

A FIRST COURSE ON BURNSIDE-TYPE PROBLEMS IN GROUPS, LIE ALGEBRAS AND ASSOCIATIVE ALGEBRAS

CHRISTIAN D'ELBÉE

DESCRIPTION OF THE COURSE

Those are the notes of a course given by the author to several places, including the *Nesin Matematik Köyü* in September 2025 and September 2026. It was also given and the *Master de Mathématiques option Algèbre* of the *Université des Sciences et Techniques de Masuku* in March 2026. The goal of the course is to introduce the classical results and methods revolving around Burnside-type questions. In the first part of the course, we introduce associative algebras of finite nil exponent, n -Engel Lie algebras and groups of finite exponent. In the second part of the course, we detailed the construction of the associated Lie algebra of a group, as well as the reduction of the Restricted Burnside Problem to the local nilpotency of the associated Lie algebra of a group of bounded exponent.

CONTENTS

Description of the course	1
1. Introduction	2
Groups of exponent 2 and 3	2
2. Associative algebras and the Nagata-Higman-Dubnov-Ivanov Theorem	4
3. Lie algebras and enveloping algebras	7
4. First footsteps with n -Engel Lie algebras	9
5. The restricted Burnside problem and n -Engel Lie algebras	11
5.1. The associated Lie algebra of a group of exponent p	11
5.2. Varieties of groups, free Burnside group	13
5.3. Connection with the restricted Burnside problem	14
6. Problem	17
Appendix A. More on n -Engel Lie algebras	18
A.1. Varieties of Lie algebras, free Lie algebras	18
A.2. Higgins Theorem	19
A.3. $\mathbb{Z}_2$ -gradings on Engel Lie algebras	20
A.4. Group algebra, Young tableau and representation theory of the symmetric group	20
A.5. A theorem of Traustason on 4-Engel Lie algebras	22
Appendix B. Zelmanov Characteristic 0 Theorem	25
B.1. Step 1: Establishing the induction scheme	25
B.2. Step 2: First footstep in the induction step	26
B.3. Step 3: Linearization	26
B.4. Step 4: Reduction to symmetric/skew-symmetric sums	27
B.5. Remark: An escape route	31
B.6. Step 5: Symmetrization using $\mathbb{Z}_2$ -gradings	32
B.7. Step 6: Conclusion	33
References	34

1. INTRODUCTION

In 1902, William Burnside raised his famous problems [3].

ON AN UNSETTLED QUESTION IN THE THEORY OF DISCONTINUOUS GROUPS.

By W. BURNSIDE.

A STILL undecided point in the theory of discontinuous groups is whether the order of a group may be not finite, while the order of every operation it contains is finite. A special form of this question may be stated as follows:—

Let $A_1, A_2, \dots, A_m$ be a set of independent operations finite in number, and suppose that they satisfy the system of relations given by

$$S^n = 1,$$

where n is a given finite integer, while S represents in turn any and every operation which can be generated from the m given operations A .

Is the group thus defined one of finite order, and if so what is its order?

In other words:

The General Burnside Problem. *Is every finitely generated group where each element has finite order necessarily finite?*

The Burnside Problem. *Is every finitely generated group of bounded exponent necessarily finite?*

(Recall that the *exponent* of a group G is, if it exists, the smallest $n \in \mathbb{N}$ such that $g^n = 1$, for all $g \in G$.)

In 1964, Golod and Shafarevich [5] constructed a counterexample to the general Burnside problem. As for the Burnside problem, it was settled negatively a few years later in 1968 by Novikov and Adian [16, 17, 18]. It is known that for $r \geq 2$ and $n \geq 8000$, r -generated groups of exponent n are infinite in general [13, 1]. A recent result from Atkarskaya, Rips and Tent [2] yields that r -generated groups of odd exponent ≥ 557 are infinite for $r \geq 2$.

Nonetheless, it is known that the Burnside problem has a positive answer for some values of the exponent. In fact, it is known that all finitely generated group of exponent $n = 2, 3, 4, 6$ are finite. Note that it is still open whether 2-generated group of exponent 5 can be infinite (!). Let us take a closer look at groups of exponent 2 and 3.

Groups of exponent 2 and 3. Before anything, let us go through some notations and easy facts.

- A group G is *abelian* if $gh = hg$ for all $g, h \in G$. Equivalently, $g^{-1}h^{-1}gh = 1$ for all $g, h \in G$. We denote $[g, h] := g^{-1}h^{-1}gh$ the *commutator* of g and h . Of course, g and h commute if $[g, h] = 1$.
- (Exercise) An abelian group of exponent 2 is a vector space over the field $\mathbb{F}_2$.

Proposition 1.1. *A finitely generated group of exponent 2 is finite.*

Proof. Let us check first that a group of exponent 2 is abelian. If G is of exponent 2, then for any $a, b \in G$, we have $(ab)^2 = 1$ hence $ab = b^{-1}a^{-1}$. Then $[a, b] = a^{-1}b^{-1}ab = a^{-1}b^{-1}b^{-1}a^{-1} = a^{-1}(b^{-1})^2a^{-1} = (a^{-1})^2 = 1$. Therefore G is abelian. Now an abelian group of exponent 2 is a vector space over the finite field $\mathbb{F}_2$, which, if r -generated, has order $\leq 2^r$. We conclude that a finitely generated group of exponent 2 is finite. $\square$

We will see later in the course why finitely generated groups of exponent 3 are finite. This will be linked to the n -Engel identity, which will be at the core of this course, and that we introduce now.

- Let G be a group and $g, h \in G$. Then $g^h := h^{-1}gh$ is called the *conjugate* of g by h . We also denote a^{-b} the element $b^{-1}a^{-1}b$. Note that $(a^b)^{-1} = a^{-b}$.
- The map $x \mapsto x^h$ is an automorphism of G .

- We define inductively the notation $[x_1, \dots, x_n] := [[x_1, \dots, x_{n-1}], x_n]$. If $x_1 = x$ and $x_2 = x_3 = \dots = x_n$, we use the notation:

$$[x, y, \dots, y] = [x, y^{(n)}].$$

- An *identity* of a group is an equation satisfied by every element of the group. For instance, a group is abelian if and only if it satisfies the identity $[x, y] = 1$.
- The identity $[x, y^{(n)}] = 1$ is called the n -Engel identity.

