ON THE COLLECTION PROCESS[1]

EUGENE F. KRAUSE

I. Introduction. Let g_1, g_2 be free generators of a free group G. Form a set C of basic commutators from g_1 and g_2 with an ordering such that $g_1 < g_2$ and $(g_2, g_1; w-1)$ is the first basic commutator of weight w (see [2, pp. 165–166]). As usual

$$(u, v; 1) = u^{-1}v^{-1}uv \quad \text{and} \quad (u, v; n) = ((u, v; n-1), v).$$

For notational convenience we shall often denote, for example, g_1 by 1 or (1), $((g_2, g_1), g_1)$ by 211, and $(((g_2, g_1), g_1), (g_2, g_1))$ by 211; 21. Thus the ordered set $(C, <)$ begins

$$1 < 2 < 21 < 211 < 212 < 2111 < \cdots.$$

The following fundamental theorem is due to Philip Hall [3]. (See also [2, p. 175].)

THEOREM 1. *For any positive integer n,*

$$(g_1g_2)^n = \prod_{u \in C} u^{e_u(n)}, \tag{1}$$

where (i) *the product is ordered by the rule, $s^{e_s(n)}$ is to the left of $t^{e_t(n)}$ if and only if $s < t$;* (ii) *the "$=$" signifies that a valid congruence will obtain* mod G_k *for any positive integer k;* (iii) *$e_u(n)$ is a uniquely determined integer producing polynomial of degree $\leqq$ weight of u.*

REMARK. It has also been observed [2, p. 326] that

$$e_{\underset{K}{211\cdots1}}(n) = \binom{n}{K+1}, \qquad K \geqq 0, \tag{2}$$

where

$$\underset{K}{211 \cdots 1}$$

denotes the commutator $(g_2, g_1; K)$.

Equation (2) has been used to prove that a group, G, of prime exponent, p, satisfies the $(p-1)$th Engel congruence:

$$(y, x; p-1) \in G_{p+1}.$$

Presented to the Society, April 19, 1963; received by the editors February 28, 1963.

[1] This paper is essentially Chapter 2 of the author's doctoral thesis written at the University of Wisconsin under the direction of Professor R. H. Bruck.

Kostrikin, in turn, used this congruence to solve the restricted Burnside problem for prime exponent [5]. In this light, equation (2) appears to be of great value.

The purpose of this paper is to determine the $e_u(n)$ for a much larger class of basic commutators u than that covered by (2). We shall outline an algorithm for computing $e_u(n)$ for all $u \in C$ which, though very tedious in general, is nevertheless practical for certain u's having an uncomplicated structure. Specifically, we shall compute

$$e_{\underbrace{211\cdots1}_{K}\underbrace{22\cdots2}_{L}}(n) = \sum_{m=0}^{n-1} \binom{m}{K}\binom{m}{L} \qquad \text{for } K \geqq 1,\ L \geqq 0, \tag{3}$$

$$e_{\underbrace{211\cdots1}_{K}\underbrace{22\cdots2}_{L};21}(n) = \sum_{m=0}^{n-1} \binom{m}{L}\left\{\binom{m}{2}\binom{m}{K} + \binom{m}{K+1}\right\} \tag{4}$$

$$\text{for } K \geqq 1,\ L \geqq 0 \text{ where, } K \geqq 2 \text{ if } L = 0,$$

$$e_{212;211}(n) = \sum_{m=0}^{n-1} m\left\{2\binom{m}{2} + (m+2)\binom{m}{3}\right\}. \tag{5}$$

Note that when $L=0$, (3) reduces to (2) by a well-known identity.

Note also that formulas (3), (4), (5) provide us with $e_u(n)$ for all u of weight $\leqq 6$. Using these formulas we have been able to prove that any group of exponent 8 satisfies the 14th Engel congruence. This result is stronger than that guaranteed by Sanov's Theorem [4] (23rd Engel congruence), and even stronger than that predicted in a conjecture of R. H. Bruck [1] (15th Engel congruence).

II. **The plan for computing the** $e_u(n)$. Theorem 1 tells us, on the one hand, that

$$\begin{aligned}[(1)(2)]^{n+1} &= (1)(2)[(1)(2)]^n \\ &= (1)(2)(1)^{e_1(n)}(2)^{e_2(n)}21^{e_{21}(n)}\cdots,\end{aligned} \tag{6}$$

and, on the other hand, that

$$[(1)(2)]^{n+1} = (1)^{e_1(n+1)}(2)^{e_2(n+1)}21^{e_{21}(n+1)}\cdots.$$

Obviously $e_u(n+1) = e_u(n)$ plus the number of new u's introduced during the collection of (6) [2, p. 165]. Let $d_u(n)$ be that number; i.e.,

$$d_u(n) = e_u(n+1) - e_u(n). \tag{7}$$

It follows that

$$e_u(n) = \sum_{m=0}^{n-1} d_u(m), \tag{8}$$

and hence the problem of computing the $e_u(n)$ is reduced to that of computing the $d_u(n)$.

