Journal of the Iranian Mathematical Society
ISSN (on-line): 2717-1612
Vol. 1 No. 3 (2020), pp. 173-188

SANOV'S THEOREM ON LIE RELATORS IN GROUPS OF EXPONENT p

M. VAUGHAN-LEE

Abstract. I give a proof of Sanov's theorem that the Lie relators of weight at most $2p-2$ in groups of exponent p are consequences of the identity $px = 0$ and the $(p-1)$-Engel identity. This implies that the order of the class $2p-2$ quotient of the Burnside group $B(m,p)$ is the same as the order of the class $2p-2$ quotient of the free m generator $(p-1)$-Engel Lie algebra over GF(p). To make the proof self-contained I have also included a derivation of Hausdorff's formulation of the Baker Campbell Hausdorff formula.

1. Introduction

The theory of Lie relators in groups of prime-power exponent has been immensely useful in understanding these groups. Probably the first significant result in this direction is a theorem due to Magnus [11], where he proves that the Lie relators of weight at most $p-1$ in groups of exponent p are all consequences of the identity $px = 0$. In this note I give a proof of Sanov's theorem [15] that the Lie relators of weight at most $2p-2$ in groups of exponent p are all consequences of the identity $px = 0$ and the $(p-1)$-Engel identity $[x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$. Perhaps a brief reminder of the definition of Lie relators is appropriate here. Let $B(m,p)$ be the free m generator Burnside group of exponent p, and let

$$\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n \geq \cdots$$

Communicated by Alireza Abdollahi
MSC(2020): Primary: 20D15; Secondary: 20F40.
Keywords: Sanov's theorem; Lie relators; groups of exponent p.
Received: 13 May 2020, Accepted: 20 July 2020.
DOI: https://dx.doi.org/10.30504/jims.2020.110856

be the descending central series of $B(m,p)$. For $n = 1, 2, \ldots$ we set L_n equal to the quotient group γ_n/γ_{n+1}. So L_n is an abelian group, and we think of it as a $\mathbb{Z}$-module and set

$$L(m,p) = \bigoplus_{n=1}^{\infty} L_n.$$

We turn $L(m,p)$ into a Lie ring (the associated Lie ring of $B(m,p)$) as follows. If $a = g\gamma_{i+1} \in L_i$ and $b = h\gamma_{j+1} \in L_j$ then we define the Lie product $[a,b]$ to be $[g,h]\gamma_{i+j+1} \in L_{i+j}$, where $[g,h]$ is the group commutator $g^{-1}h^{-1}gh$. We extend this Lie product to the whole of $L(m,p)$ by linearity, and this turns $L(m,p)$ into a Lie ring. Let $g_1, g_2, \ldots, g_m$ be free generators of $B(m,p)$ and let $a_i = g_i\gamma_2 \in L_1$. Then $a_1, a_2, \ldots, a_m$ generate $L(m,p)$. Furthermore $L(m,p)$ has a natural grading since $[L_i, L_j] \leq L_{i+j}$, and L_n is the linear span of Lie products of weight n in the generators $a_1, a_2, \ldots, a_m$. We say that an element $a \in L(m,p)$ is homogeneous of weight n if $a \in L_n$. Now let Λ_m be the free Lie ring with free generators $x_1, x_2, \ldots, x_m$. Then Λ_m is also graded by weight:

$$\Lambda_m = \bigoplus_{n=1}^{\infty} \Lambda_{m,n},$$

where $\Lambda_{m,n}$ is spanned by Lie products of weight n in the free generators. We let $\pi : \Lambda_m \to L(m,p)$ be the homomorphism mapping x_i to a_i for $i = 1, 2, \ldots, m$, and we let $I = \ker \pi$. Note that $\Lambda_{m,n}\pi = L_n$, so that I is also graded and

$$I = \bigoplus_{n=1}^{\infty} I_n,$$

with $I_n \leq \Lambda_{m,n}$. Elements of I are called Lie relators of $B(m,p)$, and if $a \in I_n$ then a is called a (homogeneous) Lie relator of weight n. As an abelian group, $\gamma_n/\gamma_{n+1} = L_n \cong \Lambda_{m,n}/I_n$. Furthermore the lower central factors γ_n/γ_{n+1} are finite, and so if we know the Lie relators of weight n then we can compute the order of γ_n/γ_{n+1}. All this is discussed in some detail in my book [16], where it is shown that as well as having characteristic p, the associated Lie rings of groups of exponent p satisfy a sequence of multilinear identities $K_n = 0$, with one identity for each $n \geq p$. It is also shown in [16] that the $(p-1)$-Engel identity is equivalent in characteristic p to the multilinear identity $K_p = 0$. However Sanov's theorem implies that the identities $K_n = 0$ for $p < n \leq 2p-2$ are all consequences of the identity $K_p = 0$ in characteristic p.

So Magnus's theorem implies that the class $p-1$ quotient $B(m,p)/\gamma_p$ has the same order as the class $p-1$ quotient of $\Lambda_m/p\Lambda_m$. Similarly Sanov's theorem implies that the class $2p-2$ quotient $B(m,p)/\gamma_{2p-1}$ has the same order as the class $2p-2$ quotient of the free m generator Lie ring in the variety of Lie rings defined by the identities $px = 0$ and $[x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$. Note that the free m generator Lie ring in this variety can be identified with the free $(p-1)$-Engel Lie algebra of rank m over GF(p). Of course the close connection between a group and its associated Lie ring means that much more information than just orders can be deduced from these theorems. There are various reasons which have led me to writing up a proof of Sanov's theorem. Firstly, I wanted to understand the proof myself. But I found that Sanov's original paper (written in Russian of course) has not been

DOI: https://dx.doi.org/10.30504/jims.2020.110856

translated into English. Furthermore Sanov's proof relies heavily on Hausdorff's formulation of the Baker-Campbell-Hausdorff formula given in his 1906 paper [5]. Hausdorff's formulation gives much more detailed information about the formula than can be obtained from more modern treatments such as may be found in Jacobson's book *Lie Algebras* [9], or in my book [16]. Hausdorff's paper was written in German, and appeared in a German physics journal which is no longer published. It has not been reviewed in Mathematical Reviews, and I was unable to obtain a copy of the paper. A full statement of Hausdorff's theorem is given in Sanov's paper, and a somewhat clearer statement can be found in Kostrikin's paper [10]. But neither of these papers give proofs. However I was able to find a proof of the first part of Hausdorff's theorem in a paper by Baker [4], and I was able to complete the proof myself. In the light of all my difficulties in tracking down a self-contained proof of Sanov's theorem, I thought it would be worthwhile committing my researches to paper.

But my main reason for writing this note, and in fact the reason that I wanted to understand the proof of Sanov's theorem in the first place, is more complicated! Over a period of many years Seymour Bachmuth circulated various drafts of a short paper in which he claimed to prove that the two generator Burnside group $B(2,q)$ is finite for all prime powers q. As well as running counter to Adjan's and Olschanskii's negative solutions of the Burnside problem [1, 14], the claims in Seymour's paper also run counter to computer calculations of the largest finite quotients, $R(2,5)$ and $R(2,7)$, of the Burnside groups $B(2,5)$ and $B(2,7)$.

