_n = V_n/M_n$. Hence $\mu(n)$ is the rank of $M_n$.

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 $u_1, \ldots, u_m$ in $F$, one has the "Leibniz rule"

**Proposition 3.1.**

$$D_c(u_1 \cdots u_m) = \sum D_c(u_1) \cdots D_c(u_m),$$

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 $D_c(u_k) = 1$.

Let the terms in (3.1) be grouped according to the number $r$ of non-empty 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

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

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

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

(3.32)  $D_c(u^p) \equiv \prod_{1 \leq k \leq p} D_c(u) \pmod{p}$ for $n \geq p$.

**Corollaries 3.4.** For $c$ of length $n$ and $p$ prime:

(3.41) If $u$ is in $F_m$ and $pm > n$, then

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
\[ D_c(u^p) \equiv 0. \]

(3.42) If \( u \equiv v \pmod{F_{n-p+2}} \), 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 \( p \) (non-empty) 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 \( c = c_1 c_2^* \) 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 \( h \leq k \), suppose that
\[ \Delta w = \omega_1 + \omega_2 + \cdots; \]
then, provided that \( 2 \leq h \leq k < 2p \), there exists \( w' = wr \) in \( F_k \), where \( r \) is in \( R \), such that
\[ \Delta w' = \omega'_1 + \omega'_{k+1} + \cdots, \]
with \( \omega'_k \equiv \omega_k, \cdots, \omega'_{2p-1} \equiv \omega_k \).

The case \( h = k \) is trivial, while the general case follows by iteration of the case \( k = h + 1 \). Since \( w \) is in \( F_h \), \( \omega_h \) is a Lie element; and \( \omega_h \equiv 0 \) implies that \( \omega_h = -p \xi \) where \( \xi \) is again a Lie element of degree \( h \). Then \( \xi \) is the leading term of \( \Delta z \) for some \( z \) in \( F_k \). Taking \( r = 2p \), \( w' = wr \) is in \( F_{k+1} \), with \( \omega'_k = 0 \). 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 \( \omega'_k \equiv \omega_k \) for \( n < 2p \).

(Remark: The same argument can be applied in the general situation \( a \leq h \leq k \leq a p \).)

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 \( F_h \), \( h \leq p \). 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 \( u v \equiv u, v \equiv 1 \pmod{F_h} \), \( D_c((uv)^p) \equiv D_c(u^p) \) and \( D_c(v^p) \equiv 0 \) for \( h \geq n - p + 2 \), hence for \( n < h + p - 1 \). Therefore \( D_c(w) \equiv 0 \) for \( n < h + p - 1 \), and \( \Delta w \equiv \omega_{h+p} + \omega_{h+p-1} + \cdots \). Applying now (3.5) and noting that \( w \) in \( R \) implies \( w' = wr \) is in \( R \), one has
Proposition 3.6. Let \( w = (uv)^p u^{-pv} \) where \( u \) is in \( F \) and \( v \) in \( F_h \), \( h \leq p \). Then \( \Delta w = \omega_{h+p-1} + \omega_{h+p} + \cdots \) and there exists \( w' \) in \( R \) such that \( \Delta w' = \omega_{h+p-1} + \omega_{h+p} + \cdots \) where \( \omega_{h+p-1} = \omega_{h+k-1} \).

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

For \( n < p \), all \( D_c(u_i^p) = 0 \) by (3.31), whence \( N_n \), and so \( M_n \), is a 0-matrix. Thus

Theorem 1. \( \mu(n) = 0 \) 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 \( N_p \) with elements \( D_c(u^p) \).

Lemma 5.1. The linear combination \( L = \sum \lambda_i [u(t)] = \sum \lambda_i [\alpha(t)_1, \cdots, \alpha(t)_q] \) belongs to the row space of \( M_p \) if and only if

\[
\sum \lambda_i \alpha(t)_k \equiv 0 \quad \text{for} \quad k = 1, 2, \cdots, q.
\]

To prove this, first remark that \( L \) belongs to (the row space of) \( M_p \) if and only if there exists some \( r = \prod u_i^{p_i} \) (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

\[
r \equiv \prod \prod k \ x_{k}^{\alpha(t)_k p_i} = \left[ \prod k \ x_{k}^{\alpha(t)_k} \right]^{p} \equiv 1 \pmod{F_2},
\]

