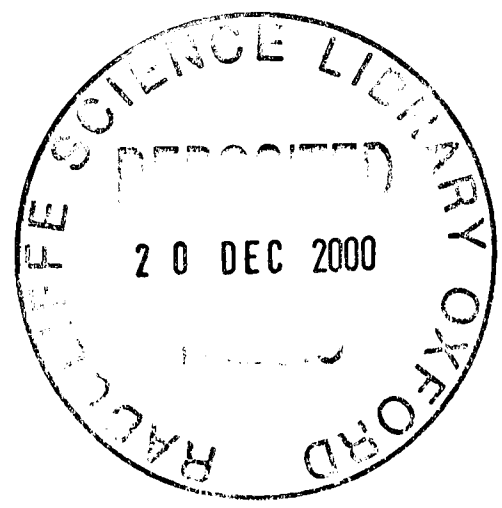
Problems in Lie rings and groups

Daniel Groves
University College
Oxford

Trinity Term, 2000
A thesis submitted in partial fulfilment of the degree of Doctor of Philosophy at the University of Oxford



Abstract

We construct a Lie relator which is not an identical Lie relator in the variety $\mathfrak{B}_4\mathfrak{B}_2$. This is the first known example of a non-identical Lie relator.

Next we consider the existence of torsion in outer commutator groups. Let L be a free Lie ring. Suppose that $1 < i \leq j \leq 2i$ and $i \leq k \leq i+j+1$. We prove that $L/[L^j, L^i, L^k]$ is torsion free. Also, we prove that if $1 < i \leq j \leq 2i$ and $j \leq k \leq l \leq i+j$ then $L/[L^j, L^i, L^k, L^l]$ is torsion free. We then prove that the analogous groups, namely $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ and $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F), \gamma_l(F)]$ (under the same conditions for i, j, k and i, j, k, l respectively), are residually nilpotent and torsion free. We prove the existence of 2-torsion in the Lie rings $L/[L^j, L^i, L^k]$ when $1 \leq k < i, j \leq 5$, and thus show that our methods do not work in these cases.

Finally, we consider the order of finite groups of exponent 8. For $m \geq 2$, we define the function $T(m, n)$ by $T(m, 1) = m$ and $T(m, k+1) = m^{T(m,k)}$. We prove that if G is a finite m-generator group of exponent 8 then $|G| \leq T(m, 7^{471})$, improving upon the best previously known bound of $T(m, 8^{8^8})$.

Acknowledgements

I would like first and foremost to express my deepest gratitude to Professor Michael Vaughan-Lee for his encouragement, guidance, inspiration and patience during the whole process of the construction of this thesis. Without him it would not exist at all. I would also like to acknowledge the proofreading he performed for me, though of course all the mistakes which remain are my own.

I would like to thank the British Council and the Association of Commonwealth Universities, for providing me with funding for my time at Oxford, and University College, for providing me with accommodation and the experience of an Oxford college.

I would like to thank my family, for providing me with support, as always, even from the other side of the world, my fellow algebraists for a very stimulating and relaxed environment to work in, and my friends both new and old for a great amount of fun and often too much to drink.

Contents

Notation

We write $\mathbb{Z}$ for the ring of integers, and $\mathbb{Q}$ for the field of rational numbers. We write $\mathbb{F}_p$ for the field of p elements, where p is a prime number.

F will indicate a free group, freely generated by a set $\{x_i\}_{i\in I}$ of some cardinality (which will generally not be fixed).

L will indicate the free Lie ring of the same rank as F, freely generated by the set $\{z_i\}_{i\in I}$. As far as possible, we will not use any of these symbols for any other objects.

We use the upper-case German font for varieties of groups and to indicate relatively free objects. In particular, $\mathfrak{B}$ will be used for Burnside varieties and free Burnside groups. We will denote the free Burnside group of rank m and exponent n by $\mathfrak{B}(m,n)$ and the variety of all groups of exponent dividing d (the so-called *Burnside variety of exponent d*) by $\mathfrak{B}_d$. We will denote the largest finite m-generator group of exponent n by $\mathfrak{R}(m,n)$.

We save the letter $\mathfrak{L}$ from the upper-case German font to denote Lie rings. Usually, we use $\mathfrak{L}$ to denote the functor taking a group G to its associated Lie ring $\mathfrak{L}(G)$. However, given an N-series $\{H_i\}$ of subgroups of G, we also write $\mathfrak{L}\{H_i\}$ for the Lie ring associated with $\{H_i\}$. Finally, we will denote the Lie ring associated with $\mathfrak{B}(m,n)$ by $\mathfrak{L}(m,n)$ (this is also the associated Lie ring of $\mathfrak{R}(m,n)$).

The symmetric group on n letters is denoted S_n and assumed, unless stated to the contrary, to act upon the set $\{1,\ldots,n\}$ on the right.

We write the action of our homomorphisms on the right.

We use the symbol $\otimes$ for the tensor product.

Finally, we use the symbol ■ to denote the end of a proof.

We hope the reader will be forgiving when we accidentally stray from any of these conventions.

Chapter 1

Introduction

The idea of studying a group by using the structure of fully invariant subgroups goes back at least to Philip Hall [13], who in particular studied the *lower central series*. The commutator structure of a group, when considered from an appropriate point of view, closely resembles the structure of a Lie algebra. This approach was formalised during the next 20 years, with the introduction of the *associated Lie ring* of a group, which is constructed from the lower central series. Particularly instrumental in the early stages was Magnus, who proved that the associated Lie ring of a free group was free (see Theorem 2.6.1). A discussion of the historical background to Lie methods in group theory, and further references, may be found in Chandler and Magnus [4].

Probably the most widespread, and certainly the most important, use of the associated Lie ring has been in the study, and solution, of the *Restricted Burnside Problem*. This originates in questions posed by Burnside [3] in 1902, problems into which an enormous amount of effort has been invested. The solution of the Restricted Burnside Problem for prime exponent, due to Kostrikin [23], and for prime-power exponent, due to Zelmanov [53], [54], both used the associated Lie ring as the major tool.

The Burnside Problem can be stated thus: for what values of m and n are all m-generator groups of exponent n finite? This is a modern restatement of a question asked by Burnside in 1902. The first examples of finitely generated infinite groups of finite exponent were found by Novikov and Adjan [34], [35], [36]. They found such examples of exponent n where n is odd and at least 4381. This result has since been improved and infinite Burnside groups of large even exponent have also been constructed, see [20] and [28], which

means that there are finitely generated infinite groups of any sufficiently large exponent.

The Burnside Problem for exponent 2 is straightforward, and Burnside himself proved that finitely generated groups of exponent 3 are finite. Sanov [39] proved that finitely generated groups of exponent 4 are finite, while M. Hall, Jr. [12] did the same for exponent 6 (see [31] for a simpler proof of this). The Burnside Problem for exponent 5 remains unsolved.

The *Restricted Burnside Problem* concerns itself with finite groups alone. This amounts to asking whether there is a largest finite m-generator group of exponent n, for a given m and n. This was solved in 1959 for prime exponent by Kostrikin [23], and for prime-power exponent by Zelmanov [53], [54]. Given results of Hall and Higman [15], and the classification of finite simple groups (see [6] for a description of the finite simple groups), this completed the positive solution to the Restricted Burnside Problem for arbitrary exponent. In other words, for any positive integers m and n, there is a largest finite group which can be generated by m elements and has exponent n. This result is sometimes known as *Zelmanov's Theorem* and we will occasionally use this nomenclature. Based on Kostrikin's solution to the Restricted Burnside Problem for groups of exponent p, Adjan and Razborov [1] found a primitive recursive bound for the order of finite m generator groups of exponent p in terms of m and p. Later, using the methods of Zelmanov's proof, Vaughan-Lee and Zelmanov [48], [49] obtained an explicit bound for arbitrary exponent.

A third and related question, also asked (in a different form) by Burnside [3], is: what is the order of the largest finite m-generator group of exponent n? We address this question in the case of exponent 8 and arbitrary m in Chapter 5.

This thesis investigates the connection between a group and its associated Lie ring in a number of ways. Primarily, we are concerned with relatively free groups. This thesis tackles three separate problems. The first problem, the existence of non-identical Lie relators, treats the associated Lie ring as an object of study in its own right, whilst the second and third, torsion in outer commutator varieties and the order of finite groups of exponent 8, use the associated Lie ring as a tool for studying groups.

In Chapter 3, we study the notion of *non-identical Lie relators.* This terminology comes from Wall [52], and we largely use the same notation as he does. The question is whether the associated Lie ring of a relatively free group is necessarily relatively free. An equivalent question is that of the

existence of non-identical Lie relators. Surprisingly few associated Lie rings have actually been computed. Magnus proved that the associated Lie ring of a free group is free (see [27] or Subsection 2.6 of this thesis for a proof). The associated Lie rings of several outer commutator groups have been known for a long time, such as those of free nilpotent groups (of a fixed class) and free soluble groups (of a fixed derived length). Newman and Vaughan-Lee [33] calculated the associated Lie rings of the relatively free m-generator Engel-4 groups of exponent 5. The Lie rings of some of the smaller free Burnside groups have also been calculated. Havas, Wall and Wamsley [17] computed the associated Lie ring of $\mathfrak{B}(2,5)$ (the free 2-generator Burnside group of exponent 5). They used this calculation in one of their proofs that the largest finite 2-generator group of exponent 5 has order 5^{34}. It is still an open question whether all 2-generator groups of exponent 5 are finite (or in fact m-generator groups of exponent 5 for any $m \geq 2$).

Havas, Newman and Vaughan-Lee [18] computed the largest 3-generator Lie ring satisfying all the multilinear Lie relators satisfied by $\mathfrak{B}_5$. Vaughan-Lee [45] subsequently proved that this Lie ring is the associated Lie ring of $\mathfrak{B}(3,5)$. We will denote the associated Lie ring of $\mathfrak{B}(m,n)$ by $\mathfrak{L}(m,n)$. $\mathfrak{L}(2,4)$, $\mathfrak{L}(3,4)$ and $\mathfrak{L}(4,4)$ are used in [46], Chapter 6. In [37] it is proved that $\mathfrak{L}(2,4)$ is relatively free, while the same is proved in [7] for $\mathfrak{L}(3,4)$. Almost certainly $\mathfrak{L}(2,7)$ has recently been computed. See Section 3.1 for a very brief discussion of the implications of this calculation.

We also calculate the Lie rings of a number of outer commutator groups (which are relatively free) in Chapter 4, using a result of Kuz'min and Shapiro [26]. We mention a number of previous results about torsion-free outer commutator groups. Any example of an outer commutator group which is torsion-free (such as free polynilpotent groups) immediately gives an example of a relatively free group with relatively free associated Lie ring. However, there seems no reason why the existence of torsion might indicate that an associated Lie ring is not relatively free, so the existence of torsion tells us little about this particular problem.

All of the above groups are relatively free, and all of their associated Lie rings, except possibly $\mathfrak{L}(4,4)$ and $\mathfrak{L}(2,7)$, are relatively free. This naturally raises the question of whether the associated Lie ring of a relatively free group is necessarily relatively free (see, for example, [51], [52] and [46]). We provide a negative answer to this question by proving that the variety $\mathfrak{B}_4\mathfrak{B}_2$ has non-identical Lie relators. The substance of Chapter 3 has appeared in [8].

In Chapter 4 we use the associated Lie ring as a tool for investigating groups, rather than as an object of study in its own right. So-called *outer commutator-subgroup functions* were first defined by Philip Hall [14]. These will be defined later. A group which is the quotient of a free group F by $\psi(F)$ for some outer commutator-subgroup function ψ will be called an *outer commutator group*. There is still much that is not known about these groups. We concentrate on whether they are torsion-free and/or residually nilpotent.

Gruenberg [9] proved that the free polynilpotent groups are residually nilpotent, while Smel'kin [41] proved that the lower central factors of free polynilpotent groups are torsion-free. In the terminology of Gorchakov [5], a group satisfying these two properties is called a *Mal'cev group*.

The first example of torsion in an outer commutator group was discovered by Gupta [10], who found torsion in the free centre-by-metabelian groups, for rank at least 4. Gupta also proved that the free centre-by-metabelian groups are residually nilpotent. Since this result, a number of instances of torsion in central extensions have been found, and the torsion subgroups classified (see [42] for details and further references).

There are a number of results about groups of the form $F/[\gamma_j(F), \gamma_i(F)]$ for free groups F. The case $i = j$ is covered by the results of Gruenberg and Smel'kin mentioned above. Kikodze [22] proved that $F/[\gamma_{i+1}(F), \gamma_i(F)]$ is torsion-free. Gorchakov [5] extended this to $i \leq j \leq 2i + 1$ and Ushakov [44] to $i \leq j \leq 4i + 1$, both proving that their groups were residually nilpotent. Kuz'min [25] proved that $F/[\gamma_j(F), \gamma_i(F)]$ is residually nilpotent for all i and j. Finally, Tasić and Vaughan-Lee [43] proved that $F/[\gamma_j(F), \gamma_i(F)]$ is torsion-free for all i and j.

Here we tackle the next obvious cases, namely $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ and $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F), \gamma_l(F)]$. We cannot hope to find that these are residually nilpotent and torsion-free for all i, j, k, because of Gupta's result. However, we are able to obtain a significant number of results nonetheless. We prove results about the analogous relatively free Lie rings and then use the associated Lie ring to deduce results about the groups. We also find a number of occurrences of torsion in outer commutator Lie rings, but our methods do not allow us to translate this into torsion in groups.

In Chapter 5 we consider the order of finite groups of exponent 8. The aim is to improve the known bounds on the order of finite m-generator groups of exponent 8. We follow the overall method of Zelmanov's solution of the Restricted Burnside Problem, [53], [54], but the specific method and some of our results come from [46], [47] and [48].

Suppose that $m \geq 2$. We define $T(m,n)$ recursively by $T(m,1) = m$ and $T(m,n+1) = m^{T(m,n)}$. The main result in Chapter 5 is that a finite m-generator group of exponent 8 has order at most $T(m,7^{471})$. The general result for exponent n and m generators from [48] obtains a bound of $T(m,n^{n^n})$, so our result is a substantial improvement. However, the numbers involved are still unimaginably large, and very unlikely to be realistic bounds, so this result is somewhat disappointing. If more powerful computers were available, then it might be possible to reduce the above bound to approximately $T(m,7^{240})$. This depends on a calculation which I have been unable to complete due to the amount of computer memory required for it.

Despite this disappointment, in the process of proving our bound, we make a number of improvements which may be of interest in their own right. Particularly, we have improved upon the known bounds on the nilpotency class of m-generator Engel-7 sandwich algebras over fields of characteristic 2. Our bound is $m^{m^{2m}}$, which improves upon the bound of $T(m,20)$, found in [48].

Chapter 2

Preliminary results

In this chapter we provide a selection of the preliminary results needed throughout this thesis. None of the results in this chapter are original, and most of our definitions and proofs are taken from elsewhere.

This chapter is intended as a reference, rather than to be read in a linear fashion.

2.1 Lie algebras

Definition 2.1.1 *Let k be a commutative ring. A* Lie algebra *J over k is a k-module on which there is a product $(a, b) \mapsto [a, b]$ satisfying the following conditions:*

$$\begin{aligned} [\lambda x + \mu y, z] &= \lambda[x, z] + \mu[y, z] \quad \forall \lambda, \mu \in k,\ \forall x, y, z \in J \\ [x, \lambda y + \mu z] &= \lambda[x, y] + \mu[x, z] \quad \forall \lambda, \mu \in k,\ \forall x, y, z \in J \\ [x, x] &= 0 \quad \forall x \in J \\ [[x, y], z] + [[y, z], x] + [[z, x], y] &= 0 \quad \forall x, y, z \in J \end{aligned} \tag{2.1}$$

A Lie algebra over $\mathbb{Z}$ is called a Lie ring.

The identity (2.1) is known as the *Jacobi Identity*.

Our definition of Lie algebras differs from the usual one in that we do not assume the ground ring is a field. We do this only so that we can define Lie rings at the same time as Lie algebras. In the sequel, our ground ring will be either $\mathbb{Z}$ and we will use the terminology *Lie ring*, or a field, where we will say *Lie algebra*. Thus our subsequent use of the terms will be standard.

The Lie product is not associative, so we adopt a convention for repeated Lie products. We use the *left-normed* convention.

Definition 2.1.2 *Let J be a Lie algebra, and $y_1, \ldots, y_n \in J$, where $n > 2$. We define inductively*

$$[y_1, y_2, \ldots, y_n] = [[y_1, y_2, \ldots, y_{n-1}], y_n].$$

A product of the form $[y_1, \ldots, y_n]$ is called *left-normed.*

Suppose that J is a Lie algebra and that $y, x_1, x_2, \ldots \in J$. Let $S = \{i, j, \ldots, r\}$ be a non-empty set of positive integers, with $i < j < \cdots < r$. Then we define

$$[y, x_S] = [y, x_i, x_j, \ldots, x_r].$$

The most common examples of Lie algebras come from associative algebras. If A is an associative algebra, then A can be turned into a Lie algebra by defining $[x, y] = xy - yx$. It is straightforward to check that this product satisfies the axioms for a Lie product. We denote this Lie algebra by A_L.

We also define *restricted Lie algebras*, which we will need in Chapter 5. First we define the elements $s_i(a, b)$, which are necessary for the definition of restricted Lie algebras.

Definition 2.1.3 *Let L be a free Lie algebra, freely generated by a and b. Let λ be a central algebraically independent indeterminate. Because the product in a Lie algebra is multilinear, the expression $a\Big(\mathrm{ad}(\lambda a+b)\Big)^{p-1}$ may be written as a polynomial in λ. We define the elements $s_i(a, b)$ by*

$$a\Big(\mathrm{ad}(\lambda a + b)\Big)^{p-1} = \sum_{i=1}^{p-1} i s_i(a, b)\lambda^{i-1}.$$

We can then define the elements $s_i(a, b)$ for any a and b in any Lie algebra by substitution.

This definition appears arbitrary at first, but will seem less arbitrary when we have defined restricted Lie algebras and noted that the above elements are exactly what we need when we turn an associative algebra into a restricted Lie algebra.

Definition 2.1.4 *Let k be a field of characteristic p, for p a prime. A* restricted Lie algebra of characteristic p, *L, is a Lie algebra over k which has an operation, $x \mapsto x^{[p]}$, satisfying the following conditions:*

$$(x+y)^{[p]} = x^{[p]} + y^{[p]} + \sum_{i=1}^{p-1} s_i(x,y) \quad \forall x, y \in L \tag{2.2}$$

$$(\alpha x)^{[p]} = \alpha^p x^{[p]} \quad \forall \alpha \in k \text{ and } x \in L \tag{2.3}$$

$$[a, b^{[p]}] = a(\mathrm{ad}b)^p \tag{2.4}$$

A *restricted Lie ideal* of a restricted Lie algebra will be a Lie ideal of the restricted Lie algebra which is also closed under the p^{th} power operation. The operation $x \mapsto x^{[p]}$ is intended to mimic the p^{th} powering in an associative algebra.

We will use restricted Lie algebras of characteristic 2 in Chapter 5. In the case of characteristic 2, the identity 2.1.4.2.2 reduces to

$$(a+b)^{[2]} = a^{[2]} + b^{[2]} + [a, b].$$

2.2 Varieties and (relatively) free things

The notion of a free group is absolutely fundamental to group theory, and we will not repeat a definition here. Almost any text on group theory will define a free group, and we mention just two; in [29] the free group is defined from a combinatorial point of view, as a group presented with no relations, while in [30] the emphasis is on the universal properties of a free group. However, there is not really a large difference between these two approaches.

A free Lie ring (or algebra) is an analogous object, and equally well-known. For a definition we point the reader to [2].

The real purpose of this first part of this section is to fix notation for the remainder of the thesis.

We will let F be a free group and L the free Lie ring of the same rank as F. We fix notation for the generators of F and L and will not use this for anything else. F will be freely generated by the set $\{x_i\}_{i \in I}$ while L will be freely generated by $\{z_i\}_{i \in I}$.

In Chapter 3 we will assume that F and L are of countably infinite rank, and we will take I to be the natural numbers in this case. However, in Chapter 4 we will generally have no restrictions on the rank.

We now give a brief list of definitions and state (without proof) some theorems about the very basic notions of varieties of groups. All of these notions, and much more, can be found in [30].

Suppose now that F is the free group of countably infinite rank, freely generated by $\{x_1, x_2, \ldots\}$. Then we refer to the elements of F as *words*. If G is a group, and $\alpha : F \to G$ is a homomorphism, then the image of a word w under α is called a *value* of w in G. The word w is a *law* (or *identical relation*) in the group G if the only value of w in G is 1, meaning that $w\alpha = 1$ for all homomorphisms $\alpha : F \to G$.

Definition 2.2.1 *Let* $\mathfrak{w}$ *be a subset of* F. *Then the* verbal subgroup $\mathfrak{w}(G)$ *of a group* G *corresponding to* $\mathfrak{w}$ *is the subgroup generated by all values in* G *of the words in* $\mathfrak{w}$.

We have the following results from [30, p.5].

Theorem 2.2.2

1. *All the words* $w \in \mathfrak{w}$ *are laws in* G *if and only if* $\mathfrak{w}(G) = \{1\}$.
2. *Every verbal subgroup of a group is fully invariant in that group (ie is invariant under all endomorphisms).*

Definition 2.2.3 *A group is called relatively free if it possesses a generating set such that every mapping of this generating set into the group can be extended to an endomorphism.*

Theorem 2.2.4 *[30, p.9] The group* G *is relatively free if and only if it has one of the following properties:*

1. G *possesses a set of generators such that every relator of these generators is a law in* G.
2. G *has a representation* $G \cong F/\mathfrak{w}(F)$ *as the factor group of a free group by a verbal subgroup of* F.

3. *G has a representation $G \cong F/R$ such that every endomorphism of F induces an endomorphism of G.*

For our purposes 2.2.4.2 will be the most useful characterisation of a relatively free group.

Definition 2.2.5 *A* variety *of groups is the class of all groups satisfying each one of a given set of laws.*

The variety defined by the set $\mathfrak{w}$ will be denoted by $\mathfrak{W}$.

Theorem 2.2.6 *[30, p.12]*

1. *A group G belongs to the variety $\mathfrak{W}$ if and only if $\mathfrak{w}(G) = \{1\}$.*

2. *A group G belongs to $\mathfrak{W}$ if and only if it is (isomorphic to) a factor group of $F/\mathfrak{w}(F)$ for a free group F of suitable rank.*

Definition 2.2.7 *If $\mathfrak{U}$ and $\mathfrak{V}$ are varieties of groups, then the product variety $\mathfrak{U}\mathfrak{V}$ is the variety of all groups which are extensions of a group in $\mathfrak{U}$ by a group in $\mathfrak{V}$.*

In other words, $G \in \mathfrak{U}\mathfrak{V}$ if and only if there is a normal subgroup $N \trianglelefteq G$ such that $N \in \mathfrak{U}$ and $G/N \in \mathfrak{V}$.

All of these notions can be defined also for Lie rings, and this is done in precisely analogous ways. We will not do this here, but direct the interested reader to [2].

2.3 Gradings and homogeneous elements

We repeat the definition of a graded algebra from [2], p.34.

Definition 2.3.1 *Let k be a commutative ring with unity and let R be a k-algebra. Also let Ξ be a commutative semigroup (which we write additively). A collection of k-submodules $\{R_\xi\}_{\xi\in\Xi}$ is called a Ξ-grading of R if*

1. $R = \bigoplus_{\xi \in \Xi} R_\xi$,

2. $R_\xi R_\eta \subseteq R_{\xi+\eta}$

Generally we will use a *grading by weight*, by which we mean that Ξ is the set of natural numbers under addition and that R_1 generates R. However, we will also use *multihomogeneous gradings*, which will be gradings by n-tuples of non-negative integers (with n varying), where addition is defined pointwise with the obvious embedding of $\mathbb{Z}^n$ in $\mathbb{Z}^m$ for $n < m$ via the first n components. Specifically, we will work with a free Lie ring L and fix a set of generators $\{z_1, z_2, \ldots\}$. We then define the multihomogeneous component of multiweight $(j_1, j_2, \ldots, j_k)$ to be the subspace spanned by all products which have j_i occurrences of z_i for $1 \leq i \leq k$ and no occurrences of z_i for $i > k$. We will see this grading later.

Definition 2.3.2 *A nonzero $r \in R_\xi$ is called* homogeneous of degree ξ. *When we do not need to specify an R_ξ, we merely say that r is* homogeneous, *and when we are using the multihomogeneous grading, we will say that r is* multihomogeneous.

The submodule R_ξ is called a homogeneous component of R. *Similarly, we define* multihomogeneous components.

2.4 Series of subgroups and ideals

Definition 2.4.1 *Let G be a group, and $x, y \in G$. Then the* commutator *of x and y, denoted $[x, y]$, is the element $x^{-1}y^{-1}xy$.*

If $H, K \leq G$, the commutator subgroup *of H and K, denoted $[H, K]$, is the subgroup of G generated by all elements $[h, k]$, where $h \in H$ and $k \in K$.*

Definition 2.4.2 *Let G be a group, and $x, y \in G$. Then the* conjugate of x by y, *denoted x^y, is the element $y^{-1}xy$.*

In a precisely analogous way as for Lie algebras, we adopt the *left-normed convention* for repeated commutators.

Definition 2.4.3 *Let G be a group. We define the* lower central series *of G, $\{\gamma_n(G)\}_{n \geq 1}$, inductively as follows.*

$\gamma_1(G) = G$ and,

if $\gamma_i(G)$ is already defined, $\gamma_{i+1}(G) = [\gamma_i(G), G]$.

Definition 2.4.4 *Let G be a group. We define the* derived series *of G, $\{G^{(n)}\}_{n\geq 0}$, inductively as follows.*
$G^{(0)} = G$ *and,*
if $G^{(i)}$ is already defined, $G^{(i+1)} = [G^{(i)}, G^{(i)}]$.

We also define here the notion of an N-series.
An *N-series* is a series of subgroups of G, $\{H_i\}_{i=1}^{\infty}$, such that

1. $G = H_1$,

2. $[H_i, H_j] \leq H_{i+j}$ for all i and j.

It is easy to see that 2 implies that $H_i \trianglelefteq G$ for all i. We note that for any N-series, $\{H_i\}_{i=1}^{\infty}$, $\gamma_i(G) \leq H_i$ for all i.

One N-series, which we use in Chapter 5, is the *Zassenhaus filtration*, $\{G_i\}_{i\geq 1}$. It is defined as follows. Fix a prime p. Let G_i be the subgroup of G generated by all powers of commutators

$$[g_1, \ldots, g_k]^{p^j},$$

where $g_1, \ldots, g_k \in G$ and $kp^j \geq i$. It can easily be proved that $\{G_i\}$ is an N-series (see [27]).

We also define the lower central and derived series for Lie algebras.

Definition 2.4.5 *Suppose that J is a Lie algebra. Then the* lower central series, *$\{J^i\}_{i\geq 1}$ is defined by $J^1 = J$ and $J^{i+1} = [J^i, J]$. The* derived series, *$\{J^{(i)}\}_{i\geq 0}$ of J is defined by $J^{(0)} = J$ and $J^{(i+1)} = [J^{(i)}, J^{(i)}]$.*

2.5 The associated Lie ring

There are many standard results about commutators and the lower central series. We state, without proof, only a few which motivate the definition of the associated Lie ring.

Theorem 2.5.1 *Let G be a group, and $\{\gamma_i(G)\}_{n\geq 1}$ its lower central series. Let $x, y, z \in G$. Then*

1. $[x, yz] = [x, z][x, y]^z$

2. $[xy, z] = [x, z]^y[y, z]$

3. $[x, y^{-1}, z]^y [y, z^{-1}, x]^z [z, x^{-1}, y]^x = 1$

4. $\gamma_i(G)$ *is a fully invariant subgroup of* G *for all* i.

5. $[\gamma_i(G), \gamma_j(G)] \leq \gamma_{i+j}(G)$, *for all* i *and* j.