Proposition 1.2. Any group of exponent 3 satisfies the 2-Engel identity:

$$[x, y, y] = 1.$$

Proof. If G is of exponent 3, then for arbitrary $a, b \in G$ we have $1 = (ba)^3 = bababa = b^3b^{-2}ababa = b^3b^{-2}ab^2b^{-1}aba$. We conclude:

$$a^{b^2}a^ba = 1. \quad (\dagger)$$

By replacing b by b^2 in $(\dagger)$, we get

$$a^ba^{b^2}a = 1. \quad (\ddagger)$$

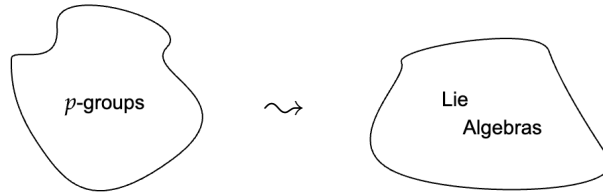
From $(\dagger)$ we obtain $a^{b^2} = a^{-1}a^{-b}$ which, plugged into $(\ddagger)$ yields $a^ba^{-1}a^{-b}a = 1$, i.e. $[a^{-b}, a] = 1$. We conclude that a^{-b} and a commute. As a trivially commutes with itself, a commutes with $a^{-b}a = [b, a]$. It follows that $[[b, a], a] = 1$. As a, b were arbitrary, we conclude that $[x, y, y] = 1$ is an identity of G . $\square$

In the 1930s, mathematicians started asking a more reasonable variant of the Burnside problem.

The Restricted Burnside Problem. For fixed $r, n \in \mathbb{N}$, are there only finitely many finite r -generated groups of exponent n ?

Using a result of Hall and Higman [7] (and the classification of finite simple groups), the restricted Burnside problem reduces to the case of p -groups, i.e. groups where every element is of order a power of a prime p . This was solved positively in 1959 by Kostrikin [10, 11, 12] (announced in 1958 in [11]) for groups of prime exponents and by Zelmanov [26, 27] in 1991 for the general case of p -groups. For that work, Zelmanov was awarded a Fields Medal at the International Congress of Mathematicians in Zürich in 1994.

Both Kostrikin and Zelmanov's result are in fact theorems about Lie algebras. The solution of the RBP is a perfect illustration of a beautiful method of reduction of group-theoretic questions to Lie algebra questions, usually called *Lie methods in group theory*.



Definition 1.3. An algebra $(V, +, 0, *)$ over a field $\mathbb{F}$ is an $\mathbb{F}$ -vector space $(V, +)$ equipped with a bilinear map $(x, y) \mapsto x * y$, that is, which satisfies

$$(\lambda x + \mu y) * z = \lambda x * z + \mu y * z \quad z * (\lambda x + \mu y) = \lambda z * x + \mu z * y$$

Definition 1.4. A Lie algebra is an algebra $(L, +, 0, [\cdot, \cdot])$ in which the bilinear map $(x, y) \mapsto [x, y]$ satisfies $[x, x] = 0$ and the Jacobi identity

$$[x, y, z] + [y, z, x] + [z, x, y] = 0.$$

The basic idea is to associate to a given p -group G a certain Lie algebra $L(G)$:

$$G \rightsquigarrow L(G)$$

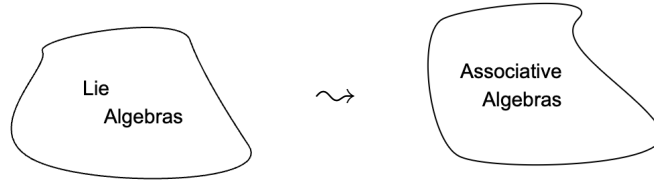
and to prove that the RBP holds for groups of prime exponent if and only if the associated Lie algebra of such groups is nilpotent. Then the turning point is given by Theorem 5.17: the associated Lie algebra of a group of exponent p satisfies the $(p-1)$ -Engel identity, which is the following:

$$\underbrace{[\dots [x, y], y], \dots, y]}_{(p-1) \text{ times}} = 0.$$

At this point, we will be able to make apparent the link between the RBP and Engel Lie algebras, which is at the core of the solution of the RBP

To solve the RBP for groups of exponent p , it suffices to prove that finitely generated $(p-1)$ -Engel Lie algebras over $\mathbb{F}_p$ are nilpotent.

We will therefore concentrate our course on understanding n -Engel Lie algebras. We will carry out the computing that 2-Engel Lie algebras and 3-Engel Lie algebras of characteristic different from 2 and 5 are nilpotent, which are due to Higgins [8]. In fact, in those cases, the finitely generated assumption is superfluous. We will also prove various results on n -Engel Lie algebras, such as another result of Higgins which states that nilpotency and solvability coincide for n -Engel Lie algebras of characteristic $p > n$. In order to palliate the lack of associativity of a Lie algebra L , we will work in the algebra of linear endomorphisms $\text{End}(L)$, which is an associative algebra. This will be done by considering enveloping algebras. Therefore, before starting the study of Lie algebras, we will



start with basics on associative algebras.

At the very core of the study of Engel Lie algebras lies Zelmanov's *characteristic 0* theorem [25].

Theorem (Zelmanov, 1988). *Every n -Engel Lie algebra over a field of characteristic 0 is nilpotent.*

This result yields an asymptotic solution to the RBP: for every $n \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that every n -Engel Lie algebra of characteristic $p > N$ is nilpotent. A proof of this theorem is given in Appendix B.

2. ASSOCIATIVE ALGEBRAS AND THE NAGATA-HIGMAN-DUBNOV-IVANOV THEOREM

Definition 2.1. An algebra $(A, +, \cdot)$ is *associative* if for all $a, b, c \in A$ we have $(ab)c = a(bc)$.

Example 2.2.

- Any ring of square matrices over a field $\mathbb{F}$ is an associative algebra.
- The polynomial ring $\mathbb{F}[x_1, \dots, x_k]$ is an $\mathbb{F}$ -algebra, which is, furthermore, commutative.

Some further notations and conventions:

- An *identity* in an algebra is an equation satisfied by all elements.
- The *commutator* in an associative algebra is the function $[x, y] = xy - yx$. Commutative algebras satisfy the identity $[x, y] = 0$.
- The *nilexponent* of an associative algebra A is, if it exists, the smallest $n \in \mathbb{N}$ such that $x^n = 0$ is an identity.
- An associative algebra A is *nilpotent* if there is k such that $x_1 \cdots x_{k+1} = 0$. The smallest such k is called the *nilpotency class* of A .
- A two-sided ideal I of an associative algebra is a subalgebra $I \leq A$ such that for all $a, b \in A$ we have $aIb \subseteq I$. The vector space quotient $(A/I, +)$ can be equipped with a multiplication: $(a+I) \cdot (b+I) = ab+I$, which turns A/I into an associative algebra, for which the quotient map $\pi : A \rightarrow A/I$ is an associative algebra homomorphism.

Exercise 2.3. Check the last point above.

Exercise 2.4. Let us prove that the ring $M_2(\mathbb{F})$ of matrices over a field $\mathbb{F}$ satisfy the identity $[[x, y]^2, z] = 0$.

- (1) Prove that for any $M, N \in M_2(\mathbb{F})$, the commutator $[M, N]$ has trace 0.
- (2) Using the Cayley-Hamilton Theorem, prove that if the trace of M is 0, then M^2 is a diagonal matrix.
- (3) Conclude.