It should be noted that the application to (6) of the collection process described in [2, p. 165] yields only basic commutators since all the exponents in (6) are positive.

If $u=1$ or $u=2$ then obviously $d_u(n)=1$ and $e_u(n)=n$. If u is not 1 or 2 then u has the form

$$u = (s, t; r), \tag{9}$$

where s, t are basic commutators, $r \geqq 1$, and r is maximal in the sense that if $s=(s_1, s_2)$ then $s_2 \neq t$; i.e., $s_2 < t$. We shall refer to (9) as the canonical form of u.

As we collect (6) a new u of the form (9) can be introduced only during the collection of the t's. We observe that after the collection of all basic commutators $<t$, (6) has assumed the form

$$[(1)(2)]^{n+1} = \left(\prod_{v<t} v^{e_v(n+1)}\right) \pi_t t^{e_t(n)} \cdots s^{e_s(n)} \cdots, \tag{10}$$

where π_t is a product of basic commutators $\geqq t$, but no basic commutator of form (w, t) appears in π_t. Hence all $d_u(n)$ new u's are introduced by the collection of the t's in (10). This suggests some definitions.

Let z be a basic commutator. We observe, as before, that after the collection of all basic commutators $<z$, (6) has assumed the form

$$[(1)(2)]^{n+1} = \left(\prod_{v<z} v^{e_v(n+1)}\right) \pi_z \left(\prod_{v \geqq z} v^{e_v(n)}\right), \tag{11}$$

where π_z is a product of basic commutators $\geqq z$, but no commutator of form (w_1, w_2) appears in π_z for $w_2 \geqq z$. (Note that (10) is (11) with $z=t$.)

Definition. Let x, y be basic commutators such that $z \leqq x \leqq y$. $P(y, i, x, z)$ denotes the number of y's preceding the ith x in (11)—numbering the x's from left to right. $F(x, i, y, z)$ denotes the number of x's following the ith y in (11)—numbering the y's from left to right.

Definition. Let $N(a, b, c, y, x)$ denote the number of $(y, x; b)$'s introduced as the x's are collected past the y's in $y^c x^a$.

Lemma 1. *For all positive integers a, b, c,*

$$N(a, b, c, y, x) = c\binom{a}{b}.$$

PROOF. The lemma is obviously true when $a=1$. Assume inductively that it is true for a and consider $N(a+1, b, c, y, x)$. When $b=1$,

$$\begin{aligned} N(a+1, 1, c, y, x) &= N(a, 1, c, y, x) + N(1, 1, c, y, x) \\ &= c\binom{a}{1} + c\binom{1}{1} = c\binom{a+1}{1}. \end{aligned}$$

When $b \geqq 2$,

$$\begin{aligned} N(a+1, b, c, y, x) &= N(a, b-1, c, y, x) + N(a, b, c, y, x) \\ &= c\binom{a}{b-1} + c\binom{a}{b} = c\binom{a+1}{b}. \end{aligned}$$

NOTATION. $P_i = P(s, i, t, t)$; i.e., P_i is the number of s's preceding the ith t in (10), where $i=1, 2, \cdots, e_t(n+1)$.

DEFINITION. Define a sequence of functions, f_k, by

$$\begin{aligned} f_1(i) &= \sum_{j=1}^{i} P_j, && i = 1, 2, \cdots, e_t(n+1); \\ f_{k+1}(1) &= 0, && k \geqq 1; \\ f_{k+1}(i) &= \sum_{j=1}^{i-1} f_k(j), && k \geqq 1;\ i = 2, 3, \cdots, e_t(n+1). \end{aligned}$$

LEMMA 2. *The number of* $(s, t; r)$*'s introduced by the collection of the first* h t*'s in* (10) *is* $f_r(h)$ *for* $1 \leqq h \leqq e_t(n+1)$. *In particular, for* u *of form* (9),

$$d_u(n) = f_r(e_t(n+1)). \tag{12}$$

PROOF. The lemma is obviously true when $h=1$ so we assume inductively that it is true for h. If $r=1$, the number of (s, t)'s introduced by the collection of the first $h+1$ t's = the number introduced by the first h t's $+ P(s, h+1, t, t) = f_1(h) + P_{h+1} = f_1(h+1)$. If $r \geqq 2$, the number of $(s, t; r)$'s introduced by the collection of the first $h+1$ t's = the number of $(s, t; r)$'s introduced by the collection of the first h t's + the number of $(s, t; r-1)$'s introduced by the collection of the first h t's $= f_r(h) + f_{r-1}(h) = f_r(h+1)$.