■

All of these are proved in [11, Chapter 10] (or almost any other standard group theory text).

Using the above theorem, and other straightforward results, we construct the *associated Lie ring* of G, denoted $\mathfrak{L}(G)$. The *lower central factors*, $\gamma_i(G)/\gamma_{i+1}(G)$ $i \geq 1$, are abelian groups, and so can be naturally considered as $\mathbb{Z}$-modules. We use their direct sum to form the associated Lie ring.

Definition 2.5.2 *Let* $\mathfrak{L}(G) = \bigoplus_{i \geq 1} \gamma_i(G)/\gamma_{i+1}(G)$. *We define a* bracket product *on* $\mathfrak{L}(G)$ *in the following way. Suppose that* $x = g\gamma_{i+1}(G)$ *and* $y = h\gamma_{j+1}(G)$ *are homogeneous elements of* $\mathfrak{L}(G)$. *Then we define*

$$[x, y] = [g, h]\gamma_{i+j+1}(G).$$

We extend this product to the whole of $\mathfrak{L}(G)$ *linearly.*

To prove that this definition turns $\mathfrak{L}(G)$ into a Lie ring is elementary. The identity 2.5.1.5 ensures that the Lie product is well-defined. The identities 2.5.1.1 and 2.5.1.2 correspond to the linearity of the Lie product, while 2.5.1.3 corresponds to the Jacobi identity (see [11, Chapter 18]).

The subspaces $\mathfrak{L}_i := \gamma_i(G)/\gamma_{i+1}(G)$ form a natural grading for $\mathfrak{L}(G)$.

In fact, given any N-series $\{H_i\}$, we can define a Lie ring associated with $\{H_i\}$ in an analogous way to the definition of the associated Lie ring, using the factors H_i/H_{i+1}. We denote this Lie ring by $\mathfrak{L}\{H_i\}$.

In Chapters 3 and 4, we use the standard associated Lie ring, as defined in Definition 2.5.2 from the lower central series. Whenever we use the term *associated Lie ring*, we mean the Lie ring associated with the lower central series. We will explicitly specify when we are using another series.

2.6 Magnus' theorem

The first question we investigate in this thesis is whether the associated Lie ring of a relatively free group is necessarily relatively free. The first obvious case to investigate is that of an absolutely free group. We prove the following theorem, due to Magnus.

Theorem 2.6.1 (Magnus) *The associated Lie ring of an absolutely free group is absolutely free. Moreover, if F is such a free group, $\mathfrak{L}(F)$ is freely generated by the images of the free generators of F in $F/\gamma_2(F)$.*

We prove Theorem 2.6.1 only for the free group of countably infinite rank. However, the proof is essentially the same for any rank. Our proof closely follows that of Lazard [27]. First we need the notion of the *universal enveloping algebra* of a Lie algebra.

Definition 2.6.2 *Let J be a Lie algebra. Then the pair $\{U, i\}$, where U is an associative algebra and $i : J \to U_L$, is a* universal enveloping algebra *for J if the following universal property holds: If B is any associative algebra, and φ is any Lie homomorphism from J to B_L, then there exists a unique homomorphism $\varphi' : U \to B$ such that $i\varphi' = \varphi$.*

For a thorough exposition of the theory of universal enveloping algebras, see [21, Chapter 5]. Jacobson proves the existence of universal enveloping algebras for all Lie algebras, something which we will assume. Other than this, we need little more than the definition of a universal enveloping algebra. However, we state without proof the following theorem from [21, p.168].

Theorem 2.6.3 *The universal enveloping algebra of the free Lie ring of countably infinite rank, L, is the free $\mathbb{Z}$-algebra of countably infinite rank, freely generated by the images under i of the free generators for L.*

We consider any Lie algebra, J, to be contained in its universal enveloping algebra, since the map $i : J \to U_L$ is a monomorphism.

Thus, if U is the universal enveloping algebra of L, we consider U to be the free $\mathbb{Z}$-algebra with free generating set $\{z_1, z_2, \ldots\}$.

Let A be the free associative $\mathbb{Z}$-algebra, freely generated by $\{y_1, y_2, \ldots\}$. Define a metric on A in the following way. If $x = u + v$ where $u \neq 0$ is a sum of monomials of degree n and v is a sum of monomials of degree strictly

greater than n, then $|x| = 2^{-n}$. Define $|0| = 0$. It is straightforward to check that this defines a norm on A. Let $\tilde{A}$ be the completion of A with respect to the corresponding metric. We can identify $\tilde{A}$ with the ring of power series in $\{y_1, y_2, \ldots\}$ with coefficients from $\mathbb{Z}$ (see [27]).

If $x \in \tilde{A}$, $x \neq 0$ and $|x| < 1$, then $1 + x$ has a multiplicative inverse in $\tilde{A}$, namely $1 - x + x^2 - \cdots$. Define T to be the multiplicative group generated in $\tilde{A}$ by the elements $\{1 + y_1, 1 + y_2, \ldots\}$.

Theorem 2.6.4 *$F \cong T$ under the homomorphism which extends $\varphi : x_i \mapsto 1 + y_i$.*

Proof: It suffices to show that T is absolutely free of countably infinite rank, with free generating set $\{1+y_1, 1+y_2, \ldots\}$. From the comments above, T is a group, so we need only show that no non-trivial relations hold in T. We follow the proof of [29, pp.310–311]. Let

$$W = (1 + y_{i_1})^{\epsilon_1} \ldots (1 + y_{i_k})^{\epsilon_k}$$

be a non-trivial freely reduced word in the generators of T. That is to say, $k \geq 1$, the ϵ_i are non-zero integers and $i_j \neq i_{j+1}$ for $1 \leq j \leq k - 1$. It is straightforward to show by induction on n that, for any x with zero constant term,

$$(1 + x)^n = 1 + nx + x^2 f_n(x),$$

where $f_n(x)$ is a power series in x, for all integers n. Then

$$W = (1 + \epsilon_1 y_{i_1} + y_{i_1}^2 f_{\epsilon_1}(y_{i_1})) \cdots (1 + \epsilon_k y_{i_k} + y_{i_k}^2 f_{\epsilon_k}(y_{i_k})).$$

Upon expanding W, we see that the coefficient of the monomial $y_{i_1} \ldots y_{i_k}$ is $\epsilon_1 \ldots \epsilon_k$. Since none of the ϵ_j are zero, this part of W is not zero. The identity element of T is 1, so we know that $W \neq 1$. This completes the proof of Theorem 2.6.4. ■

Proof of Theorem 2.6.1: We consider an embedding of $\mathfrak{L}(F)$ in the ring $\tilde{A}$. We abuse notation by assuming that F is equal to T, which we can do by Theorem 2.6.4.

Let J be the Lie algebra generated by $\{y_1, y_2, \ldots\}$ in $\tilde{A}$. We prove that J is free. Let U be the universal enveloping algebra of L. We know by Theorem 2.6.3 that $U \cong A$, which we consider to be contained in $\tilde{A}$. Suppose that this isomorphism is ψ, and that $\psi : z_i \mapsto y_i$ for all i. Then it is clear that the restriction of ψ to L exhibits a homomorphism from L onto J, which

must be an isomorphism because ψ is an algebra monomorphism. Thus, J is absolutely free, freely generated by $\{y_1, y_2, \ldots\}$. We form a series $\{H_i\}$ of subgroups of T. The subgroup H_i is defined to consist of all elements of T whose component of degree k is zero for $1 < k < i$. It is easy to check that $\{H_i\}$ is an N-series.

By the comment in Section 2.4 (after the definition of N-series), $\gamma_i(T) \leq H_i$ for all i. This embedding of $\gamma_i(T)$ into H_i induces an embedding of $\mathfrak{L}(T)$ into $\mathfrak{L}\{H_i\}$. Let M be the image of $\mathfrak{L}(T)$ under this embedding.

Using the definition of the lower central series of a group, it is easy to see that, for any group G, $\mathfrak{L}(G)$ is generated by its elements of weight 1. Thus, it is not difficult to see that $\mathfrak{L}(T)$ is generated by the set

$$\{(1 + y_1)\gamma_2(T), (1 + y_2)\gamma_2(T), \ldots\}.$$

Therefore, M is generated by $\{(1 + y_1)H_2, (1 + y_2)H_2, \ldots\}$.

We can map M into A by the mapping $(1 + y_i)H_2 \mapsto y_i$. It is a matter of routine to check that this map may be extended to a Lie homomorphism, $\varphi : M \to A_L$, and also to see that $M\varphi = J$. Thus, φ is a monomorphism. All of this is fully proved in [27].

Hence M is free, freely generated by the set $\{(1 + y_1)H_2, (1 + y_2)H_2, \ldots\}$, which gives the result of the theorem. It turns out that $\gamma_i(T) = H_i$ for all i, but we do not need this result here (see [27]). ■

2.7 Basic commutators

Basic commutators are fundamental to most of what we do in Chapter 4, and also to much of what we do in Chapter 5. They provide us with both a basis for a free Lie ring and a basis for a free group, modulo any term of the lower central series.

Let Γ be the free groupoid (see [2, Section 2.2.1]), freely generated by the set $\{a_i\}_{i \in I}$, where I is some well-ordered set. We define *basic commutators in* Γ in the following way.

Definition 2.7.1 *The basic commutators of weight* 1 *are the* a_i *for all* $i \in I$, *ordered by the ordering of* I. *If we have already defined and given an order to the basic commutators of weight less than* k, *then the basic commutators of weight* k *are products of the form* uv, *where* (i) u *is of weight* l, v *is of weight* m *and* $l + m = k$, (ii) $u > v$, *and* (iii) *if* $u = u_1u_2$ *then* $v \geq u_2$.

The basic commutators of weight k are well-ordered arbitrarily but the well-ordering is defined so that it respects weight (so if u has smaller weight than v then $u < v$).

If x is a basic commutator on Γ of weight n, then we write $\mathrm{wt}(x) = n$. We will assume similar notation for basic commutators in groups and Lie rings.

Given this definition of basic commutators, we define basic commutators for groups and for Lie rings.

Definition 2.7.2 *Let G be a group, generated by $\{g_i\}_{i\in I}$. Then we define a groupoid homomorphism $\theta : \Gamma \to G$ by $\theta : a_i \mapsto g_i$ for all $i \in I$ and by $(uv)\theta = [u\theta, v\theta]$. This uniquely defines θ. Elements of G which are images under θ of basic commutators in Γ are called* basic commutators in G, relative to the generating set $\{g_i\}_{i\in I}$.

If $G = F$, a free group, and $\{g_i\}_{i\in I}$ is a free generating set for F then this definition coincides with that of Marshall Hall ([11, Chapter 11]). The following theorem may be found in [11, Chapter 11],

Theorem 2.7.3 *[11, p.175] If F is a free group then, for any $n \geq 1$, the basic commutators of weight n form a basis modulo $\gamma_{n+1}(F)$ for the free abelian group $\gamma_n(F)/\gamma_{n+1}(F)$.*

Marshall Hall, Jr. [11] also defines and describes the *commutator collection process*, originally due to Philip Hall [13] (see Section 2.8). This is a means of uniquely expressing any word as an ordered product of basic commutators modulo $\gamma_n(F)$, for any n. The two-sided collection process described in Section 4.3 is a variation of Hall's technique, which allows us to collect words modulo $F^{(n)}$, for any n.

Definition 2.7.4 *Let K be a Lie ring, generated by $\{k_i\}_{i\in I}$. We define $\epsilon : \Gamma \to K$ by $\epsilon : a_i \mapsto k_i$ for all $i \in I$ and by $(uv)\epsilon = [u\epsilon, v\epsilon]$. Elements of K which are images under θ of basic commutators in Γ are called* basic commutators in K.

We now consider the case where K is a free Lie ring.

Each element of Γ has a natural weight. We denote the subset of Γ of all elements of weight n by Γ_n. If L is a free Lie ring and L_n is the subspace of L of homogeneous elements of weight n, then L_n is spanned by $\epsilon(\Gamma_n)$. In fact, if R_n is the set of all basic commutators of weight n in Γ then $R_n \subseteq \Gamma_n$ and $\epsilon(R_n)$ is a basis for L_n. By virtue of Magnus' isomorphism above (Theorem 2.6.1), this is just a restatement of Theorem 2.7.3.

2.8 The Hall collection process

We include this section in order to give a flavour of how the Hall collection process works, in order to make Section 4.3 clearer.

The collection process is a method of taking an arbitrary word in a free group and finding its representation in the form given in Theorem 2.8.2 below. The idea is to find the least element in the word (in terms of the ordering of basic commutators) and move it to the front of the word using the identity $uv = vu[u, v]$. This process is described in more detail in [11, Chapter 11]. In particular, it is explained why only basic commutators arise in the Hall collection process. The same reasoning holds later in the case of two-sided collection to show that only basic commutators arise.

Definition 2.8.1 *A product of basic commutators $c_1 \cdots c_k$ is said to be in* collected form *if $c_1 \leq \cdots \leq c_k$.*

We state, without proof, the following theorem from [11, p.175].

Theorem 2.8.2 *Let F be a free group of finite rank and n a positive integer. Then an arbitrary word $w \in F$ can be uniquely expressed in the form*

$$w = c_1^{e_1} c_2^{e_2} \cdots c_k^{e_k} \text{ modulo } \gamma_n(F),$$

where $c_1, c_2, \ldots, c_k$ are basic commutators of weight less than n, $c_1 < c_2 < \cdots < c_k$ and $e_1, \ldots, e_k$ are nonzero integers.

We start with an example of the Hall collection process. Suppose that our basic commutators of weight 1 are x and y, where $x < y$, and suppose that we wish to collect the word yx^2. Then we use the following identity $uv = vu[u, v]$ and make the following calculation.

$$\begin{aligned} yx^2 &= xy[y, x]x \\ &= xyx[y, x][y, x, x] \\ &= x^2y[y, x]^2[y, x, x], \end{aligned}$$

which is in collected form.

It is relatively easy to perform the Hall collection process on positive words (ie words where all exponents are positive), though in general the collection process will not terminate even for positive words. We briefly

describe the process for general words. The identities we use are $vu^{-1} = u^{-1}v[v,u^{-1}]$ and $[v,u^{-1}] = [v,u,u^{-1}]^{-1}[v,u]^{-1}$. We can then express

$$[v,u,u^{-1}] = [v,u,u,u^{-1}]^{-1}[v,u,u]^{-1}.$$

Since we are collecting modulo $\gamma_n(F)$, for some n, eventually we will be able to express vu^{-1} as a product of commutators which involve only u and v (and not u^{-1}), modulo $\gamma_n(F)$. We also need to be able to collect the expressions $v^{-1}u^{-1}$ and $v^{-1}u$. These are dealt with similarly (see [11, p.167]). So eventually, modulo $\gamma_n(F)$, we have a word consisting of a product of basic commutators which we can collect modulo $\gamma_n(F)$. Thus, using the Hall collection process, we can collect any word modulo any term of the lower central series.

It is not hard to see that we can collect positive words modulo any term of the derived series, but as soon as we have something with negative exponent, the Hall collection process breaks down if we are calculating modulo terms of the derived series. This is because we have to work with the terms $[v,u,\ldots,u,u^{-1}]$, which are not necessarily contained in any $F^{(m)}$ for $m > 1$. We will wish to collect modulo terms of the derived series in Chapter 4. In order to do this we need the two-sided collection process of Gorchakov [5], which we describe in Section 4.3. The two-sided collection process allows us to collect arbitrary words modulo any term of the derived series. The downside of the two-sided collection process is that we lose the uniqueness from Theorem 2.8.2, but this will not affect the results we achieve.

Chapter 3

A non-identical Lie relator

3.1 Introduction

Our purpose in this chapter is to investigate the connection between a group G and its associated Lie ring $\mathfrak{L}(G)$. The associated Lie ring of a group is an object which has proved enormously important in the study of the Burnside Problem, as well as in other areas. In particular, Kostrikin [23] and Zelmanov [53],[54] used it in the solution of the *Restricted Burnside Problem* (see [46]). As is the nature of things, the associated Lie ring has become an object of study in its own right (see [52]). This is the approach we take here.

The associated Lie ring has been calculated for surprisingly few groups (see Chapter 1 for an overview). Until now, whenever the Lie ring of a relatively free group has been calculated, it has been relatively free. This naturally leads to the question of whether the Lie ring of a relatively free group is necessarily relatively free. We provide a negative answer to this question by showing that the variety $\mathfrak{B}_4\mathfrak{B}_2$ contains non-identical Lie relators.

Let F be the free group of countably infinite rank, and L the free Lie ring of countably infinite rank. Given a variety $\mathfrak{V}$, we can express its relatively free group of countably infinite rank $F(\mathfrak{V})$ as a quotient of F by a fully invariant subgroup $\mathfrak{V}(F)$. So we obtain the exact sequence (see Subsection 3.2, or [52])

$$1 \to \mathfrak{V}(F) \to F \to F(\mathfrak{V}) \to 1 \tag{3.1}$$

If we let $\mathfrak{L}(F(\mathfrak{V}))$ be the associated Lie ring of $F(\mathfrak{V})$, then we can express $\mathfrak{L}(F(\mathfrak{V}))$ as a quotient of L. This gives the exact sequence

$$0 \to \mathfrak{V}(L) \to L \to \mathfrak{L}(F(\mathfrak{V})) \to 0. \tag{3.2}$$

Clearly, $\mathfrak{L}(F(\mathfrak{V}))$ is relatively free if and only if $\mathfrak{V}(L)$ is fully invariant. The existence of non-identical Lie relators for $\mathfrak{V}$ is equivalent to $\mathfrak{V}(L)$ not being fully invariant (see Subsection 3.2).

For any integer n, let $\mathfrak{B}_n$ be the variety of all groups of exponent dividing n (the *Burnside variety of exponent n*). Then, the variety $\mathfrak{B}_4\mathfrak{B}_2$ consists of all of those groups which are extensions of groups of exponent 4 by groups of exponent 2 (see Section 2.2). We investigate the variety $\mathfrak{B}_4\mathfrak{B}_2$, and show that it has Lie relators which are not identical Lie relators.

There are still many questions left unanswered by this result. Since much of the motivation for using the associated Lie ring comes from the Burnside problem, perhaps the most interesting examples to look at are the free Burnside groups. Whether or not the associated Lie ring of a free Burnside group is always relatively free remains an open question (see [46, p.39]). Vaughan-Lee conjectured that all Lie relators for $\mathfrak{B}_p$ are consequences of the multilinear ones, for all primes p (see [24, p.173]). Since all multilinear Lie relators are identical (see Theorem 3.3.2), this would mean that $\mathfrak{L}(\infty, p)$ is relatively free. Newman and Vaughan-Lee [32] calculated the largest 2-generator Lie ring which satisfies all the multilinear Lie relators satisfied by groups of exponent 7. Indications from group calculations made by Vaughan-Lee and Eamonn O'Brien [50] are that this Lie ring is strictly larger than $\mathfrak{L}(2,7)$. This would mean that there are Lie relators in the free 2-generator Burnside group of exponent 7 which are not consequences of the multilinear relators, answering Vaughan-Lee's conjecture in the negative. However, whether or not the associated Lie ring of a free Burnside group need be relatively free remains an open problem.

3.2 Varieties of groups and the functor $\mathfrak{L}$

As usual, the background for varieties of groups can be found in [30]. We briefly describe this background in Section 2.2. The definitions for varieties of Lie rings are analogous, and can be found in [2]. The following approach to the associated Lie ring is adapted (only slightly) from [52], from which we take much of our inspiration.

Let F be the absolutely free group of countably infinite rank. Given a variety of groups $\mathfrak{V}$ and the relatively free group of $\mathfrak{V}$ of countably infinite rank $F(\mathfrak{V})$, we have the following exact sequence,

$$1 \to \mathfrak{V}(F) \to F \to F(\mathfrak{V}) \to 1, \tag{3.3}$$

where $\mathfrak{V}(F)$ is a fully invariant subgroup of F. The operation which takes a group G to its associated Lie ring $\mathfrak{L}(G)$ gives rise to a functor from the category of groups to the category of Lie rings. We denote this functor by $\mathfrak{L}$. We are investigating the properties of this functor, specifically whether $\mathfrak{L}$ preserves the property of being relatively free. We have two important properties of the functor $\mathfrak{L}$ on which to build. The first is that $\mathfrak{L}$ preserves the surjectivity of homomorphisms. The second is Theorem 2.6.1.

We have the following exact sequence of graded Lie rings,

$$0 \to \mathfrak{V}(L) \to L \to \mathfrak{L}(F(\mathfrak{V})) \to 0, \tag{3.4}$$

where L is the absolutely free Lie ring of countably infinite rank and $\mathfrak{V}(L)$ is an ideal of L. Elements of the ideal $\mathfrak{V}(L)$ are called *Lie relators* for $\mathfrak{V}$. An element in L which is mapped to zero by every homomorphism from L into $\mathfrak{L}(F(\mathfrak{V}))$ is called an *identical Lie relator* for $\mathfrak{V}$. It is clear that the set of identical Lie relators forms a fully invariant ideal of L. We denote this ideal by $\mathfrak{V}^{id}(L)$. Then $\mathfrak{V}^{id}(L)$ is the maximal fully invariant ideal of L contained in $\mathfrak{V}(L)$ (see [52]). It is also clear that

$$\mathfrak{V}^{id}(L) \subseteq \mathfrak{V}(L). \tag{3.5}$$

Our result is that

$$(\mathfrak{B}_4\mathfrak{B}_2)^{id}(L) \neq (\mathfrak{B}_4\mathfrak{B}_2)(L). \tag{3.6}$$

This is the first variety for which the inclusion (3.5) is known to be strict.

3.3 Multilinear Lie relators

Definition 3.3.1 *An element of L is* multilinear *if it is a linear combination of Lie products of the form*

$$[q_{1\sigma}, q_{2\sigma}, \ldots, q_{n\sigma}] \quad (\sigma \in S_n),$$

for some distinct $q_1, \ldots, q_n \in \{z_1, z_2, \ldots\}$. We say that the element has support $\{q_1, \ldots, q_n\}$.

This definition of multilinear elements of L implies that multilinear elements are homogeneous.

Theorem 3.3.2 *Let $\mathfrak{V}$ be a variety of groups. Then, any multilinear Lie relator for $\mathfrak{V}$ is an identical Lie relator.*

Proof: We follow the proof of [33]. Without loss of generality, assume that the Lie relator has support $\{z_1, \ldots, z_n\}$. If $w(z_1, \ldots, z_n)$ is the Lie relator, then we can write

$$w(z_1, \ldots, z_n) = \sum_{\sigma \in S_n} \alpha_\sigma [z_{1\sigma}, \ldots, z_{n\sigma}],$$

for some $\alpha_\sigma \in \mathbb{Z}$.

Let $F(\mathfrak{V})$ be the relatively free group of countably infinite rank of $\mathfrak{V}$, and let $\{f_1, f_2, \ldots\}$ be a free generating set for $F(\mathfrak{V})$. Then, $\mathfrak{L}(F(\mathfrak{V}))$ is generated by $\{a_1, a_2, \ldots\}$, where $a_i = f_i \gamma_2(F(\mathfrak{V}))$.

Since $w(z_1, \ldots, z_n)$ is a Lie relator for $\mathfrak{V}$,

$$\sum_{\sigma \in S_n} \alpha_\sigma [a_{1\sigma}, \ldots, a_{n\sigma}] = 0.$$

This is equivalent to

$$\prod_{\sigma \in S_n} [f_{1\sigma}, \ldots, f_{n\sigma}]^{\alpha_\sigma} \in \gamma_{n+1}(F(\mathfrak{V})).$$

This, in turn, is equivalent to saying that

$$\prod_{\sigma \in S_n} [f_{1\sigma}, \ldots, f_{n\sigma}]^{\alpha_\sigma} = c_1 \ldots c_k, \tag{3.7}$$

where the c_i are commutators in $\{f_1, \ldots, f_n, f_1^{-1}, \ldots, f_n^{-1}\}$ of weight at least $n+1$, such that, for each $i \in \{1, \ldots, k\}$ and each $j \in \{1, \ldots, n\}$, c_i contains either f_j or f_j^{-1} (see Section 3.4.2). Now, let $b_1, \ldots, b_n$ be homogeneous elements of $\mathfrak{L}(F(\mathfrak{V}))$. Then, there exist $g_1, \ldots, g_n \in F(\mathfrak{V})$ such that $b_i = g_i \gamma_{j_i+1}(F(\mathfrak{V}))$ for $1 \leq i \leq n$ and some $j_1, \ldots, j_n$. Let $\theta : F(\mathfrak{V}) \to F(\mathfrak{V})$ be the endomorphism which extends the mapping $f_i \mapsto g_i$ $(1 \leq i \leq n)$. This endomorphism exists because $F(\mathfrak{V})$ is relatively free.

Now, let $j = j_1 + \cdots + j_n$ and apply the mapping θ to both sides of (3.7). Then,

$$\prod_{\sigma \in S_n} [g_{1\sigma}, \ldots, g_{n\sigma}]^{\alpha_\sigma} \in \gamma_{j+1}(F(\mathfrak{V})).$$

Passing to the Lie ring, $\mathfrak{L}(F(\mathfrak{V}))$, we obtain

$$\sum_{\sigma \in S_n} \alpha_\sigma [b_{1\sigma}, \ldots, b_{n\sigma}] = 0.$$

This proves the theorem for homogeneous elements. The theorem in the general case follows from the fact that the relator is linear in all of its variables (hence the term *multilinear*). ■

Vaughan-Lee completely characterised all of the multilinear Lie relators for the varieties $\mathfrak{B}_q$, for all prime powers q (see [46], Theorem 2.5.1).

Expanding this result to all varieties of groups, Wall [52], completely characterised the multilinear relators for any variety of groups, given a set of laws which determine the variety.

3.4 Gradings of $\mathfrak{L}(G)$

We mentioned in Section 2.5 that the subspaces $\mathfrak{L}_i$ form a natural grading of $\mathfrak{L}(G)$. This is called the grading by *degree*. There are two other gradings which we will consider, both of which are discussed in [52].

Definition 3.4.1 *Given a grading of L, write an arbitrary element y of L as*

$$y = \sum y_i,$$

where the y_i are each in a distinct homogeneous component of L with respect to the grading, and call the y_i the homogeneous components *of y. An additive subgroup A of L is called* graded *if, for all $y \in A$, each of the homogeneous components of y are also in A.*

It is clear that $\mathfrak{V}(L)$ is graded with respect to the grading by degree.

3.4.1 The uniform grading

Definition 3.4.2 *We call an element, z, of L* uniform *if all of its constituent Lie monomials contain exactly the same set of variables. If this set of variables is $\{z_{i_1}, \ldots, z_{i_k}\}$, then we say that z is* uniform with support $\{z_{i_1}, \ldots, z_{i_k}\}$.

The Lie algebra L is graded by taking each homogeneous component to be the set of uniform elements of L with a particular support. This is called the *uniform grading* of L. It is easy to see that any element can be uniquely written as a sum of uniform elements with given (distinct) supports.

Theorem 3.4.3 *Let $\mathfrak{V}$ be a variety of groups. Then $\mathfrak{V}(L)$ is uniformly graded.*

Proof: If N is a fully invariant subgroup of F, then we can use the method described in Section 3.1 to associate with N an ideal of L. Take the associated Lie ring of F/N and take the kernel of the natural map from L to $\mathfrak{L}(F/N)$. Denote this ideal by $\psi(N)$. It is not difficult to see that any endomorphism of L which maps all weight 1 elements of L to weight 1 elements must map $\psi(N)$ to itself (see [52]).

Let $S = \{z_{i_1}, \ldots, z_{i_k}\}$ be some set of variables. We prove that, for any homogeneous $a \in \mathfrak{V}(L)$, the uniform component of a with support S is also contained in $\mathfrak{V}(L)$. Define the endomorphism φ of L by $z_i\varphi = 0$ if $i \notin S$ and $z_i\varphi = z_i$ if $i \in S$. Then the uniform component of a with support S is equal to $a\varphi$ and is contained in $\mathfrak{V}(L)$ since $\psi(\mathfrak{V}(F)) = \mathfrak{V}(L)$, by definition. ■

3.4.2 Higman's argument

In this section we present an argument used in [19]. This is an argument about Lie relators and a possible method of proving that they are identical.