Assume that an associative algebra A over a field of characteristic p has nilexponent 2, i.e. satisfies the identity $x^2 = 0$. Then

$$0 = (x + y)^2 = x^2 + xy + yx + y^2 = xy + yx.$$

Therefore, it satisfies $xy = -yx$. In particular, we also have

$$xyz = x(yz) = -xzy = -(xz)y = zxy = z(xy) = -xyz$$

If $p \neq 2$, it follows that $xyz = 0$, hence A is nilpotent. If $p = 2$, then A is commutative, but it is not nilpotent in general (see Exercise 2.7). The following theorem was for a long time known as the Nagata-Higman theorem, but it was later discovered that Dubnov and Ivanov had it proved first.

Theorem 2.5 (Nagata-Higman-Dubnov-Ivanov). *Let A be an associative algebra over a field of characteristic p satisfying the identity $x^n = 0$ with $p > n$. Then A is nilpotent of class $< 2^n - 1$.*

Proof. The proof is by induction on n . The case $n = 2$ above is our initialisation. Let us assume that any associative algebra of nilindex $k < n$ is nilpotent of class $< 2^k - 1$.

$$\begin{aligned} 0 &= (x + \lambda y)^n = (x + \lambda y)(x + \lambda y) \dots (x + \lambda y) \\ &= x^n \\ &\quad + \lambda (x^{n-1}y + x^{n-2}yx + \dots + xyx^{n-2} + yx^{n-1}) \\ &\quad + \lambda^2 (x^{n-2}y^2 + \dots) \\ &\quad \vdots \\ &\quad \lambda^n y^n \end{aligned}$$

Let $A_1 = x^{n-1}y + x^{n-2}yx + \dots + xyx^{n-2} + yx^{n-1}$, $A_2 = x^{n-2}y^2 + \dots$, ..., $A_{n-1} = \dots$. We have

$$\lambda A_1 + \lambda^2 A_2 + \dots + \lambda^{n-1} A_{n-1} = 0$$

Let $\lambda_1, \dots, \lambda_{n-1}$ be nonzero elements of $\mathbb{F}$. Then

$$\begin{bmatrix} \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n-1} & \lambda_{n-1}^2 & \dots & \lambda_{n-1}^{n-1} \end{bmatrix} \cdot \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_{n-1} \end{bmatrix} = 0$$

This is a Vandermonde matrix, classically its determinant is $\prod_{i,j} (\lambda_i - \lambda_j)$ which is nonzero if λ_i are all distinct, which we can choose because $p > n$. So the matrix is invertible and hence $A_1 = A_2 = \dots = 0$. We define $f(x, y) = x^{n-1}y + x^{n-2}yx + \dots + xyx^{n-2} + yx^{n-1} = \sum_{i=0}^{n-1} x^{n-1-i} y x^i$, then $f(a, b) = 0$.

Now let's look at the following quantity:

$$g(a, b, c) = \sum_{i,j=0}^{n-1} a^{n-1-i} c b^{n-1-j} a^i b^j$$

On one side we have:

$$g(a, b, c) = \sum_j \left(\sum_{i=0}^{n-1} a^{n-1-i} c b^{n-1-j} a^i \right) b^j = \sum_j f(a, c b^{n-1-j}) b^j = 0$$

On the other hand, we also have

$$g(a, b, c) = \sum_{i=0}^{n-1} a^{n-1-i} c \left(\sum_j b^{n-1-j} a^i b^j \right) = \sum_{i=1}^{n-1} a^{n-1-i} c f(b, a^i) + a^{n-1} c \sum_{i=0}^{n-1} b^{n-1}$$

We conclude

$$n a^{n-1} c b^{n-1} = 0$$

hence as $p > n$ we have $x^{n-1} y z^{n-1} = 0$.

Let I be the two sided ideal generated by $\{a^{n-1} \mid a \in A\}$. Two elements of I are sums $\sum_i u_i a_i^{n-1} v_i$ and $\sum_i u'_i a_i'^{n-1} v'_i$ hence by the identity above, for all $c \in A$:

$$\left(\sum_i u_i a_i^{n-1} v_i \right) c \left(\sum_j u'_j a_j'^{n-1} v'_j \right) = \sum_{i,j} u_i (a_i^{n-1} (v_i c u'_j) a_j'^{n-1}) v'_j = 0.$$

Therefore, I is nilpotent of class 2. Now, the quotient A/I is a nilalgebra of nilindex $(n-1)$. By induction, A/I is nilpotent ($p > n > n-1$) of class $< k = 2^{n-1} - 1$. This implies that for all $a_1, \dots, a_k, b_1, \dots, b_k, c \in A$, we have $a_1 \dots a_k, b_1 \dots b_k \in I$. So $a_1 \dots a_k c b_1 \dots b_k = 0$ and A is nilpotent of class $< 2k + 1 = 2 \cdot (2^{n-1} - 1) + 1 = 2^n - 1$. $\square$

Remark 2.6.

- For the induction step, we only used that the field is of size large enough and that p does not divide n . However, in order to apply the induction hypothesis, we need that p does not divide $n-1$, then $n-2$, etc so we indeed need $p > n$.
- A nilalgebra does not have an identity element. Therefore, the identity $x^{n-1} y z^{n-1} = 0$ does not imply $x^{n-1} y^{n-1} = 0$, hence the presence of c in the last lines of the proof above, in case elements v_i, u'_j do not occur.

Exercise 2.7. Let $A := \mathbb{F}_2[(x_i)_{i \in \mathbb{N}}]$ be the ring of polynomials with commutative variable. Let I be the ideal generated by $S := \{x_i^2 \mid i \in \mathbb{N}\}$ and $\pi : A \rightarrow A/I$. Let J be ideal generated by $\pi(S)$ in A/I .

- (1) Prove that J is of nilexponent 2.
- (2) Prove that J is not nilpotent.
- (3) Let $(\mathbb{Z}, +)$ be given the trivial multiplication $xy = 0$ and set $K = J \oplus \mathbb{Z}$ be the direct sum of rings. Deduce that, even in characteristic 0, the Nagata-Higman-Dubnov-Ivanov does not generalize from algebras to rings.

We end this chapter by considerations on the associative algebra analogous of the Burnside problem, the *Kurosh-Levitzki problem*. Recall that an associative algebra A is *locally nilpotent* if every finitely generated subalgebra of A is nilpotent.

The Kurosh-Levitzki Problem. *Is every associative algebra of bounded nilexponent locally nilpotent?*

The answer to this question is positive. This implies that, regardless of the characteristics, any finitely generated associative algebra of bounded nilexponent is nilpotent. In particular, any counter-example to the Nagata-Higman-Dubnov-Ivanov must be infinitely generated.

Definition 2.8. Given an associative algebra A , the *Levitzki radical* $\text{rad}(A)$ is the sum of all locally nilpotent ideals of A .

Exercise 2.9. Let I be a (two-sided) ideal of A . If both I and A/I are locally nilpotent, then A is locally nilpotent.