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

Next consider the columns of \( N_p \). For \( c = c_1 \cdots c_p \) of length \( p \), (3.32) yields \( D_c(u^p) = D_{c_1}(u) \cdots D_{c_p}(u) = \alpha_{c_1}^h \cdots \alpha_{c_p}^h \) where \( h_1, \cdots, h_q \) are the frequencies of the symbols 1, \( \cdots, q \) in the sequence \( c \). Write \( \phi_c(u) = \alpha_{c_1}^h \cdots \alpha_{c_p}^h \), and, for \( L = \sum \lambda_i [u(t)] \), write \( \phi_c(L) = \sum \lambda_i \phi_c(u(t)) \). The column space of \( N_p \), hence of \( M_p \), is thus spanned by columns given by the \( \phi_c \) 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 \( \leq h_k \leq p \), the sequence \( c \)
obtained by arranging the prescribed number of symbols 1, ⋅ ⋅ ⋅ , q in non-descending order belongs to $S_p$. The number of distinct $\phi_e$ is therefore

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

That the $\phi_e$, clearly independent over $\mathcal{N}_p$, are independent over $\mathcal{M}_p$ follows from homogeneity considerations (§6). Or, directly, if any combination $\sum \nu_e \phi_e$ 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 $\nu_e \equiv 0$.

**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, $B_2 = 1^{(*)}$. Since it follows that, for $p = 2$, $Q_n = 1$ for all $n \geq 2$, 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 $\sum h_k = n$, $0 \leq h_k < n$, 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 $L(x_1^{e_1}, \cdots, x_q^{e_q}) = e_1^{h_1} \cdots e_q^{h_q} \cdot L(x_1, \cdots, x_q)$. 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

(*) Elementary; see Burnside [2].
$(h) = (1, p, 0, \cdots, 0), \ (h') = (p, 1, 0, \cdots, 0).$ For $n = p + 2$, provided $g = 2$, 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 $\xi_n = (x_1, x_2, \cdots, x_n)$ ($n - 1$ symbols $x_2$). The proof of Theorem II shows that, for $n = p$, $M(h)$ has dimension 1, hence $M(h) = V(h)$, and $\xi_p$ lies in $R \cap F_{p+1}$. Since $\xi_{n+1} = (\xi_n, x_2)$, it follows inductively that $\xi_n$ lies in $R \cap F_{n+1}$ for all $n \geq p$, 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 $N(h), M(h)$ be the submatrices of $S_n, S_n$ consisting of those columns indexed by $c$ in $C(h)$, and let $\mu(h)$ be the rank of $N(h)$. 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 $F_q$, and hence only upon the numbers, taken modulo $p$, $D_c(u) = \gamma_{ij}$ for $1 \leq i < j \leq q$. Therefore we write $[u] = [\alpha_1, \cdots, \alpha_q; \gamma_{12}, \cdots, \gamma_{q-1,q}]$ for the row of $N_{p+1}$ whose entries are $D_c(u^p)$.

**Lemma 7.1.** The linear combination $L = \sum \lambda(t) \alpha(t)$ belongs to the row space of $M_{p+1}$ if and only if $\eta(L) = 0$ for every form $\eta(\alpha_1, \cdots, \alpha_q)$ homogeneous of total degree $p$ in the $\alpha_k$.

If $L$ corresponds to some $r = \prod u(t)^{\alpha_k}$ in $R \cap F_{p+1}$, then, since $r$ is in $R \cap F_p$, all $\lambda(t) \in C_p$, whence $D_c(r) = \sum \lambda(t) \alpha(t)^{h_k} \equiv 0$ for all solutions of $\sum h_k = p, 0 \leq h_k < p$. In the excluded cases, where some $h_k = p$, with the remaining $h_i = 0$, one has $\sum \lambda(t)^p \equiv 0$. Hence $\eta(L) = 0$ for all $\eta$.

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

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

For this, let $\gamma'_u = \gamma_{uv} + \gamma''_u$ and choose $u$ and $v$ such that $[u] = [\alpha_u; \gamma_{uv}]$ and $[v] = [0; \gamma'_u]$; more precisely, $v = \prod_{i < j} (x_i, x_j)^{\gamma''_u}$. 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 $D_c(w') = D_c(w)$ for all $c$ in $C_{p+1}$. Thus $w'$ gives rise to a row $[uv] - [u] - [v]$ in $\mathcal{M}_{p+1}$. Since $v$ is in $F_2$, $D_c(wp) = 0$ 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 $\mathcal{M}_{p+1}$, as required.