Let $\mathfrak{V}$ be a variety of groups, and $F(\mathfrak{V})$ its relatively free group of countably infinite rank.

Suppose that $w(z_1, z_2, \ldots, z_n)$ is a homogeneous Lie relator for $\mathfrak{V}$, of weight m, and that $\tilde{w}(x_1, x_2, \ldots, x_n)$ is the corresponding group word in F. That is to say, if we write out w as a sum of Lie products of the z_i, then $\tilde{w}$ is the word obtained in F by replacing each z_i with x_i, each Lie product with a group commutator, each minus sign by group inversion, and each addition with a group multiplication. This is something we will do regularly, and we try to use the symbol '˜' to denote this operation. Note that this operation is not well-defined, because addition is commutative in L, but multiplication is not commutative in F. However, this choice will not affect any of our results, so we ignore this problem.

We can assume, by the previous section, that w is a uniform element of L with support $\{z_1, \ldots, z_n\}$.

Suppose that $F(\mathfrak{V})$ is relatively freely generated by $\{f_1, f_2, \ldots\}$.

Then, $\tilde{w}(f_1, \ldots, f_n) \in \gamma_{m+1}(F(\mathfrak{V}))$ and so, in $F(\mathfrak{V})$, we have an identity of the following form:

$$\tilde{w}(f_1, f_2, \ldots, f_n) = c_1 c_2 \cdots c_k,$$

where $c_1, \ldots, c_k$ are commutators of weight at least $m+1$ involving only $f_1, f_2, \ldots, f_n$ and their inverses. For the remainder of this section, we will say 'contains f_i' where we mean 'contains either f_i or f_i^{-1}'.

We now prove that we can assume that all of the variables $f_1, \ldots, f_n$ occur in each of the commutators $c_1, \ldots, c_k$. Suppose that $f_1, \ldots, f_i$ occur in all of $c_1, \ldots, c_k$, but that f_{i+1} does not (this includes the case $i = 0$). We use the process of *commutator collection* (see [11], Chapter 11) to collect all of the commutators without f_{i+1} in them to the left. Suppose that we have a minimal $j, j+1$ such that c_j contains f_{i+1}, but c_{j+1} does not. Then we use the identity $c_j c_{j+1} = c_{j+1} c_j [c_j, c_{j+1}]$ to swap their order, at the expense of adding commutators which involve f_{i+1}. Repeating this procedure, we can place all of the commutators which do not contain f_{i+1} on the left, and get an expression

$$\tilde{w}(f_1, \ldots, f_n) = d_1 \cdots d_t, \tag{3.8}$$

where $d_1, \ldots, d_t$ all contain $f_1, \ldots, f_i$, $d_1, \ldots, d_s$ do not contain f_{i+1}, for some s and $d_{s+1}, \ldots, d_t$ do contain f_{i+1}. Now, let $\varphi : F(\mathfrak{V}) \to F(\mathfrak{V})$ be the endomorphism which extends the mapping $f_k \mapsto f_k$, if $k \neq i+1$, and $f_{i+1} \mapsto 1$. Applying φ to both sides of 3.8 gives

$$1 = d_1 \cdots d_s,$$

from which it follows that

$$\tilde{w}(f_1, \ldots, f_n) = d_{s+1} \cdots d_t.$$

We repeat this argument for $0 \leq i \leq n-1$. Thus, we can assume that all commutators contain each of $f_1, \ldots, f_n$.

We can now consider substituting all possible combinations of elements of $\mathfrak{L}(F(\mathfrak{V}))$ for $z_1, \ldots, z_n$ and seeing what the commutator identity above tells us. This is what we do when we try to discover if a given Lie relator is identical.

3.4.3 The multihomogeneous grading

Other than the grading by degree and the uniform grading, we consider one other, the *multihomogeneous* grading. This means that the element is uniform, and for a fixed variable, the weight of that variable in each monomial is the same. A subspace is called *multigraded* if it is graded under the multihomogeneous grading.

It is not true that $\mathfrak{V}(L)$ is necessarily multigraded. For example, the variety $\mathfrak{B}_4$ has the element

$$\begin{aligned}&[z_1,z_2,z_3,z_3]+[z_1,z_3,z_2,z_3]+[z_1,z_3,z_3,z_2]+\\&[z_2,z_3,z_1,z_1]+[z_2,z_1,z_3,z_1]+[z_2,z_1,z_1,z_3]+\\&[z_3,z_1,z_2,z_2]+[z_3,z_2,z_1,z_2]+[z_3,z_2,z_2,z_1]\end{aligned}$$

as a Lie relator (see [46, p.145]), whilst none of the three multihomogeneous components of this element are Lie relators. Thus, $\mathfrak{B}_4(L)$ is not multigraded.

Given Higman's argument, we know that $\mathfrak{V}(L)$ is graded by the uniform grading. We also know that it is graded by degree, from the construction of the associated Lie ring. However, it is not necessarily multigraded, as the above example shows. This helps us to direct our search for a non-identical Lie relator somewhat, for we can restrict our search to homogeneous elements which are also uniform. We also know that a non-identical Lie relator cannot be multilinear. Our example, as given in Section 3.6 will not be multihomogeneous either.

3.5 Calculating the associated Lie ring

We use this section to describe how we go about calculating the associated Lie ring in particular cases. We use the program GAP [38] to make our calculations. Within GAP, we use the p-quotient algorithm (see [16] for a theoretical background). Since our group will a 2-group, we will calculate the lower central 2-series (see Section 3.5.2 below). First we describe *power-commutator presentations*, a means for effectively describing and computing with finite p-groups. Most of what we say in this section is relevant to computing the associated Lie ring of any finite p-group for which the lower central series coincides with the lower central p-series (by replacing '2' with 'p').

3.5.1 Power-commutator presentations

Definition 3.5.1 *Let p be a prime number, and G a finite p-group. A* power-commutator presentation (or PCP) *for G is a presentation consisting of a set of generators $g_1, \ldots, g_n$ and a set of relations*

$$g_i^p = g_{i+1}^{\alpha(i,i+1)} \cdots g_n^{\alpha(i,n)} \quad (1 \leq i \leq n), \tag{3.9}$$

$$[g_i, g_j] = g_{i+1}^{\beta(i,j,i+1)} \cdots g_n^{\beta(i,j,n)} \quad (1 \leq j < i \leq n), \tag{3.10}$$

where $0 \leq \alpha(i,k), \beta(i,j,k) < p$ for all i, j and k.

For a theoretical background to PCPs see [16].

3.5.2 The lower central p-series

Definition 3.5.2 *Let $G_1 := G$, and, if G_i is already defined,*

$$G_{i+1} := [G_i, G]G_i^p.$$

We call the series of subgroups $\{G_i\}$ the lower central p-series *of G.*

Lemma 3.5.3 1. $[G_i, G_j] \leq G_{i+j}$,

2. *G_i/G_{i+1} is an elementary abelian p-group, and is a central factor,*

3. *$\gamma_i(G) \leq G_i$ for all $i \geq 1$.*

Proof: First, 3.5.3.1 is proved in a similar way to 2.5.1.5 (see [46, pp.28–29]). 3.5.3.2 follows immediately from the definition of the series. 3.5.3.3 follows from the definitions of $\gamma_i(G)$ and G_i by the easiest of inductions. ■

We are using GAP, (see [38]), to calculate the lower central p-factors of our groups, using the p-quotient algorithm (see [16]). In the group we will calculate, the lower central series coincides with the lower central 2-series. Thus when we calculate the lower central 2-series, we can use this to calculate the associated Lie ring. The PCP which we obtain allows us to obtain the associated Lie ring almost directly.

3.5.3 Using GAP to calculate the associated Lie ring

We now address the question of how to calculate the associated Lie ring of a group. We implicitly assume here that all of our groups are nilpotent. Later in the chapter when we make explicit computations, this will be true. It is much quicker and easier to calculate the lower central p-series of a group than the lower central series itself. In addition to this, when we use the ANU p-quotient program, GAP gives us a PCP which gives a concrete description of the group in terms of the lower central p-series, and makes calculations very easy also.

In our group $J \in \mathfrak{B}_4\mathfrak{B}_2$, which exhibits the non-identical Lie relator, the generators are all of order 2, so the lower central 2-series is the same as the lower central series. To see this, use Lemma 2.2.3 of [46] and simple commutator identities.

Suppose that we wish to compute the associated Lie ring of a finitely generated group G, for which the lower central series coincides with the lower central 2-series. The basic commutators form a graded basis for the free Lie ring, and we may consider the associated Lie ring of any group as a graded quotient of the free Lie ring. Thus, to calculate the homogeneous component of $\mathfrak{L}(G)$ of any weight n, we need only find all the (additive) relations that hold in $\mathfrak{L}(G)$ between basic commutators of weight n. This involves working only in $\gamma_n(G)/\gamma_{n+1}(G)$. We are assuming that we can use the lower central 2-series. So we calculate $H = G/\gamma_{n+1}(G)$ using the ANU p-quotient algorithm and obtain a finite group.

We then calculate the values of the basic commutators of weight n in H. This gives us elements of $\gamma_n(G)/\gamma_{n+1}(G)$, which is an elementary abelian $\mathbb{F}_2$-space, and thus a vector space. Hence it is a straightforward matter to calculate all additive relations between the basic commutators. This gives us an additive basis for the relations of $\mathfrak{L}(G)$ of weight n.

In the case we work with the group, and hence the associated Lie ring, is nilpotent. Therefore this method of calculating the relations one weight at a time will terminate.

Given that we have calculated the associated Lie ring, our next task is to determine whether this Lie ring is relatively free. In the variety $\mathfrak{B}_4\mathfrak{B}_2$, the associated Lie ring is not relatively free, so all we need to do to prove this is exhibit a non-identical Lie relator. If the associated Lie ring is relatively free, then it becomes a more complicated process to establish this, but this will not concern us here.

3.6 Constructing a non-identical Lie relator

This section contains the main result of this chapter.

We still assume that F is the absolutely free group of countably infinite rank, with free generating set $\{x_1, x_2, \ldots\}$, and that L is the absolutely free Lie ring of countably infinite rank, with free generating set $\{z_1, z_2, \ldots\}$. We implicitly identify L with $\mathfrak{L}(F)$ and z_i with $x_i\gamma_2(F)$ for each i.

For any group G and any Lie ring K we define expressions $\tilde{w}$ and w, respectively. If $a, b, c \in G$, then

$$\begin{aligned}\tilde{w}(a,b,c) &= [a,b,c,c][a,c,b,c][a,c,c,b][b,c,a,a][b,a,c,a][b,a,a,c] \\ &\quad [c,a,b,b][c,b,a,b][c,b,b,a]. \end{aligned} \tag{3.11}$$

Similarly, if $x, y, z \in K$, then

$$\begin{aligned}w(x,y,z) &= [x,y,z,z] + [x,z,y,z] + [x,z,z,y] \\ &\quad + [y,z,x,x] + [y,x,z,x] + [y,x,x,z] \\ &\quad + [z,x,y,y] + [z,y,x,y] + [z,y,y,x]. \end{aligned} \tag{3.12}$$

The Lie relator w is certainly a possible candidate for one which is non-identical (see [52]). The main reason for this is that it is not multi-homogeneous. Let G be a relatively free group of countably infinite rank, relatively freely generated by $g_1, g_2, \ldots$, such that $w(z_1, z_2, z_3)$ is a Lie relator for G. This is equivalent to $\tilde{w}(g_1, g_2, g_3) = c$, for some $c \in \gamma_5(G)$. We can assume that $c = c(g_1, g_2, g_3)$ is a product of commutators in g_1, g_2, g_3 (and their inverses), each of weight at least 5, and each involving all of g_1, g_2, g_3 (see Subsection 3.4.2). The fact that G satisfies the Lie relator $w(z_1, z_2, z_3)$, does not immediately imply that it satisfies the Lie relator $w(z_1, z_2, [z_3, z_4])$ (in fact this remains an open question). To see this, let

$$\begin{aligned}w_1 &= [z_2,[z_3,z_4],z_1,z_1] + [z_2,z_1,[z_3,z_4],z_1] + [z_2,z_1,z_1,[z_3,z_4]] \\ &\quad + [[z_3,z_4],z_1,z_2,z_2] + [[z_3,z_4],z_2,z_1,z_2] + [[z_3,z_4],z_2,z_2,z_1],\end{aligned}$$

and

$$\begin{aligned}w_2 &= [z_1,z_2,[z_3,z_4],[z_3,z_4]] + [z_1,[z_3,z_4],z_2,[z_3,z_4]] \\ &\quad + [z_1,[z_3,z_4],[z_3,z_4],z_2]. \end{aligned} \tag{3.13}$$

Then $w(z_1, z_2, [z_3, z_4]) = w_1 + w_2$, and w_1 is homogeneous of degree 5, while w_2 is homogeneous of degree 6. Now, let

$$\begin{aligned}\tilde{w}_1 &= [g_2,[g_3,g_4],g_1,g_1][g_2,g_1,[g_3,g_4],g_1][g_2,g_1,g_1,[g_3,g_4]] \\ &\quad [[g_3,g_4],g_1,g_2,g_2][[g_3,g_4],g_2,g_1,g_2][[g_3,g_4],g_2,g_2,g_1], \qquad (3.14)\end{aligned}$$

and

$$\begin{aligned}\tilde{w}_2 &= [g_1,g_2,[g_3,g_4],[g_3,g_4]][g_1,[g_3,g_4],g_2,[g_3,g_4]] \\ &\quad [g_1,[g_3,g_4],[g_3,g_4],g_2]. \qquad (3.15)\end{aligned}$$

Then $\tilde{w}(g_1, g_2, [g_3, g_4]) = \tilde{w}_2\tilde{w}_1$. Now, since $\tilde{w}(g_1, g_2, g_3) = c(g_1, g_2, g_3)$, and since G is relatively free,

$$\tilde{w}_2\tilde{w}_1 = \tilde{w}(g_1, g_2, [g_3, g_4]) = c(g_1, g_2, [g_3, g_4]) \in \gamma_6(G), \qquad (3.16)$$

and so $\tilde{w}_1 \in \gamma_6(G)$. This is enough to ensure that w_1 is a Lie relator, but not enough to ensure that w_2 is a Lie relator.

Despite all this, it is a well known result (see, for example, [46, p.145]) that if $B \in \mathfrak{B}_4$, and $x, y, z \in \mathfrak{L}(B)$, then

$$w(x, y, z) = 0. \qquad (3.17)$$

So w is an identical Lie relator for $\mathfrak{B}_4$. It is also known (in fact it is a consequence of the above) that for all $a, b, c \in B$,

$$\tilde{w}(a, b, c) \in \gamma_5(B). \qquad (3.18)$$

However, it is unknown whether there may be other varieties for which w is a Lie relator which is not identical.

We use the above approach, with a slightly more complicated substitution, to prove that $\mathfrak{B}_4\mathfrak{B}_2$ has non-identical Lie relators.

We now prove that $w([z_1, z_2], [z_3, z_4], [z_5, z_6])$ is a Lie relator for the variety $\mathfrak{B}_4\mathfrak{B}_2$.

Lemma 3.6.1 *Let F_6 be the relatively free group of rank 6 of the variety $\mathfrak{B}_4\mathfrak{B}_2$ and let F_6 be generated by $\{f_1, f_2, f_3, f_4, f_5, f_6\}$. For $1 \leq i \leq 6$, let $y_i = f_i\gamma_2(F_6)$. Then*

$$w([y_1, y_2], [y_3, y_4], [y_5, y_6]) = 0, \qquad (3.19)$$

in $\mathfrak{L}(F_6)$. *Thus,* $w([z_1, z_2], [z_3, z_4], [z_5, z_6])$ *is a Lie relator for* $\mathfrak{B}_4\mathfrak{B}_2$.

Proof: F_6' has exponent 4 because $\mathfrak{B}_2$ is abelian. This means that, for any $h_1, h_2, h_3 \in F_6'$, $\tilde{w}(h_1, h_2, h_3) \in \gamma_5(F_6')$. A trivial inductive proof using 2.5.1.5 shows that $\gamma_5(F_6') \leq \gamma_{10}(F_6)$. Thus, in particular,

$$\tilde{w}([f_1, f_2], [f_3, f_4], [f_5, f_6]) \in \gamma_{10}(F_6). \tag{3.20}$$

Now, if non-trivial, $w([y_1, y_2], [y_3, y_4], [y_5, y_6])$ is a homogeneous element of weight 8 in $\mathfrak{L}(F_6)$. However,

$$\begin{aligned} w([y_1, y_2], [y_3, y_4], [y_5, y_6]) &= \tilde{w}([f_1, f_2], [f_3, f_4], [f_5, f_6])\gamma_9(F_6) \\ &= \gamma_9(F_6). \end{aligned} \tag{3.21}$$

This completes the proof of the lemma. ■

We prove that $w([x_1, x_2], [x_3, x_4], [x_5, x_6])$ is not an *identical* Lie relator for $\mathfrak{B}_4\mathfrak{B}_2$ by explicitly constructing a particular group $J \in \mathfrak{B}_4\mathfrak{B}_2$, and proving that $w([z_2, z_1], [z_3, z_1], [z_2, z_1, z_3])$ is not a Lie relator for J.

Let $G = \langle a, b, c \mid a^2 = b^2 = c^2 = 1 \rangle$, and define $H := G/((G')^4\gamma_5(G'))$. It is clear that $H \in \mathfrak{B}_4\mathfrak{B}_2$ since H' has exponent 4 and the natural generators of H have order 2. We examined various finite quotients of H (which are also in $\mathfrak{B}_4\mathfrak{B}_2$). We eventually settled on the following group:

$$\begin{aligned} J = \Big\langle a, b, c \;\Big|\;\; & a^2 = b^2 = c^2 = [c, b] = 1, \\ & [c, a]^4 = [b, a]^4 = ([b, a][c, a])^4 = 1, \\ & [[b, a, c], [b, a, a]] = [b, a, c, a, b, a, b][b, a, c, a, b, a, b, c, a], \\ & [[c, a, a], [b, a, c]] = [b, a, c, a, c, a, c], [[c, a, a], [b, a, a]] = 1, \\ & [b, a, c, a, c, a, b][b, a, c, a, b, a, b, c] = 1, \\ & [b, a, c, a]^2 = [b, a, c, a, c, a, c, a], \\ & [b, a, c, a, c]^2 = 1, [b, a, c, a, c, a, b, c] = 1, \\ & \text{nilpotent class } 10 \Big\rangle . \end{aligned} \tag{3.22}$$

To compute J from within GAP, it is sufficient to define a group satisfying the 14 relations and then run the p-quotient algorithm to class 10 at the prime 2. Note that since the generators are of order 2, the lower central 2-series and the lower central series coincide. GAP tells us immediately that $|J| = 2^{21}$.

It is straightforward to verify that J' has exponent 4 and nilpotency class 4. To do this using GAP once J has been computed, we use the following commands:

```
gap > K := DerivedSubgroup(J);
gap > L := LowerCentralSeries(K);
gap > Exponent(K);
```

The second command calculates the lower central series for J', from which it is easy to see that the fifth term is trivial. The third command yields the answer 4. Thus, J is a finite quotient of H. In Appendix A we present a PCP for J, from which all calculations can be read off.

We show that

$$w([b\gamma_2(J), a\gamma_2(J)], [c\gamma_2(J), a\gamma_2(J)], [b\gamma_2(J), a\gamma_2(J), c\gamma_2(J)]) \neq 0 \qquad (3.23)$$

in $\mathfrak{L}(J)$.

First, $\tilde{w}([b,c],[c,a],[b,a,c]) \in \gamma_5(J') = \{1\}$ (because J' has exponent 4). Then we have

$$w([z_2,z_1],[z_3,z_1],[z_2,z_1,z_3]) = u(z_1,z_2,z_3) + v(z_1,z_2,z_3), \qquad (3.24)$$

where

$$\begin{aligned} u(z_1,z_2,z_3) \;:=\; & [[z_3,z_1],[z_2,z_1,z_3],[z_2,z_1],[z_2,z_1]] \\ & + [[z_3,z_1],[z_2,z_1],[z_2,z_1,z_3],[z_2,z_1]] \\ & + [[z_3,z_1],[z_2,z_1],[z_2,z_1],[z_2,z_1,z_3]] \\ & + [[z_2,z_1,z_3],[z_2,z_1],[z_3,z_1],[z_3,z_1]] \\ & + [[z_2,z_1,z_3],[z_3,z_1],[z_2,z_1],[z_3,z_1]] \\ & + [[z_2,z_1,z_3],[z_3,z_1],[z_3,z_1],[z_2,z_1]]\,, \end{aligned} \qquad (3.25)$$

is homogeneous of weight 9, and

$$\begin{aligned} v(z_1,z_2,z_3) \;:=\; & [[z_2,z_1],[z_3,z_1],[z_2,z_1,z_3],[z_2,z_1,z_3]] \\ & + [[z_2,z_1],[z_2,z_1,z_3],[z_3,z_1],[z_2,z_1,z_3]] \\ & + [[z_2,z_1],[z_2,z_1,z_3],[z_2,z_1,z_3],[z_3,z_1]]\,, \end{aligned} \qquad (3.26)$$

is homogeneous of weight 10.

For any group G and any $g, h, k \in G$, define

$$\begin{aligned}\tilde{v}(g,h,k) &= [[h,g],[k,g],[h,g,k],[h,g,k]]\,[[h,g],[h,g,k],[k,g],[h,g,k]] \\ &\quad [[h,g],[h,g,k],[h,g,k],[k,g]] \qquad (3.27)\end{aligned}$$

and,

$$\begin{aligned}\tilde{u}(g,h,k) &= [[k,g],[h,g,k],[h,g],[h,g]]\,[[k,g],[h,g],[h,g,k],[h,g]] \\ &\quad [[k,g],[h,g],[h,g],[h,g,k]]\,[[h,g,k],[k,g],[h,g],[h,g]] \\ &\quad [[h,g,k],[h,g],[k,g],[h,g]]\,[[h,g,k],[h,g],[h,g],[k,g]] \qquad (3.28)\end{aligned}$$

Then, in J,

$$\tilde{v}(a,b,c)\tilde{u}(a,b,c) = \tilde{w}([b,a],[c,a],[b,a,c]) = 1. \qquad (3.29)$$

Since $\tilde{v}(a,b,c) \in \gamma_{10}(J)$, this implies that $\tilde{u}(a,b,c) \in \gamma_{10}(J)$, and hence that $u(z_1,z_2,z_3)$ is a Lie relator of J. So $w([z_2,z_1],[z_3,z_1,[z_2,z_1,z_3])$ will only be a Lie relator of J if $v(z_1,z_2,z_3)$ is a Lie relator. But, for $v(z_1,z_2,z_3)$ to be a Lie relator of J we need $\tilde{v}(a,b,c) \in \gamma_{11}(J) = \{1\}$. From the PCP for J given in Appendix A, it is easy to see that $\tilde{v}(a,b,c) = j_{21} = [b,a,c,a,b,a,b,c,a,c] \notin \gamma_{11}(J) = \{1\}$.

Lemma 3.6.1 proves that the element $w([z_1,z_2],[z_3,z_4],[z_5,z_6])$ is a Lie relator of $F(\mathfrak{B}_4\mathfrak{B}_2)$. We now show that $w([z_2,z_1],[z_3,z_1],[z_2,z_1,z_3])$ is not a Lie relator of $F(\mathfrak{B}_4\mathfrak{B}_2)$. To see this, note that

$$w([z_2,z_1],[z_3,z_1],[z_2,z_1,z_3]) = u(z_1,z_2,z_3) + v(z_1,z_2,z_3). \qquad (3.30)$$

The same argument as that given above for J shows that $u(z_1,z_2,z_3)$ is a Lie relator of $F(\mathfrak{B}_4\mathfrak{B}_2)$. Thus we need only show that $v(z_1,z_2,z_3)$ is not a Lie relator of $F(\mathfrak{B}_4\mathfrak{B}_2)$. If it were then the variety $\mathfrak{B}_4\mathfrak{B}_2$ would satisfy an identity

$$\tilde{v}(x_1,x_2,x_3) = d(x_1,x_2,x_3), \qquad (3.31)$$

where $d \in \gamma_{11}(F)$. Since $J \in \mathfrak{B}_4\mathfrak{B}_2$, J would have to satisfy this identity, which would imply that

$$\tilde{v}(a,b,c) = d(a,b,c) \in \gamma_{11}(J), \qquad (3.32)$$

which we have already shown is not the case. Hence $w([z_2,z_1],[z_3,z_1],[z_4,z_3])$ is a non-identical Lie relator for the variety $\mathfrak{B}_4\mathfrak{B}_2$ and we have the following:

Theorem 3.6.2 *There are non-identical Lie relators in the variety* $\mathfrak{B}_4\mathfrak{B}_2$. *Hence* $(\mathfrak{B}_4\mathfrak{B}_2)(L)$ *is not fully invariant, so*

$$(\mathfrak{B}_4\mathfrak{B}_2)^{id}(L) \neq (\mathfrak{B}_4\mathfrak{B}_2)(L).$$

Chapter 4

Torsion in outer commutator varieties

4.1 Introduction

In this chapter we define outer commutator groups and outer commutator Lie rings, after Hall [14]. We wish to investigate whether outer commutator groups are torsion-free or residually nilpotent (or both). Our method of attack comes from outer commutator Lie rings and from considering basic commutators within outer commutator Lie rings and the associated outer commutator ideals.

We prove results about certain outer commutator Lie rings and use a result of Kuz'min and Shapiro [26] to translate these into results about groups. We then prove that these groups are residually nilpotent, using our results about Lie rings and the *two-sided collection process* of Gorchakov [5] (see Section 4.3).

Whether a group is residually nilpotent and/or torsion-free appear, at first sight, to be unrelated issues. However, in order to prove that a group is torsion-free using Lie ring methods, the group must be residually nilpotent. Conversely, in our proof that the groups under consideration are residually nilpotent, the fact that the groups have torsion-free lower central factors is crucial. Thus, it appears that these two seemingly unrelated questions are actually intricately linked.

In Section 4.6 we exhibit torsion in a number of outer commutator Lie rings. Here our method breaks down and we can say nothing about the

analogous outer commutator groups. However, the existence of torsion in the outer commutator Lie rings may be a result of interest in its own right.

In this chapter (except for Section 4.6), F will be a free group of arbitrary rank, and L will be a free Lie ring of the same rank. However, any element of L which we consider will be contained in a single multihomogeneous component of L, which is finite dimensional.

4.2 Outer commutators

We define commutator-subgroup functions after Hall [14], and then move on to outer commutator identities.

Definition 4.2.1 *Suppose φ and ψ are functions defined for all groups G whose values are subgroups of G, denoted $\varphi(G)$ and $\psi(G)$ respectively. Then $[\varphi, \psi]$ is the function whose value for G is the subgroup $[\varphi(G), \psi(G)]$.*

Definition 4.2.2 *The class $\mathfrak{C}$ of* commutator-subgroup functions *is the intersection of all classes $\mathfrak{D}$ which consist of functions whose values are subgroups and which satisfy (i) $\mathfrak{D}$ contains the identity function, γ_1, and (ii) for all $\varphi, \psi \in \mathfrak{D}$, $[\varphi, \psi] \in \mathfrak{D}$.*

It is clear that a group G is soluble if and only if $\varphi(G) = 1$ for some $\varphi \in \mathfrak{C}$.

Definition 4.2.3 *An* outer commutator group *is a group which is isomorphic to $F/\varphi(F)$ for some free group F and some $\varphi \in \mathfrak{C}$.*

We will be largely concerned with groups $F/\varphi(F)$ where $\varphi = [\gamma_j, \gamma_i, \gamma_k]$ or $\varphi = [\gamma_j, \gamma_i, \gamma_k, \gamma_l]$, where γ_n is the function taking any group G to $\gamma_n(G)$.