Lemma 2.10. *For any associative algebra A , the Levitzki radical $\text{rad}(A)$ is a locally nilpotent ideal. Further, $\text{rad}(A/\text{rad}(A)) = \{0\}$.*

Proof. First, observe that the sum $I + J$ of two locally nilpotent ideals I and J of A is locally nilpotent: use $(I + J)/I \cong J/(I \cap J)$ and Exercise 2.9. Therefore, $\text{rad}(A)$ is locally nilpotent. Assume now that $a_1 + \text{rad}(A), \dots, a_n + \text{rad}(A) \in \text{rad}(A/\text{rad}(A))$ generate a nilpotent associative algebra, say of class $< r$. Then, every monomial $a_{i_1} \cdots a_{i_r}$ belong to $\text{rad}(A)$ for any $i_1, \dots, i_r \in \{1, \dots, n\}$. Those monomials generate a nilpotent algebra, say of class $< s$, so that any product of s of those monomials $a_{i_1} \cdots a_{i_r}$ vanishes. Therefore, any product of rs of the $a_1, \dots, a_n$ also vanish, and so $a_1, \dots, a_n$ generated a nilpotent algebra. It follows that $a_1, \dots, a_n \in \text{rad}(A)$. $\square$

Let A be an associative algebra of bounded nilexponent n . Define the operation $a \circ b := a + b + ab$. Then the following is easy to check:

- $(A, \circ)$ is a group, called the *adjoint group* of A ;
- The $\circ$ -inverse of $a \in A$ is given by $a^{-1} = -a + a^2 - a^3 + \cdots + (-a)^{n-1}$;
- the unit of $(A, \circ)$ is the element 0.

Exercise 2.11. Check the previous claims.

Lemma 2.12. *Let A be an associative algebra of nilexponent n . Then A satisfies the identity $x^{n-1}yx^{n-1} = 0$.*

Proof. By developing $(a + ba^{n-1})^n = 0$, we obtain

$$\begin{aligned} 0 &= a^n + a^{n-1}ba^{n-1} + a^{n-2}(ba^{n-1})(ba^{n-1}) + a^{n-3}(ba^{n-1})^3 + \cdots + a(ba^{n-1})^{n-1} + (ba^{n-1})^n \\ &= a^{n-1}ba^{n-1} + ca^{n-1}ba^{n-1} \end{aligned}$$

for $c = a^{n-2}b + a^{n-3}ba^{n-1}b + \cdots + a(ba^{n-1})^{n-2}b$. Let $d = a^{n-1}ba^{n-1}$. We have

$$0 = d + cd = c + d + cd - c = c \circ d - c$$

It follows that $c = c \circ d$, hence by $\circ$ -multiplying on the left by $c^{-1} = \sum_{i=1}^{n-1} (-c)^i$, we obtain $d = 0$. $\square$

Theorem 2.13. *Let A be an associative algebra of bounded nilexponent, then A is locally nilpotent.*

Proof. By induction on n , assume that the theorem holds for nilexponent $< n$ and assume that A is of nilexponent n . We prove that $A = \text{rad}(A)$. Assume not, then $B := A/\text{rad}(A)$ is nonzero. If $b^{n-1} = 0$ for all $b \in B$, then B is of nilexponent $< n$ thus locally nilpotent by induction. By Exercise 2.9 and Lemma 2.10, A is also locally nilpotent. Otherwise, there exists $b \in B$ such that $b^{n-1} \neq 0$ and by Lemma 2.12, b^{n-1} generates an abelian ideal. It follows that $b \in \text{rad}(B)$ which contradicts Lemma 2.10. $\square$

For a given set S we denote by $\mathfrak{S}(S)$ the group of permutations of S . We will also use the notation $\mathfrak{S}(n)$ or $\mathfrak{S}_n$ for the group of permutations of $\{1, \dots, n\}$.

Exercise 2.14. (1) Prove that in any associative algebra, we have the following equality:

$$\sum_{\sigma \in \mathfrak{S}_n} y_{1\sigma} \cdots y_{n\sigma} = \sum_{\emptyset \neq S \subseteq \{1, \dots, n\}} (-1)^{n-|S|} \left(\sum_{i \in S} y_i \right)^{(n)} = \sum_{\substack{1 \leq s \leq n \\ 1 \leq i_1 < \dots < i_s \leq n}} (-1)^{n-s} (y_{i_1} + \dots + y_{i_s})^{(n)}$$

Hint. It suffices to develop the right hand side. To convince yourself, do it for $n = 3$:

$$(y_1 + y_2 + y_3)^3 = \sum_{\sigma \in \mathfrak{S}_3} y_{1\sigma} y_{2\sigma} y_{3\sigma} + (y_1 + y_2)^3 + (y_1 + y_3)^3 + (y_2 + y_3)^3 - y_1^3 - y_2^3 - y_3^3.$$

(2) Deduce that in any associative algebra of nilpotent n , we have

$$\sum_{\sigma \in \mathfrak{S}_n} y_{1\sigma} \cdots y_{n\sigma} = 0$$

3. LIE ALGEBRAS AND ENVELOPING ALGEBRAS

As mentioned in the introduction, a Lie algebra $(L, +, [\cdot, \cdot])$ over a field $\mathbb{F}$ is an $\mathbb{F}$ -algebra which further satisfies the following properties:

- (alternative) $[a, a] = 0$,
- (Jacobi) $[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$.

A Lie algebra always satisfy the property antisymmetry $[a, b] = -[b, a]$ but that property is weaker than alternativity (think of characteristic 2).

Example 3.1. • Given any associative algebra $(A, +, \cdot)$, the bracket $[a, b] = ab - ba$ turns $(A, +, [\cdot, \cdot])$ into a Lie algebra.

- The space $\mathcal{M}_n(\mathbb{F})$ of $n \times n$ matrices over $\mathbb{F}$. The subspace of matrices of trace 0 is a Lie subalgebra.
- The cross-product in 3-dimensional vector spaces.

Here are some further notations and definitions.

- A Lie subalgebra $M \subseteq L$ is a Lie algebra which is a vector subspace such that for all $a, b \in M$ we have $[a, b] \in M$. A Lie algebra L is *abelian* if $[a, b] = 0$ for all $a, b \in L$.
- An ideal I of a Lie algebra is a Lie subalgebra such that for all $a \in I$ and $b \in L$ we have $[a, b] \in I$. If L is a Lie algebra with an ideal $I \trianglelefteq L$, the group quotient L/I is again a Lie algebra for the bracket $[a + I, b + I] = [a, b] + I$ and the quotient map $\pi : L \rightarrow L/I$ is a Lie algebra homomorphism. The three isomorphism theorems hold for Lie algebras.
- A *monomial (or commutator)* in $x_1, \dots, x_n$ is a term obtained using iteration of the bracket, for instance $[[x_1, [x_3, x_2]], x_4]$ or $[x_1, [x_2, [x_4, x_3]]]$. The length of a monomial is the number of variables occurring. For instance, $[[x_1, [x_3, x_2]], x_4]$ and $[x_1, [x_2, [x_4, x_3]]]$ have length 4. A distinguished kind of monomials is given by *left-normed commutators* which are inductively defined as $[x_1, \dots, x_n] = [[x_1, \dots, x_{n-1}], x_n]$. We also use the notation:

$$[\underbrace{\dots [x, y], y], \dots, y]}_{n \text{ times}} = [x, y^{(n)}].$$

- Given a subset $S \subseteq L$, the Lie $\mathbb{F}$ -algebra *generated by* S is the smallest Lie subalgebra of L containing S . It is the $\mathbb{F}$ -span of all possible commutators in elements of S . According to Exercise 3.2, it is the span of all left-normed commutators evaluated in elements of S .