The original draft of Seymour's paper was only about 10 pages long, but over the years it grew slightly. In 2008 Seymour posted a 14 page version of his paper on the arXiv [2], and in 2016 he posted a 22 page version on the arXiv [3]. In most versions of his paper he takes about 10 pages to construct a certain two generator group $F(\mathcal{S}[t,t^{-1}])$ and to prove that it is solvable. All this is essentially correct, though it is rather sloppily written and hence not all that easy to follow. In fact it takes less than a page to define his group, which is generated by two elements M_1 and M_2T, and it takes less than two pages to show that the normal closure of the generator M_1 is nilpotent. Solvability of $F(\mathcal{S}[t,t^{-1}])$ follows immediately from this. In the case when q is prime the normal closure of M_1 has class $q-1$. (In fact it was I who first pointed out to Seymour that his group $F(\mathcal{S}[t,t^{-1}])$ is solvable. Before that he only claimed to have a new proof of the restricted Burnside problem.) For a short definition of $F(\mathcal{S}[t,t^{-1}])$, and a short proof that it is solvable see http://users.ox.ac.uk/~vlee/bachmuth.pdf.

After constructing his group $F(\mathcal{S}[t,t^{-1}])$ he makes the claim that it is a preimage of the Burnside group $B(2,q)$, and he deduces that $B(2,q)$ is finite. In early drafts he just stated that $F(\mathcal{S}[t,t^{-1}])$ is a preimage of $B(2,q)$ as a fact, without any attempt at justification. In later drafts he did make some attempt to justify this claim, but these attempts all seemed to me to be meaningless gobbledygook. I corresponded with Seymour over this point for a year or more, but I never managed to convince him that there was a gap in his argument. Seymour kept saying that since his paper was so short and straightforward and so obviously correct then Adjan's and Olschanskii's work, and all the computer calculations must be wrong. Now of course I have a lot more faith in Adjan and Olschanskii and in all their students than I do in Seymour. But I don't understand their work *at all*. However I do understand the computer calculations of $R(2,5)$ and $R(2,7)$ ([6, 13]) in considerable detail. If

DOI: https://dx.doi.org/10.30504/jims.2020.110856

Seymour's claims were correct then the normal closure of one of the free generators of $B(2,5)$ would have class at most 4. This would imply that both free generators of $B(2,5)$ have normal closures of class at most 4, so that $B(2,5)$ (and hence also $R(2,5)$) would have class at most 8. Similarly, if Seymour's claims were correct then $R(2,7)$ would have class at most 12. However the computer calculations that the actual classes of these groups are 12 and 28 are not all that easy to justify. Most of the calculations involve the p-quotient algorithm, and as far as I know all the implementations of that algorithm in use today are based on George Havas's original program [7] consisting of 7000 lines of Fortran. There are a good number of reputable mathematicians around the world who can vouch for the correctness of the p-quotient algorithm, but only a handful who understand George's implementation. I do understand his implementation in considerable detail and I am convinced that his approach is valid. But I certainly could *not* prove that there is no bug in his 7000 lines of code. Seymour said that he just didn't believe that the p-quotient algorithm was producing correct results, and that he would need to see a permutation representation or a matrix representation of $R(2,5)$ before he would believe the claim that it has class 12 and order 5^{34}. Now $R(2,5)$ has a core free subgroup of index 5^9, so it is actually quite easy to obtain a faithful permutation representation of degree 5^9. But I didn't think that Seymour would think much of that. So I looked for a matrix representation, and I was able to find two 66×66 upper unitriangular matrices over GF(5) which generate $R(2,5)$. (The matrices can be found on my website http://users.ox.ac.uk/~vlee/selected.htm.) Of course you can't multiply these matrices by hand, but you can look at them! And as long as you believe that computers know how to multiply matrices, you can subject the two matrices to any tests you like to verify that they generate a group of exponent 5 which has order 5^{34} and class 12. Seymour, of course, wouldn't have it, and just said that he was unable to verify that the group has exponent 5. (This does take a minimal amount of theory due to Higman [8], who showed that to prove that a group which is nilpotent of class c has exponent n it is only necessary to check that $g^n = 1$ for every element g in the group which can be written as a product of length at most c in the generators.) It was some while after I stopped corresponding with Seymour that I realized that Sanov's theorem together with a few lines of hand calculation imply that there are non-trivial commutators of weight 8 in $R(2,5)$ which have weight 5 in one generator, and weight 3 in the other generator. This runs against Seymour's claim that the normal closure of a generator in $B(2,p)$ (and hence in $R(2,p)$) has class $p-1$. Similarly a few lines of hand calculation show that if $p > 5$ then there are non-trivial commutators of weight $p+2$ in $R(2,p)$ which have weight p in one generator, and weight 2 in the other. These calculations are given at the end of this note. It was this observation which led me on my quest to understand Sanov's theorem, and to give a self-contained proof of the theorem.

2. The Baker Campbell Hausdorff formula

In this section we derive Hausdorff's version of the Baker Campbell Hausdorff formula. The setting is as follows. We let A be the free associative algebra over the rationals $\mathbb{Q}$, freely generated by $x_1, x_2, \ldots$. The algebra A has a natural grading by weight: if we set A_n equal to the vector space over

DOI: https://dx.doi.org/10.30504/jims.2020.110856

$\mathbb{Q}$ spanned by the products of weight n in the generators then

$$A = A_1 \oplus A_2 \oplus \ldots \oplus A_n \oplus \ldots,$$

and $A_m A_n = A_{m+n}$.

We turn A into a Lie algebra over $\mathbb{Q}$ by defining a Lie product $[,]$ on A setting $[a, b] = ab - ba$. We let L be the Lie subalgebra of A generated by $x_1, x_2, \ldots$. Then L is a free Lie algebra over $\mathbb{Q}$, freely generated by $x_1, x_2, \ldots$. (See [12].) We identify the free Lie ring Λ of countably infinite rank with the Lie subring of A generated by $x_1, x_2, \ldots$.

We let P be the ring of formal power series

$$a_0 + a_1 + \ldots + a_n + \ldots,$$

with $a_0 \in \mathbb{Q}$ and $a_n \in A_n$ for $n = 1, 2, \ldots$. If $a = a_0 + a_1 + \ldots + a_n + \ldots \in P$, and if $a_0 = a_1 = \ldots = a_{n-1} = 0$, but $a_n \neq 0$ then we say that a_n is the *leading term* of a, and we say that the leading term of a has weight n. More generally we call a_m the homogeneous component of a of weight m. If $a \in P$, and if the leading term of a has weight at least 1, then we set

$$\mathrm{e}^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \ldots.$$

A key observation is that the elements e^{x_1}, e^{x_2}, e^{x_3}, $\ldots$ are free generators of a free group. (See, for example, pages 41 and 42 of [16].) Note that the group inverse of e^{x_i} is e^{-x_i}. Note also that since the generators of A do not commute the familiar formula

$$\mathrm{e}^x \mathrm{e}^y = \mathrm{e}^{x+y}$$

no longer holds true. The Baker Campbell Hausdorff formula is what replaces this formula. If $a \in P$ has leading term 1 then we can write $a = 1 + u$ where u has leading term with weight at least 1, and we define

$$\log a = u - \frac{u^2}{2} + \frac{u^3}{3} - \ldots.$$

If x and y have leading terms with weight at least 1 then $\mathrm{e}^x \mathrm{e}^y = \mathrm{e}^z$ where $z = \log(\mathrm{e}^x \mathrm{e}^y)$. The main content of the Baker Campbell Hausdorff formula is the remarkable (even amazing) fact that if x and y are free generators of A then the homogeneous components of z are Lie elements of A (i.e. elements of L). More generally, if we let F be the free group generated by e^{x_1}, e^{x_2}, $\ldots$, and if $\mathrm{e}^z \in F$ then the homogenous components of z are Lie elements of A. In addition the leading term of z is a $\mathbb{Z}$-linear combination of Lie products of the free generators of A, in other words an element of Λ.