We also make analogous definitions for Lie rings.

Definition 4.2.4 *The class $\mathfrak{P}$ of* commutator-subring functions *is the intersection of all classes $\mathfrak{Q}$ which consist of functions whose values are subrings which satisfy (i) $\mathfrak{Q}$ contains the identity function and (ii) for all $\varphi, \psi \in \mathfrak{Q}$, $[\varphi, \psi] \in \mathfrak{Q}$.*

Definition 4.2.5 *An* outer commutator Lie ring *is a Lie ring which is isomorphic to $L/\psi(L)$ for some free Lie ring L and some $\psi \in \mathfrak{P}$.*

For any given commutator subgroup function there is an obvious analogous commutator subring function and vice versa.

4.3 The two-sided collection process

The commutator collection process of Philip Hall is deficient for our purposes in one major respect. Whenever a negative exponent appears in the word being collected the collection process breaks down unless we are working modulo a term of the lower central series. We wish to collect modulo terms of the derived series. It is to overcome this difficulty that we turn to the two-sided collection process of Gorchakov [5].

The principle behind two-sided collection is collecting elements with positive exponent to the left of a word and elements with negative exponent to the right, leaving the unarranged part in the middle. This relies on the following elementary formulae about commutators:

$$\begin{aligned} vu &= uv[v,u], && (4.1)\\ u^{-1}v &= v[u,v]^{-1}u^{-1} && (4.2)\\ &= v[v,u]u^{-1}, && (4.3)\\ u^{-1}v^{-1} &= [v,u]^{-1}v^{-1}u^{-1}. && (4.4) \end{aligned}$$

A word

$$c_1 \cdots c_m c_{m+1}^{\epsilon_1} \cdots c_{m+k}^{\epsilon_k} c_{m+k+1}^{-1} \cdots c_n^{-1}$$

where the c_i are basic commutators and $\epsilon_j = \pm 1$ for $1 \leq j \leq k$, generally contains a *positive arranged part*

$$c_1 \cdots c_m,$$

where $c_1 \leq \ldots \leq c_m$ and $c_m \leq c_j$, for all $j \in \{m+1, \ldots, m+k\}$, a *negative arranged part*

$$c_{m+k+1}^{-1} \cdots c_n^{-1},$$

where $c_{m+k+1} \geq \cdots \geq c_n$ and $c_{m+k+1} \leq c_j$, for all $j \in \{m+1, \ldots, m+k\}$, and an *unarranged part*

$$c_{m+1}^{\epsilon_1} \cdots c_{m+k}^{\epsilon_k},$$

where if $\epsilon_1 = 1$, c_{m+1} is not the least of the c_j,for $m+1 \leq j \leq m+k$, and if $e_k = -1$, c_{m+k} is not the least of the same c_i. If the unarranged part is empty, we say the word is *arranged.*

Two-sided collection using identities (4.1)–(4.4) yields only basic commutators (see [11] for an explanation of this for positive words, or [5] in the general case).

As with the collection process of Hall, the two-sided collection process does not in general terminate. However, it does terminate modulo any term of the derived series, and hence modulo any outer commutator subgroup.

Theorem 4.3.1 (Gorchakov) *Let $N \trianglelefteq F$ be a subgroup of a free group F such that $F^{(m)} \subseteq N$ for some m. Then, modulo N, any word in F can be collected to a word in arranged form in a finite number of steps.*

This theorem is not difficult to prove. A drawback with the two-sided collection process is that it does not yield a unique expression of each element as an arranged word modulo a term of the derived series (see [5] for an example). However, this does not cause problems in our situation.

We finish this section with an example of two-sided collection, in order to make the notion more concrete. Let $f = x^{-1}y^{-1}x^2y$. Then

$$\begin{aligned} x^{-1}y^{-1}x^2y &= x^{-1}x[y,x]^{-1}y^{-1}xy \\ &= [y,x]^{-1}x[y,x]^{-1}y^{-1}y \\ &= x[y,x,x]^{-1}[y,x]^{-2}, \end{aligned}$$

which is in collected form. More complicated words get much more complicated very quickly. In [5, p.758], the example

$$\begin{aligned} x^{-1}y^{-1}xy &= xy[y,x,x,y]^{-1}[y,x,y]^{-1}[y,x,x]^{-1}[y,x]^{-1}y^{-1}x^{-1} \bmod F^{(2)} \\ &= xxy[y,x,x,x,y]^{-1}[y,x,x,y]^{-1}[y,x,x,y]^{-2}[y,x,x,x]^{-1} \\ &\quad [y,x,y]^{-1}[y,x,x]^{-2}[y,x]^{-1}y^{-1}x^{-1}x^{-1} \bmod F^{(2)}, \end{aligned}$$

was used to show that the two-sided collection process is not unique, even modulo a term of the derived series. Of course, $f = [y,x]^{-1}$ is also in collected form. One source of this lack of uniqueness is the indecision as to whether xx^{-1} should be cancelled or collected to opposite sides of the word. The above example comes from not cancelling, and from writing $f = x^{-1}y^{-1}xxx^{-1}y$.

4.4 The main results for Lie rings

In this section we prove that certain Lie rings defined by outer commutator identities are torsion-free. In Sections 4.8 and 4.9 we use these results, and the two-sided collection process, to prove that the analogous groups are both residually nilpotent and torsion-free.

Throughout this section we will be considering how elements are expressed as sums of basic commutators. In Section 2.3, the concept of the *multihomogeneous components* of a free Lie ring are introduced. We will use linearity to assume that we start with an element lying in a single multihomogeneous component. The process of transforming a multihomogeneous element into a linear combination of basic commutators uses only anti-commutativity and the Jacobi identity, and so we never move outside of the given multihomogeneous component.

The upshot of this is that whatever the rank of F and L may be, we are effectively working in the case of finite rank. A second point is that only a finite number of variables occur with non-zero weight in any given multihomogeneous component, and so there are only a finite number of basic commutators in each multihomogeneous component. We use the ordering of basic commutators to produce a recursive method which must eventually terminate. What we mean by this is that given an element $[a, b]$ with a and b basic commutators satisfying a number of properties, we construct an element $[c, d]$ such that c and d are basic commutators, c and d satisfy the same properties as a and b, respectively, and such that $d > b$. Since we are in a single multihomogeneous component, this process can only happen a finite number of times. We refer to this process as *reverse induction.*

The following two lemmas are well-known (see [2], Chapter 2 or [43]).

Lemma 4.4.1 *The basic commutators c such that $\mathrm{wt}(c) \geq n$ form a basis for L^n as a free $\mathbb{Z}$-module.*

Lemma 4.4.2 *The ideal L^n of a free Lie ring L is a free Lie ring, freely generated by the set of basic commutators c such that (i) $\mathrm{wt}(c) \geq n$ and (ii) if $c = [c_1, c_2]$ then $\mathrm{wt}(c_2) < n$.*

Given these two results we have the following theorem.

Theorem 4.4.3 *Suppose that L is a free Lie ring and that i and j are positive integers. If $x \in (L^i)^j$ then x can be written as a linear combination of basic commutators, each of which lies in $(L^i)^j$.*

Proof: We have a basis for L^i which is given to us by Lemma 4.4.2. Let us call this set of basic commutators C. We can then use Lemma 4.4.1 to see that $(L^i)^j$ is spanned by products of at least j elements of C (with some bracketing). We also know that we can express each of these products as a

linear combination of basic commutators, considered as an element of L. We show that each of the basic commutators which appear in some such linear combination is in fact an element of $(L^i)^j$. We do this by showing that each of these basic commutators is in fact a product of at least j elements from C.

Consider an element a which is a product of r elements from C, where $r \geq j$. Let $S = \{c_1, \ldots, c_k\}$ be the set of elements from C which occur in a. Suppose, inductively, that if x is a product of s elements from S, where $s < r$, then x can be expressed as a linear combination of basic commutators, each of which consists of a product of s elements from S.

This is clear when $s = 1$, since then x itself is a basic commutator. We prove that a may be expressed as a linear combination of basic commutators, each of which is a product of r elements from S.

Now, let $a = [d, e]$ where d and e are products of elements from S (the case where there is no such representation of a is covered by the base case of the induction). By the inductive hypothesis and linearity, we may assume that d and e are also basic commutators. We may assume without loss of generality that $d > e$. If d and e are themselves elements of S (and so elements of C), then condition (ii) from Lemma 4.4.2 ensures that a is a basic commutator. Thus in this case there is nothing more to prove.

So suppose that one of d and e is not contained in C. If $d \in C$ then condition (ii) of Lemma 4.4.2 again tells us that $[d, e]$ is a basic commutator. Thus suppose that $d \notin C$.

If $[d, e]$ is not a basic commutator, then $d = [d_1, d_2]$ where $d_1 > d_2 > e$, and d_1 and d_2 are both products of elements of S and basic commutators themselves. Then

$$[d, e] = [d_1, e, d_2] - [d_2, e, d_1],$$

by the Jacobi identity. By induction, we may replace $[d_1, e]$ and $[d_2, e]$ by basic commutators f_1 and f_2 which consist of products of the same number of elements of S as $[d_1, e]$ and $[d_2, e]$ respectively. It is possible that $d_1 > f_2$ (since it is possible that $d_1 > [d_2, e]$). We now prove the result for $[f_1, d_2]$ and $[f_2, d_1]$ (or $[d_1, f_2]$ if $d_1 > f_2$). However, in order to do this, we may repeat the above procedure. Since we are always in the same multihomogeneous component of L, which is finite-dimensional, and since at each stage of this process we replace a product $[u, v]$ with a linear combination of products $[u_i, v_i]$ where $v_i > v$ for each i, this process cannot continue indefinitely. Therefore eventually we must be left with basic commutators, each of which

will be a product of r elements from S, and therefore is clearly contained in $(L^i)^j$. This completes the induction and also the proof of the theorem. ■

The following lemmas would allow us to prove that $F/[\gamma_j(F), \gamma_i(F)]$ is torsion-free and residually nilpotent for $1 < i \leq j \leq 2i+1$. However, this has already been proved for all i and j by Tasić and Vaughan-Lee [43]. The result below and its proof are quite similar to one proved by Gorchakov [5], though we prove the result in Lie rings while Gorchakov proves it in groups.

Lemma 4.4.4 *Let L be a free Lie ring and suppose $1 < i \leq j \leq 2i$. If $x \in [L^j, L^i]$, then x can be expressed as a linear combination of commutators $[u, v]$ where (i) $u \in L^j$ and $v \in L^i$ and (ii) $[u, v]$ is a basic commutator.*

Proof: By Lemma 4.4.1 and linearity, it suffices to show that if $a \in L^j$ and $b \in L^i$ are basic commutators, then $[a, b]$ can be expressed as a linear combination of basic commutators of the required type.

The case where $i = j$ is covered by Theorem 4.4.3. So suppose that $j > i$. If $\mathrm{wt}(b) \geq j$, then $[a, b] \in [L^j, L^j]$ and so, by Theorem 4.4.3, can be written as a linear combination of basic commutators, each of which lies in $[L^j, L^j]$, a subset of $[L^j, L^i]$. Thus, we can assume that $\mathrm{wt}(b) < j$. Therefore, $a > b$.

If $[a, b]$ is not a basic commutator, then $a = [a_1, a_2]$ where $a_1 > a_2 > b$. Therefore $[a, b] \in [L^i, L^i, L^i]$ and by Theorem 4.4.3 can be written as a linear combination of basic commutators, each of which lies in $[L^i, L^i, L^i]$. Since $j \leq 2i$, $[L^i, L^i, L^i] \subseteq [L^j, L^i]$. This completes the proof of the lemma. ■

Lemma 4.4.5 *Suppose that $x \in [L^{2i+1}, L^i]$. Then x can be expressed as a linear combination of basic commutators, each of which lies in $[L^{2i+1}, L^i]$.*

Proof: We need only consider the case $x = [a, b]$, where a and b are basic commutators, $a > b$, $\mathrm{wt}(a) \geq 2i+1$ and $\mathrm{wt}(b) \geq i$.

Suppose that $[a, b]$ is not a basic commutator. Then $a = [a_1, a_2]$ where $a_1 > a_2 > b$. Then

$$[a_1, a_2, b] = [a_1, [a_2, b]] + [a_1, b, a_2].$$

We are always working in the same multihomogeneous component of L, which is finite-dimensional. Therefore, by reverse induction, $[a_1, b, a_2]$ can be expressed as a linear combination of basic commutators, each of which lies in $[L^{2i+1}, L^i]$. This is because $\mathrm{wt}(a_1) \geq i+1$, so $[a_1, b] \in L^{2i+1}$ and $a_2 > b$. Therefore, we need only prove that $[a_1, [a_2, b]]$ can be expressed as a linear combination of basic commutators, each of which lies in $[L^{2i+1}, L^i]$.

First suppose that $[a_2, b] \in L^{2i+1}$. Then again we can prove by reverse induction that $[a_2, b, a_1]$ may be expressed as a linear combination of basic commutators, each of which lies in $[L^{2i+1}, L^i]$, since $a_1 > b$.

Therefore, suppose that $[a_2, b] \notin L^{2i+1}$. Then $\mathrm{wt}(a_2) = \mathrm{wt}(b) = i$, and since $a_2 > b$, $[a_2, b]$ is a basic commutator. Suppose that $[a_2, b] > a_1$. Then $[a_2, b, a_1]$ is a basic commutator, since $a_1 > b$.

Lastly, suppose that $a_1 > [a_2, b]$. Then let $a_1 = [a_3, a_4]$. Well, since $[a_1, a_2]$ is a basic commutator, $a_2 \geq a_4$. Thus $[a_2, b] > a_4$ and so $[a_1, [a_2, b]]$ is a basic commutator.

In either case, we can see that either $[a_1, [a_2, b]]$ is a basic commutator, or $[a_2, b, a_1]$ is a basic commutator. Now

$$[a_2, b, a_1] = [a_1, b, a_2] - [a_1, a_2, b],$$

so it is clear that $[a_2, b, a_1] \in [L^{2i+1}, L^i]$, completing the proof of the lemma. ■

We build on Lemma 4.4.4 to prove the following result.

Theorem 4.4.6 *Let L be a free Lie ring. Suppose $1 < i \leq j \leq 2i$ and $j \leq k \leq i + j$. If $x \in [L^j, L^i, L^k]$ then x can be expressed as a sum of commutators of the form $[u, v, w]$ where (i) $u \in L^j$, $v \in L^i$ and $w \in L^k$ and (ii) Either $[u, v, w]$ or $[w, [u, v]]$ is a basic commutator*

Proof: By Lemmas 4.4.4 and 4.4.1 and linearity, it suffices to show the result for commutators of the form $[a_1, a_2, b]$ where $a_1 \in L^j, a_2 \in L^i$ and $b \in L^k$ are all basic, and $[a_1, a_2]$ is also basic. Let $a = [a_1, a_2]$.

Suppose that $b > a$. If $[b, a]$ is not basic then $b = [b_1, b_2]$ where $b_1 > b_2 > a$. Thus $[b, a] \in [L^k, L^k, L^k]$ because $a \in L^k$. Therefore, by Theorem 4.4.3, $[b, a]$ can be written as a linear combination of basic commutators which are in $[L^k, L^k, L^k]$.

Now suppose that $a > b$. If $[a, b]$ is not a basic commutator, then $a = [a_1, a_2]$ where $a_1 > a_2 > b$. Thus $[a, b] \in [L^k, L^k, L^k]$ and so can, by Theorem 4.4.3, be written as a linear combination of basic commutators each of which lies in $[L^k, L^k, L^k]$. This completes the proof of the theorem. ■

We use a different idea to extend the result to $k = i + j + 1$.

Theorem 4.4.7 *Let L be a free Lie ring. Suppose that $1 < i \leq j \leq 2i$. If $x \in [L^j, L^i, L^{i+j+1}]$ then x can be written as a linear combination of commutators*

of the form $[a,b]$ *where* (i) $a=[a_1,a_2]$ *and* $a_1 \in L^j$, $a_2 \in L^i$, (ii) $b \in L^{i+j+1}$ *and* (iii) *either* $[a,b]$ *or* $[b,a]$ *is a basic commutator.*

Proof: First, suppose that when the basic commutators of weight $i+j$ were ordered, we ensured that if $[a,b]$ and $[c,d]$ are basic commutators of weight $i+j$ such that $\mathrm{wt}(a)=j$ and $\mathrm{wt}(c)\neq j$ then $[a,b]>[c,d]$. Since the ordering of basic commutators of weight $i+j$ is arbitrary, this is certainly possible.

By Lemmas 4.4.4 and 4.4.1 and linearity, it is sufficient to prove the result when $x=[u,v]$ where u and v are basic commutators, $u=[u_1,u_2]$, $u_1 \in L^j$, $u_2 \in L^i$ and $v \in L^{i+j+1}$.

Suppose that $u>v$. If $[u,v]$ is not a basic commutator, then $u=[u_1,u_2]$ where $u_1>u_2>v$. Then $[u,v] \in [L^{i+j+1},L^{i+j+1},L^{i+j+1}]$ and so by Theorem 4.4.6, we know that $[u,v]$ is a linear combination of basic commutators of the form $[[r,s],t]$ where $r,s,t \in L^{i+j+1}$. Now, $[[r,s],t]$ satisfies (i), (ii) and (iii) from the statement of the theorem. Thus, we may assume that $v>u$.

If $[v,u]$ is not basic then $v=[v_1,v_2]$ and $v_1>v_2>u$. Because of the way we ordered the basic commutators of weight $i+j$, we have either $v_1 \in [L^j,L^i]$ or $v_1 \in L^{i+j+1}$ and either $v_2 \in [L^j,L^i]$ or $v_2 \in L^{i+j+1}$. Now,

$$\begin{aligned}[v,u] &= [v_1,v_2,u] \\ &= [v_1,u,v_2]+[u,v_2,v_1]\end{aligned}$$

If $v_1 \in [L^j,L^i]$, then since $[u,v_2] \in L^{i+j+1}$ we can write $[u,v_2]$ as a linear combination of basic commutators $c \in L^{i+j+1}$. Then $[v_1,c]$ satisfies (i) and (ii) from the statement of the theorem. If $v_1 \notin [L^j,L^i]$, then $v_1 \in L^{i+j+1}$, and $[u,v_2] \in [L^j,L^i]$ so we can write $[u,v_2]$ as a linear combination of basic commutators $[r,s]$ with $r \in L^j$ and $s \in L^i$. Then $[r,s,v_1]$ satisfies (i) and (ii) from the statement of the theorem. Similarly, $[v_1,u,v_2]$ can be written as a linear combination of elements $[v_2,c']$ or $[r',s',v_2]$ which satisfy (i) and (ii). Thus, we have represented $[v,u]$ as a linear combination of elements satisfying (i) and (ii). Since $v_1,v_2,c,c',[r,s],[r',s']>u$, we can repeat this process and eventually yield commutators satisfying (i), (ii) and (iii) (it must stop because there are only finitely many basic commutators in the multihomogeneous component in which everything takes place). This completes the proof of the theorem. ■

Corollary 4.4.8 *Let* L *be a free Lie ring,* $1<i\leq j\leq 2i$ *and* $j\leq k\leq i+j+1$. *Then the Lie ring* $L/[L^j,L^i,L^k]$ *is torsion-free.*

Proof: The set of all basic commutators form a basis for L as a free $\mathbb{Z}$-module. We proved in Theorems 4.4.6 and 4.4.7 that $[L^j, L^i, L^k]$ has a set of basic commutators as a spanning set. Thus, the images of the basic commutators which are not in this spanning set form a basis for $L/[L^j, L^i, L^k]$ as a free $\mathbb{Z}$-module. Thus $L/[L^j, L^i, L^k]$ must be torsion-free. ■

Theorem 4.4.9 *Suppose* $1 < i < j \leq 2i$. *If* $x \in [L^j, L^i, L^i]$ *then* x *can be expressed as a linear combination of basic commutators, each of which lies in* $[L^j, L^i, L^i]$.

Proof: By Lemma 4.4.4, Lemma 4.4.1 and linearity, we need only consider the case where $x = [a, b, c]$, where $[a, b]$ is a basic commutator, c is a basic commutator, $\mathrm{wt}(a) \geq j$ and $\mathrm{wt}(b), \mathrm{wt}(c) \geq i$. Suppose that $[a, b, c]$ is not a basic commutator. Then $c < b$ or $[a, b] < c$.

Suppose that $[a, b] < c$ and $[c, [a, b]]$ is not a basic commutator. In this case $c = [c_1, c_2]$ where $c_1 > c_2 > [a, b]$. Then $\mathrm{wt}(c_1) \geq \mathrm{wt}(c_2) \geq \mathrm{wt}([a, b]) \geq i + j$, so $[c_1, c_2, [a, b]] \in (L^{i+j})^3$ and we may apply Theorem 4.4.3.

Suppose then that $c < b$. Then

$$[a, b, c] = [a, [b, c]] + [a, c, b].$$

We first prove that $[a, [b, c]]$ can be written as a linear combination of basic commutators, each of which lies in $[L^j, L^i, L^i]$.

Suppose that either $[a, [b, c]]$ or $[b, c, a]$ is a basic commutator. Then since $[a, [b, c]] = [a, b, c] - [a, c, b]$, we have $[a, [b, c]] \in [L^j, L^i, L^i]$.

Thus suppose that neither $[a, [b, c]]$ nor $[b, c, a]$ is a basic commutator. If $\mathrm{wt}(c) \geq j$ then $[a, b, c] \in (L^j)^3$ and we can use Theorem 4.4.3. Therefore we may assume that $\mathrm{wt}(c) < j$. Then $\mathrm{wt}(a) > \frac{1}{2}(\mathrm{wt}(b) + \mathrm{wt}(c))$, since $a > b$.

Suppose that $[b, c]$ is not a basic commutator. Then we have $b = [b_1, b_2]$ where $b_1 > b_2 > c$. Therefore $[b, c] \in (L^i)^3 \subseteq [L^j, L^i]$. Thus, by Lemma 4.4.4, we may write $[b, c] = \sum_{s \in S} \alpha_s [d_s, e_s]$ where $\alpha_s \in \mathbb{Z}$, $[d_s, e_s]$ is a basic commutator, $\mathrm{wt}(d_s) \geq j$ and $\mathrm{wt}(e_s) \geq i$. We prove that for any $s \in S$ we can write $[d_s, e_s, a]$ as a linear combination of basic commutators, each of which lies in $[L^j, L^i, L^i]$.

Suppose first that $a > [d_s, e_s]$. If $[a, [d_s, e_s]]$ is not a basic commutator, then $a = [a_1, a_2]$ with $a_1 > a_2 > [d_s, e_s]$. Then $[a_1, a_2, [d_s, e_s]] \in (L^{i+j})^3$ and we have Theorem 4.4.3.

Thus suppose that $[d_s, e_s] > a$. Then $\mathrm{wt}(e_s) \leq \frac{1}{2}(\mathrm{wt}(d_s) + \mathrm{wt}(e_s)) = \frac{1}{2}(\mathrm{wt}(b) + \mathrm{wt}(c)) < \mathrm{wt}(a)$. Therefore $a > e_s$, so $[d_s, e_s, a]$ is a basic commutator.

Now suppose that $[b,c]$ is a basic commutator. If $a > [b,c]$ then $[a,[b,c]]$ is a basic commutator. To see this, note that $[a,b]$ is a basic commutator, so if $a = [a_1,a_2]$ then $b \geq a_2$. Therefore $[b,c] > a_2$. Alternatively, suppose that $a < [b,c]$. Then clearly $[b,c,a]$ is a basic commutator (since $a > b > c$). Therefore $[a,[b,c]]$ can be expressed as a linear combination of basic commutators, each of which lies in $[L^j,L^i,L^i]$.

Now we consider $[a,c,b]$. Suppose that $[a,c,b]$ is not a basic commutator. Then since $b > c$, we must have that $[a,c]$ is not basic. Well $[a,c] \in [L^j,L^i]$ so by Lemma 4.4.4 we may express $[a,c]$ as a linear combination of basic commutators, each of which lies in $[L^j,L^i]$. Let $[a,c] = \sum_{r\in R} \beta_r [p_r,q_r]$ where $\beta_r \in \mathbb{Z}$, $[p_r,q_r]$ is a basic commutator, $\mathrm{wt}(p_r) \geq j$, $\mathrm{wt}(q_r) \geq i$ and R is some finite set. Then we prove that $[p_r,q_r,b]$ can be expressed as a linear combination of basic commutators, each of which lies in $[L^j,L^i,L^i]$. Now $a > b$ and $\mathrm{wt}([p_r,q_r]) = \mathrm{wt}([a,c]) > \mathrm{wt}(a)$, so $[p_r,q_r] > b$. Suppose that $[p_r,q_r,b]$ is not a basic commutator. Then $q_r > b$.

Thus we may write

$$[p_r,q_r,b] = [p_r,[q_r,b]] + [p_r,b,q_r].$$

By reverse induction we can see that $[p_r,b,q_r]$ can be written as a linear combination of basic commutators, each of which lies in $[L^j,L^i,L^i]$. The same proof that proved that $[a,[b,c]]$ can be written as a linear combination of basic commutators which all lie in $[L^j,L^i,L^i]$ proves the result for $[p_r,[q_r,b]]$. This completes the proof of the theorem. ■

Corollary 4.4.10 *Suppose that* $1 < i \leq k \leq j \leq 2i$. *If* $x \in [L^j,L^i,L^k]$ *then* x *can be expressed as a linear combination of basic commutators, each of which lies in* $[L^j,L^i,L^k]$.

Proof: As usual, we may assume that $x = [a,b,c]$ where $[a,b]$ and c are basic commutators, $\mathrm{wt}(a) \geq j$, $\mathrm{wt}(b) \geq i$ and $\mathrm{wt}(c) \geq k$.

Suppose that $[a,b,c]$ is not a basic commutator. Then either $c > [a,b]$ or $c < b$.

If $c > [a,b]$ and $[c,[a,b]]$ is not a basic commutator, then $c = [c_1,c_2]$ and $c_1 > c_2 > [a,b]$. Then $[c_1,c_2,[a,b]] \in (L^{i+j})^3$ and so can be expressed as a linear combination of basic commutators, each of which lies in $(L^{i+j})^3$, by Theorem 4.4.3. Since $(L^{i+j})^3 \subseteq [L^j,L^i,L^k]$, we are done.

So suppose that $c < b$. Then $b \in L^k$, so $[a,b,c] \in [L^j,L^k,L^k]$ and we may apply Theorem 4.4.9. This completes the proof of the corollary. ■

The same proof that gave us Corollary 4.4.8 yields the following result.

Corollary 4.4.11 *Suppose that* $1 < i \leq k \leq j \leq 2i$. *Then the Lie ring* $L/[L^j, L^i, L^k]$ *is torsion-free.*

We also use similar ideas as above to prove the following result.

Theorem 4.4.12 *Let* L *be a free Lie ring. Suppose that* $1 < i \leq j \leq 2i$, $j \leq k \leq l \leq i + j$. *Then any element of* $[L^j, L^i, L^k, L^l]$ *can we written as a linear combination of basic commutators, each of which is an element of* $[L^j, L^i, L^k, L^l]$.

Proof: By Theorem 4.4.6 and linearity, we need only prove the result for elements of L of the form $[a, b, c, d]$ where a, b, c and d are basic commutators, $a \in L^j$, $b \in L^i$, $c \in L^k$, $d \in L^l$ and where $[a, b, c]$ is a basic commutator and elements of the form $[c, [a, b], d]$ where a, b, c, d are basic commutators, where $a \in L^j$, $b \in L^i$, $c \in L^k$ and $d \in L^l$ and where $[c, [a, b]]$ is a basic commutator.