Exercise 3.2. Any commutator in $x_1, \dots, x_n$ can be written as a sum of left-normed commutators in $x_1, \dots, x_n$. (*Hint.* Every monomial of length ≥ 2 is written as a product $[u, v]$, where u, v are monomials. Do an induction on the length of v .)

Exercise 3.3. Prove that in a non-abelian Lie algebra L , there is no element 1 such that $[a, 1] = a$.

Given a Lie algebra L over a field $\mathbb{F}$, we denote by $\text{End}(L)$ the $\mathbb{F}$ -algebra of linear endomorphisms of L . For each a in a Lie $\mathbb{F}$ -algebra L , we will denote by ad_a the map $x \mapsto [x, a]$. By bilinearity of $[\cdot, \cdot]$, it is an element of $\text{End}(L)$.

In this course, we will use the right-sided notation for the application of functions

For instance $(a) \text{ad}_b = [a, b]$.

It might seem a bit odd at first but it is very consistent with our convention of left-normed bracketing. Elements $f, g \in \text{End}(L)$ are linear endomorphisms of L , and the addition and multiplication in $\text{End}(L)$ is given by

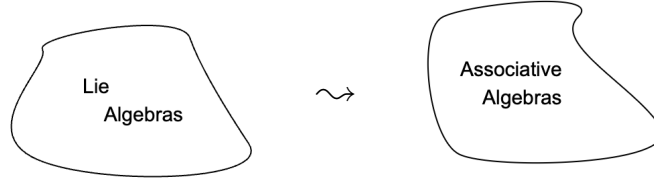
$$(x)(f + g) = (x)f + (x)g \quad (x)(f \circ g) = ((x)f)g.$$

In particular, the right-sided notation of the composition means that $f \circ g$ consist in first applying f then applying g , which is quite natural when you think about it. The composition $\text{ad}_{b_1} \circ \dots \circ \text{ad}_{b_s}$ applied at an element $a \in L$ gives

$$(a)(\text{ad}_{b_1} \circ \dots \circ \text{ad}_{b_s}) = ([a, b_1]) \text{ad}_{b_2} \circ \dots \circ \text{ad}_{b_s} = [a, b_1, \dots, b_s]$$

Most of the time we will use juxtaposition instead of the symbols $\circ$ for the composition of maps. The notation $[a, b^{(n)}]$ corresponds to evaluating $\text{ad}_b^n \in \text{End}(L)$ at a .

Let us now realize the following mapping.



Adjoint representation and enveloping algebra. As we saw above, as any other associative algebra, the associative algebra $(\text{End}(L), +, \cdot)$ has an underlying Lie algebra $(\text{End}(L), +, [\cdot, \cdot])$ where the bracket is given by $[f, g] = fg - gf$. The Jacobi identity in L yields for all $a, b, c \in L$:

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

so that $\text{ad}_{[b, c]} = \text{ad}_b \text{ad}_c - \text{ad}_c \text{ad}_b = [\text{ad}_b, \text{ad}_c]$, where the last bracket is the associative algebra commutator in $\text{End}(L)$. This means that the map

$$\text{ad} : (L, +, [\cdot, \cdot]) \rightarrow (\text{End}(L), +, [\cdot, \cdot])$$

is a *Lie algebra homomorphism*.

We denote $\text{ad}L$ the image of L via the map ad . Within $\text{End}(L)$, the algebra generated by $\text{ad}L$ is the smallest associative subalgebra of $\text{End}(L)$ which contains the Lie algebra $\text{ad}L$. We make the following definition.

- The *enveloping algebra of L* is the subalgebra of $\text{End}(L)$ generated by $\text{ad}L$. It is denoted by $A(L)$.
- Any element of $A(L)$ of the form ad_a is called a *Lie element* of $A(L)$.
- An equation of $A(L)$ which is satisfied only by Lie elements is called a *weak identity* of $A(L)$.

We encounter our first fundamental principle.

Principle 1. Any (weak) identity of $A(L)$ yields an identity of L .

This principle works as follows. If $f = \sum_i \text{ad}_{y_{1,i}} \cdots \text{ad}_{y_{s,i}}$ is a polynomial in variables $(y_{i,j})$ such that $f = 0$ is a weak identity of $A(L)$. Then for all $\vec{b} = (b_{i,j}) \in \text{ad}(L)$ we have the equation $f(\vec{b}) = \sum_i \text{ad}_{b_{1,i}} \cdots \text{ad}_{b_{s,i}} = 0$. If $f(\vec{b}) = 0$ as an element of $A(L)$, means that, as an endomorphism, it is equal to the zero map. This implies that for all $a \in L$:

$$\sum_i [a, b_{1,i}, \dots, b_{s,i}] = 0$$

whence L satisfies the identity $\sum_i [x, y_{1,i}, \dots, y_{s,i}] = 0$.

Exercise 3.4. Prove that any Lie algebra satisfies the identity

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

Let us turn now to our second principle, whose statement is a direct consequence of the definitions.

Principle 2. Any identity of L of the form $\sum_i [x, y_{1,i}, \dots, y_{s,i}] = 0$ yields a weak identity of $A(L)$, of the form $\sum_i \text{ad}_{y_{1,i}} \cdots \text{ad}_{y_{s,i}} = 0$.

Remark 3.5 (Computing in the enveloping algebra). The simple fact above that ad is a Lie algebra homomorphism allows to quickly compute in Lie algebra. It is just a matter of having the right viewpoint. For instance, given any Lie monomial of the form $[\dots, [a, b], \dots]$, one should see the term $[\dots, [a, b]]$ as an instance of $\text{ad}_{[a, b]}$ applied to the term " $\dots$ ". As $\text{ad}_{[a, b]} = \text{ad}_a \text{ad}_b - \text{ad}_b \text{ad}_a$, we obtain $[\dots, [a, b]] = [\dots, a, b] - [\dots, b, a]$. Then the map $x \rightarrow [x, \dots]$ being linear, we conclude:

$$[\dots, [a, b], \dots] = [\dots, a, b, \dots] - [\dots, b, a, \dots]$$

Conversely, as $ab = ba - [b, a]$ in $A(L)$, we immediately see the extra term obtained by commuting a and b in $[\dots, a, b, \dots]$:

$$[\dots, a, b, \dots] = [\dots, b, a, \dots] - [\dots, [b, a], \dots].$$

Exercise 3.6. Prove that in any Lie algebra, we have $[x, y, x, y] = [x, y, y, z]$. (*Hint.* What is $[x, y, [x, y]]$?)