A key idea exploited by Hausdorff in his proof of this result is the notion of Hausdorff differentiation. This is defined as follows. Let $a \in A$, and let x be one of the free generators of A. We let t be an indeterminate scalar in $\mathbb{Q}$ and let $\pi : A \to A$ be the homomorphism which maps $x \mapsto x + ta$, and which maps every other free generator of A to itself. Then if $b \in A$ we can express

$$\pi b = b + tb_1 + t^2 b_2 + \ldots + t^k b_k$$

for some $b_1, b_2, \ldots, b_k \in A$. We define the operator $a\frac{\partial}{\partial x} : A \to A$ by setting $a\frac{\partial}{\partial x}(b) = b_1$. Perhaps a more intuitive way of defining $a\frac{\partial}{\partial x}$ is the following. Suppose b is a product of the free generators of

DOI: https://dx.doi.org/10.30504/jims.2020.110856

A with k entries in the product equal to x. Then $a\frac{\partial}{\partial x}(b)$ is a sum of k distinct products where these k products are obtained from b by successively replacing each of the k occurrences of x by a. For example if x, y, z are free generators of A, and if $b = xxyxzxy$ then

$$a\frac{\partial}{\partial x}(b) = axyxzxy + xayxzxy + xxyazxy + xxyxzay.$$

We then extend $a\frac{\partial}{\partial x}$ to the whole of A by linearity. (If $b \in \mathbb{Q}$ then we set $a\frac{\partial}{\partial x}(b) = 0$.) We can of course extend this operation to P, allowing a to be an element of P.

Lemma 2.1. *If $a \in P$ and if x is a free generator of A, then*

$$\begin{aligned} &a\frac{\partial}{\partial x}(x^m) \\ &= ax^{m-1} + xax^{m-2} + x^2ax^{m-3} + \ldots + x^{m-1}a \\ &= [a, \underbrace{x, \ldots, x}_{m-1}] + mx[a, \underbrace{x, \ldots, x}_{m-2}] + \binom{m}{2}x^2[a, \underbrace{x, \ldots, x}_{m-3}] + \ldots + \binom{m}{m-1}x^{m-1}a. \end{aligned}$$

Proof. The proof is by induction on m, the cases $m = 1, 2$ being easy to check. So suppose the result is true for m. Then

$$\begin{aligned} &a\frac{\partial}{\partial x}(x^{m+1}) \\ &= a\frac{\partial}{\partial x}(x^m)x + x^m a \\ &= [a, \underbrace{x, \ldots, x}_{m-1}]x + mx[a, \underbrace{x, \ldots, x}_{m-2}]x + \ldots + \binom{m}{m-1}x^{m-1}ax + x^m a. \end{aligned}$$

Now

$$[a, \underbrace{x, \ldots, x}_{r}]x = x[a, \underbrace{x, \ldots, x}_{r}] + [a, \underbrace{x, \ldots, x}_{r+1}]$$

for any $r \geq 0$, and so $a\frac{\partial}{\partial x}(x^{m+1})$ equals

$$[a, \underbrace{x, \ldots, x}_{m}] + \ldots + \left(\binom{m}{r} + \binom{m}{r+1}\right)x^{r+1}[a, \underbrace{x, \ldots, x}_{m-r-1}] + \ldots + (m+1)x^m a.$$

The lemma follows immediately, since

$$\binom{m}{r} + \binom{m}{r+1} = \binom{m+1}{r+1}.$$

□

Corollary 2.2. $a\frac{\partial}{\partial x}(\mathrm{e}^x) = \mathrm{e}^x(a + \frac{1}{2!}[a, x] + \frac{1}{3!}[a, x, x] + \ldots + \frac{1}{n!}[a, \underbrace{x, \ldots, x}_{n-1}] + \ldots)$.

Proof. If we set $X = x^m$ then Lemma 2.1 can be expressed in the form

$$a\frac{\partial}{\partial x}(X) = X'a + \frac{X''}{2!}[a, x] + \frac{X'''}{3!}[a, x, x] + \ldots + \frac{X^{(n)}}{n!}[a, \underbrace{x, x, \ldots, x}_{n-1}] + \ldots.$$

We obtain the corollary by substituting e^x for X in this expression. □

DOI: https://dx.doi.org/10.30504/jims.2020.110856

We now consider the inverse of the power series $1 + \frac{x}{2!} + \frac{x^2}{3!} + \ldots + \frac{x^n}{(n+1)!} + \ldots$.

$$(1 + \frac{x}{2!} + \frac{x^2}{3!} + \ldots + \frac{x^n}{(n+1)!} + \ldots)^{-1} = 1 - \frac{x}{2} + \frac{x^2}{12} - \frac{x^4}{720} + \frac{x^6}{30\,240} + \cdots .$$

The Bernoulli numbers $B_0, B_1, B_2, \ldots$ are defined in terms of this inverse by setting

$$(1 + \frac{x}{2!} + \frac{x^2}{3!} + \ldots + \frac{x^n}{(n+1)!} + \ldots)^{-1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}.$$

Thus $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$, $B_5 = 0$, $B_6 = \frac{1}{42}$. In fact it is known that $B_{2n+1} = 0$ for all $n > 0$, but we do not need to use this fact.

Corollary 2.3. *If* $a\frac{\partial}{\partial x}(\mathrm{e}^x) = b\frac{\partial}{\partial x}(\mathrm{e}^x)$ *then* $a = b$.

Proof. Let the linear map $X : P \to P$ be defined by $aX = [a, x]$, and let

$$Z = 1 + \frac{X}{2!} + \frac{X^2}{3!} + \ldots + \frac{X^n}{(n+1)!} + \ldots .$$

Note that Z is invertible, with inverse

$$\sum_{n=0}^{\infty} B_n \frac{X^n}{n!}.$$

Then Corollary 2.2 implies that $\mathrm{e}^x(aZ) = \mathrm{e}^x(bZ)$ and so $a = b$. □

Lemma 2.4. *If* x *and* y *are free generators of* A*, and if we set*

$$a = y - \frac{1}{2}[y, x] + \frac{B_2}{2!}[y, x, x] + \ldots + \frac{B_n}{n!}[y, \underbrace{x, x, \ldots, x}_{n}] + \ldots$$

then $a\frac{\partial}{\partial x}(\mathrm{e}^x) = \mathrm{e}^x y$.