We first consider $[a, b, c, d]$. Suppose first that $[a, b, c] > d$ but that $[a, b, c, d]$ is not a basic commutator. Then $d < c$. Now,

$$[a, b, c, d] = [a, b, d, c] - [c, d, [a, b]].$$

Since $d < c$, we have $c \in L^l$ and thus

$$[a, b, d, c] \in [L^j, L^i, L^l, L^l] \subseteq [L^j, L^i, L^k, L^l].$$

Also,

$$-[c, d, [a, b]] = [a, b, c, d] - [a, b, d, c],$$

and so $-[c, d, [a, b]] \in [L^j, L^i, L^k, L^l]$. If $[c, d]$ is not a basic commutator, then $c = [c_1, c_2]$ where $c_1 > c_2 > d$. Then $[c, d, [a, b]] \in [L^l, L^l, L^l, L^l]$ (since $l \leq i+j$) and so can be written as a linear combination of basic commutators which lie in $[L^l, L^l, L^l, L^l]$ (by Theorem 4.4.3). Thus we may assume that $[c, d]$ is a basic commutator. Suppose that $[c, d] > [a, b]$. Then since $d < c < [a, b]$, $[c, d, [a, b]]$ is a basic commutator. Suppose then that $[a, b] > [c, d]$. Then it is clear that $[c, d] > b$ (since $c \geq b$) and so $[a, b, [c, d]]$ is a basic commutator. Thus we now need to express $[a, b, d, c]$ as a linear combination of basic commutators which are in $[L^j, L^i, L^k, L^l]$.

Suppose then that $[a, b, d, c]$ is not a basic commutator. Then $[a, b, d, c] \in [L^l, L^l, L^l, L^l]$ and so we can express $[a, b, d, c]$ as a linear combination of basic commutators in $[L^l, L^l, L^l, L^l]$ (by Theorem 4.4.3), which is clearly sufficient.

Now suppose that $d > [a,b,c]$. If $[d,[a,b,c]]$ is not a basic commutator, then $d = [d_1, d_2]$ where $d_1 > d_2 > [a,b,c]$. Then

$$\begin{aligned}[d,[a,b,c]] &= [d_1,d_2,[a,b,c]] \\ &= [d_1,[a,b,c],d_2] - [d_2,[a,b,c],d_1].\end{aligned}$$

Note that $[d_1,[a,b,c],d_2], [d_2,[a,b,c],d_1] \in [L^j, L^i, L^k, L^l]$, so if they are basic then we are done. Thus, suppose that $[d_1,[a,b,c],d_2]$ is not basic. This means that $[d_1,[a,b,c]]$ is not basic. Thus $d_1 = [d_3, d_4]$ where $d_3 > d_4 > [a,b,c]$. Then

$$[d_1,[a,b,c],d_2] = [d_3,d_4,[a,b,c],d_2] \in [L^l, L^l, L^l, L^l],$$

(since $l < i+j+k$) which means that $[d_1,[a,b,c],d_2]$ can be written as a linear combination of basic commutators which lie in $[L^l, L^l, L^l, L^l]$ (by Theorem 4.4.3). Similar considerations for $[d_2,[a,b,c],d_1]$ show that $[d,[a,b,c]]$ can be written as a linear combination of basic commutators which lie in $[L^j, L^i, L^k, L^l]$.

Now consider the case where $[c,[a,b]]$ is a basic commutator. Suppose that $[c,[a,b]] > d$ and that $[c,[a,b],d]$ is not a basic commutator. Then $d < [a,b] < c$. Now

$$[c,[a,b],d] = [c,d,[a,b]] - [a,b,d,c].$$

Also, $c > [a,b]$ so $[c,d] > [a,b]$. Now the same considerations as above show that $[c,[a,b],d]$ can be written as a linear combination of basic commutators which all lie in $[L^j, L^i, L^k, L^l]$.

Finally, it remains to consider the case where $d > [c,[a,b]]$. The proof that we can express $[d,[c,[a,b]]]$ as a linear combination of basic commutators which all lie in $[L^j, L^i, L^k, L^l]$ is similar to the case $d > [a,b,c]$ above. This completes the proof of the theorem. ■

Again, the same proof which gave us Corollary 4.4.8 yields

Corollary 4.4.13 *Let L be a free Lie ring, $1 < i \leq j \leq 2i$ and $j \leq k \leq l \leq i+j$. Then the Lie ring $L/[L^j, L^i, L^k, L^l]$ is torsion-free.*

4.5 Limitations of methods

In this section we show why our proof that $L/[L^j, L^i, L^k]$ is torsion-free for the case $1 < i \leq j \leq k$ works only when $k \leq i+j+1$, by considering the Lie

ring $L/[L^2, L^2, L^6]$, where L is the free Lie ring on 3 generators, which we will denote by x, y and z. We show that there exists an element of $L/[L^2, L^2, L^6]$ such that when it is expressed as a sum of basic commutators these basic commutators are not all in $L/[L^2, L^2, L^6]$.

Let L be freely generated by x, y and z. Let

$$c_1 = [z, y, y, y, y],$$

$$c_2 = [z, x, y, y, y], \text{ and}$$

$$c_3 = [z, y, [y, x]],$$

and consider the commutator $a = [c_3, [c_1, c_2]]$. Clearly $a \in [L^2, L^2, L^6]$, and we express a as a sum of basic commutators.

$$\begin{aligned} a &= [z, y, [y, x], [[z, y, y, y, y], [z, x, y, y, y]]] \\ &= -[z, y, y, y, y, [z, x, y, y, y], [z, y, [y, x]]] \\ &= [z, x, y, y, y, [z, y, [y, x]], [z, y, y, y, y]] - \\ &\quad [z, y, y, y, y, [z, y, [y, x]], [z, x, y, y, y]]. \end{aligned}$$

We prove that $[z, x, y, y, y, [z, y, [y, x]], [z, y, y, y, y]] \notin [L^2, L^2, L^6]$. We do this using the 'lienq' program developed in Canberra. We use modifications of routines written by M.R. Vaughan-Lee to enforce the identity

$$[y_1, y_2, [y_3, y_4], [y_5, y_6, y_7, y_8, y_9, y_{10}]] = 0. \tag{4.5}$$

We calculate the weight 14, $(2, 9, 3)$-multiweight quotient of L at the prime $p = 2$ while enforcing the identity (4.5).

This gives us a Lie algebra of order 2^{4547}. In this Lie algebra, we calculate the value of $[z, x, y, y, y, [z, y, [y, x]], [z, y, y, y, y]]$ and find that it is not zero. This is enough to show that $[z, x, y, y, y, [z, y, [y, x]], [z, y, y, y, y]] \notin [L^2, L^2, L^6]$. Thus, our method of proof cannot be applied to the Lie ring $L/[L^2, L^2, L^6]$.

We now show that Lemma 4.4.5 is best possible, in a sense. Tasić and Vaughan-Lee [43] proved that $L/[L^j, L^i]$ is torsion-free for all i and j, so we cannot expect to find that $L/[L^{2i+2}, L^i]$ contains torsion. However, using the same method as above, and enforcing the identity

$$[y_1, y_2, y_3, y_4, y_5, y_6, [y_7, y_8]] = 0$$

in L, we calculate the value of $[z,x,y,[z,y],[y,x,x]]$ and find that it is not zero. Since

$$[z,x,y,[y,x,x],[z,y]] = [z,x,y,[z,y],[y,x,x]] - [y,x,x,[z,y],[z,x,y]],$$

we have an element of $[L^6,L^2]$ which when expressed as a linear combination of basic commutators, has at least one of these basic commutators not contained in $[L^6,L^2]$. Therefore, Lemma 4.4.5 is best possible.

Obviously, since $L/[L^j,L^i]$ is torsion-free for all i and j, the fact that our method fails does not indicate the existence of torsion. Thus we can say nothing about the existence of torsion in $L/[L^j,L^i,L^k]$ when $k > i+j+1$.

4.6 Occurrences of torsion

In this section we describe some occurrences of torsion in outer commutator Lie rings, and conjecture the existence of torsion in a general class of outer commutator Lie rings.

Theorem 4.6.1 *Let L be a free Lie ring. Let $1 \leq k < i \leq j$. Let s be an integer such that $k \leq s$ and $2s \geq i$. Suppose that x and y are elements of L of weight s and z is an element of weight j. Then $2[[z,y],[z,x]] \in [L^j,L^i,L^k]$.*

Proof:

$$\begin{aligned} 2[[z,y],[z,x]] &= ([z,y,z,x]-[z,y,x,z])+([z,x,y,z]-[z,x,z,y]) \\ &= ([z,y,z,x]-[z,x,y,z]-[z,[y,x],z]) \\ &\quad +([z,x,y,z]-[z,x,z,y]) \\ &= [z,y,z,x]-[z,x,z,y]-[z,[y,x],z]. \end{aligned}$$

Now, $[z,y,z,x],[z,x,z,y] \in [L^j,L^s,L^j,L^s] \subseteq [L^j,L^i,L^k]$, and $[z,[y,x],z] \in [L^j,[L^s,L^s],L^j] \subseteq [L^j,L^i,L^k]$. This completes the proof of the theorem. ■

We can use this theorem to come up with many likely candidates for torsion elements in different outer commutator Lie rings. We then explicitly calculate some nilpotent quotient of these Lie rings and check whether the candidates are in fact torsion elements. We will see below that in many cases they are, but that in some they are not. However, in the cases when we do not have an explicit candidate for torsion, we can deduce the existence of torsion anyway.

Lie ring	Torsion Element
$L/[L^2, L^2, L]$	$[4,1,3,[4,1,2]]$
$L/[L^3, L^2, L]$	$[4,1,1,3,[4,1,1,2]]$
$L/[L^3, L^3, L^2]$	$[4,1,1,[3,1],[4,1,1,[2,1]]]$
$L/[L^4, L^2, L]$	$[4,1,1,1,3,[4,1,1,1,2]]$
$L/[L^4, L^3, L^2]$ $L/[L^4, L^4, L]$ $L/[L^4, L^4, L^2]$	$[4,1,1,1,[3,1],[4,1,1,1,[2,1]]]$
$L/[L^4, L^4, L^3]$	$[4,1,1,1,[3,1,1],[4,1,1,1,[2,1,1]]]$
$L/[L^5, L^2, L]$	$[4,1,1,1,1,3,[4,1,1,1,1,2]]$
$L/[L^5, L^3, L^2]$ $L/[L^5, L^4, L^2]$	$[4,1,1,1,1,[3,1],[4,1,1,1,1,[2,1]]]$
$L/[L^5, L^4, L^3]$ $L/[L^5, L^5, L^3]$	$[4,1,1,1,1,[3,1,1],[4,1,1,1,1,[2,1,1]]]$
$L/[L^5, L^5, L^4]$	$[3,1,1,1,1,[3,1,1,1],[3,1,1,1,1,[2,1,1,1]]]$

Table 4.1: Some explicit occurrences of torsion

Theorem 4.6.2 *Suppose that L is a free Lie ring of rank at least 4. If $1 \leq k < i \leq j \leq 5$ and $i \leq 2k$ then there is torsion in $L/[L^j, L^i, L^k]$.*

Proof: First we calculate a number of examples based upon Theorem 4.6.1. Table 4.1 lists a number of examples. In Table 4.1 we use the convention of writing 'i' instead of 'z_i' in order to save space. For each of the Lie rings listed in Table 4.1, an element which satisfies the hypotheses of Theorem 4.6.1 is shown. The nilpotent quotient algorithm was then used to show that this element is non-trivial in the appropriate outer commutator Lie ring (the method for doing this is the same as that described in Section 4.5). Thus, the element listed is an example of torsion in the appropriate Lie ring.

Table 4.2 lists the remaining Lie rings $L/[L^j, L^i, L^k]$ where $i, j, k \leq 5$ and $k < i, j$. In each of these, the obvious element from Theorem 4.6.1 was trivial, and so did not give an example of torsion. However, by considering multiweight components in Lie algebras over a field of characteristic 2 and then of characteristic 3, we are able to deduce the existence of torsion. If we use the Lie nilpotent quotient algorithm in multiweight $\underline{w} = (w_1, \ldots, w_n)$ at the prime p, we are effectively calculating the multiweight component of $J \otimes_{\mathbb{Z}} \mathbb{F}_p$ where J is the Lie ring in question: $L/[L^j, L^i, L^k]$. Thus if x is a

Lie ring	Multiweight	Dimension at $p = 2$	Dimension at $p = 3$
$L/[L^3, L^3, L]$	$(4, 1, 1, 2)$	317	316
$L/[L^4, L^3, L]$	$(6, 1, 1, 2)$	580	575
$L/[L^5, L^3, L]$	$(8, 1, 1, 2)$	967	963
$L/[L^5, L^4, L]$	$(8, 1, 1, 2)$	1521	1515
$L/[L^5, L^5, L]$	$(10, 1, 1, 2)$	3129	3125
$L/[L^5, L^5, L^2]$	$(12, 1, 1, 2)$	6386	6381

Table 4.2: Other occurrences of torsion

nonzero element of J that is zero at the prime $p = 2$ but not at the prime $p = 3$, then 2 divides the order of x and 3 does not. Thus, if the same calculation at $p = 2$ gives a larger Lie algebra than at $p = 3$ this indicates 2-torsion in the Lie ring. Table 4.2 gives a summary of the calculations performed, and finishes the proof of the theorem.

■

Note that the existence of 2-torsion in $[(L^i)^2, L]$ for $i = 2, 3, 4, 5$ above seems to concur with the results of Stöhr [42]. Stöhr proves a number of results about groups of the form $F/[\gamma_c(R), F]$, where $R \trianglelefteq F$. He proves, among other things, that the order of torsion elements in these groups must divide c (unless $c = 2$ when it must divide 4). We are considering the case when $R = \gamma_i(F)$, and $c = 2$. The fact that $F/\gamma_i(F)$ is torsion-free allows us to quote another of Stöhr's results to say that the order of torsion elements in $F/[\gamma_i(F)', F]$ must be 2.

The occurrences of torsion in this section do not immediately furnish us with examples of occurrences of torsion in outer commutator groups. They seem to rely heavily on the Jacobi identity, which is not exact in groups. Choosing the obvious elements in the analogous outer commutator groups and trying to prove that they are torsion does not seem to work.

We finish this section with a conjecture.

Conjecture 1 *If k, i and j are integers such that $1 < k \leq i \leq j$, then the Lie ring $L/[L^j, L^i, L^k]$ is not torsion-free.*

It is difficult to see how one might prove this conjecture. It involves finding an element $x_{i,j,k}$ for each value of i, j and k which is a torsion element. It seems relatively easy to find elements for which multiples of them are trivial (as in Theorem 4.6.1), but to show that the element itself is nontrivial involves either a case by case analysis, or some new method.

4.7 Kuz'min and Shapiro's result

Suppose that $S = \{\varphi_i\}_{i\in I}$ is a set of commutator-subgroup functions. Let $\mathfrak{V}$ be the variety of all groups G such that $\varphi_i(G) = 1$ for all $i \in I$. Let $F(\mathfrak{V})$ be a free group of $\mathfrak{V}$, and $\mathfrak{L}(F(\mathfrak{V}))$ its associated Lie ring.

Let $T = \{\psi_i\}_{i\in I}$ be the set of commutator subring functions which correspond to those in S under the correspondence mentioned in Subsection 4.2. Let $\mathfrak{V}_L$ be the variety of all Lie rings J such that $\psi_i(J) = 0$ for all $i \in I$, and let $L(\mathfrak{V}_L)$ be the free Lie ring of $\mathfrak{V}_L$, of the same rank as $F(\mathfrak{V})$.

We have the following theorem due to Kuz'min and Shapiro.

Theorem 4.7.1 (Kuz'min and Shapiro, [26]) *Suppose that the Lie ring $L(\mathfrak{V}_L)$ is torsion-free. Then we have an isomorphism $L(\mathfrak{V}_L) \cong \mathfrak{L}(F(\mathfrak{V}))$.*

By considering the proof of this theorem, we see that this isomorphism is graded, takes a natural generating set for $L(\mathfrak{V}_L)$ to a natural generating set for $\mathfrak{L}(F(\mathfrak{V}))$, and thus maps basic commutators in $L(\mathfrak{V}_L)$ to the corresponding basic commutators in the homogeneous components of $\mathfrak{L}(F(\mathfrak{V}))$.

Importantly, given Corollary 4.4.8 and Theorem 4.7.1, the associated Lie ring of $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ (where the limitations on i, j and k are given by the results proved in section 4.4) is torsion-free, and the statements in the proofs of Corollarys 4.4.8 and 4.4.11 also hold for these Lie rings. The same is true for $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F), \gamma_l(F)]$ and Corollary 4.4.13.

It is worth noting here that Kuz'min and Shapiro also prove that we always have $L(\mathfrak{V}_L)\otimes\mathbb{Q} \cong \mathfrak{L}(F(\mathfrak{V}))\otimes\mathbb{Q}$. However, this ignores any occurrences of torsion, and does not seem of any use to us here.

4.8 Residual nilpotence

In this section we prove that the groups under consideration are residually nilpotent.

Theorem 4.8.1 *Let F be a free group. Suppose $1 < i \leq j \leq 2i$ and $j \leq k \leq i + j + 1$. Then $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ is residually nilpotent.*

Proof: The terms of the lower central series of $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ are

$$\gamma_n(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)] \Big/ [\gamma_j(F), \gamma_i(F), \gamma_k(F)].$$

Thus, we need to show that if

$$x \in \bigcap_{n=1}^{\infty} \Big(\gamma_n(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)]\Big),$$

then $x \in [\gamma_j(F), \gamma_i(F), \gamma_k(F)]$. So suppose that

$$x \in \bigcap_{n=1}^{\infty} \Big(\gamma_n(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)]\Big).$$

Note that, for some m, $F^{(m)} \subseteq [\gamma_j(F), \gamma_i(F), \gamma_k(F)]$. Thus by Theorem 4.3.1, we can write

$$x = c_1 \cdots c_r d_t^{-1} \cdots d_1^{-1} g_x, \tag{4.6}$$

where $c_1, \ldots, c_r$, $d_1, \ldots, d_t$ are basic commutators, where $c_1 \leq \cdots \leq c_k$ and $d_1 \leq \cdots \leq d_t$, and where $g_x \in [\gamma_j(F), \gamma_i(F), \gamma_k(F)]$. We assume, without loss of generality, that none of $c_1, \ldots, c_r, d_1 \ldots, d_t$ are contained in $[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$.

Let s be the least weight of any of $c_1, \ldots, c_r$, $d_1, \ldots, d_t$, and suppose that $c_1, \ldots, c_p, d_1, \ldots, d_q$ are the basic commutators which are of weight exactly s. Then,

$$x \equiv c_1 \cdots c_p d_q^{-1} \cdots d_1^{-1} \bmod \gamma_{s+1}(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)].$$

Therefore, since $x \in \gamma_{s+1}(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$,

$$c_1 \cdots c_p d_q^{-1} \cdots d_1^{-1} \equiv 1 \bmod \gamma_{s+1}(F)[\gamma_j(F), \gamma_k(F), \gamma_k(F)].$$

Now, given Theorems 4.4.6, 4.4.7 and 4.7.1 and the discussion in Section 4.7, we can see that the commutators of weight s above are independent modulo $\gamma_{s+1}(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ (recalling that we assumed that none of the basic commutators in the expression (4.6) for x lie in $[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$). Therefore, we must have that $p = q$ and $c_a = d_a$ for $1 \leq a \leq p$. We now have

$$x = (c_{p+1} \cdots c_r d_t^{-1} \cdots d_{p+1}^{-1})^{(c_1 \cdots c_p)^{-1}} g_x.$$

Since $\bigcap_{n=1}^{\infty} \Big(\gamma_n(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)]\Big) \trianglelefteq F$,

$$c_{p+1} \cdots c_r d_t^{-1} \cdots d_{p+1}^{-1} \in \bigcap_{n=1}^{\infty} \Big(\gamma_n(F)[\gamma_j(F), \gamma_i(F), \gamma_k(F)]\Big).$$

Thus, we can now repeat this argument on $c_{p+1} \cdots c_r d_t^{-1} \cdots d_{p+1}^{-1}$, where the lowest weight commutators are at least of weight $s+1$. Since the original expression for x is finite, we eventually see that

$$x = c_1 \cdots c_r c_r^{-1} \cdots c_1^{-1} g_x,$$

which shows that

$$x \in [\gamma_j(F), \gamma_i(F), \gamma_k(F)].$$

This completes the proof of the theorem. ■

The same proof yields the following theorems.

Theorem 4.8.2 *Let F be a free group. Suppose $1 < i \leq k \leq j \leq 2i$. Then $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ is residually nilpotent.* ■

Theorem 4.8.3 *Let F be a free group. Suppose $1 < i \leq j \leq 2i$ and $j \leq k \leq l \leq i+j$. Then $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F), \gamma_l(F)]$ is residually nilpotent.* ■

4.9 The groups are torsion-free

Finally, we prove that the groups under consideration are torsion-free.

Theorem 4.9.1 *Let F be a free group. Suppose $1 < i \leq j \leq 2i$ and $j \leq k \leq i+j+1$. Then $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ is torsion-free.*

Proof: Let L be a free Lie ring of the same rank as F. By Theorems 4.4.6 and 4.4.7, $L/[L^j, L^i, L^k]$ is torsion-free. Therefore, by Theorem 4.7.1, this Lie ring is isomorphic to the associated Lie ring of $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$. By Theorem 4.8.1 $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ is residually nilpotent, so it must be torsion-free. This completes the proof of the theorem. ■

Once again, the same proof gives

Theorem 4.9.2 *Let F be a free group. Suppose that $1 < i \leq k \leq j \leq 2i$. Then $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$ is torsion-free.*

and

Theorem 4.9.3 *Let F be a free group. Suppose $1 < i \leq j \leq 2i$ and $j \leq k \leq l \leq i+j$. Then $F/[\gamma_j(F), \gamma_i(F), \gamma_k(F), \gamma_l(F)]$ is torsion-free.*

In fact, in the terminology of Gorchakov, we have proved that (with the appropriate restrictions on i, j, k and l)

$$F/[\gamma_j(F), \gamma_i(F), \gamma_k(F)]$$

and

$$F/[\gamma_j(F), \gamma_i(F), \gamma_k(F), \gamma_l(F)]$$

are *Mal'cev Groups.*

Chapter 5

Finite groups of exponent 8

5.1 Introduction

The aim in this chapter is to improve upon the known bounds on the order of finite groups of exponent 8, in terms of the number of generators. We adapt the methods and results of [46], [47] and [48], while the overall approach to the problem comes from [53] and [54]. We re-prove results from these works only if we need them in a different form, or if our result is an improvement upon the general result. Otherwise, we just restate or quote them.

We proceed via the concept of sandwich algebras. We first obtain a bound on the nilpotency class of Lie algebras over fields of characteristic 2 which satisfy the Engel-7 identity and are generated by m sandwiches.

The bound we obtain is $m^{m^{2m}}$, whereas without the assumption on the characteristic or the Engel-7 identity, Vaughan-Lee and Zelmanov found a bound of $m^{m^{\cdot^{\cdot^{\cdot^{m}}}}}$, where there are 20 m's in this tower.

Unfortunately this reduction for sandwich algebras does not filter through to a reduction in the bound on the order of finite groups of exponent 8, since the height of the tower for sandwich algebras effectively contributes a constant to the final bound, which is absorbed by the enormousness of the numbers involved. Nonetheless, the reduction in the class of sandwich algebras may be a result of interest in itself.

Continuing to chart a parallel course to [48], we adapt the remainder of the proof of their bound for the general case in the case of exponent 8.

After [48], we inductively define the function $T(m, n)$ for positive integers m and n (with $m \geq 2$) by $T(m, 1) = m$, $T(m, n + 1) = m^{T(m,n)}$. The result

from [49] is the following (the case for prime-power exponent, which we are interested in, was proved in [48]).

Theorem 5.1.1 [49] *Let G be a finite m generator group of exponent n, with $m > 1$. Then*

$$|G| \leq T(m, n^{n^n}).$$

Our main result is

Theorem 5.1.2 *Let G be a finite m generator group of exponent 8, with $m > 1$. Then*

$$|G| \leq T(m, 7^{471}).$$

Thus our achievement consists of reducing 8^{8^8} to 7^{471} which is obviously an enormous reduction. Unfortunately, the numbers involved are still unimaginably large and it is very unlikely that these bounds are in any way realistic.

5.2 Sandwiches and things Engel

Definition 5.2.1 *An element a in a Lie algebra J is a* sandwich *if, for all $b, c \in J$, $[b, a, a] = [b, a, c, a] = 0$.*

We call J a sandwich algebra *if it is generated by sandwiches.*

We note the following identity.

$$[b, [c, a, a]] = [b, c, a, a] - 2[b, a, c, a] + [b, a, a, c].$$

Therefore, unless the characteristic of the underlying field is 2, an element a of J is a sandwich if and only if $[b, a, a] = 0$ for all $b \in J$. Unfortunately, in this chapter we work exclusively over fields of characteristic 2, and so this reduction is of no use to us.

The following theorem is originally due to Kostrikin and Zelmanov [55], and a proof may be found in [46].

Theorem 5.2.2 *Let J be a finitely generated sandwich algebra over a field K. Then J is nilpotent.*

Definition 5.2.3 *Suppose that n is a positive integer. Then the* Engel-n identity, *E_n, is the identity*

$$[x, \underbrace{y, \ldots, y}_{n}] = 0,$$

and we say that J is an Engel Lie algebra *if E_n is an identical relation for some n, and an* Engel-n Lie algebra *if E_n is an identical relation.*

Definition 5.2.4 *Suppose that n is a positive integer. Then the* linearised Engel-n identity *is the identity*

$$\sum_{\sigma \in S_n} [x_{n+1}, x_{1\sigma}, \ldots, x_{n\sigma}] = 0.$$

In any Lie algebra or Lie ring the Engel-n identity implies the linearised Engel-n identity (see [46, Lemma 1.2.7 and Lemma 1.2.8]). If $n!$ is a unit in the underlying ring, then the linearised Engel-n identity implies the Engel-n identity, and so they are equivalent. Also, in the Lie algebra we will be working with later in this chapter, namely $\tilde{J}$, the two identities will be equivalent.

In Section 5.4, we find a bound on the nilpotency class of an Engel-7 sandwich Lie algebra J over a field of characteristic 2 in terms of the number of sandwiches needed to generate J.

5.3 Outline of approach

This section is included in order to give an overview of the method of approaching this chapter. We first include a construction which we will need throughout this chapter, and a definition of objects known as *U-polynomials*. At least a knowledge of this definition is needed to understand Subsection 5.3.3.

5.3.1 The construction of $\tilde{J}$

We recall the following construction from [46, p.162].

Let J be a Lie algebra over a field K, generated by $a_1, \ldots, a_r$. Suppose that J satisfies the following identities

$$\sum_{\sigma \in S_7} [y, y_{1\sigma}, \ldots, y_{7\sigma}] = 0 \text{ for all } y, y_1, \ldots, y_7 \in J, \tag{5.1}$$

$$[x, \underbrace{y, \ldots, y}_{8}] = 0 \text{ for all } x \in J \text{ and for all homogeneous } y \in J. \tag{5.2}$$

Recall that the Lie ring associated with the Zassenhaus filtration of a finite group of exponent 8 satisfies both of these relations.

Let E be the associative algebra over K generated by the infinite set $\{e_1, e_2, \ldots\}$, subject to the relations $e_i e_j = e_j e_i$, $e_i^2 = 0$ for $i, j \geq 1$. If π is a finite set of positive integers, and $\pi = \{i, j, \ldots, k\}$, with $i < j < \cdots < k$, then we let $e_\pi = e_i e_j \cdots e_k$. Clearly E has a K-basis $\{e_\pi\}$.

We define $\tilde{J} = J \otimes_K E$. Every element of $\tilde{J}$ can be expressed uniquely in the form $\sum u_\pi \otimes e_\pi$, for some $u_\pi \in J$, where $u_\pi = 0$ for all but finitely many π. We turn $\tilde{J}$ into a Lie algebra over K by setting

$$\left[\sum u_\pi \otimes e_\pi, \sum v_\sigma \otimes e_\sigma\right] = \sum [u_\pi, v_\sigma] \otimes e_\pi e_\sigma.$$

It is straightforward to check that this turns $\tilde{J}$ into a Lie algebra over K. Note that E is locally nilpotent as an associative algebra, which implies that $\tilde{J}$ is locally nilpotent as a Lie algebra. It is also straightforward to check that if J satisfies the linearised version of an identity, then $\tilde{J}$ satisfies the identity itself. Thus, since J satisfies (5.1), the *linearised Engel-7 identity*, $\tilde{J}$ satisfies the Engel-7 identity

$$[y, \underbrace{x, \ldots, x}_{7}] = 0.$$

We also note that in $\tilde{J}$, any identity is equivalent to its linearised version (in a general Lie algebra, an identity implies, but is not equivalent to, its linearised version).