In view of (12) we have found $d_u(n)$ in terms of $P(s, i, t, t)$ for $i=1, 2, \cdots, e_t(n+1)$. But in view of (8) we shall be able to compute $e_t(n+1)$ once we know $d_t(n)$. This provides us with the key to the algorithm (Theorem 4).

III. **The main theorems, with examples.** Unless u, of form (9), has the additional properties that

$$
\begin{aligned}
u &= (s, t; r) \text{ with } r \geqq 1 \text{ in canonical form, and} \\
s &= (s_1, \sigma; a) \text{ with } a \geqq 1 \text{ in canonical form, and} \qquad (13)\\
t &= (t_1, \tau; b) \text{ with } b \geqq 1 \text{ in canonical form,}
\end{aligned}
$$

we can compute $P(s, i, t, t)$ for $i=1, \cdots, e_t(n+1)$ and hence can compute $d_u(n)$ by (12) and $e_u(n)$ by (8). Specifically:

THEOREM 2.

$$
\begin{aligned}
P(2, i, 1, 1) &= 0 && \textit{for } i = 1, \\
&= 1 && \textit{for } i = 2, \cdots, n+1. \\
P(211\underset{a}{\cdots}1, i, 2, 2) &= 0 && \textit{for } i = 1, \qquad a > 0, \\
&= \binom{n}{a} && \textit{for } i = 2, \cdots, n+1.
\end{aligned}
$$

$$
d_{211\underset{a}{\cdots}1\,22\underset{r}{\cdots}2}(n) = \binom{n}{a}\binom{n}{r} \qquad a > 0, r \geqq 0, \tag{14}
$$

and (3) *follows from* (14) *by* (8).

PROOF. All u's of form (9) but not of form (13) fall into two classes: those with $t=1$, those with $t=2$. Use Lemmas 1 and 2.

Before studying the u's of form (13) we need a definition.

DEFINITION. $A(K, L, m, i, y, x)$ denotes the number of $(y, x; K)$'s preceding the ith $(y, x; L)$ in the form yx^m assumes after the collection of the x's. Since $A(K, L, m, i, y, x)$ is independent of x and y we shall write it as $A(K, L, m, i)$.

THEOREM 3. HYPOTHESES. (i) *u is a basic commutator of form* (13); (ii) *n_0 is a positive integer*; (iii) *for all $n \leqq n_0$ we have the following situation*:

1.1. *If $\tau < \sigma$ then we know all of $d_t(n)$; $d_\sigma(n)$; $d_{s_1}(n)$; $P(s_1, i, t, t)$ for $i=1, \cdots, e_t(n+1)$ if $s_1 \geqq t$; $P(t, i, s_1, s_1)$ for $i=1, \cdots, e_{s_1}(n+1)$ if $s_1 < t$; and $P(s_1, i, \sigma, \sigma)$ for $i=1, \cdots, e_\sigma(n+1)$.*

1.2. *If $\tau > \sigma$ then we know all of $d_t(n)$; $d_\tau(n)$; $d_{t_1}(n)$; $P(s, i, t_1, t_1)$ for $i=1, \cdots, e_{t_1}(n+1)$; and $P(t_1, i, \tau, \tau)$ for $i=1, \cdots, e_\tau(n+1)$.*

2.1. *If $\tau = \sigma$ but $s_1 \neq t_1$ then we know all of $d_t(n)$; $d_s(n)$; $d_{t_1}(n)$; $d_{s_1}(n)$; $d_\sigma(n)$; $P(s_1, i, t_1, t_1)$ for $i=1, \cdots, e_{t_1}(n+1)$ if $s_1 > t_1$; $P(t_1, i, s_1, s_1)$ for $i=1, \cdots, e_{s_1}(n+1)$ if $s_1 < t_1$; $P(s_1, i, \sigma, \sigma)$ for $i=1, \cdots, e_\sigma(n+1)$; and $P(t_1, i, \tau, \tau)$ for $i=1, \cdots, e_\tau(n+1)$.*