Proof. Note that $a = yZ^{-1}$, where Z is as defined in the proof of Corollary 2.3. Corollary 2.2 implies that $a\frac{\partial}{\partial x}(\mathrm{e}^x) = \mathrm{e}^x(aZ)$ and so $a\frac{\partial}{\partial x}(\mathrm{e}^x) = \mathrm{e}^x y$. □

Now let a be as defined in Lemma 2.4, set $a_1 = a$, and inductively define $a_{n+1} = a\frac{\partial}{\partial x}(a_n)$ for $n \geq 1$. Note that a_1 is an infinite sum of Lie elements which are homogeneous of degree 1 in y. For each $n \geq 1$, a_n is an infinite sum of Lie elements which are homogeneous of degree n in y. We are at last in a position to state Hausdorff's theorem.

Theorem 2.5 (Hausdorff [5])**.** $\mathrm{e}^x\mathrm{e}^y = \mathrm{e}^z$ *where* $z = x + a_1 + \frac{a_2}{2!} + \ldots + \frac{a_n}{n!} + \ldots$.

Proof. We express z in the form $z = x + b_1 + \frac{b_2}{2!} + \ldots + \frac{b_n}{n!} + \ldots$ where, for each n, b_n is an infinite sum of terms each of which is homogeneous of degree n in y. We show by induction on n that $b_n = a_n$ for all n.

First consider the case $n = 1$. If we pick out the terms of degree 1 in y from

$$\mathrm{e}^z = 1 + (x + b_1 + \frac{b_2}{2!} + \ldots) + \frac{1}{2!}(x + b_1 + \frac{b_2}{2!} + \ldots)^2 + \ldots$$

then we obtain $b_1\frac{\partial}{\partial x}(\mathrm{e}^x)$. On the other hand, if we pick out the terms of degree 1 in y from $\mathrm{e}^x\mathrm{e}^y$ then we obtain $\mathrm{e}^x y$. But by Lemma 2.4, $\mathrm{e}^x y = a\frac{\partial}{\partial x}(\mathrm{e}^x)$, and so by Corollary 2.3 $b_1 = a = a_1$.

DOI: https://dx.doi.org/10.30504/jims.2020.110856

So let $n > 1$ and assume by induction $b_k = a_k$ for all $k < n$. If we pick out the terms of degree n in y from z^m then we obtain a sum of all possible products of the form

$$x \dots x \frac{b_i}{i!} x \dots \frac{b_j}{j!} \dots \frac{b_k}{k!} \dots x, \tag{2.1}$$

with m terms from the set $\{x, b_1, \frac{b_2}{2!}, \frac{b_3}{3!}, \dots\}$ in each product, and with $i + j + \dots + k = n$. Note that the sum of the products of this form involving b_n is $\frac{b_n}{n!}\frac{\partial}{\partial x}(x^m)$. On the other hand, if we pick out the terms of degree n in y from $\mathrm{e}^x\mathrm{e}^y$ then we obtain $\mathrm{e}^x\frac{y^n}{n!}$ and by repeated application of Lemma 2.4 we see that $\mathrm{e}^x\frac{y^n}{n!} = \frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n(\mathrm{e}^x)$. Now

$$\frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n(\mathrm{e}^x) = \sum_{m=0}^{\infty}\frac{1}{m!n!}\left(a\frac{\partial}{\partial x}\right)^n(x^m),$$

and we prove that $b_n = a_n$ by comparing $\frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n(x^m)$ with the sum of the terms

$$x \dots x \frac{b_i}{i!} x \dots \frac{b_j}{j!} \dots \frac{b_k}{k!} \dots x$$

of degree n in y from the expansion of z^m.

It may help to compute $\left(a\frac{\partial}{\partial x}\right)^n(x^m)$ in full detail in the case when $m = n = 3$. Using the fact that $a = a_1$ we have

$$a\frac{\partial}{\partial x}(x^3) = a_1xx + xa_1x + xxa_1,$$

$$\begin{aligned}
&\left(a\frac{\partial}{\partial x}\right)^2(x^3)\\
&= a_2xx + a_1a_1x + a_1xa_1 + a_1a_1x + xa_2x + xa_1a_1 + a_1xa_1 + xa_1a_1 + xxa_2\\
&= a_2\frac{\partial}{\partial x}(x^3) + 2a_1a_1x + 2a_1xa_1 + 2xa_1a_1.
\end{aligned}$$

$$\begin{aligned}
&\left(a\frac{\partial}{\partial x}\right)^3(x^3)\\
&= a_3\frac{\partial}{\partial x}(x^3) + a_2a_1x + a_2xa_1 + a_1a_2x + xa_2a_1 + a_1xa_2 + xa_1a_2\\
&+ 2a_2a_1x + 2a_1a_2x + 2a_1a_1a_1 + 2a_2xa_1 + 2a_1a_1a_1 + 2a_1xa_2\\
&+ 2a_1a_1a_1 + 2xa_2a_1 + 2xa_1a_2\\
&= a_3\frac{\partial}{\partial x}(x^3) + 3a_2a_1x + 3a_2xa_1 + 3a_1a_2x + 3xa_2a_1 + 3a_1xa_2 + 3xa_1a_2\\
&+ 6a_1a_1a_1.
\end{aligned}$$

Now consider $\frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n(x^m)$ for general m and n. We can express this as a sum of terms of the form

$$\frac{1}{n!}x \dots xa_ix \dots a_j \dots a_k \dots x,$$

DOI: https://dx.doi.org/10.30504/jims.2020.110856

where $i + j + \ldots + k = n$. (Compare this expression with (2.1).) We claim that each term $\frac{1}{n!}x \ldots xa_ix \ldots a_j \ldots a_k \ldots x$ appears exactly $\frac{n!}{i!j!\ldots k!}$ times. Accepting this claim for the moment we see that $\frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n (x^m)$ is the sum of all possible distinct products of the form

$$x \ldots x\frac{a_i}{i!}x \ldots \frac{a_j}{j!} \ldots \frac{a_k}{k!} \ldots x, \tag{2.2}$$

with $i+j+\ldots+k = n$. Comparing (2.1) with (2.2), and using our inductive hypothesis that $b_k = a_k$ for $k < n$, we see that the terms in (2.1) which do not involve b_n exactly match the terms from (2.2) which do not involve a_n. The terms from (2.1) which involve b_n sum to $\frac{b_n}{n!}\frac{\partial}{\partial x}(x^m)$, and the terms from (2.2) which involve a_n sum to $\frac{a_n}{n!}\frac{\partial}{\partial x}(x^m)$. So comparing the terms of degree n in y from the expansion of e^z with the expansion of $\frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n (\mathrm{e}^x)$, and using our inductive hypothesis, we see that

$$\frac{b_n}{n!}\frac{\partial}{\partial x}(\mathrm{e}^x) = \frac{a_n}{n!}\frac{\partial}{\partial x}(\mathrm{e}^x),$$

so that Corollary 2.3 implies $b_n = a_n$ as claimed.