The importance of $\tilde{J}$ in our work rests upon the following theorem from [48, p.131], where we are assuming that the underlying field K is infinite of characteristic 2.

Theorem 5.3.1 (Proposition 2, p.131 [48]) *Let V be a variety of Lie algebras which satisfy the Engel-n identity. Suppose that for any $m \geq 1$, an m generator algebra in V is nilpotent of class at most $T(m, r)$, where r depends only on V, and is independent of m, except that $T(2, r) \geq m$.*

Let $J = \langle a_1, \ldots, a_r \rangle$ be a Lie algebra such that $J \otimes_K E \in V$, and suppose that for an arbitrary commutator ρ in $a_1, \ldots, a_r$ we have $\mathrm{ad}(\rho)^{n+1} = 0$. Then the algebra J is nilpotent of class at most $T(m, (r+6)n)$. ■

This allows us to calculate with $\tilde{J}$ rather than J, and then translate our results back to J. In particular, we may assume that all of our Lie algebras are Engel-7.

5.3.2 Definition of U-polynomials

We repeat the definition of [48, pp.136–37].

To define U-polynomials we first define U-words, which are abstract words, each associated with a linear operator on $\tilde{J}$.

We remarked above that each element $a \in \tilde{J}$ can be expressed uniquely in the form $a = \sum u_\pi \otimes e_\pi$. We will call this the *canonical decomposition of* a and write $a_\pi = u_\pi \otimes e_\pi$ so that $a = \sum a_\pi$.

First, we let V_i denote the ideal of $\tilde{J}$ consisting of the elements $a = \sum a_\pi$ such that $a_\pi = 0$ whenever i does not occur in π. It is clear that $V_i^2 = 0$ and that $\tilde{J} = \sum_{i \geq 1} V_i$. Let I be an ordered set and let $A = \{a_i, i \in I\}$ be a family of elements of $\tilde{J}$ such that

$$\text{Every element } a_i \text{ in } A \text{ lies in one of the ideals } V_{j(i)}, \text{ and} \tag{5.3}$$

$$[a_i, a_j] = 0 \text{ for all } a_i, a_j \in A \tag{5.4}$$

Given such a set A, we define the linear operator $U_k(A)$ on $\tilde{J}$ by setting

$$U_k(A) = \sum \mathrm{ad}(a_{i_1}) \cdots \mathrm{ad}(a_{i_k}),$$

where the summation is over all k-element subsets $\{i_1, \ldots, i_k\} \subseteq I$ where $i_1 < \cdots < i_k$.

For a given Lie algebra J, we define U-words to be words in the alphabet y_i, U_k, (,), ad, [,]. If the letters $y_1, \ldots, y_r$ are the only letters appearing in a U-word W then we write $W = W(y_1, \ldots, y_r)$ (note that we do not assume that all of $y_1, \ldots, y_r$ appear in W). If the variables y_i are associated with values $a_i \in \tilde{J}$, then the U-word W is associated with the linear operator $W(a_1, \ldots, a_r)$ acting on $\tilde{J}$. We now define U-words inductively as follows.

If ρ is a commutator in $y_1, \ldots, y_r$ with an arbitrary arrangement of brackets, then the word $W = \mathrm{ad}(\rho)$ is a U-word. If the y_i are associated with

$a_i \in \tilde{J}$, then W is associated with the linear operator $W(a_1, \ldots, a_r) = \mathrm{ad}(\rho(a_1, \ldots, a_r))$.

If $W = W(y_1, \ldots, y_r)$ is a U-word which does not involve the letter y_i then $W' = \mathrm{ad}(y_i W)$ is a U-word. If $y_1 \ldots, y_r$ are associated with $a_1, \ldots, a_r \in \tilde{J}$ and y_i is associated with $a_i \in \tilde{J}$ then W' is associated with the linear operator $\mathrm{ad}(a_i W(a_1, \ldots, a_r))$.

If $W = W(y_1, \ldots, y_r)$ and $V = V(y_1, \ldots, y_r)$ are U-words then their product WV is a U-word. The linear operator which is associated with WV is $W(a_1, \ldots, a_r)V(a_1, \ldots, a_r)$.

Suppose that $W = W(y_1, \ldots, y_r)$ is a U-word such that

$$[aW(a_1, \ldots, a_r), bW(a_1, \ldots, a_r)] = 0$$

for all $a, b, a_1, \ldots, a_r \in \tilde{J}$, and that y_i does not appear in the expression for W. Then $W' = U_k(y_i W)$ is a U-word. If the letters $y_1, \ldots, y_r$ and y_i are associated with elements $a_1, \ldots, a_r$ and a respectively, and if $a = \sum a_\pi$ is the canonical decomposition of a, then the set $A = \{a_\pi W(a_1, \ldots, a_r)\}$ satisfies conditions (5.3) and (5.4) above. The U-word W' is then associated with the linear operator $U_k(A)$.

This gives an inductive definition of U-words. A U-polynomial is a word $y_i W$ where W is a U-word and y_i is a variable. The degree of a U-polynomial is naturally defined, with y_j having degree 1 for each j.

5.3.3 Outline

We are attempting to find a bound on the order of a finite group of exponent 8, in terms of the number of generators for the group.

We will use G to denote a finite group of exponent 8. What we actually calculate is a bound on the nilpotency class of a particular Lie ring associated with G. This Lie ring is not the standard associated Lie ring, but the one associated with the Zassenhaus filtration of G, defined as follows [1]. Let p be a prime and let G_i be the subgroup of G generated by all powers of commutators $[g_1, \ldots, g_k]^{p^j}$ such that $g_1, \ldots, g_k \in G$ and $kp^j \geq i$. The series $\{G_i\}$ is the fastest descending series of subgroups of G such that $[G_i, G_j] \leq G_{i+j}$ and $G_i^p \leq G_{pi}$ for all i and j. It is not difficult to prove that $\{G_i\}_{i \geq 1}$

[1] The subgroups which form the Zassenhaus filtration are also called the *modular dimension subgroups*. The names of Lazard and Jennings are also sometimes associated with this series

is an N-series. Thus we can associate with this series a Lie ring $\mathfrak{L}\{G_i\}$ in the usual manner. The Lie ring $\mathfrak{L}\{G_i\}$ satisfies the identity $px = 0$, so can be considered to be a Lie algebra over $\mathbb{F}_p$, the field with p elements. For this chapter only, we will write $L(G) = \mathfrak{L}\{G_i\}$, where the subgroups G_i are defined using $p = 2$.

We can turn $L(G)$ into a restricted Lie algebra of characteristic 2 in the following way. If $a_\pi \otimes e_\pi \in L(G)$ then $(a_\pi \otimes e_\pi)^2 = 0$. We then extend the squaring operation to all of $L(G)$ by using the identity 2.2 from Definition 2.1.4. Thus

$$\begin{aligned}(a_\pi \otimes e_\pi + a_\rho \otimes e_\rho)^2 &= [a_\pi \otimes e_\pi, a_\rho \otimes e_\rho] \\ &= [a_\pi, a_\rho] \otimes e_\pi e_\rho.\end{aligned}$$

We will use the fact that $L(G)$ is a restricted Lie algebra later in the chapter.

In [40], it is proved that $L(G)$ satisfies the following relations

$$\sum_{\sigma \in S_7} [y, y_{1\sigma}, \ldots, y_{7\sigma}], \quad \text{for all } y, y_1, \ldots y_7 \in L(G), \text{ and} \tag{5.5}$$

$$\mathrm{ad}(a)^8 = 0 \quad \text{for any homogeneous element } a \in G_i/G_{i+1} \subseteq L(G). \tag{5.6}$$

Zelmanov's original solution of the restricted Burnside theorem from [54] used these two identities. We follow part of another proof from [47] which also uses a further identity. We define, for any positive integer n,

$$R_n(x_1, x_2, \ldots, x_{4n-1}) = \sum [x_{4n-1}, x_{S_1}, \ldots, x_{S_n}],$$

where the summation is taken over all partitions of $\{1, 2, \ldots, 4n-2\}$ into an ordered sequence of disjoint non-empty subsets $S_1, S_2, \ldots, S_n$ (recall the definition of $[y, x_S]$ for arbitrary y and S a non-empty set of positive integers from Section 2.1).

There is a series of multilinear expressions K_n for $n \geq 2$ such that $K_n = 0$ is an identity in the ordinary associated Lie ring of groups of exponent 8 for all n (see [46, Theorem 2.4.7]). In fact all multilinear identities which hold in the associated Lie ring of all groups of exponent 8 are consequences of these identities.

K_n is of weight n, and multilinear, so to say that it is an identity is equivalent to saying that

$$K_n(x_1, x_2, \ldots, x_n) \in \gamma_{n+1}(G),$$

for all $x_1, \ldots, x_n \in G$. Now, since $\{G_i\}$ is a central series, $\gamma_{n+1}(G) \subseteq G_{n+1}$. Therefore, K_n is an identity for $L(G)$. In characteristic 2, K_{27} is equivalent to R_7 (see [47, p.82]). Therefore, $R_7 = 0$ is an identity for $L(G)$.

We describe a method of constructing a particular series of U-polynomials $W_1, \ldots, W_t$ within $L(\tilde{G})$. From $W_1, \ldots, W_t$ are determined their linearisations, which we will call $f_1, \ldots f_t$. The method of constructing the linearisation of a U-polynomial comes from [53] and is described in [46, Lemma 9.5.21]. We then let

$$f_0 = \sum_{\sigma \in S_7} [y, y_{1\sigma}, \ldots, y_{7\sigma}].$$

Then the U-polynomials $W_1, \ldots, W_t$ satisfy the following property: for $1 \leq i \leq t$, if J satisfies the identities $f_0 = 0, f_1 = 0, \ldots, f_{i-1} = 0$, then all values of W_i are sandwiches in $\tilde{J}$.

The construction of $W_1, \ldots, W_t$ follows the approach of [46] and [47]. First we describe a method of constructing so-called *nil*-3 *elements*, and then we use the method of [47] to show how to construct sandwich elements from nil-3 elements. It is at this stage that we use the fact that $L(G)$ satisfies the identity $R_7 = 0$.

The main improvement we make at this stage is bounding the value of t. If we used the general theory, then we would obtain a value of $t \leq 8^8 - 5$, whereas the value we obtain is $t \leq 437$.

In [53] the following result is proved. Suppose that W is a U-polynomial of degree d. If f is the linearisation of W then

$$\begin{aligned} &\text{The values of } W \text{ in } \tilde{J} \text{ span the same subspace as the} \\ &\text{values of } f \text{ in } \tilde{J}, && (5.7) \\ &\text{Every value of } f \text{ in } \tilde{J} \text{ can be represented as a} \\ &\text{linear combination of at most } 2^d \text{ values of } W. && (5.8) \end{aligned}$$

We have already remarked that our Lie algebra $L(G)$ satisfies $f_0 = 0$, and it is clear that only the trivial Lie algebra satisfies $f_t = 0$.

We now fix a value of i and suppose that all values of $W_0, \ldots, W_{i-1}$ are zero in some Lie algebra J_{i-1} (or equivalently that all values of $f_0, \ldots f_{i-1}$ are zero). Since the values of f_i are in a subspace spanned by a bounded number of values of W_i which are sandwiches, we know that the subspace spanned by all values of f_i, which will be an ideal, is a sandwich algebra generated by a bounded number of sandwiches. Thus we can bound the nilpotency class

of this ideal using the results of Section 5.4. We then use Theorem 5.3.1 to transfer from $\tilde{J}$ back to J and then use induction to bound the nilpotency class of $L(G)$.

More specifically, d_i is defined so that an m-generator Lie algebra which satisfies the identities $f_0 = 0, \ldots, f_i = 0$ is nilpotent of class at most $T(m, d_i)$. We know that we can take $d_t = 1$, and we derive a bound for d_{i-1} in terms of d_i. This allows us to find a bound for d_0, which immediately gives us a bound for the nilpotency class for $L(G)$. From there it is straightforward to convert this into a bound on the order of G.

5.4 Sandwich algebras

This section stands alone from the rest of the chapter in many ways. We prove a single result which does not rely on the rest of the result in this chapter, but is of use in the latter half of the chapter. The result that is proved in this section, namely a new bound on the nilpotency class of m-generator Engel-7 Lie algebras over fields of characteristic 2, may be of some interest in its own right. However, our motivation for finding this bound is practical.

In this section we prove the following theorem, which forms one of two major parts in the proof of Theorem 5.1.2.

Theorem 5.4.1 *Let K be a field of characteristic 2, and suppose that J is an Engel-7 Lie algebra over K which is generated by m sandwiches. Then J is nilpotent of class less than $m^{m^{2m}}$.*

We consider the cases $1 \leq m \leq 5$ separately, since we can calculate the exact bounds for these. We use the Nilpotent Quotient Algorithm for Lie rings (originally developed in Canberra in 1976), with specially written subroutines which make any number of the generators into sandwiches. We then calculate the Lie ring over a field of characteristic 2 which is generated by m sandwiches, for $m = 1, 2, 3$, which satisfies the Engel-7 identity. Obviously, if $m = 1$, then the Lie ring is abelian, so nilpotent of class 1. If $m = 2$, we find that J is nilpotent of class at most 2, and for $m = 3$, we find J is nilpotent of class at most 3. We can also calculate bounds for the nilpotency class when $m = 4$ (where the exact bound is 5) and $m = 5$ (where the exact bound is 10). All of these numbers easily fit into the bound given in Theorem 5.4.1. Beyond 5, the calculations become much longer, and clearly no general

bound can be found from calculating specific cases. Thus, from now on, we assume that $m \geq 6$.

First of all, if J is a Lie algebra over K generated by m sandwiches (we will introduce the assumption that J is Engel-7 later), then we define $N = \mathrm{ad}(J)$, and let A be the algebra in $End_K(J)$ generated by N. If J is generated by $y_1, \ldots, y_m$, which are sandwiches, then N is generated by $b_i = \mathrm{ad}(y_i)$ for $1 \leq i \leq m$. Notice that the fact that y_i is a sandwich implies that $b_i^2 = 0$ and $b_i N b_i = \{0\}$ in A.

Let $V_i = [N, b_i] + Kb_i$ for $1 \leq i \leq k$. It is clear that $N = \sum_{i=1}^{m} V_i$.

Let $\hat{A} = K.1 + A$, for a formal unit 1.

For $j \geq 0$, let $N^{[j]}$ be the subspace of $\hat{A}$ spanned by 1 and by products of the form

$$a_1 \cdots a_i,$$

where $i \leq j$ and $a_1, \ldots, a_i \in N$.

We prove the following lemma, which is crucial in our proof of Theorem 5.4.1.

Lemma 5.4.2 $N^{[m]} = \hat{A}$

Proof: It suffices to show that any product $\rho_1 \cdots \rho_{m+1}$ can be written as a linear combination of products of length at most m.

We use the identity $\rho\nu = \nu\rho + [\rho, \nu]$ to see that, modulo products of greater length, any two elements ρ_i, ρ_j in a product $\rho_1 \cdots \rho_k$ commute with each other.

So, we assume that $\rho\nu = \nu\rho$ for any $\rho, \nu \in N$ and show that, with this assumption, any product of $m+1$ elements from N is zero.

Recall the expression $N = \sum_{i=1}^{m} V_i$. We write an arbitrary product of $m+1$ elements of N as a linear combination of products of length $m+1$ in elements of $V_1, \ldots, V_m$, using linearity, and show that all of these products are zero.

Suppose that we have $v_1, \ldots, v_{m+1}$ where $v_j \in V_{i_j}$ for all $j \in \{1, \ldots, m+1\}$ and some $i_j \in \{1, \ldots, m\}$. We know that the homogeneous component of any Lie algebra is spanned by left-normed words in the generators (see [46, Section 1.2]). Thus by linearity we can assume that each $v_{i_j} = b_{i_j}$ or $[q_j, b_{i_j}]$ for some $q_j \in N$. Choose $r \neq s$ such that $i_r = i_s = t$, say.

There are four cases to consider. The first is when $v_{i_r} = v_{i_s} = b_t$. Then clearly $v_{i_r} v_{i_s} = 0$, and since we assumed that the v_i commute with each other, $v_1 \cdots v_{m+1} = 0$.

The second case is where $v_{i_r} = [q_r, b_t]$ and $v_{i_s} = b_t$. Then

$$\begin{aligned} v_{i_r}v_{i_s} &= [q_r, b_t]b_t \\ &= q_r b_t b_t - b_t q_r b_t \\ &= 0, \end{aligned}$$

since the b_t come from sandwiches in J. Therefore, again $v_1 \cdots v_{m+1} = 0$. The third case is when $v_{i_r} = b_t$ and $v_{i_s} = [q_s, b_t]$. Since $v_{i_r}v_{i_s} = v_{i_s}v_{i_r}$, the fact that $v_{i_r}v_{i_s} = 0$ follows from the second case.

The final case to consider is when $v_{i_r} = [q_r, b_t]$ and $v_{i_s} = [q_s, b_t]$. Then

$$\begin{aligned} v_{i_r}v_{i_s} &= [q_r, b_t][q_s, b_t] \\ &= q_r b_t q_s b_t - q_r b_t b_t q_s - b_t q_r q_s b_t + b_t q_r b_t q_s \\ &= -b_t q_r q_s b_t \\ &= [q_r, b_t, q_s]b_t \end{aligned} \tag{5.9}$$

Now, if we express $[q_r, b_t, q_s]$ as a linear combination of elements of N of the form $[y_w, b_w]$, then we can replace v_{i_r} by $\sum_w [y_w, b_w]$ and v_{i_s} by b_t in the original product. We now have fewer v_a of the form $[q_u, b_u]$, so we can choose $r' \neq s'$ such that $i^{r'} = i^{s'}$, and repeat the above process if necessary. Since at each stage we reduce the number of elements v_{i_v} which are of the form $[q_w, b_w]$, this process must terminate eventually, so we must eventually end up in one of the first three cases. Thus, our original product $v_1 \cdots v_{m+1}$ is zero. If we reach the case where all v_{i_j} are of the form $v_{i_j} = b_{i_j}$ then $v_1 \cdots v_{m+1} = 0$ by the case $v_{i_r} = v_{i_s} = b_t$. This completes the proof of the lemma. ■

We consider the lower central series and derived series for N. Let N^i be the i-th term of the lower central series of N (where $N^1 = N$), and let $N^{(j)}$ be the j-th term of the derived series of N (where $N^{(0)} = N$).

Lemma 5.4.3 *For any $r \geq 1$,*

$$A^{m(r-1)+1} \subseteq Id_A(N^r),$$

the ideal generated in A by N^r.

Proof: Suppose that we have a product $b_{i(1)} \ldots b_{i(t)}$ in A. By Lemma 5.4.2, it can be expressed as a linear combination of elements of the form $\rho_1 \ldots \rho_k$, where $k \leq m$ and the ρ_i are left-normed commutators in $b_1, \ldots, b_m$, the lengths of which add up to t.

So, if $t \geq m(r-1)+1$, then at least one of $\rho_1, \ldots, \rho_k$ must have length at least r. This completes the proof of the lemma. ■

Lemma 5.4.4 *For any $k \geq 0$,*

$$J^{m^k} \subseteq J^{(k)}.$$

Proof: The proof goes by induction on k. The case $k = 0$ is trivial, since $J \subseteq J$, and $k = 1$ is also trivial, since $m \geq 2$. Thus we suppose that the result is true for $i = k$ and consider the case $i = k + 1$. Let $r = m^k$, so that $J^r \subseteq J^{(k)}$. Then

$$\begin{aligned} J^{r+m(r-1)+1} &= [J^r, \underbrace{J, \ldots, J}_{m(r-1)+1}] \\ &\subseteq J^r.A^{m(r-1)+1} \\ &\subseteq J^r.\mathrm{Id}_A N^r \text{ by Lemma 5.4.3} \\ &\subseteq J^{(k+1)}, \end{aligned}$$

by the inductive hypothesis, and the fact that $N^r = \mathrm{ad}(J^r)$. Now, we can in fact trim this to $J^{r+m(r-1)} \subseteq J^{(k+1)}$. We do this by following the details of the proofs of Lemmas 5.4.3 and 5.4.2. Consider an element of $J^{r+m(r-1)}$. We have an element of J^r, acted on by an element of $A^{m(r-1)}$. The only way for this element of $A^{m(r-1)}$ to not be in N^r is if it is a product $\rho_1 \cdots \rho_m$, where each of $\rho_1, \ldots, \rho_m$ is a left-normed commutator in the generators of N of weight exactly $r - 1$ and each of which ends in a different generator. To see this, we consider the proof of Lemma 5.4.2. If two of the ρ_i end in the same generator we can replace one by a generator and the other by a product of weight $2r - 1$ (see (5.9)). But then the element of weight $2r - 1$ is in N^r, and so the whole expression is contained in $J^{(k+1)}$. Therefore, we can assume that each of $\rho_1, \ldots, \rho_m$ end in a different generator. However, this means that one of $\rho_1, \ldots, \rho_m$ must end in the same generator as ρ_0, the element of weight r. Then, modulo $J^r.Id_A(N^r)$, all of $\rho_1, \ldots, \rho_m$ commute in the product $[\rho_0, \rho_1 \ldots, \rho_m]$, so we can assume that ρ_1 ends in the same generator as ρ_0. It is then a matter of a trivial calculation to see that $[\rho_0, \rho_1] = 0$.

Thus $J^r \subseteq J^{(k)}$ implies $J^{mr} \subseteq J^{(k+1)}$. This completes the induction and also the proof of the lemma. ■

Lemma 5.4.4 immediately implies the following theorem.

Theorem 5.4.5 *Let J be a Lie algebra over K which is generated by m sandwiches. Then for any $k \geq 1$ the algebra $J/J^{(k)}$ is nilpotent of class at most m^k.*

We now find a bound on the nilpotency class of $J^{(k)}$ (for some appropriate k) and then show how the nilpotency of $J^{(k)}$ and of $J/J^{(k)}$ can be put together to yield a bound on the nilpotency class of J.

We know that J is nilpotent (by [46, Theorem 3.2.1], for example), so our task is just a matter of finding a bound on the nilpotency class.

By Theorem 5.4.5 we know that $J/J^{(k+1)}$ is nilpotent of class at most m^{k+1}, so it follows that $J/J^{(k+1)}$ has dimension as a vector space at most $m^{m^{k+1}}$ (to see an analogous argument, see Subsection 5.6). Thus $J^{(k)}/J^{(k+1)}$ has dimension less than $m^{m^{k+1}}$. Since we know that $J^{(k)}$ is nilpotent, we can see that $J^{(k)}$ is generated as a subalgebra by less than $m^{m^{k+1}}$ elements. The only question is whether we can take these generators to be derived commutators of the generators of J which lie in $J^{(k)}$. We now prove that we can.

Lemma 5.4.6 *$J^{(k)}$ is generated by less than $m^{m^{k+1}}$ derived commutators in the generators of J.*

Proof: We can certainly say that $J^{(k)}/J^{(k+1)}$ is generated by less than $m^{m^{k+1}}$ cosets of linear combinations of derived commutators of the required sort (by the linearity of the Lie product, the above comments and the definition of $J^{(k)}$).

Let S be the set of all such derived commutators appearing in some such coset. Then take a subset of S such that the cosets of elements of this subset form a basis for $J^{(k)}/J^{(k+1)}$ as a vector space. This subset of S is of size less than $m^{m^{k+1}}$, and the elements will generate $J^{(k)}$ as a subalgebra, since they generate $J^{(k)}$ modulo its derived subalgebra, and since $J^{(k)}$ is nilpotent. This completes the proof of the lemma. ■

Let $k = 2m - 3$. We prove that $J^{(2m-3)}$ is nilpotent of class less than $2m^{m^{2m-2}}$.

We include without proof and subsequently use without mention the following result from [48, pp.112–113].

Lemma 5.4.7 *If $k \geq 3$, then for an arbitrary subset $S \subseteq N$ we have*

$$\begin{aligned}
N^{[k]}S &\subseteq SN^{[k]} + N^{[k]},\\
SN^{[k]} &\subseteq N^{[k]}S + N^{[k]},\\
N^{[k]}S &\subseteq N^{[1]}SN^{[k-1]} + SN^{[k]} + N^{[k-1]},\\
SN^{[k]} &\subseteq N^{[k-1]}SN^{[1]} + N^{[k]}S + N^{[k-1]},
\end{aligned}$$

$$\begin{aligned} N^{[k]}S &\subseteq N^{[2]}SN^{[k-2]} + N^{[1]}SN^{[k-1]} + SN^{[k]} + N^{[k-2]}, \\ SN^{[k]} &\subseteq N^{[k-2]}SN^{[2]} + N^{[k-1]}SN^{[1]} + N^{[k]} + N^{[k-2]}. \end{aligned}$$

Lemma 5.4.8 *If s and t are elements of N which arise from sandwiches in J (ie $ss = tt = 0$ and $sNs = tNt = \{0\}$) and, for some $k, r \geq 2$, $sN^{[k]}sN^{[r]}s = tN^{[k]}tN^{[r]}t = 0$, then*

$$[s,t]N^{[k+1]}[s,t]N^{[r]}[s,t] = \{0\}.$$

Proof:

$$\begin{aligned} [s,t]N^{[k+1]}[s,t]N^{[r]}[s,t] &\subseteq stN^{[k+1]}[s,t]N^{[r]}[s,t] \\ &\quad + tsN^{[k+1]}[s,t]N^{[r]}[s,t]. \end{aligned}$$

If the first of the terms on the right hand side above is zero, then the second is also zero by symmetry.

Thus we consider only $stN^{[k+1]}[s,t]N^{[r]}[s,t]$. Then,

$$\begin{aligned} stN^{[k+1]}[s,t]N^{[r]}[s,t] &\subseteq stN^{[k+1]}stN^{[r]}st + stN^{[k+1]}stN^{[r]}ts \\ &\quad + stN^{[k+1]}tsN^{[r]}st + stN^{[k+1]}tsN^{[r]}ts, \end{aligned}$$

and we separately prove that each of these terms is zero.