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

$$D_c(wp) = \sum_{k=1}^{p} D_{c_1}(u) \cdots D_{c_{k-1}}(u) D_{c_k}(v) D_{c_{k+1}}(u) \cdots D_{c_{p+1}}(u)$$

$$= A \sum D_{c_1 c_{k+1}}(u)/\alpha_k \alpha_{k+1}$$

where $A = \alpha_1^{h_1} \cdots \alpha_q^{h_q}$. For $i < j$, we defined $\gamma_{ij} = D_{ij}(u)$. The shuffle relations $D_i : D_j = D_{ij} + D_{ji}$ ($i \neq j$) and $D_i : D_i = 2 D_{ii} + D_{ii}$ give

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

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

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

Then, for $i < j$,

$$D_{ij}(u) = \gamma_{ij} = \alpha_i \alpha_j \theta_{ij} + \alpha_i \alpha_j/2,$$

$$D_{ji}(u) = \alpha_i \alpha_j - \gamma_{ij} = - \alpha_i \alpha_j \theta_{ij} + \alpha_i \alpha_j/2 = \alpha_j \alpha_i \theta_{ji} + \alpha_j \alpha_i/2,$$

$$D_{ii}(u) = \alpha_i/2 - \alpha_i/2 = \alpha_i \alpha_i \theta_{ii} + \alpha_i \alpha_i/2 - \alpha_i/2.$$

In this notation,

$$D_c(wp) = A \sum_{1 \leq k \leq p} \left( \theta_{c_k c_{k+1}} + \frac{1}{2} \right) + \eta(\alpha_1, \cdots, \alpha_q)$$

where $\eta$ is a form of total degree $p$ in the $\alpha_k$, and by (7.1) may be neglected in investigating the columns of $\mathcal{M}_{p+1}$. If, for $1 \leq i, j \leq q$, 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

$$\phi_c(\alpha_k; \gamma_{ij}) = A \sum_{i,j} h_{ij} \theta_{ij} + \frac{1}{2} p A$$

$$= A \sum h_{ij} \theta_{ij}.$$

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 $c = kc_2 \cdots c_{p-1}ij$ and $c' = jkc_2 \cdots c_{p-1}i$. Comparing the $h_{ij}$ and $h'_{j'i}$ gives
\[ \psi_{ijk} = \phi_o - \phi_{i'} \equiv A(\theta_{ij} - \theta_{jk}). \]

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

\[ \psi_{ij} = \psi_{ijk} + \psi_{ijk'} - \psi_{kjk} \]

\[ \equiv A(\theta_{ij} - \theta_{jk} + \theta_{ij} - \theta_{jk} + \theta_{kj}) \]

\[ \equiv 2A\theta_{ij}. \]

From this it follows that the columns given by the \( \psi_{ij} \), for \( i < j \), span the column space of \( N(h) \) and so that of \( M(h) \). We shall show that the \( \psi_{ij} \) give independent columns of \( M(h) \). 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 \( M_{p+1} \), by Lemma 7.1. But \( \psi_{st}(L_{st}) = +1 \), while all other \( \psi_{ij}(L_{st}) = -1 \).

It follows that the rank of \( M_{p+1} \), \( \mu(p+1) = \sum \mu(h) \), 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, \( B_4 = 1 \). Since it follows that, for \( p = 3 \), all \( Q_n = 1, n \geq 4 \), 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 \( p \geq 5 \). 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 \( D_c(uv) \) modulo \( p \) depends upon \( u \) only through the numbers \( \alpha, \beta, \gamma \) and \( \sigma = D_{112}(u), \ \tau = D_{122}(u) \). We write \([u] = [\alpha, \beta, \gamma, \sigma, \tau] \) for the row of \( N_{p+2} \) given by the \( D_c(uv) \).

**Lemma 8.1.** The combination \( L = \sum \lambda_i [u_i] \) belongs to the row space of \( M_{p+2} \) if and only if

\begin{align*}
(8.1) & \quad \eta(L) = 0 \text{ for all forms } \eta(\alpha, \beta) \text{ of total degree } p, \\
(8.2) & \quad \sum \lambda_i \alpha_i^h \beta_i^k = 2 \sum \lambda_i \alpha_i^{h-1} \beta_i^{k-1} \gamma_i \end{align*}

for all \( 1 \leq h \leq p, \ k = p+1-h \).

(*) See Burnside [2], Levi-van der Waerden [7].
Observing that, for $q = 2$, the columns of $M_{p+1}$ 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 $N(h)$ and $M(h)$, 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 $c_i c_{i+1} c_{i+2}$, the rest $c_r$;

(ii) two segments $c_i c_{i+1}$ and $c_j c_{j+1}$, 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),

\[
D_c(u^\nu) = A \sum (ijk)D_{ijk}(u)/\alpha_i \alpha_j \alpha_k
+ A \sum (ij/\nu)D_{ij}(u)D_{\nu}(u)/\alpha_i \alpha_j \alpha_k
\]

with summation over all distinct partition types.