2.2. *If $\tau = \sigma$ and $s_1 = t_1$ (and hence $a > b$) then we know all of $d_t(n)$; $d_s(n)$; $d_\sigma(n)$; $d_{s_1}(n)$; $P(s_1, i, \sigma, \sigma)$ for $i=1, \cdots, e_\sigma(n+1)$; and $A(a, b, m, i)$ for $m=1, \cdots, e_\sigma(n+1)$ and*

$$i = 1, \cdots, \binom{e_\sigma(n+1)}{b}.$$

CONCLUSION. We can compute $P(s, i, t, t)$ for $i=1, \cdots, e_t(n+1)$ for all $n \leqq n_0$ by means of formulas (15), (16), (17), and (18)—to follow—which only involve numbers obtainable from the hypotheses. Hence, since we also know r and $d_t(n)$ for all $n \leqq n_0$, we can compute $d_u(n)$ for all $n \leqq n_0$ by (12).

$$(15)\quad \begin{cases} \text{If } \tau < \sigma \text{ then for } i = 1, \cdots, e_t(n+1), \\ P(s, i, t, t) = \displaystyle\sum_{j=1}^{P(s_1, i, t, \sigma)} \binom{F(\sigma, j, s_1, \sigma)}{a} \text{ if } P(s_1, i, t, \sigma) \geqq 1, \\ \qquad = 0 \text{ if } P(s_1, i, t, \sigma) = 0. \\ \text{Notice that } P(s_1, i, t, \sigma) = P(s_1, i, t, t) \text{ if } s_1 \geqq t, \\ \qquad\qquad = P(s_1, i, t, s_1) \text{ if } s_1 < t. \end{cases}$$

$$(16)\quad \begin{cases} \text{If } \tau > \sigma, \\ P(s, (k_0 + \cdots + k_v) + i, t, t) = P(s, v+1, t_1, \tau) \\ \text{for } v = 0, \cdots, d_{t_1}(n) - 1; \text{ where } k_0 = 0, \\ \qquad k_h = \binom{F(\tau, h, t_1, \tau)}{b} \text{ for } 1 \leqq h \leqq d_{t_1}(n), \\ \qquad i = 1, \cdots, k_{v+1} \text{ where if } k_{v+1} = 0 \text{ we simply ignore} \\ \qquad\quad (16) \text{ for that value of } v, \text{ and} \\ P(s, d_t(n) + j, t, t) = P(s, d_{t_1}(n) + 1, t_1, t_1) = d_s(n) \\ \text{for } j = 1, \cdots, e_t(n). \text{ Notice that } P(s, i, t_1, \tau) = P(s, i, t_1, t_1). \end{cases}$$

$$(17)\quad \begin{cases} \text{If } \tau = \sigma \text{ but } s_1 \neq t_1, \\ P(s, (k_0 + \cdots + k_v) + i, t, t) = \displaystyle\sum_{j=1}^{P(s_1, v+1, t_1, \sigma)} \binom{F(\sigma, j, s_1, \sigma)}{a} \\ \text{for } v = 0, \cdots, d_{t_1}(n) - 1; \text{ where } k_0 = 0, \\ \qquad k_h = \binom{F(\sigma, h, t_1, \sigma)}{b} \text{ for } 1 \leqq h \leqq d_{t_1}(n), \\ \qquad i = 1, \cdots, k_{v+1} \text{ where we ignore (17) for any } v \text{ such} \\ \qquad\quad \text{that } k_{v+1} = 0, \text{ and} \\ P(s, d_t(n) + j, t, t) = d_s(n) \text{ for } j = 1, \cdots, e_t(n). \\ \text{Notice that } P(s_1, i, t_1, \sigma) = P(s_1, i, t_1, t_1) \text{ if } s_1 > t_1, \\ \qquad\qquad = P(s_1, i, t_1, s_1) \text{ if } s_1 < t_1. \end{cases}$$

$$
(18)\quad \begin{cases}
\text{If } \tau = \sigma,\ s_1 = t_1, \text{ and hence } a > b, \\
P(s, (k_0 + \cdots + k_v) + i, t, t) = \displaystyle\sum_{j=1}^{v} \binom{F(\sigma, j, s_1, \sigma)}{a} \\
\qquad\qquad + A(a, b, F(\sigma, v+1, s_1, \sigma), i) \\
\text{for } v = 0 \cdots, d_{s_1}(n) - 1; \text{ where, } k_0 = 0 \\
\qquad k_h = \dbinom{F(\sigma, h, s_1, \sigma)}{b} \text{ for } 1 \leqq h \leqq d_{s_1}(n), \\
\qquad i = 1, \cdots, k_{v+1} \text{ where we ignore (18) for any } v \text{ such} \\
\qquad\quad \text{that } k_{v+1} = 0, \text{ and} \\
P(s, d_t(n) + j, t, t) = d_s(n) \text{ for } 1 \leqq j \leqq e_t(n).
\end{cases}
$$

PROOF. Use Lemma 1 and properties of the canonical forms in (13).