We use the convention of underlining something which we know to be zero. Now,

$$\begin{aligned} stN^{[k+1]}stN^{[r]}st &\subseteq sN^{[k]}tN^{[1]}stN^{[r]}st + sN^{[k+1]}\underline{tst}N^{[r]}st \\ &\quad + sN^{[k]}stN^{[r]}st, \text{ and} \\ sN^{[k]}tN^{[1]}stN^{[r]}st &\subseteq sN^{[k]}\underline{tN^{[1]}stN^{[r]}t}, \end{aligned}$$

(the last underlining being there because $k \geq 2$ and so $N^{[2]} \subseteq N^{[k]}$). Also,

$$sN^{[k]}stN^{[r]}st \subseteq \underline{sN^{[k]}sN^{[r]}st},$$

so that we have $stN^{[k+1]}stN^{[r]}st = \{0\}$. Next,

$$\begin{aligned} stN^{[k+1]}stN^{[r]}ts &\subseteq sN^{[k]}\underline{tN^{[1]}stN^{[r]}t}s + sN^{[k+1]}\underline{tst}N^{[r]}ts \\ &\quad + sN^{[k]}stN^{[r]}ts \\ &\subseteq sN^{[k]}sN^{[r-1]}\underline{tN^{[1]}t}s + sN^{[k]}sN^{[r]}\underline{tt}s \\ &\quad + \underline{sN^{[k]}sN^{[r-1]}ts} \\ &= \{0\}. \end{aligned}$$

Also,

$$\begin{aligned} stN^{[k+1]}tsN^{[r]}st &\subseteq sN^{[k-1]}tN^{[2]}tsN^{[r]}st + sN^{[k]}\underline{tN^{[1]}t}sN^{[r]}st \\ &\quad + sN^{[k+1]}\underline{tt}sN^{[r]}st + \underline{sN^{[k-1]}tsN^{[r]}st} \\ &\subseteq sN^{[k-1]}tN^{[2]}tN^{[r-1]}\underline{sN^{[1]}st} + sN^{[k-1]}tN^{[2]}tN^{[r]}\underline{sst} \\ &\quad + sN^{[k-1]}\underline{tN^{[2]}tN^{[r-1]}st}. \end{aligned}$$

Finally,

$$\begin{aligned} stN^{[k+1]}tsN^{[r]}ts &\subseteq sN^{[k-1]}tN^{[2]}tsN^{[r]}ts + sN^{[k-1]}\underline{tN^{[1]}t}sN^{[r]}ts \\ &\quad + sN^{[k+1]}\underline{tt}sN^{[r]}ts + sN^{[k-1]}tsN^{[r]}ts \\ &\subseteq sN^{[k-1]}tN^{[2]}tN^{[r]}\underline{sts} + sN^{[k-1]}\underline{tN^{[2]}tN^{[r]}t}s \\ &\quad + sN^{[k-1]}\underline{tst}N^{[r]}s + \underline{sN^{[k-1]}tsN^{[r]}s} \\ &= \{0\}. \end{aligned}$$

This completes the proof of the lemma. ■

Note that by Lemma 5.4.8 and symmetry we have the following

Corollary 5.4.9 *If s and t are elements of N which arise from sandwiches in J (ie $ss = tt = 0$ and $sNs = tNt = \{0\}$) and, for some $k, r \geq 2$, $sN^{[k]}sN^{[r]}s = tN^{[k]}tN^{[r]}t = \{0\}$, then*

$$[s,t]N^{[k]}[s,t]N^{[r+1]}[s,t] = \{0\}.$$

We now introduce the assumption that our Lie algebra is Engel-7, in order to facilitate the following calculation. Again we use the Nilpotent Quotient Algorithm for Lie rings, with the routines to make generators sandwiches. Suppose that M is the largest Lie algebra subject to the constraints that M is Engel-7 over a field of characteristic 2, M is generated by the set $\{x, y, a, b, c, d, e\}$, and x and y are sandwich elements in M. We calculated the multiweight $(3, 3, 1, 1, 1, 1, 1)$ component of M (being weight at most 3 in x and y and linear in each other generator), and found that

$$[a, [x, y], b, c, [x, y], d, e, [x, y]] = 0.$$

This is enough to show that, in our algebra A, the following identity holds:

$$[s,t]N^{[2]}[s,t]N^{[2]}[s,t] = \{0\}, \tag{5.10}$$

whenever s and t satisfy $ss = tt = 0$ and $sNs = tNt = \{0\}$.

Definition 5.4.10 *Let L be a free Lie ring, freely generated by $z_1, z_2, \ldots$. Then define words $\delta_i(z_1, \ldots, z_{2_i})$ by $\delta_0(z_1) = z_1$ and*

$$\delta_{k+1}(z_1, \ldots, z_{2^{k+1}} = [\delta_k(z_1, \ldots, z_{2^k}), \delta_k(z_{2^k+1}, \ldots, z_{2^{k+1}})].$$

A derived commutator *is a value of δ_i for some i.*

We use Corollary 5.4.9, Lemma 5.4.8 and (5.10) to prove the following result.

Lemma 5.4.11 *Suppose that x is a derived commutator that lies in $N^{(2m-3)}$. Then*

$$xN^{[m]}xN^{[m]}x = \{0\},$$

so that, in particular, if $a, b \in \hat{A}$, $xaxbx = 0$.

Proof: The lemma follows immediately from Lemma 5.4.8, Corollary 5.4.9, the identity (5.10) and induction on m. ■

We are now in a position to prove the following result.

Theorem 5.4.12 *Let J be an Engel-7 Lie algebra over K which is generated by m sandwiches. Then $J^{(2m-3)}$ is nilpotent of class less than $2m^{m^{2m-2}}$.*

Proof: Suppose that x is a left-normed product of generators of $J^{(2m-3)}$ of length at least $2m^{m^{2m-2}}$. Now, we know that these generators are derived commutators of the generators, each of which lies in $J^{(2m-3)}$. Then, at least one of the generators of $J^{(2m-3)}$ must appear at least three times, by Lemma 5.4.6. Suppose that this generator is y. We then write

$$x = [a_1, \ldots, a_k, y, a_{k+1}, \ldots, a_j, y, , a_{j+1}, \ldots, a_i, y, a_{i+1}, a_r],$$

where $r+3$ is the length of the commutator. This element is equal to a linear combination of elements of the form $[a_1, \ldots, a_k]\mathrm{ad}(y)\alpha\mathrm{ad}(y)\beta\mathrm{ad}(y)\gamma$ for some $\alpha, \beta, \gamma \in N^{[m]}$, by Lemma 5.4.2. By Lemma 5.4.11,

$$\mathrm{ad}(y)\alpha\mathrm{ad}(y)\beta\mathrm{ad}(y) = 0,$$

so $x = 0$. This completes the proof of the theorem. ■

Given Theorem 5.4.12, we are in a position to prove Theorem 5.4.1.

Proof of Theorem 5.4.1: Let $b = 2m^{m^{2m-2}}$ and $c = m^{2m-2}$ and suppose that we have a left-normed product x in J of length

$$d = c + 1 + (mc + 1)b.$$

Write $x = [a_1, a_2, \ldots, a_d]$, say. Then, by Theorem 5.4.5, $a = [a_1, \ldots, a_{c+1}] \in J^{(2m-3)}$. We thus write

$$x = [a, a_{c+2}, \ldots, a_{mc+c+2}, a_{mc+c+3}, \ldots, a_{2mc+c+3}, \ldots, a_d].$$

Consider,

$$\begin{aligned} [a, a_{c+2}, \ldots, a_{mc+c+2}] &= a.\mathrm{ad}(a_{c+2}) \cdots \mathrm{ad}(a_{mc+c+2}) \\ &= \sum a.\mathrm{ad}(d_1) \cdots \mathrm{ad}(d_k), \end{aligned}$$

where the d_i are commutators in the a_j of total weight $mc+1$, and where $k \leq m$ (by Lemma 5.4.2). Thus, one of the d_i must have weight at least $c+1$, which means that $d_i \in J^{(2m-3)}$, by Lemma 5.4.5. Therefore $[a_1, \ldots, a_{mc+c+2}] \in (J^{(2m-3)})^2$. Repeating this argument on each $mc+1$ terms of x in turn gives us that

$$[a_1, \ldots, a_d] \in (J^{(2m-3)})^{2m^{m^{2m-2}}+1} = \{0\}.$$

Therefore, any product of length d in J will be zero and so J is nilpotent of class less than d.

Now,

$$\begin{aligned} d &= c + 1 + (mc+1)b \\ &= m^{2m-2} + 1 + (m^{2m-1}+1)(2m^{m^{2m-2}}) \\ &< 2m^{m^{2m-2}+2m-1} + 2m^{m^{2m-2}} + m^{2m-2} + 1 \\ &< m^{m^{2m-2}}(2m^{2m-1}+3) \\ &< m^{m^{2m}}. \end{aligned}$$

This completes the proof of Theorem 5.4.1. ■

5.5 U-polynomials and sandwich identities

5.5.1 From U-polynomials to multilinear words

As we mentioned earlier, in [53] the following is proved. Suppose that f is a U-polynomial of degree d. Then there exists a multilinear Lie polynomial f' (called the *linearisation* of f) such that

The values of f in $\tilde{J}$ span the same subspace as the

values of f' in $\tilde{J}$, (5.11)

Every value of f' in $\tilde{J}$ can be represented as a linear combination of at most 2^d values of f. (5.12)

It is not difficult to calculate that the U-polynomials of which we will describe the construction in the next two sections will have weight at most 370000. In fact the maximum possible weight is 369967.

We can now use the above result to say that every value of f' in $\tilde{J}$ can be represented as a linear combination of at most 2^{370000} values of f, something which we will use later. The proof of the above result of [53] is slightly different in our case, but can be easily adapted from the proof in [46, pp.196–200]. However, we will not prove this here, since the general theory tells us that we can use $2^{8^{8^2}}$ instead of 2^{370000}, and doing this in the proof of Lemma 5.6.1 makes no difference to the final result.

5.5.2 Nil-3 elements

It is easy to construct a series of multilinear Lie polynomials starting at $f_t = y$ and finishing at a linearised version of the Engel-6 identity (which will be f_{t-4}) such that any value of f_i in any Lie algebra over an infinite field of characteristic 2 will be a sandwich modulo the ideal generated by all values of $f_0, \ldots, f_{i-1}$. For details of this construction, see Subsection 5.5.4. Therefore, in this subsection and in Subsection 5.5.3 we will try to construct the series $f_0, \ldots, f_{t-5}$, and will stop if we find that the Lie algebra we are dealing with is Engel-6.

The construction of $f_0, \ldots, f_{t-5}$ occurs in two stages. The first is to construct a word which satisfies (5.13), (5.14) and (5.15) below. Such a word is said to be *nil*-3. Then we follow the approach of [47] to find an element which is not identically zero but whose values are all sandwiches. We add this as an identical relation to the Lie algebra, and repeat the process.

Consider the following conditions:

$$w(a_1, \ldots, a_r) \neq 0 \text{ for some } a_1, \ldots, a_r \in \tilde{J}; \quad (5.13)$$

$$[bw(a_1, \ldots, a_r), cw(a_1, \ldots, a_r)] = 0 \text{ for all } b, c, a_1, \ldots, a_r \in \tilde{J}; \quad (5.14)$$

$$U_k(a_{r+1}w(a_1, \ldots, a_r)) = 0 \text{ for all } a_1, \ldots, a_{r+1} \in \tilde{J}, \text{ and } k \geq 3 \quad (5.15)$$

Let $v(y_1) = \mathrm{ad}(y_1)^6$. We know (see [46, p172]) that v satisfies (5.13), (5.14) and (5.15) for $k \geq 7$. We also know, by direct calculation using the

Nilpotent Quotient Algorithm for Lie rings, that v does *not* satisfy (5.15) for $k = 5$. See Section 5.7 for a brief discussion of the possibilities for the case $k = 6$.

We follow [46, pp.172–73] to construct from v an element that satisfies (5.15) for $k \geq 3$, as well as (5.13) and (5.14).

Suppose that we have a word $w = w(y_1, \ldots, y_r)$ which satisfies (5.13), (5.14) and (5.15) for $k \geq m$ where $3 < m \leq 7$. We construct a word which satisfies (5.13), (5.14) and (5.15) for $k \geq m - 1$.

Define

$$w^{(1)} = U_{m-1}(y_{r+1}w(y_1, \ldots, y_r)),$$

and also define

$$w^{(2)} = \mathrm{ad}(y_{r+m+1}w^{(1)}) \cdots \mathrm{ad}(y_{r+2m-1}w^{(1)}).$$

Theorem 5.5.1 *For any word w, as above, if $m = 4, 5, 6$ or 7, then $w^{(2)} = 0$.*

Proof: The calculations for this are long, and not especially illuminating, so we omit them here. They appear in Appendix B. ■

It is this result which allows us to achieve most of the improvement in our final bound on the order of a finite group of exponent 8, as we shall shortly see. In [46] we need $w^{(1)}, w^{(2)}, \ldots, w^{(t)}$ for some t much larger than 1 (see [46, pp.172–3]).

From [46, p.173], $w^{(1)}$ satisfies (5.13) and (5.14). Theorem 5.5.1 implies that $w^{(1)}$ satisfies (5.15) for $k \geq m - 1$. We now repeat this procedure, starting with $w^{(1)}$, and eventually obtain an element which satisfies (5.13), (5.14) and (5.15) for $k \geq 3$. That is to say we will obtain a nil-3 element.

Once we have a nil-3 element, we follow the procedure in Section 5.5.3 to find a U-polynomial which has non-zero values which are all sandwiches. We now calculate how many possible nil-3 elements we may need to consider in Section 5.5.3.

We will always have $v(y_1) = \mathrm{ad}(y_1)^6$. We made the comment above that in a free Engel-7 Lie algebra v does not satisfy (5.15) for $k = 5$. But we are working in a Lie algebra satisfying $f_0 = 0, f_1 = 0, \ldots, f_{i-1} = 0$. In this Lie algebra v may satisfy (5.15) for $k \geq n$ for some $n < 7$.

Recall that the definition of $w^{(1)}$ is dependent on the value of m. To make this clear, we will write $w_m^{(1)}$ when we have used m in the definition of $w^{(1)}$.

If we have a word u which satisfies (5.15) for $k \geq n$, then we know that $u_n^{(1)}$ satisfies (5.15) for $k \geq n - 1$. However, it may also satisfy (5.15) for

$k \geq n-2$. Thus, either $u_n^{(1)}$ will satisfy (5.15) for $k \geq 3$ or our next word will be $(u_n^{(1)})_r^{(1)}$ for some $4 \leq r \leq n-1$. Therefore, the words which we use will be of the form

$$(\ldots((v_{m_1}^{(1)})_{m_2}^{(1)})\ldots)_{m_k}^{(1)},$$

where $7 \geq m_1 > m_2 > \ldots > m_k \geq 4$, and $k \geq 0$. Thus, the number of words we are considering is equal to the number of subsets of $\{7,6,5,4\}$, which is 16. Theoretically, there seems to be no reason why we may not need to use all 16 of these words.

5.5.3 From nil-3 elements to sandwiches

Let W be a value of the U-polynomial w, the nil-3 element which we constructed in the previous section. Then $W \in \mathrm{ad}(\tilde{J})$ is also a nil-3 element, in that for any $x \in \tilde{J}$, $xW^3 = 0$.

We follow the argument from [47], starting with a Lemma (which we repeat without proof) from page 81.

Lemma 5.5.2 *Let J be a Lie algebra over an infinite field K, and suppose that J satisfies then identity $R_n = 0$ for some positive integer n. Let $W \in \mathrm{ad}(J)$ be nil-3. If $a, b \in J$ and $k \geq 0$ we let*

$$G_k(a,b,W) = U_2(a,W).(U_2(b).U_2(a,W))^k.$$

1. *$G_k(a,b,W)$ is nil-3 for all $a, b \in \tilde{J}$ and for all $k \geq 0$.*
2. *$G_{n-1}(a,b,W) = 0$ for all $a, b \in \tilde{J}$.*
3. *If $G_{2m}(a,b,W)$ is a sandwich then $aW(U_2(b).U_2(a,W))^m$ is a sandwich.*
4. *If $G_{2m+1}(a,b,W) = 0$ then $bG_m(a,b,W)$ is a sandwich.*
5. *If $M \geq 1$ and if $G_{m+1}(a,b,W) = 0$ and $aW(U_2(b).U_2(a,W))^m = 0$ for all $a, b \in \tilde{J}$, then $G_{\lfloor (m+1)/2 \rfloor}(c,d,G_m(a,b,W)) = 0$ for all $a,b,c,d \in \tilde{J}$.*

Since J satisfies R_7, this lemma tells us that $G_6(a,b,W) = 0$ for all $a, b \in \tilde{J}$.

We now follow the argument from [47, p.82]. If $G_0(a,b,W) = 0$, then aW is a sandwich by 5.5.2.3. Also, aW takes non-zero values because W satisfies (5.13). If $G_0(a,b,W) \neq 0$ for some a, b and W, but $G_1(a,b,W) = 0$ for all

a, b and W, then $bU_2(a, W)$ is a sandwich, by 5.5.2.4. Again, $bU_2(a, W)$ takes non-zero values because $G_0(a, b, W) \neq 0$ for some a, b and W.

Suppose that $G_1(a, b, W) \neq 0$ for some a, b and W. Then find q such that $G_q(a, b, W) \neq 0$ for some a, b and W, but $G_{q+1}(a, b, W) = 0$ for all a, b and W (note: $q = 1, 2, 3, 4$ or 5). Then 5.5.2.3 implies that $aW(U_2(b)U_2(a, W))^q$ is a sandwich for all a, b, W. If there exist $a, b \in \tilde{J}$ and W value of w such that $aW(U_2(b)U_s(a, W))^q$ is nonzero then the required sandwich-valued polynomial is

$$x_{r+1}w(U_2(x_{r+2})U_2(x_{r+1}w))^q.$$

However, if $aW(U_2(b)U_2(a, W))^q = 0$ for all a, b and W then we define

$$v = U_2(x_{r+1}w).(U_2(x_{r+2}).U_2(x_{r+1}w))^q.$$

Then we can see that v is a U-polynomial with the property that all values of v are nil-3 and such that some value of v is nonzero (by 5.5.2.1 and the assumption $G_q(a, b, W) \neq 0$ for some a, b and W). By 5.5.2.5, $G_{\lfloor \frac{m+1}{2} \rfloor}(c, d, V) = 0$ for all $c, d \in \tilde{J}$ and all values V of v. We can continue in this way to eventually obtain a sandwich valued U-polynomial f, where some value of f is nonzero.

We now need to calculate how many such U-polynomials we may need to add for each w (as constructed in the previous section) before we find the sandwich valued U-polynomial which has nonzero values. The largest possible value for $\lfloor \frac{m+1}{2} \rfloor$ is 3. We consider the separate cases in turn.

Case 1: $G_0(a, b, W) = 0$. Then the sandwich polynomial we require is aW. This gives us one possible word.

Case 2: $G_0(a, b, W) \neq 0$ but $G_1(a, b, W) = 0$. Then the sandwich polynomial is $bU_2(a, W)$. This also gives us one possible word.

Case 3: $G_1(a, b, W) \neq 0$ but $G_2(a, b, W) = 0$. Then we either have that $aW(U_2(b)U_2(a, W))$ is the required sandwich, or else we define

$$v := U_2(y_{r+1}w)(U_2(y_{r+2})U_2(y_{r+1}w)),$$

and let V be a value of v. Then $\lfloor \frac{2}{2} \rfloor = 1$, so $G_1(c, d, V) = 0$ for all c, d and V, so we have two cases, where $G_0(c, d, V) = 0$ and where $G_0(c, d, V) \neq 0$. Thus Case 3 yields 3 possible words.

Case 4: $G_2(a, b, W) \neq 0$ but $G_3(a, b, W) = 0$. Then we either have that $aW(U_2(b)U_2(a, W))^2$ is the required sandwich, or we define

$$v := U_2(y_{r+1}W)(U_2(y_{r+2})U_2(y_{r+1}w))^2,$$

and let V be a value of v. In this case $\lfloor \frac{3}{2} \rfloor = 1$, so $G_1(c,d,V) = 0$ for all c, d and V. The possibilities for $G_0(c,d,V) = 0$ and $G_0(c,d,V) \neq 0$ yield another 2 possible words. Thus Case 4 yields 3 possible words.

Case 5: $G_3(a,b,W) \neq 0$ but $G_4(a,b,W) = 0$. This leads to the possibility of $aW(U_2(b)U_2(a,W))^3$ as a sandwich or yields a V such that $G_2(c,d,V) = 0$. This possibility reverts to **Case 1**, **Case 2** or **Case 3** and so gives 5 possible words, so that Case 5 yields 6 possible words.

Case 6: $G_4(a,b,W) \neq 0$ but $G_5(a,b,W) = 0$. Then we either have that $aW(U_2(b)U_2(a,W))^4$ is the sandwich polynomial, or we get a V such that $G_2(c,d,V) = 0$. This leads to 6 possibilities for Case 6 in total.

Case 7: $G_5(a,b,W) \neq 0$ but $G_6(a,b,W) = 0$. Case 7 yields 9 possible sandwich polynomials, corresponding to whether $aW(U_2(b)U_2(a,W))^5$ is our sandwich or whether the construction of v puts us in **Case 1,2,3** or **4**.

Adding all of these possibilities together gives us 29 possible sandwich-valued polynomials.

To work out the length of the string of sandwich-valued polynomials to get from Engel-6 to Engel-7, we calculate as follows. We construct a word W as in the previous subsection. We then construct a sandwich-valued polynomial as above. We add this as a relation and repeat the process. We can perform this process a total of 29 times before we must find that the original W is trivial in the Lie algebra. Then we find another W, which can be any one of the other 15 possibilities. This means that the maximum length of the string of polynomials is $16 \times 29 = 464$.

5.5.4 From Engel-6 to the universal identity

We mentioned earlier that $\tilde{J}$ is a restricted Lie algebra. Thus, if $x \in \tilde{J}$ then so is x^2.

For each of the words we use, we describe them as we actually use them, rather than in their multilinear form. It is always the case that any identity implies a multilinear identity. The converse is not generally the case, however since we are working within $\tilde{J}$ the converse is true and any identity is equivalent to a multilinear one. Thus we can easily turn any of our words into multilinear Lie polynomials.

Our first word will be $f_{t-1} = x^2$. The identity $x^2 = 0$ is equivalent, by the identity $(x+y)^2 = [x,y]$, to the identity $[x,y] = 0$. Therefore, if we enforce the identity $f_{t-1} = 0$, all elements are clearly sandwiches, and thus all values of $f_t = y$ are sandwiches.

Our second word is $f_{t-2} = x^4$. We prove that x^2 is a sandwich when we enforce the identity $x^4 = 0$.

We have the following identities

$$\begin{aligned} [y, x^4] &= [y, x, x, x, x] \\ [y, x^2] &= [y, x, x], \end{aligned}$$

which follow immediately from the definition of restricted Lie algebras. These identities show that $[y, x^2, x^2] = [y, x, x, x, x] = [y, x^4] = 0$.

We now prove that in $\tilde{J}$ the identity $x^4 = 0$ is equivalent to the linearised Engel-3 identity. Suppose $x^4 = 0$ is an identity in $\tilde{J}$, and let $x_i = a_{\pi_i} \otimes e_{\pi_i}$ for $i = 1, 2, 3, 4$. Then

$$(x_1 + x_2 + x_3 + x_4)^2 = [x_4, x_1] + [x_4, x_2] + [x_4, x_3] + [x_3, x_1] + [x_3, x_2] + [x_2, x_1],$$

and so

$$\begin{aligned} (x_1 + x_2 + x_3 + x_4)^4 &= [[x_4, x_3], [x_2, x_1]] + [[x_4, x_2], [x_3, x_1]] \\ &\quad + [[x_4, x_1], [x_3, x_2]] \\ &= \sum_{\sigma \in S_3} [x_4, x_{1\sigma}, x_{2\sigma}, x_{3\sigma}], \end{aligned}$$

by the Jacobi identity and the fact that the ground field has characteristic 2. It now follows that in $\tilde{J}$ the identity $x^4 = 0$ is equivalent to the linearised Engel-3 identity.

Over an infinite field, the Engel-3 identity implies the following identity

$$[a, b, b, c] + [a, b, c, b] + [a, c, b, b] = 0,$$

for all a, b, c. Thus, if we enforce the identity $x^4 = 0$ then

$$\begin{aligned} [y, x^2, z, x^2] &= [y, x, x, z, x, x] \\ &= [y, x, z, x, x, x] + [y, z, x, x, x, x] \\ &= 0, \end{aligned}$$

since we have the Engel-3 identity. Therefore, if $x^4 = 0$, then x^2 is a sandwich.

Our third word is $f_{t-3} = [y, x, x, x, x, x]$. We now prove that if we enforce the Engel-5 identity, then x^4 is a sandwich. First of all,

$$\begin{aligned} [y, x^4, x^4] &= [y, x, x, x, x, x, x, x, x] \\ &= 0. \end{aligned}$$

Also, over an infinite field the Engel-5 identity implies the following identity

$$[a,b,b,b,b,c]+[a,b,b,b,c,b]+[a,b,b,c,b,b]+[a,b,c,b,b,b]+[a,c,b,b,b,b] = 0.$$

Therefore

$$\begin{aligned}[y,x^4,z,x^4] &= [y,x,x,x,x,z,x,x,x,x] \\ &= [y,x,x,x,z,x,x,x,x,x]+[y,x,x,z,x,x,x,x,x,x] \\ &\quad +[y,x,z,x,x,x,x,x,x,x]+[y,z,x,x,x,x,x,x,x,x] \\ &= 0.\end{aligned}$$

Thus, if we enforce the Engel-5 identity, x^4 is a sandwich.

Our fourth word is $f_{t-4} = [y,x,x,x,x,x,x]$. Using the 'lienq' program again, we can calculate the multiweight $(10,2,1,1)$-component of the free Engel-6 Lie algebra on 4 generators (which we call w,x,y and z here) at the prime $p = 2$ (which means using a field of characteristic 2), and check that the words

$$[z,[y,x,x,x,x,x],[y,x,x,x,x,x]],$$

and

$$[z,[y,x,x,x,x,x],w,[y,x,x,x,x,x]],$$

are zero. It turns out that they are. This means that in an Engel-6 Lie algebra over an infinite field of characteristic 2, all values of f_{t-3} are sandwiches. This gives us the string of words which takes us from Engel-6 to $f_t = y$. Combining this with the fact that we have calculated we need at most 464 words to get from Engel-7 to Engel-6, we can say that $t \leq 469$.

5.6 Tying it all together

The crucial result which allows us to tie everything together is the following lemma, adapted from Lemma 7.1 of [48, pp.142–143].

Lemma 5.6.1 *Let J_{i-1} be the free algebra of countably infinite rank determined by the identities $f_0 = 0, \ldots, f_{i-1} = 0$. Let V be the variety generated by $\tilde{J}_{i-1} = J_{i-1} \otimes_K E$. Let M be an arbitrary m-generator Lie algebra in V. Suppose that d_i is defined so that an m generator Lie algebra which satisfies the identities $f_0 = 0, \ldots, f_i = 0$ is nilpotent of class at most $T(m,d_i)$.*

Then M is nilpotent of class at most $T(m, \max\{d_i, 7\} + 7)$.

Proof: Let I be the linear span of all values of the polynomial f_i on M. By Zelmanov's theorem, M is nilpotent, since it is a finitely generated Engel-7 Lie algebra. We could prove this without using Zelmanov's theorem by proving that M/I and I are locally nilpotent and using [46, Lemma 1.2.13] (see [48, p.142]), but since we have Zelmanov's theorem, there is no harm in using it. The minimal number of values of f_i which generate I as an algebra is less than $\dim_K(M/I^2)$. To estimate this, we first estimate the nilpotency class of M/I^2. Every algebra in the variety V satisfies the Engel-7 identity (since $\tilde{J}_{i-1}$ satisfies the identity $f_0 = 0$). In addition, M/I satisfies the identity $f_i = 0$ and so the algebra M/I is nilpotent of class at most $T(m, d_i)$. Let $q = 1 + T(m, d_i)$, let $a_1, \ldots, a_m$ be generators of M and let $r = 4qm^q$. By [48, Lemma 5.1], every product $\mathrm{ad}(a_{i_1}) \cdots \mathrm{ad}(a_{i_r})$ can be represented as a linear combination of terms of the form $v.\mathrm{ad}(\rho).w$, where v and w and words in $\mathrm{ad}(a_1), \ldots, \mathrm{ad}(a_m)$ and ρ is a commutator of length q. Hence,

$$M^{q+r} = M^q \mathrm{ad}(M)^r \subseteq I\mathrm{ad}(M)^r \subseteq I^2.$$

The same argument gives that $M^{q+rs} \subseteq I^{1+s}$ for any positive integer $s \geq 0$. Clearly, $q + r < 5qm^q$ and $\dim_K(M/I^2) < m^{q+r} < m^{5qm^q}$. Every value of the polynomial f_i on $\tilde{J}_{i-1}$ is a sum of at most 2^{370000} sandwiches. Thus, for any k values $a_1, \ldots, a_k$ of the polynomial f_i on $\tilde{J}_{i-1}$, the subalgebra $\langle a_1, \ldots, a_k \rangle$ is nilpotent of class less than $T(2^{370000}k, 4)$, by Theorem 5.4.1. Therefore the algebra I is nilpotent of class less than $T(2^{370000}m^{5qm^q}, 4)$, and M is nilpotent of class less than $q + r.T(2^{370000}m^{5qm^q}, 4)$. Let $b = T(m, \max\{d_i, 7\})$. It is easy to see that

$$\begin{aligned} 2^{370000}m^{5qm^q} &\leq 2^{370000}.m^{5(1+b).m^{1+b}} \\ &< m^{m^{2b}}. \end{aligned}$$

Now M is nilpotent of class at most

$$\begin{aligned} q + r.T(m^{m^{2b}}, 4) &< (r+1).T(m^{m^{2b}}, 4) \text{ by Lemma [48, Lemma 5.5]} \\ &< ((r+1)/m^{4b}).(m \hat{} m \hat{} \cdots^5 \cdots \hat{} m \hat{} m^{2b}) \\ &< m \hat{} m \hat{} \cdots^5 \cdots \hat{} m \hat{} (m \hat{} m \hat{} b) \\ &< T(m, \max\{d_i, 7\} + 7). \end{aligned}$$

This completes the proof of the lemma. ■

Now, from Lemma 5.6.1 and Theorem 5.3.1 we conclude that

$$d_{i-1} \leq 7(\max\{d_i, 7\} + 13). \tag{5.16}$$

We can clearly take $d_t = 1$, however given (5.16) it makes no difference to take $d_t = 7$. Then

$$\begin{aligned} d_0 &\leq 7^{t+1} + 13(7^{t+1} - 7)/(7-1) \\ &< 7^{t+2} - 2 \end{aligned}$$

Given that $t \leq 469$, we have the following theorem.