Exercise 3.7. Prove that if $[x, y^{(n)}] = 0$ is an identity of L then $y^n = 0$ is a weak identity of $A(L)$.

In some lucky cases, an identity of L translates into a (full) identity of $A(L)$:

Exercise 3.8. Prove that if L is a Lie algebra over a field of characteristic 2 which satisfies the identity $[x, y^2] = 0$, then $A(L)$ satisfies the identity $xy - yx = 0$, i.e. $A(L)$ is commutative.

Definition 3.9. For A, B subsets of a Lie algebra L , we denote by $[A, B]$ to be the span of the set $\{[a, b] \mid a \in A, b \in B\}$. We define the *lower central series* as the series defined inductively by $L^1 = L$ and $L^{n+1} = [L^n, L]$. A Lie algebra is *nilpotent* if there exists k such that $L^{k+1} = 0$. The smallest such k is called the *nilpotency index* of L .

Proposition 3.10. Let L be a Lie algebra, the following are equivalent.

- (1) L is nilpotent of class $< k$
- (2) L satisfies the identity $m(x_1, \dots, x_k) = 0$, for all variables $x_1, \dots, x_k$ and monomials $m(x_1, \dots, x_k)$ of length k .
- (3) L satisfies the identity $[x_1, \dots, x_k] = 0$.

Proof. It is clear that (1) and (3) are equivalent and that both follow from (2). The fact that (2) implies (3) is Exercise 3.2. $\square$

Exercise 3.11. Prove that for a Lie algebra L , if $A(L)$ is nilpotent of class $\leq k$ then L is nilpotent of class $\leq k - 1$.

4. FIRST FOOTSTEPS WITH n -ENGEL LIE ALGEBRAS

Definition 4.1. A Lie algebra is *n -Engel* if it satisfies the identity

$$0 = [x, \underbrace{y, y, \dots, y}_{n \text{ times}}] = [x, y^{(n)}].$$

Equivalently, the enveloping algebra $A(L)$ satisfies the weak identity $y^n = 0$.

Lemma 4.2. Let L be a Lie algebra over any field $\mathbb{F}$. If L is n -Engel then L satisfies the identity

$$\sum_{\sigma \in \mathfrak{S}_n} [x, y_{1\sigma}, \dots, y_{n\sigma}] = 0$$

Furthermore, the converse holds if $\mathbb{F}$ has characteristic 0 or $p > n$.

Proof. Applying Principle 1 to Exercise 2.14, we have the following:

Claim 1.

$$\sum_{\sigma \in \mathfrak{S}_n} [x, y_{1\sigma}, \dots, y_{n\sigma}] = \sum_{\emptyset \neq S \subseteq \{1, \dots, n\}} (-1)^{n-|S|} [x, (\sum_{i \in S} y_i)^{(n)}]$$

If L is n -Engel, each term on the right of the expression in Claim 1 vanishes, so we are done. Conversely, if the identity

$$\sum_{\sigma \in \mathfrak{S}_n} [x, y_1, \dots, y_n] = 0$$

is satisfied by L , then by putting $y_i = y_j = y$ we get $n![x, y^{(n)}] = 0$. Then if $\mathbb{F}$ has characteristic 0 or $p > n$, we conclude $[x, y^{(n)}] = 0$. $\square$

Lemma 4.3. Let L be an n -Engel Lie algebra over $\mathbb{F}_p$ with $p < n$. Then for all $c, a_1, \dots, a_m$ with $m \geq n$, there exists some $k \in \mathbb{N}$ and $(b_{i,j})_{1 \leq i \leq k, 1 \leq j \leq n_i} \in \langle a_1, \dots, a_m \rangle$ with $n_i < n$ for all $1 \leq i \leq k$, such that

$$[c, a_1, \dots, a_m] = \sum_{i=1}^k [c, b_{i,1}, \dots, b_{i,n_i}]$$

Proof. We do an induction on $m - n$, starting with the case $m = n$. Using the principle that $[\dots, a, b, \dots] = [\dots, b, a, \dots] - [\dots, [b, a], \dots]$, one sees that every monomial of the form $[c, a_{1\sigma}, \dots, a_{n\sigma}]$ can be written as $[c, a_1, \dots, a_n] + u$ where u is a sum of monomials of the form $[c, u_1, \dots, u_s]$ with $u_i \in \langle a_1, \dots, a_n \rangle$ and $s < n$. From $\sum_{\sigma \in \mathfrak{S}_n} [c, a_{1\sigma}, \dots, a_{n\sigma}] = 0$, we obtain that $n![c, a_1, \dots, a_n]$ is a sum of monomials of the right form, and we conclude the case $m = n$ as $p > n$. For $m > n$ and $c, a_1, \dots, a_m \in L$, we have that $[c, a_1, \dots, a_m] = [[c, a_1, \dots, a_n], a_{n+1}, \dots, a_m]$, and $[c, a_1, \dots, a_n] = \sum_{i=1}^k u_i$ where $u_i = [c, d_1, \dots, d_s]$ are left-normed monomials with $s < n$ with $d_i \in \langle a_1, \dots, a_n \rangle$. We may therefore apply the induction hypothesis to each $[u_i, a_{n+1}, \dots, a_m]$ in order to conclude. $\square$

Corollary 4.4. Let L be an n -Engel Lie algebra over $\mathbb{F}_p$ with $n < p$. Then every element of $A(L)$ is of the form

$$\sum_{i=1}^k f_1 \cdots f_{i,n_i}$$

with $n_i < n$.

We now look at particular case of nilpotency of n -Engel Lie algebras. This section draws from Higgins [8] and Traustason [19].

Theorem 4.5. Let L be a 2-Engel Lie $\mathbb{F}_p$ -algebras. If $p \neq 3$, then $L^3 = 0$ and if $p = 3$ then $L^4 = 0$.

Proof. Let L be a 2-Engel Lie $\mathbb{F}_p$ -algebra. For convenience, we will write $A, B, C, \dots$ for $\text{ad}_a, \text{ad}_b, \text{ad}_c, \dots$. As L is 2-Engel, we have $[c, a^{(2)}] = [c, b^{(2)}] = 0$, so that, in $A(L)$ we have $A^2 = B^2 = 0$. Because $A + B$ is also a Lie element (it is ad_{a+b}) we also have:

$$0 = (A + B)^2 = AB + BA. \quad (\star)$$

The case $p \neq 3$. In particular, for all $c \in L$ we have $(c)AB + (c)BA = 0$ hence using antisymmetry and Jacobi:

$$\begin{aligned} 0 &= [c, a, b] + [c, b, a] = -[a, c, b] - [a, [c, b]] \\ &= -[a, c, b] - ([a, c, b] - [a, b, c]) \\ &= [a, b, c] - 2[a, c, b]. \end{aligned}$$

As a, b, c are arbitrary, it follows that

$$BC - 2CB = 0 \quad (\star\star)$$

Together, $(\star)$ and $(\star\star)$ yield $3BC = 0$ so that as $p \neq 3$ we have $[a, b, c] = 0$ for all $a, b, c \in L$, i.e. $L^3 = 0$.