It remains to justify our claim that the term

$$\frac{1}{n!}x \ldots xa_ix \ldots a_j \ldots a_k \ldots x$$

appears $\frac{n!}{i!j!\ldots k!}$ times in the expansion of $\frac{1}{n!}\left(a\frac{\partial}{\partial x}\right)^n (x^m)$. Let us return to our expansion of $\left(a\frac{\partial}{\partial x}\right)^3 (x^3)$. On the first application of the operator the product xxx is split into a sum of three products. On the second application, each of these three products is split into a sum of another three products, and so on. After three applications of the operator we have 27 products in the sum. We assign "pedigrees" to each of the products which appear in this process as follows. The first application of the operator maps xxx to $a_1xx + xa_1x + xxa_1$ and we assign pedigrees (1), (2), (3) to the three products in this sum, so that the pedigree of a_1xx is (1) , the pedigree of xa_1x is (2) and the pedigree of xxa_1 is (3). If we apply the operator to a_1xx then we obtain $a_2xx + a_1a_1x + a_1xa_1$ and we assign pedigrees (1,1), (1,2), (1,3) to the three products in this sum in the order in which they appear in the sum. Similarly, if we apply the operator to xa_1x then we obtain $a_1a_1x + xa_2x + xa_1a_1$ and we assign pedigrees (2,1), (2,2), (2,3) to the three products in this sum. If we take xa_2x for example, with pedigree (2,2), and we apply the operator once more then we obtain $a_1a_2x + xa_3x + xa_2a_1$ and we assign pedigrees (2,2,1), (2,2,2), (2,2,3) to the three products in this sum. So there are 27 possible pedigrees (r, s, t) with $1 \le r, s, t \le 3$ for the 27 products in the final sum. A product a_2a_1x, for example, must have pedigree (r, s, t) where two of r, s, t are equal to 1 and one of r, s, t is equal to 2. There are three such pedigrees: (1,1,2), (1,2,1) and (2,1,1) and so the product a_2a_1x occurs three times in the final sum.

Now return to the expansion of $\left(a\frac{\partial}{\partial x}\right)^n (x^m)$. This will be a sum of m^n products with pedigrees $(r_1, r_2, \ldots, r_n)$ with $1 \le r_i \le m$ for $i = 1, 2, \ldots, n$. Consider a particular product $x \ldots xa_ix \ldots a_j \ldots a_k \ldots x$ from this sum with a_i in the r^{th} position in the product, a_j in the s^{th} position, $\ldots$, and with a_k in the t^{th} position. Then the pedigree of the product must be a sequence $(r_1, r_2, \ldots, r_n)$ where i of the integers $r_1, r_2, \ldots, r_n$ are equal to r , j of the integers $r_1, r_2, \ldots, r_n$ are equal to s, $\ldots$, and k of the integers $r_1, r_2, \ldots, r_n$ are equal to t. The total number of pedigrees of this form is $\frac{n!}{i!j!\ldots k!}$, as claimed. □

DOI: https://dx.doi.org/10.30504/jims.2020.110856

3. Some properties of the BCH formula

We want to study the terms of degree at most $2p-2$ which appear in Hausdorff's formula $z = x + a_1 + \frac{a_2}{2!} + \ldots + \frac{a_n}{n!} + \ldots$ in some detail. As an illustration, take $p = 5$ and then consider the expression for $a = a_1$ up to weight 8.

$$a = y - \frac{1}{2}[y,x] + \frac{1}{12}[y,x,x] - \frac{1}{720}[y,x,x,x,x] + \frac{1}{30240}[y,x,x,x,x,x,x] + \ldots.$$

The denominators of the coefficients of the terms with weight less than 5 in the expression for a are coprime to 5. The denominators of the coefficients of the terms with weight between 5 and 8 are divisible by 5, but not by 5^2. The next term in the expression for a has weight 9, and is $-\frac{1}{1209600}[y,x,x,x,x,x,x,x,x]$. The denominator of the coefficient of this term is divisible by 5^2. The same pattern applies to a general prime p. The definition of a in terms of the inverse of the power series $1+\sum_{n=1}^{\infty}\frac{x^n}{(n+1)!}$ makes it clear that the coefficients of the terms of weight at most $p-1$ in the expression for a have denominators which are coprime to p. However the coefficient of the term $[y,\underbrace{x,x,\ldots,x}_{p-1}]$ has denominator which is divisible by p. It seems that the denominators of the coefficients of the terms $[y,\underbrace{x,x,\ldots,x}_{r}]$ with $p-1 < r < 2p-2$ are also divisible by p, though this is not that obvious to me. This need not concern us however — what is obvious is that the denominators of the coefficients of these terms are *not* divisible by p^2. The coefficient of $[y,\underbrace{x,x,\ldots,x}_{2p-2}]$ does have a denominator divisible by p^2. This is because the coefficient of x^{2p-2} in

$$\left(1+\sum_{n=1}^{\infty}\frac{x^n}{(n+1)!}\right)^{-1}$$
$$= 1 - \sum_{n=1}^{\infty}\frac{x^n}{(n+1)!} + \left(\sum_{n=1}^{\infty}\frac{x^n}{(n+1)!}\right)^2 - \left(\sum_{n=1}^{\infty}\frac{x^n}{(n+1)!}\right)^3 + \ldots$$

contains a contribution of $\frac{1}{p!^2}$ from the expansion of $\left(\sum_{n=1}^{\infty}\frac{x^n}{(n+1)!}\right)^2$. To summarize, if we look at the coefficients of the terms of weight up to $p-1$ in the expression for a then their denominators are coprime to p. The coefficient of the term of weight p from the expression for a has denominator which is divisible by p, and the terms of weight p, $p+1$, $\ldots$, $2p-2$ have coefficients with denominators which are *not* divisible by p^2. Note also that the terms with weight at least p all lie in the Lie ideal generated the $(p-1)$-Engel word $[y,\underbrace{x,x,\ldots,x}_{p-1}]$.

Since we are concerned with the power of p dividing the denominators of the coefficients in the formula for z, we introduce the ring R of rational numbers of the form $\frac{m}{n}$ where n is coprime to p. Then A is a Lie algebra over R, and we let L_R be the Lie subalgebra of A over R generated by x and y. We let E_{p-1} be the Lie ideal of L_R generated by elements $[b,\underbrace{c,c,\ldots,c}_{p-1}]$ with $b,c \in L_R$. So we have shown that the terms of weight less than p in the expression for a lie in L_R and that the terms with weight between p and $2p-2$ lie in $\frac{1}{p}E_{p-1}$. Putting this together we see that the homogeneous

DOI: https://dx.doi.org/10.30504/jims.2020.110856

components of a of weight at most $2p-2$ all lie in $L_R + \frac{1}{p}E_{p-1}$. We want to establish this same result for all the homogeneous components of z of weight at most $2p-2$, but before we can do this we need to introduce a little more machinery.