Theorem 5.6.2 *Let G be a finite group of exponent 8, generated by m elements. Then G is nilpotent of class at most $T(m, 7^{471}) - 2$.*

It is now very easy to surmise Theorem 5.1.2 from Theorem 5.6.2. We know that the homogeneous component of weight i of a graded Lie algebra is spanned by left-normed commutators in the generators of weight i (again, see [46, Section 1.2]). It is easy to see that there are less than m^i of these for $i > 1$. If $c = T(m, 7^{471} - 2)$ then $L(G)^c = 0$, and we have

$$\begin{aligned} \dim_K L(G) &< \sum_{i=1}^{c-1} m^i \\ &< m^c. \end{aligned}$$

Thus,

$$\begin{aligned} |G| = |L(G)| &< m^{m^c} \\ &= T(m, 7^{471}), \end{aligned}$$

which is the content of Theorem 5.1.2.

5.7 Disappointment

The final bound obtained in Theorem 5.1.2 is somewhat disappointing. A lot of work was put into improving the numbers involved in a variety of places in the proof, without this filtering through to an improvement in the final result.

Using the above method, the final result will yield a bound on the number of m's in the tower of approximately 7^t, where t is the length of the string of sandwich-valued polynomials. We have made a reduction of 8 to 7 as well

as a reduction in the bound on t. It is the bound on t which makes the real difference, however.

The reduction from $8^8 - 5$ (which is 16777211) to 469 is obviously immense. However, the numbers involved in the final bounds are still unimaginably large. The hope was that the number of U-polynomials needed would be much smaller. Unfortunately, calculations showed that this was not going to be possible using this method, since $v = \mathrm{ad}(x_1)^6$ did not satisfy (5.15) for $k = 5$. Whether or not v satisfies (5.15) for $k = 6$ is a question which appears to be beyond the scope of computational power available to me. Thus we used the result from the general theory that (5.15) is satisfied for $k = 7$. If it turns out that (5.15) is satisfied for $k = 6$, this would lead to an immediate reduction of the length of U-polynomials between Engel-7 and Engel-6 by a factor of 2. However, this remains an open question.

Appendix A

Appendix: PCP for J from Chapter 3

We give a consistent PCP for J with 21 generators, denoted $j_1, j_2, \ldots, j_{21}$. It is easy to see that J is generated by j_1, j_2 and j_3. Any squares of generators, or commutators of the form $[j_i, j_k]$ with $i > k$, which are not listed are trivial. For ease of notation, we denote j_i by the integer i.

$4^2 = 6,\ 5^2 = 8,$

$7^2 = 9 * 11 * 15 * 16 * 19 * 21,\ 10^2 = 19,$

$[2,1] = 4\ , [3,1] = 5,\ [4,1] = 6,\ [4,2] = 6,\ [4,3] = 7,\ [5,1] = 8,$

$[5,2] = 7 * 9 * 10 * 11 * 15 * 20,$

$[5,3] = 8,\ [5,4] = 9 * 10 * 12 * 13 * 14 * 15 * 16 * 18 * 19 * 20,$

$[6,3] = 9,\ [6,5] = 13 * 19 * 20,$

$[7,1] = 10,\ [7,2] = 9 * 15 * 20,$

$[7,3] = 9 * 11 * 15 * 16 * 19 * 21,$

$[7,4] = 11 * 15 * 16 * 18 * 19,$

$[7,5] = 12 * 14 * 16 * 19 * 20 * 21,$

$[7,6] = 15 * 20,$

$[8,2] = 9 * 11 * 12 * 15 * 16 * 17 * 21,$

$[8,4] = 14 * 16 * 20 * 21,\ [8,7] = 17,$

$[9,1] = 13 * 16 * 19 * 20 * 21,\ [9,4] = 15 * 18 * 20 * 21,$

$[9,5] = 16 * 21,\ [9,7] = 18 * 21,\ [10,1] = 19,\ [10,2] = 11,\ [10,3] = 12,$

$[10,4] = 13 * 15 * 21,\ [10,5] = 14 * 17 * 19,\ [10,7] = 16 * 20 * 21,$

$[10,8] = 19,\ [11,1] = 13,$

$[11,3] = 18,\ [11,4] = 15,\ [11,5] = 16 * 20 * 21,$

$[11,7] = 18,\ [11,10] = 21,\ [12,1] = 14,\ [12,2] = 18,$

$[12,4] = 18,\ [12,5] = 17 * 19,$
$[13,2] = 15,\ [13,3] = 16,$
$[13,7] = 20 * 21,\ [14,2] = 18,\ [14,3] = 17,$
$[14,4] = 20,\ [14,5] = 19,\ [15,3] = 18,\ [15,5] = 20 * 21,$
$[16,2] = 18,\ [16,4] = 20,\ [16,7] = 21, [17,1] = 19,\ [17,4] = 21,$
$[18,1] = 20,\ [18,5] = 21, [19,2] = 21,\ [20,3] = 21.$

Appendix B

Appendix: Proof of Theorem 5.5.1

Proof of Theorem 5.5.1: We prove the theorem separately for each value of m.

Now, because $w^{(1)}$ satisfies (5.14), we will be considering elements of the form $\mathrm{ad}(aU_k(A))$ where $a_i a_j = 0$ for all $a_i, a_j \in A$. Therefore, we can use the following identity, from page 168 of [46] (noting that because the elements of A commute with each other $\bar{U}_i(A) = U_i(A)$ for all i, and that because we are in characteristic 2, we can ignore the sign):

$$\mathrm{ad}(aU_k(A)) = \sum_{i=0}^{k} U_i(A)\mathrm{ad}(a)U_{k-i}(A). \tag{B.1}$$

We also have

$$U_s(A)U_t(A) = \binom{s+t}{s} U_{s+t}(A). \tag{B.2}$$

If we assume that $U_k(A) = 0$ for all $k \geq m$ (which will also be the case for us), then

$$U_{m-1}(A).\mathrm{ad}(a).U_{m-1}(A) = 0, \tag{B.3}$$

$$U_{m-1}(A).\mathrm{ad}(a).U_{m-2}(A) = U_{m-2}(A).\mathrm{ad}(a).U_{m-1}(A), \tag{B.4}$$

and

$$\begin{aligned} &U_{m-2}(A).\mathrm{ad}(a).U_{m-2}(A) \\ &= U_{m-1}(A).\mathrm{ad}(a).U_{m-3}(A) + U_{m-3}(A).\mathrm{ad}(a).U_{m-1}(A). \end{aligned} \tag{B.5}$$

The proofs of all of these may be found in [46], Section 9.2. The following also comes from the same section:

Lemma B.0.1 *Suppose that $m \geq 4$ and that $U_k(A) = 0$ for all $k \geq m$. Let $b_1, \ldots, b_s \in \tilde{L}$ and $k_0, \ldots, k_s$ be non-negative integers such that*

$$k_0 + \cdots + k_s > (m-2)(s+1) + 1.$$

Then

$$U_{k_0}(A).\mathrm{ad}(b_1).U_{k_1}(A).\mathrm{ad}(b_2).\cdots.\mathrm{ad}(b_s)U_{k_s}(A) = 0.$$

We will use the fact that, for whatever m we are using, 5.15 holds for $k \geq m$, and the above two identities.

Now, we write $\bar{y} = y_{r+1}w(y_1, \ldots, y_r)$. Then we have, from the definition of $w^{(2)}$,

$$w^{(2)} = \sum_{i_1=0}^{m-1} \cdots \sum_{i_{m-1}=0}^{m-1} \Big(U_{i_1}(\bar{y})\mathrm{ad}(y_{r+2})U_{m-1-i_1}(\bar{y}).U_{i_2}(\bar{y})\mathrm{ad}(y_{r+3})U_{m-1-i_2}(\bar{y}) \cdots U_{i_{m-1}}(\bar{y})\mathrm{ad}(y_{r+m})U_{m-1-i_{m-1}}(\bar{y}) \Big).$$

We introduce the following notation, which will shorten our calculations tremendously. In place of

$$U_{j_0}(\bar{y})\mathrm{ad}(a_1)U_{j_2}(\bar{y})\mathrm{ad}(a_2)\cdots U_{j_{k-1}}(\bar{y})\mathrm{ad}(a_k)U_{j_k}(\bar{y}),$$

we write

$$(j_0, j_1, \ldots, j_k).$$

This ignores what the variables $a_1, \ldots, a_k$ are, but we will be able to prove the theorem anyway.

Once we are ignoring the variables $a_1, \ldots, a_k$, we can introduce a convention for multiplying expressions $(j_0, \ldots, j_k)$ and $(l_0, \ldots, l_i)$ together. We will shift the subscripts of the variables in $(l_0, \ldots, l_i)$ by k. Thus we define

$$(j_0, \ldots, j_k)(l_0, \ldots, l_i) = U_{j_0}(\bar{y})\mathrm{ad}(y_1)\cdots\mathrm{ad}(y_k)U_{j_k}(\bar{y})U_{l_0}(\bar{y})\mathrm{ad}(y_{k+1})\cdots\mathrm{ad}(y_{k+i})U_{l_i}(\bar{y}).$$

We will see below that (B.2) allows us to write the above as a multiple of $(j_0, \ldots, j_{k-1}, j_k + l_0, l_1, \ldots, l_i)$.

We now translate the above identities into our new notation. First,

$$w^{(2)} = \sum_{i_1=0}^{m-1} \cdots \sum_{i_{m-1}=0}^{m-1} (i_1, m-1-i_1) \cdots (i_{m-1}, m-1-i_{m-1}).$$

Also, by (B.2), (B.3), (B.4) and Lemma B.0.1, we have

$$(j_1, \ldots, j_k)(h_1, \ldots, h_l) = \binom{j_k + h_1}{j_k} (j_1, \ldots, j_{k-1}, j_k + h_1, h_2, \ldots, h_l), \quad \text{(B.6)}$$

$$(m-1, m-1) = 0, \quad \text{(B.7)}$$

$$(m-1, m-2) = (m-2, m-1), \quad \text{(B.8)}$$

$$(m-2, m-2) = (m-1, m-3) + (m-3, m-1), \quad \text{(B.9)}$$

and

Lemma B.0.2 *Let $k_0, \ldots, k_s$ be non-negative integers such that*

$$k_0 + \cdots + k_s > (m-2)(s+1) + 1.$$

Then

$$(k_0, k_1, \ldots, k_s) = 0. \quad \text{(B.10)}$$

We use the convention of underlining something which we know to be zero from one of the relations (B.7),(B.8),(B.9) and (B.10) or from the fact that our field is of characteristic 2 (sometimes there may be more than one reason for an element to be trivial).

We now move on to the actual calculations.

Case 1: $m = 4$.

We reduce all of our coefficients modulo 2.

$$\begin{aligned} w^{(2)} &= \sum_{i_1=0}^{3} \sum_{i_2=0}^{3} \sum_{i_3=0}^{3} (i_1, 3-i_1)(i_2, 3-i_2)(i_3, 3-i_3) \\ &= \sum_{i_1=0}^{3} \sum_{i_2=0}^{3} (i_1, 3-i_1)(i_2, 3-i_2)\Big[(0,3) + (1,2) + (2,1) + (3,0)\Big] \end{aligned}$$

$$
\begin{aligned}
= \quad & \Big[\sum_{i_1=0}^{3}(i_1, 3-i_1)\Big]. \\
& \Big[(0,3)(0,3)+(0,3)(1,2)+(0,3)(2,1)+(0,3)(3,0)+(1,2)(0,3) \\
& +(1,2)(1,2)+(1,2)(2,1)+(1,2)(3,0)+(2,1)(0,3)+(2,1)(1,2) \\
& +(2,1)(2,1)+(2,1)(3,0)+(3,0)(0,3)+(3,0)(1,2)+(3,0)(2,1) \\
& +(3,0)(3,0)\Big] \\
= \quad & \Big[\sum_{i_1=0}^{3}(i_1, 3-i_1)\Big]. \\
& \Big[(0,\underline{3,3})+\binom{4}{1}(0,\underline{4},2)+\binom{5}{2}(0,\underline{5},2)+\binom{6}{3}(0,\underline{6},0)+(1,2,3) \\
& +(1,3,2)+\binom{4}{2}(1,\underline{4},1)+\binom{5}{2}(1,\underline{5},0)+(2,1,3) \\
& +\underline{\binom{2}{1}}(2,2,2)+(2,3,1)+\binom{4}{1}(2,\underline{4},0)+(3,0,3) \\
& +(3,1,2)+(3,2,1)+(\underline{3,3},0)\Big] \\
= \quad & \Big[\sum_{i_1=0}^{3}(i_1, 3-i_1)\Big]. \\
& \Big[(1,2,3)+(1,3,2)+(2,1,3)+(2,3,1) \\
& +(3,0,3)+(3,1,2)+(3,2,1)\Big] \\
= \quad & (0,3)(1,2,3)+(0,3)(1,3,2)+(0,3)(2,1,3)+(0,3)(2,3,1) \\
& +(0,3)(3,0,3)+(0,3)(3,1,2)+(0,3)(3,2,1)+(1,2)(1,2,3) \\
& +(1,2)(1,3,2)+(1,2)(2,1,3)+(1,2)(2,3,1)+(1,2)(3,0,3) \\
& +(1,2)(3,1,2)+(1,2)(3,2,1)+(2,1)(1,2,3)+(2,1)(1,3,2) \\
& +(2,1)(2,1,3)+(2,1)(2,3,1)+(2,1)(3,0,3)+(2,1)(3,1,2) \\
& +(2,1)(3,2,1)+(3,0)(1,2,3)+(3,0)(1,3,2)+(3,0)(2,1,3) \\
& +(3,0)(2,3,1)+(3,0)(3,0,3)+(3,0)(3,1,2)+(3,0)(3,2,1) \\
= \quad & \binom{4}{1}(0,\underline{4},2,3)+\binom{4}{1}(0,\underline{4},3,2)+\binom{5}{2}(0,\underline{5},1,3)+\binom{5}{2}(0,\underline{5},3,1)
\end{aligned}
$$

$$
\begin{aligned}
&+\binom{6}{3}(0,\underline{6},0,3)+\binom{6}{3}(0,\underline{6},1,2)+\binom{6}{3}(0,\underline{6},2,1)+(1,\underline{3,2,3})\\
&+(1,\underline{3,3},2)+\binom{4}{2}(1,\underline{4},1,3)+\binom{4}{2}(1,\underline{4},3,1)+\binom{5}{2}(1,\underline{5},0,3)\\
&+\binom{5}{2}(1,\underline{5},1,2)+\binom{5}{2}(1,\underline{5},2,1)+\underline{\binom{2}{1}}(2,2,2,3)+\underline{\binom{2}{1}}(2,2,3,2)\\
&+(2,3,1,3)+(2,\underline{3,3},1)+\binom{4}{1}(2,\underline{4},0,3)+\binom{4}{1}(2,\underline{4},1,2)\\
&+\binom{4}{1}(2,\underline{4},2,1)+(3,1,2,3)+(3,1,3,2)+(3,2,1,3)\\
&+(\underline{3,2,3},1)+(\underline{3,3},0,3)+(\underline{3,3},1,2)+(\underline{3,3},2,1)\\
=\;&(2,3,1,3)+(3,1,2,3)+(3,1,3,2)+(3,2,1,3)\\
=\;&0 \text{ by (B.8).}
\end{aligned}
$$

This completes the proof for $m=4$. For $m=5,6$ and 7, we will abbreviate the proof by leaving out some of the terms which are more obviously zero. For example, if a term starts with $(m-1,0)(m-1,0)$, then it will be zero. Also, any term starting with $(0,m-1)$ will be zero, since it must either be followed by $(i,m-i-1)$ where $i\neq 0$, in which case we have $(0,\underline{m-1+i},m-1-i)$, or by $(0,m-1)$, in which case we have $(0,\underline{m-1,m-1})$.

Case 2: $m=5$.

We have omitted any terms which can be immediately seen to be zero by one or more of the above reasons, or by characteristic considerations. For example, the term

$$
\begin{aligned}
(3,1)(2,2)(2,2)(1,3) &= \binom{3}{1}.\binom{4}{2}.\binom{3}{1}(3,3,4,3,3)\\
&= 0,
\end{aligned}
$$

because $2|\binom{4}{2}$. Also,

$$
\begin{aligned}
(4,0)(3,1)(3,1)(2,2) &= \binom{4}{1}.\binom{3}{1}(4,3,4,3,2)\\
&= 0,
\end{aligned}
$$

either because $2|\binom{4}{1}$ or because $(\underline{4,3,4},3,2)=0$. We now move on to the

actual calculations of $w^{(2)}$.

$$\begin{aligned} w^{(2)} &= \sum_{i_1=0}^{4}\sum_{i_2=0}^{4}\sum_{i_3=0}^{4}\sum_{i_4=0}^{4}(i_1, 4-i_1)(i_2, 4-i_2)(i_3, 4-i_3)(i_4, 4-i_4) \\ &= (4,0)(3,1)(2,2)(0,4) + (4,0)(3,1)(2,2)(1,3) \\ &\quad +(4,0)(2,2)(1,3)(0,4) + (3,1)(2,2)(1,3)(0,4) \\ &= (4,3,3,2,4) + (4,3,3,3,3) + (4,2,3,3,4) + (3,3,3,3,4). \end{aligned}$$

Now, by (B.9),

$$\begin{aligned} (3,3,3,3,4) &= (4,2,3,3,4) + (2,\underline{4,3,3,4}), \text{ and,} \\ (4,3,3,3,3) &= (\underline{4,3,3,4},2) + (4,3,3,2,4), \end{aligned}$$

which is to say that $w^{(2)} = 0$.

Case 3: $m = 6$ or 7.

For these, any term in the expansion of $w^{(2)}$ will be zero, by dint of either characteristic considerations, or the above identities (B.6)-(B.10).

For example, when $m = 6$,

$$\begin{aligned} (5,0)(4,1)(4,1)(4,1)(2,2) &= (\underline{5,4,5},5,3,2), \\ (5,0)(4,1)(3,2)(2,3)(1,4) &= \binom{4}{1}^2 \cdot \binom{4}{2}(5,4,4,4,4,4), \\ (5,0)(4,1)(2,3)(2,3)(0,5) &= \binom{5}{2}(5,4,3,5,3,5). \end{aligned}$$

Since $2|\binom{4}{2}$ and $2|\binom{5}{2}$, the second and third of these are zero. The first is zero because of the underlined $(5,4,5)$. All other terms in the expressions are zero for similar reasons, and the case $m = 7$ works similarly.

This completes the proof of the theorem. ■

Bibliography

[1] S.I. Adjan and A.A. Razborov, Periodic groups and Lie algebras, *Uspekhi Mat. Nauk*, **42** (1987), 3-68.

[2] Yu. A. Bahturin, "Identical relations in Lie algebras", VNU Science Press, Utrecht, 1987.

[3] W. Burnside, On an unsettled question in the theory of discontinuous groups, *The Quarterly Journal of Pure and Applied Mathematics*, **33** (1902), 230-238.

[4] B. Chandler and W. Magnus, "The history of combinatorial group theory", Springer-Verlag, New York, 1982.

[5] Yu. M. Gorchakov, Commutator subgroups, *Sibirsk. Mat. Zh.* (Russian), **10** (1969), 1023-1033.

[6] D. Gorenstein, "Finite simple groups", Plenum Press, New York, 1982.

[7] D. P. Groves, Lie relators for varieties of groups, Change of status dissertation, University of Oxford, 1998.

[8] D. P. Groves, A note on non-identical Lie relators, *Journal of Algebra*, **211** (1999), 15-25.

[9] K. W. Gruenberg, Residual properties of infinite soluble groups, *Proc. London Math. Soc. (3)*, **7** (1957), 29-62.

[10] C. K. Gupta, The free centre-by-metabelian groups, *J. Austral. Math. Soc.*, **16** (1973), 294-299.

[11] M. Hall Jr., "The theory of groups", Macmillan, New York, 1959.

[12] M. Hall Jr., Solution of the Burnside problem for exponent six, *Illinois J. Math.*, **2** (1958), 764-786.

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

[14] P. Hall, Finiteness conditions for soluble groups, *Proc. London Math. Soc. (3)*, **4** (1954), 419-436.

[15] P. Hall and G. Higman, On the p-length of p-soluble groups and reduction theorems for Burnside's problem, *Proc. London Math. Soc., (3)* **6** (1956), 1-42.

[16] G. Havas and M.F. Newman, Applications of computers to questions like those of Burnside, *in*, "Lecture Notes in Mathematics", Vol. 806, Springer-Verlag, Berlin, 1980.

[17] G. Havas, G.E. Wall and J.W. Wamlsey, The two generator Burnside Group of exponent five, *Bull. Austral. Math. Soc.*, **10** (1974), 459-470.

[18] G. Havas, M.F. Newman and M.R. Vaughan-Lee, A nilpotent quotient algorithm for graded Lie rings, *J. Symbolic Computation*, **9** (1990), 653-664.

[19] G. Higman, Some remarks on varieties of groups, *Quart. J. Math. Oxford (2)*, **10** (1959), 165-78.

[20] S.V. Ivanov, The free Burnside groups of sufficiently large exponents, *Internat. J. Algebra Comput.*, **4** (1994), no.1-2, ii + 308 pp.

[21] N. Jacobson, "Lie Algebras", Interscience, New York, 1962.

[22] E. B. Kikodze, Free groups of certain varieties, *Algebra i Logika Sem.*, **5** (1966), 15-23.

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

[24] A.I. Kostrikin, "Around Burnside", Ergebnisse der Math., Springer-Verlag, Berlin, 1990, translated from the Russian by James Wiegold.

[25] Yu. V. Kuz'min, Some approximation properties of the variety $\mathfrak{UV}_c$, *Uspekhi Mat. Nauk*, **33** (1978), 217-218.

[26] Yu. V. Kuz'min and M.Z. Shapiro, The connection between varieties of groups and varieties of Lie rings, *Sibirsk. Mat. Zh.* (Russian), **28** (1987), 100-101.

[27] M. Lazard, Sur les groupes nilpotents et les anneaux de Lie, *Ann. Sci. Éc. Norm. Supér III*, **71** (1954), 101-190.

[28] I. G. Lysënok, Infinite Burnside groups of even period, *Izv. Ross. Akad. Nauk Ser. Mat.*, **60** (1996), 3-224.

[29] W. Magnus, A. Karrass and D. Solitar, "Combinatorial group theory", Interscience, New York, 1966.

[30] H. Neumann, "Varieties of groups", Ergebnisse der Math., Springer-Verlag, Berlin, 1967.

[31] M.F. Newman, Groups of exponent 6, *Computational group theory (Durham, 1982)*, 39–41, Academic Press, London-New York, 1984.

[32] M.F. Newman and M.R. Vaughan-Lee, Some Lie rings associated with Burnside groups, *Electron. Res. Announc. Amer. Math. Soc.*, **4** (1998), 1-3.

[33] M.F. Newman and M.R. Vaughan-Lee, Engel-4 groups of exponent 5 II. Orders, *Proc. London Math. Soc.*, **79** (1999), 283-317.

[34] P.S. Novikov and S.I. Adjan, Infinite periodic groups I, *Izv. Akad. Nauk. SSSR, Ser. Mat.*, **32** (1968), 212-244.

[35] P.S. Novikov and S.I. Adjan, Infinite periodic groups II, *Izv. Akad. Nauk. SSSR, Ser. Mat.*, **32** (1968), 251-524.

[36] P.S. Novikov and S.I. Adjan, Infinite periodic groups III, *Izv. Akad. Nauk. SSSR, Ser. Mat.*, **32** (1968), 709-731.

[37] M.R. Quick, Identical relations in groups of exponent four, *Change of status dissertation*, University of Oxford, 1992.

[38] M. Schönert et al., "GAP - groups, algorithms and programming", Lehrstuhl D für Mathematik, Rheinisch Westfälische Technische Hochschule, Aachen, Germany, fifth edition, 1995.

[39] I.N. Sanov, Solution of Burnside's problem for exponent four, *Leningrad State Univ. Ann. Math. Ser.*, **10** (1940), 166-170.

[40] I.N. Sanov, A certain system of relations in periodic groups of prime-power exponent, *Izv. Akad. Nauk. SSSR, Ser. Mat.*, **15** (1951), 477-502.

[41] A. L. Smel'kin, Free polynilpotent groups, *Dokl. Akad. Nauk SSSR*, **151** (1963) 73-75.

[42] R. Stöhr, On torsion in free central extensions of some torsion-free groups, *J. Pure Appl. Algebra*, **46** (1987), 249-289.

[43] V. Tasić and M. R. Vaughan-Lee, Torsion in certain relatively free groups, *Bull. London Math. Soc.*, **27** (1995), 327-333.

[44] Yu. D. Ushakov, On the free group of a commutator of two nilpotent varieties, *Sibirsk. Mat. Zh.*, **14** (1973), 636-650.

[45] M.R. Vaughan-Lee, Lie rings of groups of prime exponent, *J. Austral. Math. Soc.*, **49** (1990), 386-398.

[46] M.R. Vaughan-Lee, "The restricted Burnside problem (second edition)", Clarendon Press, Oxford, 1993

[47] M.R. Vaughan-Lee, On Zelmanov's solution of the restricted Burnside problem, *Journal of Group Theory*, **1** (1998), 65-94.

[48] M.R. Vaughan-Lee and E.I. Zelmanov, Upper bounds in the restricted Burnside problem, *Journal of Algebra*, **162** (1993), 107-145.

[49] M.R. Vaughan-Lee and E.I. Zelmanov, Upper bounds in the restricted Burnside problem II, *International Journal of Algebra and Computation*, **6** (1996), 735-744.

[50] M.R. Vaughan-Lee, Computing $R(2,7)$, Algebra Seminar, Mathematical Institute, Oxford, February 22, 2000.

[51] G.E. Wall, Lie methods in group theory, *in* "Lecture Note in Mathematics", Vol. 697, Springer-Verlag, Berlin, 1978.

[52] G.E. Wall, Multilinear Lie relators for varieties of groups, *Journal of Algebra*, **157** (1993), 341-393.

[53] E.I. Zelmanov, The solution of the restricted Burnside problem for groups of odd exponent, *Izv. Math. USSR*, **36** (1991), 41-60.

[54] E.I. Zelmanov, The solution of the restricted Burnside problem for 2-Groups, *Mat. Sb.*, **182** (1991), 568-592.

[55] E.I. Zelmanov and A.I. Kostrikin, A theorem on sandwich algebras (Russian), Translated in *Proc. Steklov Inst. Math.* (1991), no.4, 121-126.