The case $p = 3$. As equation $(\star)$ hold for any Lie element of L , it also holds for $[B, C]$ in place of B , so that

$$0 = A[B, C] + [B, C]A = ABC - ACB + BCA - CBA$$

Using $(\star)$ we have $CB = -BC$ so $ACB = -ABC$ and similarly $BCA = ABC$ and $CBA = -CAB = ACB = -ABC$ so that the above yields $4ABC = 0$ hence if $p = 3$ we have $L^4 = 0$. $\square$

Exercise 4.6. Let L be a 3-Engel Lie algebra over a field $\mathbb{F}$ of size ≥ 3 . Then $A(L)$ satisfies the weak identity $A^2B + ABA + AB^2 = 0$.

(Solution. It is an easy linearization. We develop the equation $0 = (A + \lambda B)^3$ for $a, b \in L$ and $\lambda \in \mathbb{F}$:

$$\lambda(A^2B + ABA + AB^2) + \lambda^2(AB^2 + BAB + B^2A) = 0 \quad (\heartsuit(\lambda))$$

This equation holds for all values of λ . As $|\mathbb{F}| > 2$, there exists $\lambda \in \mathbb{F}$ such that $\lambda^2 \neq \lambda$. By subtracting λ times $\heartsuit(1)$ from $\heartsuit(\lambda)$ and dividing the result by $\lambda - \lambda^2$, we obtain $AB^2 + BAB + B^2A = 0$.

Theorem 4.7. Let L be a 3-Engel Lie $\mathbb{F}_p$ -algebras. If $p \neq 2, 5$ then $L^5 = 0$.

Proof. We write $A, B, \dots$ for $\text{ad}_a, \text{ad}_b, \dots$. From Exercise 4.6, we have

$$AB^2 + BAB + B^2A = 0 \quad (\heartsuit)$$

Applying $(\heartsuit)$ to an arbitrary $c \in L$ and using antisymmetry, we have

$$\begin{aligned} 0 &= [c, a, b^{(2)}] + [c, b, a, b] + [c, b^{(2)}, a] \\ &= -([a, c, b^{(2)}] + [a, [c, b], b] + [a, [c, b^{(2)}]]) \end{aligned}$$

whence $CB^2 + [C, B]B + [C, B, B] = 0$. By developping using $[X, Y] = XY - YX$ we obtain the identity

$$3CB^2 - 3BCB + B^2C = 0 \quad (\heartsuit\heartsuit)$$

The case $\mathbb{F}_3$. In that case, $(\heartsuit\heartsuit)$ yields the identity $B^2C = 0$. Consider the vector space

$$I = \text{Span}_{\mathbb{F}_3}([a, b^{(2)}] \mid a, b \in L)$$

As $B^2C = 0$ we have $[[a, b^{(2)}], c] = [c, [a, b^{(2)}]] = 0$ hence I is in fact an ideal of L . Now L/I is a 2-Engel Lie $\mathbb{F}_3$ -algebra hence by Example $(L/I)^4 = 0$. It follows that the product of every 4 elements in L can be written as $\sum_i [a_i, b_i^{(2)}]$, so by $B^2C = 0$ we conclude $L^5 = 0$.

The case $p \neq 2, 3, 5$. From $(\heartsuit)$ and $(\heartsuit\heartsuit)$ we deduce

$$AB^2 = 2BAB \quad B^2A = -3BAB$$

In particular, we have

$$3AB^2 = -2B^2A. \quad (\heartsuit\heartsuit\heartsuit)$$

From $AB^2 = 2BAB$ we may multiply on the left by A to deduce $A^2B^2 = 2ABAB$. Further, we may exchange A and B in $B^2A = -3BAB$ to get $A^2B = -3ABA$ then multiply by B on the right the latter to get $A^2B^2 = -3ABAB$. We conclude that $5ABAB = 0$. As we assume $p \neq 5$, we have

$$ABAB = A^2B^2 = 0.$$

As above, let

$$J = \text{Span}_{\mathbb{F}_p}([a, b^{(2)}] \mid a, b \in L).$$

From $(\otimes)$, we have that for all $a, b, c, [a, b^{(2)}, c] \in J$ so $[d, c] \in J$ for any $d \in J$ and so J is an ideal. The Lie algebras L/J is 2-Engel of characteristic 3 hence by Example 4 it is nilpotent of class at most 2 hence $[a, b, c] \in J$ for all $a, b, c \in L$. It follows from $A^2B^2 = 0$ that $[a, b, c, d^2] = 0$ hence L satisfies the identity

$$ABC^2 = 0.$$

Using $(\otimes)$ we have that

$$0 = ABC^2 = \frac{-2}{3}AC^2B = \frac{4}{9}C^2AB$$

Again, as $[a, b, c] \in J$ for all $a, b, c \in L$ we conclude using C^2AB that $[a, b, c, d, e] = 0$ for all $a, b, c, d, e \in L$. We conclude that $L^5 = 0$. $\square$

Remark 4.8. The bound on the nilpotency classes are optimal in Theorem 4.5 and Theorem 4.7. The proofs make apparent that the arithmetical properties in an n -Engel Lie algebra of characteristic p are very much dependent on the values of the pair (n, p) . The exclusion case $p \neq 2, 5$ for 3-Engel Lie algebras is also optimal. The following table shows how the nilpotency class – when it exists – has quite an unpredictable behaviour. The intersection of the “ n -Engel” line and the “characteristic p ” column is an optimal bound on the nilpotency class.

n -Engel Lie algebra \ Characteristic p	2	3	5	7	> 7
2-Engel	2	3	2	2	2
3-Engel	∞	4	∞	4	4
4-Engel	∞	∞	∞	7	7
5-Engel	∞	∞	∞	∞	11

5. THE RESTRICTED BURNSIDE PROBLEM AND n -ENGEL LIE ALGEBRAS

We will assume that the reader is familiar with the notion of a group G , a subgroup $H \subseteq G$. The group of endomorphisms will be denoted $\text{End}(G)$, the group of automorphisms $\text{Aut}(G)$. Recall that inner automorphisms are of the form $x \mapsto x^g := g^{-1}xg$, and a subgroup $H \subseteq G$ is normal if it is closed under inner automorphisms of G , in which case we will write $H \trianglelefteq G$. The group generated by a set S is denoted $\langle S \rangle$, and a group G is r -generated if there exists $g_1, \dots, g_r \in G$ such that $G = \langle g_1, \dots, g_r \rangle$. The normal closure of a set $S \subseteq G$ is the smallest normal subgroup of G containing S , and it is the group generated by all conjugates of elements of S .