It is well known that the $(p-1)$-Engel identity is equivalent in characteristic p to the multilinear identity $K_p(x_1, x_2, \ldots, x_p) = 0$ where

$$K_p(x_1, x_2, \ldots, x_p) = \sum_{\sigma \in \mathrm{Sym}(p-1)} [x_p, x_{1\sigma}, x_{2\sigma}, \ldots, x_{(p-1)\sigma}].$$

To see that the identity $[y, \underbrace{x, x, \ldots, x}_{p-1}] = 0$ implies $K_p = 0$, set $y = x_p$ and set $x = x_1 + x_2 + \ldots + x_{p-1}$ in $[y, \underbrace{x, x, \ldots, x}_{p-1}]$. If we expand and pick out the terms which are multilinear in $x_1, x_2, \ldots, x_p$ then we obtain K_p. So the $(p-1)$-Engel identity implies $K_p = 0$ (in any characteristic). Conversely, if we set $x_p = y$ and set $x_1 = x_2 = \ldots = x_{p-1} = x$ in the expression for K_p, then we obtain $(p-1)![y, \underbrace{x, x, \ldots, x}_{p-1}]$. So the Lie ideal E_{p-1} of L_R is generated by elements $K_p(b_1, b_2, \ldots, b_p)$ with $b_1, b_2, \ldots, b_p \in L_R$. In fact E_{p-1} is spanned by these elements, since the Jacobi identity implies that

$$[K_p(b_1, b_2, \ldots, b_p), c] = \sum_{i=1}^{p} K_p(b_1, \ldots, [b_i, c], \ldots, b_p).$$

As mentioned above, we want to show that all the homogeneous components of z of weight at most $2p-2$ lie in $L_R + \frac{1}{p}E_{p-1}$. So consider the homogeneous components of $a_2 = a\frac{\partial}{\partial x}(a)$. We write

$$a = b + \frac{1}{p}c + \text{terms of weight at least } 2p-1,$$

where $b \in L_R$ and $c \in E_{p-1}$, so that the homogeneous components of a_2 of weight at most $2p-2$ are identical to the homogeneous components of weight at most $2p-2$ from $d\frac{\partial}{\partial x}(d)$, where $d = b + \frac{1}{p}c$. Now

$$d\frac{\partial}{\partial x}(d) = b\frac{\partial}{\partial x}(b) + \frac{1}{p}c\frac{\partial}{\partial x}(b) + \frac{1}{p}b\frac{\partial}{\partial x}(c) + \frac{1}{p^2}c\frac{\partial}{\partial x}(c).$$

It is clear that $b\frac{\partial}{\partial x}(b) \in L_R$ and that $\frac{1}{p}c\frac{\partial}{\partial x}(b) \in \frac{1}{p}E_{p-1}$. It is also clear that all the homogeneous components of $\frac{1}{p^2}c\frac{\partial}{\partial x}(c)$ have weight at least $2p-1$. So it remains to show that $\frac{1}{p}b\frac{\partial}{\partial x}(c) \in \frac{1}{p}E_{p-1}$. This follows from the fact that c is an R-linear combination of elements of the form $K_p(b_1, b_2, \ldots, b_p)$ with $b_1, b_2, \ldots, b_p \in L_R$ and the fact that

$$b\frac{\partial}{\partial x}(K_p(b_1, b_2, \ldots, b_p)) = \sum_{i=1}^{p} K_p(b_1, \ldots, b\frac{\partial}{\partial x}b_i, \ldots, b_p).$$

So the homogeneous components of a_2 of weight at most $2p-2$ lie in $L_R + \frac{1}{p}E_{p-1}$, as claimed. A similar argument shows that this also applies to a_n for all n. We have proved the following theorem.

Theorem 3.1. *If* $\mathrm{e}^x\mathrm{e}^y = \mathrm{e}^z$ *then the homogeneous components of* z *of weight at most* $2p-2$ *lie in* $L_R + \frac{1}{p}E_{p-1}$.

DOI: https://dx.doi.org/10.30504/jims.2020.110856

In the proofs of Theorem 2.5 and Theorem 3.1 we assumed that x and y were elements from the free generating set for A, but this was mainly to ensure that Hausdorff differentiation was well defined. Since Theorem 3.1 holds true when x and y are free generators of A, it holds true when x and y are arbitrary elements of the power series ring P, provided their constant terms are zero. Note however that "homogeneous component" in Theorem 3.1 refers to homogeneity in terms of x and y. If x and y are not homogeneous as elements of P, then the homogeneous components in Theorem 3.1 will not be homogeneous as elements of P.

We want to extend Theorem 3.1 to arbitrary products of the generators e^{x_1}, e^{x_2}, ... for F. We let L_R^X be the Lie subalgebra of A over R generated by $x_1, x_2, \ldots$, and we let E_{p-1}^X be the $(p-1)$-Engel ideal of L_R^X. As we showed above, E_{p-1}^X is spanned by elements $K_p(b_1, b_2, \ldots, b_p)$ with $b_1, b_2, \ldots, b_p \in L_R^X$.

Theorem 3.2. *Let e^z be an arbitrary element of F. Then the homogeneous components of z of weight at most $2p-2$ lie in $L_R^X + \frac{1}{p}E_{p-1}^X$.*

Proof. The proof is by induction on the length of e^z as a product of the generators of F and their inverses. The result is trivial for products of length 1, and Theorem 3.1 covers products of length 2. Assume by induction that the result is true for products of length at most k, and let e^x, e^y be elements of F which have length at most k as products of the generators of F and their inverses. Let $\mathrm{e}^x\mathrm{e}^y = \mathrm{e}^z$. To complete our proof we need to show that the homogeneous components of z of weight at most $2p-2$ lie in $L_R^X + \frac{1}{p}E_{p-1}^X$.

It makes the exposition easier, and clearer, if we assume that every product in A of length $2p-1$ is trivial. Formally, we replace A by A/I where I is the ideal of A generated by A_{2p-1}. Our inductive hypothesis then implies that $x, y \in L_R^X + \frac{1}{p}E_{p-1}^X$, and we need to show that this implies that $z \in L_R^X + \frac{1}{p}E_{p-1}^X$. We apply Theorem 3.1. Our assumption that products of length $2p-1$ in A are trivial implies that products of length $2p-1$ in x and y are trivial. So Theorem 3.1 implies that $z \in L_R + \frac{1}{p}E_{p-1}$, where L_R is the Lie subalgebra of A over R generated by x and y, and where E_{p-1} is the ideal of L_R spanned over R by elements $K_p(b_1, b_2, \ldots, b_p)$ with $b_1, b_2, \ldots, b_p \in L_R$. It is easy to see that the fact that $x, y \in L_R^X + \frac{1}{p}E_{p-1}^X$ implies that $L_R \leq L_R^X + \frac{1}{p}E_{p-1}^X$, and our proof will be complete if we can show that $E_{p-1} \leq E_{p-1}^X$. So consider a spanning element $K_p(b_1, b_2, \ldots, b_p)$ for E_{p-1} with $b_1, b_2, \ldots, b_p \in L_R$. Write $b_i = c_i + d_i$ for $i = 1, 2, \ldots, p$ with $c_i \in L_R^X$ and $d_i \in \frac{1}{p}E_{p-1}^X$. Then

$$K_p(b_1, \ldots, b_p) = K_p(c_1 + d_1, \ldots, c_p + d_p) = K_p(c_1, \ldots, c_p) \in E_{p-1}^X.$$

□

Corollary 3.3. *If e^z is a product of p^{th} powers in F then the homogeneous components of z of weight at most $2p-2$ lie in $pL_R^X + E_{p-1}^X$.*

Proof. As in the proof of Theorem 3.2 we assume that products of length $2p-1$ in A are trivial. The proof is by induction of the length of e^z as a product of p^{th} powers. The result for products of length 1 follows immediately from Theorem 3.2 since $(\mathrm{e}^z)^p = \mathrm{e}^{pz}$. Assume that the result is true for products of p^{th} powers of length at most k, and let e^x, e^y be products of at most k p^{th} powers. To complete our