By means of the shuffle relations, the $D_{ijk}(u)$ and $D_{ij}(u)D_{\nu}(u)$ are all expressible as polynomials in the $\alpha$, $\beta$, $\gamma$, $\sigma$, $\tau$. For example, from the shuffle relation $D_1 D_1 = D_{12} + D_{12} + D_{12} + D_{12}$ we find that

\[
(8.3) \quad \frac{D_{121}(u)}{\alpha^2 \beta} = -2 \frac{\sigma}{\alpha^2 \beta} + \frac{\gamma}{\alpha \beta} - \frac{1}{\alpha^2 \beta}.
\]

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

\[
A \left\{ K_s \frac{\sigma}{\alpha^2 \beta} + K_\tau \frac{\tau}{\alpha \beta^2} + K_\gamma \frac{\gamma^2}{\alpha^2 \beta^2} + K_{\gamma \gamma} \frac{\gamma}{\alpha \beta} + K_1 \right.
+ H_\alpha \frac{\gamma}{\alpha^2 \beta} + H_\beta \frac{\gamma}{\alpha \beta^2} + H' \frac{1}{\alpha} + H' \frac{1}{\beta} \right\} + \eta(\alpha, \beta),
\]

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

\[
\sum \lambda_i (H_\alpha^{h-2} \beta_1 \gamma_i + H'_1 \alpha^{h-1} \beta_1) \equiv \sum \lambda_i (H_\alpha + 2H'_1)\alpha_i \beta_1 \gamma_i,
\]

and it follows that, for the purpose of investigating $M_{p+2}$, we may describe $D_c(u^\nu)$ by the polynomial
\[ \phi_c = A \left\{ K_\sigma \frac{\sigma}{\alpha^2 \beta} + K_\tau \frac{\tau}{\alpha \beta^2} + K_{\gamma \gamma} \frac{\gamma^2}{\alpha^2 \beta^2} + K_\gamma \frac{\gamma}{\alpha \beta} \right. \\
+ K_1 + K_\sigma \frac{\gamma}{\alpha^2 \beta} + K_\beta \frac{\gamma}{\alpha \beta^2} \right\}, \]  

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 \( M_{p+2} \).

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 \( K_\sigma \), \(+1\) to \( K_\gamma \), \(-1\) to \( H_\alpha \) (and thus to \( K_\alpha \)), and nothing to the remaining coefficients. We tabulate the result of analogous computations for the other types of partitions in Table 1.

<table>
<thead>
<tr>
<th>( K_\sigma )</th>
<th>( K_\tau )</th>
<th>( K_{\gamma \gamma} )</th>
<th>( K_\gamma )</th>
<th>( K_1 )</th>
<th>( H_\alpha )</th>
<th>( H_\alpha' )</th>
<th>( H_\beta )</th>
<th>( H_\beta' )</th>
<th>( K_\alpha )</th>
<th>( K_\beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(111)</td>
<td>1/6</td>
<td>1/6</td>
<td>-1/2</td>
<td>-1/2</td>
<td>-1</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(222)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(112)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(121)</td>
<td>-2</td>
<td>1</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(211)</td>
<td>1</td>
<td>-1</td>
<td>1/2</td>
<td>1</td>
<td>-1/2</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(222)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(122)</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(212)</td>
<td>-2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(221)</td>
<td>1</td>
<td>-1</td>
<td>1/2</td>
<td>1</td>
<td>-1/2</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11/11)</td>
<td>1/4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11/22)</td>
<td>1/4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22/22)</td>
<td></td>
<td>1/4</td>
<td></td>
<td></td>
<td>-1/2</td>
<td>-1/2</td>
<td>-1/2</td>
<td>-1/2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11/12)</td>
<td>1/2</td>
<td>-1/2</td>
<td>1/2</td>
<td>-1/2</td>
<td>-1/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11/21)</td>
<td></td>
<td></td>
<td>1/2</td>
<td>-1/2</td>
<td>-1/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22/12)</td>
<td>1/2</td>
<td>-1/2</td>
<td>1/2</td>
<td></td>
<td>-1/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22/21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12/12)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12/21)</td>
<td>-1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(21/21)</td>
<td>1</td>
<td>-2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 $c,c+i=12$ 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 $(11/21) = (h-a)(a-1) - (b+d-1)$.