EXAMPLES OF THE USE OF THEOREM 3.

1. $$u = 211 \underset{K}{\cdots} 1\ 22 \underset{L}{\cdots} 2; 21$$

with $K>0$, $L>0$ satisfies hypothesis 1.1. Formula (15) yields

$$P(s, i, t, t) = \binom{i-1}{K}\binom{n}{L} \text{ for } 1 \leqq i \leqq n;$$

$$P(s, i, t, t) = \binom{n}{K}\binom{n}{L} \qquad \text{for } n+1 \leqq i \leqq \binom{n+1}{2}.$$

Equations (12), (8) yield (4) with the restriction $L>0$.

2. $u=212$; 211 satisfies hypothesis 1.1. Formula (15) yields

$$P(s, i, t, t) = k \text{ for } i = \binom{k-1}{2} + 1, \cdots, \binom{k}{2} \text{ where } 1 \leqq k \leqq \binom{n}{2};$$

$$P(s, i, t, t) = \binom{n+1}{2} \text{ for } \binom{n}{2} + 1 \leqq i \leqq \binom{n+1}{3}.$$

Equations (12), (8) yield (5).

3. $$u = 211 \underset{K}{\cdots} 1; 21$$

with $K \geqq 2$ satisfies hypothesis 2.2. Formula (18) yields

$$P(s, i, t, t) = \binom{i-1}{K} \text{ for } i = 1, \cdots, n;$$

$$P(s, i, t, t) = \binom{n}{K} \qquad \text{for } i = n+1, \cdots, \binom{n+1}{2}.$$

Equations (12), (8) yield (4) in the case of $L=0$.

COMMENT. The author wishes to thank the referee for pointing out that one can also obtain formulas (3), (4), and (5) from formula (8) by applying the methods of P. Hall [3] to (6). These methods include assigning the labels $1, 2, \cdots, n+1$ in order to each of the two basic commutators of weight one which appear in (6), and then determining what "precedence conditions" are satisfied by the n-tuple of labels corresponding to a basic commutator of weight n which arises during the collection of (6).

IV. The algorithm.

THEOREM 4. *Let D and n_0 be positive integers. Suppose we know all of the following values*: (i) $d_w(n)$ *for all w of weight $\leqq D$ and for all $n \leqq n_0$*; (ii) $P(s', i, t', t')$ *for $i=1, \cdots, e_{t'}(n+1)$ for all $n \leqq n_0$ for all v of weight $\leqq D$ having canonical form $v=(s', t'; r')$*; (iii) $A(a, b, m, j)$ *for all positive integers a, b, m, j. Then, by means of Theorems* 2 *and* 3 *and equation* (12), *we can compute* (i) *and* (ii) *with D replaced by $D+1$.*

PROOF. Take u of weight $D+1$—say u has form (9). If u does not have form (13), Theorem 2 gives us our result. Study the u's of form (13) in four cases: case 1.1—$\tau<\sigma$, case 1.2—$\tau>\sigma$, case 2.1—$\tau=\sigma$ but $s_1 \neq t_1$, case 2.2—$\tau=\sigma$ and $s_1=t_1$. In case 1.1 we want to apply Theorem 3 so we must have hypothesis 1.1 holding. To see this is true we observe that t, σ, and s_1 are of weight $\leqq D$; that $s_1=t$ implies $P(s_1, i, t, t)=i-1$; that $s_1>t$ implies (s_1, t) is basic and of weight $\leqq D$; that $s_1<t$ implies (t, s_1) is basic and of weight $\leqq D$; that (s_1, σ) is basic and of weight $\leqq D$. Then we apply hypotheses (i), (ii) of the present theorem. The other cases are handled similarly.

REFERENCES

1. R. H. Bruck, *On the restricted Burnside problem*, Arch. Math. **13** (1962), 179–186.

2. M. Hall, Jr., *The theory of groups*, Macmillan, New York, 1959.

3. P. Hall, *A contribution to the theory of groups of prime-power order*, Proc. London Math. Soc. **36** (1934), 29–95.

4. R. R. Struik, *Notes on a paper by Sanov*. II, Proc. Amer. Math. Soc. **12** (1961), 758–763.

5. A. I. Kostrikin, *The Burnside problem*, Izv. Akad. Nauk SSSR Ser. Mat. **23** (1959), 3–34.

UNIVERSITY OF MICHIGAN