DOI: https://dx.doi.org/10.30504/jims.2020.110856

proof we need to show that if $e^x e^y = e^z$ then $z \in pL_R^X + E_{p-1}^X$. Our inductive hypothesis implies that $x, y \in pL_R^X + E_{p-1}^X$, and Theorem 3.1 implies that $z \in L_R + \frac{1}{p}E_{p-1}$, where L_R is the Lie subalgebra of A over R generated by x and y, and where E_{p-1} is the ideal of L_R spanned by elements $K_p(b_1, b_2, \ldots, b_p)$ with $b_1, b_2, \ldots, b_p \in L_R$. The fact that $z \in pL_R^X + E_{p-1}^x$ follows easily from this, using an argument similar to the argument used in the proof of Theorem 3.2. □

4. **Sanov's Theorem**

Theorem 4.1 (Sanov [15]). *Let r be a Lie relator in groups of exponent p, and suppose that r is homogeneous of weight n for some $n \leq 2p-2$. Then the relation $r = 0$ is a consequence of the identical relations $px = 0$ and $[x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$.*

Proof. We write r as an element of the free Lie ring Λ generated by the free generators $x_1, x_2, \ldots$ for A, with Lie product defined by $[x, y] = xy - yx$:

$$r = \sum_{i=1}^{k} n_i [x_{i_1}, x_{i_2}, \ldots, x_{i_n}],$$

with $n_i \in \mathbb{Z}$ for $i = 1, 2, \ldots, k$. Let the free generators of A which appear in this expression for r lie in the set $\{x_1, x_2, \ldots, x_m\}$. We let $B(m, p)$ be the free m generator group of exponent p freely generated by $g_1, g_2, \ldots, g_m$, and we construct the associated Lie ring $L(m, p)$ of $B(m, p)$ as described in Section 1. As we showed in Section 1, $L(m, p)$ is generated by $a_1, a_2, \ldots, a_m$ where $a_i = g_i \gamma_2$ for $i = 1, 2, \ldots, m$. Since r is a Lie relator in groups of exponent p we see that

$$\sum_{i=1}^{k} n_i [a_{i_1}, a_{i_2}, \ldots, a_{i_n}] = 0. \tag{4.1}$$

Now

$$\sum_{i=1}^{k} n_i [a_{i_1}, a_{i_2}, \ldots, a_{i_n}] = \prod_{i=1}^{k} [g_{i_1}, g_{i_2}, \ldots, g_{i_n}]^{n_i} \gamma_{n+1},$$

and so equation 4.1 implies that

$$\prod_{i=1}^{k} [g_{i_1}, g_{i_2}, \ldots, g_{i_n}]^{n_i} \in \gamma_{n+1}.$$

This in turn implies that in the free group F generated by e^{x_1}, e^{x_2}, ...

$$\prod_{i=1}^{k} [e^{x_{i_1}}, e^{x_{i_2}}, \ldots, e^{x_{i_n}}]^{n_i} \in F^p \gamma_{n+1}(F).$$

Now it is well known (and easy to show) that the group commutator

$$[e^x, e^y] = e^{[x,y] + \text{higher terms}},$$

and this implies that if we let

$$e^z = \prod_{i=1}^{k} [e^{x_{i_1}}, e^{x_{i_2}}, \ldots, e^{x_{i_n}}]^{n_i},$$

DOI: https://dx.doi.org/10.30504/jims.2020.110856

then $z = r+$ higher terms. It also implies that if $e^u \in \gamma_{n+1}(F)$, then the leading term of u has weight at least $n+1$. We write

$$\prod_{i=1}^{k} [e^{x_{i_1}}, e^{x_{i_2}}, \ldots, e^{x_{i_n}}]^{n_i} = e^t e^u,$$

where $e^t \in F^p$ and $e^u \in \gamma_{n+1}(F)$. This gives $e^z e^{-u} = e^t$. The remarks above imply that r is the leading term of $e^z e^{-u}$, and Corollary 3.3 implies that $r \in pL_R^X + E_{p-1}^X$.

We want to show that $r \in p\Lambda + E_{p-1}(\Lambda)$, where Λ is the free Lie ring generated inside A by $x_1, x_2, \ldots$ and $E_{p-1}(\Lambda)$ is the ideal of Λ generated by elements $[x, \underbrace{y, y, \ldots, y}_{p-1}]$ with $x, y \in \Lambda$. However the expression for r as an element of $pL_R^X + E_{p-1}^X$ may involve rational coefficients with denominators coprime to p. Nevertheless there will be some integer k coprime to p such that $kr \in p\Lambda + E_{p-1}(\Lambda)$. Since pr, kr are in $p\Lambda + E_{p-1}(\Lambda)$ and p and k are coprime, it follows that $r \in p\Lambda + E_{p-1}(\Lambda)$. □

5. An application of Sanov's Theorem

As we mentioned in Section 1, Sanov's Theorem implies that if we let $L(2,p)$ be the associated Lie ring of the Burnside group $B(2,p)$, then the class $2p-2$ quotient of $L(2,p)$ is isomorphic to the class $2p-2$ quotient of the free 2 generator $(p-1)$-Engel Lie algebra over GF(p). We show that if $p = 5$ this implies that $L(2,p)$ has non-trivial Lie products of multiweight $(5,3)$ in its generators a_1, a_2, and if $p > 5$ then $L(2,p)$ has non-trivial Lie products of multiweight $(p,2)$ in its generators a_1, a_2. (The generators of $L(2,p)$ are defined in Section 1.) This in turn implies that if g_1, g_2 are the free generators of $B(2,p)$ then $B(2,5)$ has non-trivial commutators of multiweight $(5,3)$ in g_1, g_2, and if $p > 5$ then $B(2,p)$ has non-trivial commutators of multiweight $(p,2)$ in g_1, g_2.

To see this we proceed as follows. Let Λ_2 be the free Lie algebra over GF(p) of rank 2, with free generators x_1, x_2. Then Λ_2 is graded by multiweight. If m, n are non-negative integers we let $W_{m,n}$ be the GF(p)-span in Λ_2 of all Lie products with multiweight (m,n) in x_1, x_2. We have

$$\Lambda_2 = \bigoplus_{m+n>0} W_{m,n},$$

and

$$[W_{m,n}, W_{r,s}] \leq W_{m+r,n+s}$$

for all $m, n, r, s \geq 0$. As we have shown, the $(p-1)$-Engel ideal $E_{p-1}(\Lambda_2)$ of Λ_2 is spanned by elements $K_p(b_1, b_2, \ldots, b_p)$ with $b_i \in \Lambda_2$. By linearity we can take the entries b_i in these spanning elements to be basic Lie products in x_1, x_2. (See page 6 of [16] for the definition of basic Lie products. They are defined in analogy with basic commutators in groups.) Since K_p is symmetric in its entries modulo p (we give a proof of this fact at the end of this section), we can assume that $b_1 \leq b_2 \leq \ldots \leq b_p$ in the ordering on basic Lie products. We can also assume that $b_1 \neq b_p$, as if all the entries in K_p are equal then K_p evaluates to zero. These spanning elements are known as Kostrikin elements, and they are multihomogeneous in x_1, x_2. (In other words, each of these spanning elements lies in $W_{m,n}$ for some m, n.) This implies that the free $(p-1)$-Engel Lie algebra $\Lambda_2/E_{p-1}(\Lambda_2)$ is also graded by multiweight.