\begin{table}[h]
\begin{center}
\begin{tabular}{l}
$(111) = h-a-b$
\hline
$(222) = 2-h-a-f$
\hline
$(112) = b$
\hline
$(121) = a-f-e$
\hline
$(211) = b-d-1$
\hline
$(11/21) = h-a-b$
\hline
$(22/22) = 2-h-a-f$
\hline
$(11/21) = h-a-b$
\hline
$(22/12) = 2-h-a-f$
\hline
$(12/12) = a-(a-1)-(a-b-d)$
\hline
\end{tabular}
\end{center}
\end{table}

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

$$K_x = 2g + 2e + d - 1,$$
$$K_y = 2g + e + 2d - 1,$$
$$K_{yy} = -g - e - d + 1,$$
$$K_y = -g - e/2 - d/2,$$
$$K_1 = g/12 + 1/12,$$
$$K_a = K_x/2, \quad K_b = K_y/2,$$

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

We are now in a position to determine what polynomials $\phi_c$ correspond to columns in the matrix $M_{p+2}$. 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 $7 \leq p+2 = h+k$, odd, by symmetry we may suppose that $h > k \geq 2$ and that $h \geq 4$. Then $c$ must begin with 11, and we may henceforth suppose that $d=0$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
First, let \( h > 4 \), and \( k > 2 \). Then \( S(h) \) contains the three sequences listed below, with \( g \) and \( e \) as shown:

\[
\begin{align*}
c &= 11 \cdots 122 \cdots 22 &| & 1 & 1 & 1 & 1 & 0 \\
c' &= 11 \cdots 122 \cdots 212 &| & 2 & 1 & 1 & 0 & 1 \\
c'' &= 11 \cdots 122 \cdots 2112 &| & 2 & 2 & 1 & 1 & 1
\end{align*}
\]

If the corresponding polynomials are \( \phi, \phi', \phi'' \), evidently \( \phi_1 = \phi'' - \phi - \phi' \) has coefficients corresponding to setting \( g = \epsilon = d = 0 \) 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

\[
\begin{align*}
\phi_1 & \quad \phi_2 & \quad \phi_3 \\
K_\sigma & -1 & +2 & +2 \\
K_\tau & -1 & +2 & +1 \\
K_{\gamma \gamma} & +1 & -1 & -1
\end{align*}
\]

(9.2)

If \( h = 4 \), \( k \geq 3 \), and a similar argument applies with \( c'' \) replaced by

\[
c'' = 1122 \cdots 21212
\]

\[
\begin{cases}
3 & 1 & 1 \\
3 & 1 & 0
\end{cases}
\begin{cases}
-1 & 1 \\
-2 & 1
\end{cases}
\]

(for \( k > 3 \))

If \( k = 2 \), then \( h \geq 5 \), and one uses

\[
\begin{align*}
c &= 11 \cdots 122 &| & 1 & 1 & 1 & -1 & 0 \\
c' &= 11 \cdots 1212 &| & 2 & 1 & 0 & -1 & 1 \\
c'' &= 11 \cdots 12112 &| & 2 & 2 & 0 & 0 & 1
\end{align*}
\]

In all cases, the same \( \phi_2, \phi_3, \phi_3 \) define columns spanning \( \mathcal{M}(h) \), and it remains to show that these columns are independent.

Define three rows \( L = \sum \lambda_i [u_i] = \sum \lambda_i [\alpha, \beta, \gamma, \sigma, \tau] \) of \( \mathcal{N}_{p+2} \) as follows:

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

It is easily seen, in accordance with Lemma 8.1, that these lie in the row space of \( \mathcal{M}_{p+2} \). Applying \( \phi_c \), as given by (8.4), to \( L_1 \), one sees that all terms not containing \( \sigma \) 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 \( \Omega_\tau = [1, 1, \nu, 0, 0] - \nu [1, 1, 1, 0, 0] \); then \( \phi_c(\Omega_\tau) \) contains only terms in \( \gamma \):

\[
\phi_c(\Omega_\tau) \sim \nu K_\gamma + \nu^2 K_{\gamma \gamma} + \nu H_\alpha + \nu H_\beta.
\]
Since $L_3 = \Omega_2 - 2\Omega_1$, in $\phi_3(L_3)$ those terms that are linear in $\nu$ cancel out, leaving

$$\phi_3(L_3) \sim 2^2 \cdot K_{\gamma\gamma} - 2 \cdot 1^2 \cdot K_{\gamma\gamma} = 2K_{\gamma\gamma}.$$

Applying $\phi_1$, $\phi_2$, $\phi_3$ to $L_1$, $L_2$, $L_3$ yields essentially the matrix (9.2) as a submatrix of $\mathcal{M}(h)$; and since this matrix is clearly nonsingular, $\mu(h) = 3$.

Combining this result, for $h = 2, \ldots, p$, 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**


**Princeton University, Princeton, N. J.**