DOI: https://dx.doi.org/10.30504/jims.2020.110856

It also implies that if we let $U_{m,n}$ be the multihomogeneous component of $\Lambda_2/E_{p-1}(\Lambda_2)$ of multiweight (m,n) then

$$\dim U_{m,n} \geq \dim W_{m,n} - k,$$

where k is the number of Kostrikin elements with multiweight (m,n).

The dimension of $W(m,n)$ is the number of basic Lie products with multiweight (m,n), so it is easy to compute $\dim W_{m,n}$. For example, the basic Lie products of multiweight $(p,2)$ in x and y are

$$[y,\underbrace{x,\ldots,x}_{p},y],\ [y,\underbrace{x,\ldots,x}_{p-1},[y,x]],\ \ldots,\ [y,\underbrace{x,\ldots,x}_{p-r},[y,\underbrace{x,\ldots,x}_{r}]],\ \ldots$$

with $p-r>r$, and so $\dim W_{p,2} = \frac{p+1}{2}$. The Kostrikin elements with multiweight $(p,2)$ are

$$K_p(x,x,\ldots,x,[y,x],[y,x]),\ K_p(x,x,\ldots,x,y,[y,x,x]),\ K_p(x,x,\ldots,x,[y,x,y]),$$

and so $\dim U_{p,2} \geq \frac{p+1}{2} - 3$, which is positive for $p>5$. So Sanov's Theorem implies that if $p>5$ then $L(2,p)$ contains nontrivial Lie products of multiweight $(p,2)$ in a_1,a_2.

When $p=5$, $U_{5,2}=\{0\}$, and we need to look at $U_{5,3}$. The basic Lie products of multiweight $(5,3)$ are

$$[y,x,x,x,x,x,y,y],\ [y,x,x,x,x,y,[y,x]],\ [y,x,x,x,y,[y,x,x]],$$
$$[y,x,x,x,x,[y,x,y]],\ [y,x,x,y,[y,x,x,x]],\ [y,x,x,x,[y,x],[y,x]],$$
$$[y,x,x,[y,x],[y,x,x]],$$

and so $\dim W_{5,3}=7$. The Kostrikin elements of multiweight $(5,3)$ are

$$K_5(x,x,y,y,[y,x,x,x]),\ K_5(x,x,x,y,[y,x,x,y]),$$
$$K_5(x,x,x,x,[y,x,y,y]),\ K_5(x,x,x,[y,x],[y,x,y]),$$
$$K_5(x,x,y,[y,x],[y,x,x]),\ K_5(x,x,[y,x],[y,x],[y,x]),$$

and so $\dim U_{5,3}\geq 1$. (In fact the dimension is 1.) This shows that $L(2,5)$ has non-trivial Lie products of multiweight $(5,3)$ in a_1,a_2.

Finally we give a short proof of our claim above that K_p is symmetric in its entries modulo p.

Lemma 5.1. *If $\sigma\in Sym(p)$ then $K_p(x_1,x_2,\ldots,x_p) = K_p(x_{1\sigma},x_{2\sigma},\ldots,x_{p\sigma})$ modulo p.*

Proof. We let x and y be free generators of the associative algebra A and we let n be a positive integer. An easy induction shows that if we expand the Lie element $[y,\underbrace{x,x,\ldots,x}_{n}]$ as a sum of associative products in A we obtain

$$\sum_{r=0}^{n}(-1)^r\binom{n}{r}x^ryx^{n-r}.$$

Taking $n=p-1$ we obtain

$$[y,\underbrace{x,x,\ldots,x}_{p-1}] = \sum_{r=0}^{p-1}x^ryx^{p-1-r} \text{ modulo } p.$$

DOI: https://dx.doi.org/10.30504/jims.2020.110856

We substitute $x_1 + x_2 + \ldots + x_{p-1}$ for x and substitute x_p for y in this equation, expand, and pick out the terms on both sides which are multilinear in $x_1, x_2, \ldots, x_p$. This gives

$$K_p(x_1, x_2, \ldots, x_p) = \sum_{\sigma \in \mathrm{Sym}(p)} x_{1\sigma} x_{2\sigma} \ldots x_{p\sigma} \text{ modulo } p,$$

which proves the lemma. □

References

[1] S. I. Adjan, The Burnside Problem and Identities in Groups, Ergebnisse der Mathematik und ihrer Grenzgebiete, 95, Springer-Verlag, Berlin, 1979.

[2] S. Bachmuth, Solution to the Burnside Problem, arXiv:0803.1612, 2008.

[3] S. Bachmuth, A straightforward solution to the Burnside Problem, arXiv:1603.08421, 2016.

[4] H. E. Baker, Alternants and continuous groups, *Proc. London Math. Soc.* **3** (1905) 24–47.

[5] E. Hausdorff, Die symbolische exponentialformel in der gruppentheorie, Berichte der Saechsischen Akademie der Wissenschaften, (Math.-Phys. Kl.) Leipsig **58** (1906) 19–48.

[6] G. Havas, G. E. Wall and J. W. Wamsley, The two generator restricted Burnside group of exponent five, *Bull. Austral. Math. Soc.* **10** (1974) 459–470.

[7] G. Havas and M. F. Newman, Applications of computers to questions like those of Burnside, Lecture Notes in Mathematics, 806, Berlin, Springer-Verlag, (1974) 330–332.

[8] G. Higman, Some remarks on varieties of groups, *Quart. J. Math. Oxford Ser. (2)* **10** (1959) 165–178.

[9] N. Jacobson, Lie Algebras, Wiley-Interscience, New York, 1962.

[10] A. I. Kostrikin, On the connection between periodic groups and Lie rings (Russian), *Izv. Akad. Nauk SSSR, Ser. Mat.* **21** (1957) 289–310.

[11] W. Magnus, A connection between the Baker-Hausdorff formula and a problem of Burnside, *Ann. of Math. (2)* **52** (1950) 111–126.

[12] W. Magnus, A. Karrass and D. Solitar Combinatorial group theory, Presentations of groups in terms of generators and relations, Second Revised ed., Dover Publications, Inc., New York, 1976.

[13] E. A. O'Brien and M. Vaughan-Lee, The 2-generator restricted Burnside group of exponent 7, *Internat. J. Algebra Comput.* **12** (2002), no. 4, 575–592.

[14] A. Ju. Olschanskii, Groups of bounded period with subgroups of prime order, *Algebra and Logic* **21** (1982) 369–418.

[15] I. N. Sanov, Establishment of a connection between periodic groups with period a prime number and Lie rings, *Izv. Akad. Nauk SSSR, Ser. Mat.* **16** (1952) 23–58.

[16] M. R. Vaughan-Lee, The Restricted Burnside Problem, second ed., Oxford University Press, 1993.

Michael Vaughan-Lee

Christ Church, University of Oxford, Oxford, OX1 1DP, England.

Email: michael.vaughan-lee@chch.ox.ac.uk

DOI: https://dx.doi.org/10.30504/jims.2020.110856