5.1. The associated Lie algebra of a group of exponent p . The exposition draws mainly from [21]. Given a group G , the *commutator of two elements* is $[g, h] = g^{-1}h^{-1}gh$. Given two sets $S, T \subseteq G$, we write $[S, T]$ for the group spanned by all elements $[s, t]$ with $s \in S, t \in T$. Then the *lower central series* (LCS) is the series

$$\gamma_1 \supseteq \gamma_2 \supseteq \dots$$

defined inductively as $\gamma_1 = \gamma_1(G) = G$ and

$$\gamma_{i+1} = [\gamma_i, G].$$

Using that $[g, h]^k = [g^k, h^k]$, each γ_i is a normal subgroup of G (see Exercise 5.1). As $gh = hg[gh]$, we see that G/γ_2 is abelian and more generally, each quotient γ_i/γ_{i+1} is abelian. (See Exercise 5.1).

Exercise 5.1. Let G be a group and $(\gamma_i)_{i \in \mathbb{N}}$ its LCS.

- (1) Prove that γ_i is a normal subgroup of G .
- (2) Prove that γ_i/γ_{i+1} is a subgroup of $Z(G/\gamma_{i+1})$, in particular, γ_i/γ_{i+1} is abelian.

Exercise 5.2. Let G be a group and $g, h, k \in G$. Prove the following equations.

- (1) $[g, h] = [h, g]^{-1}$
- (2) $[g, h]^k = [g^k, h^k]$.
- (3) $[gh, k] = [g, k]^h[h, k]$.
- (4) $[g, hk] = [g, k][g, h]^k$.
- (5) $[g, h^{-1}, k]^h[h, k^{-1}, g]^k[k, g^{-1}, h]^g = 1$.

Definition 5.3. A group is *nilpotent* if there exists $k \in \mathbb{N}$ such that $\gamma_k = 1$. In this case, the smallest c such that $\gamma_{c+1} = 1$ is called the *nilpotency class* of G .

It is a classical fact that finite p -groups are nilpotent, see Exercise 5.20.

For $g \in G$, we let g^* denote a conjugate of g , where the element by which g is conjugated is not explicit. Similarly, g^{-*} denotes a conjugate of g^{-1} , which is the inverse of a conjugate of g . Revisiting the classical identities from Exercise 5.2 (using that $gh = h^*g$), we have the following:

$$[gh, k] = [g, k]^*[h, k]^* \quad (1)$$

$$[g, hk] = [g, h]^*[g, k]^* \quad (2)$$

$$[g^{-1}, h] = [g, h]^{-*} \quad (3)$$

$$[g, h^{-1}] = [g, h]^{-*} \quad (4)$$

$$[g, h, k] = [k, g^{-1}, h^{-1}]^{-*}[h^{-1}, k^{-1}, g]^{-*} \quad (5)$$

Lemma 5.4. *Let G be a group and $(\gamma_i)_{i \in \mathbb{N}}$ be the LCS. For all i, j , $[\gamma_i, \gamma_j] \subseteq \gamma_{i+j}$.*

Proof. We prove it by induction on j , the case $j = 1$ is by definition, so assume by induction that $j > 1$ and for all k and for all $j_0 < j$ we have $[\gamma_k, \gamma_{j_0}] \subseteq \gamma_{k+j_0}$. Let $a \in \gamma_i, b \in \gamma_j$. Then b is a product of commutators $[c, d]$ and their inverse with $c \in \gamma_{j-1}, d \in G$, so using (2) and (4) above, $[a, b]$ is in the normal closure of elements of the form $[a, [c, d]]$ with $c \in \gamma_{j-1}, d \in G$. Now, γ_{i+j} is normal, so it suffices to prove that $[a, [c, d]] \in \gamma_{i+j}$. By Exercise 5.2(1), we have $[a, [c, d]] = [c, d, a]^{-1}$ and by (5), we have

$$[c, d, a]^{-1} = ([a, c^{-1}, d^{-1}]^{-*}[d^{-1}, a^{-1}, c]^{-*})^{-1} = [d^{-1}, a^{-1}, c]^*[a, c^{-1}, d^{-1}]^*$$

As $a \in \gamma_i, d \in \gamma_1$ we have $[d^{-1}, a^{-1}] = [a^{-1}, d^{-1}]^{-1} \in \gamma_{i+1}$. As $c \in \gamma_{j-1}$ the induction hypothesis yields that $[d^{-1}, a^{-1}, c] \in \gamma_{i+j}$. Similarly, we deduce $[a, c^{-1}, d^{-1}] \in \gamma_{i+j}$ and we conclude the proof of the lemma. $\square$

Lemma 5.5. *Let G be a group, let $S \subseteq G$ such that $G = \langle S \rangle$. For each $n \in \mathbb{N}$, define $S_n = \{[s_1, \dots, s_n] \mid s_1, \dots, s_n \in S\}$. Then*

- (1) γ_n is the normal closure of S_n ,
- (2) If $\pi: \gamma_i \rightarrow \gamma_i/\gamma_{i+1}$, then γ_n/γ_{n+1} is generated by $S_n\pi$.

Proof. We prove it by induction on n . For $n = 1$, $S_1 = G$ so (1) and (2) are trivial. Assume $n > 1$ and let $h \in \gamma_n, g \in G$. By induction h can be expressed as a product of elements of the form $[s_1, \dots, s_n]^k$ and their inverses with $k \in G, s_i \in S$. Using (1) and (3), $[h, g]$ is in the normal closure of elements of the form $[[s_1, \dots, s_n]^k, g]$. By Exercise 5.2, we have $[[s_1, \dots, s_n]^k, g] = [[s_1, \dots, s_n], g^{k^{-1}}]^k$. As $G = \langle S \rangle$, $g^{k^{-1}}$ is a product of elements of S and their inverses, hence by (2) and (4), $[[s_1, \dots, s_n], g^{k^{-1}}]^k$ is in the normal closure of elements of the form $[[s_1, \dots, s_n], s]$, with $s \in S$, and so does $[h, g]$. As γ_{n+1} is generated by such $[h, g]$ and γ_{n+1} is normal, we conclude (1). For (2), observe that

$$a^g = [g, a^{-1}]a = [a^{-1}, g]^{-1}a$$

with $a = [s_1, \dots, s_{n+1}]$ to conclude that $[s_1, \dots, s_{n+1}]^g = [s_1, \dots, s_{n+1}]$ modulo γ_{n+2} . $\square$

Let G be a group of exponent p with LCS $(\gamma_i)_{i \in \mathbb{N}}$. We define $L_i = \gamma_i/\gamma_{i+1}$. Each L_i is an abelian group of exponent p , otherwise known as an elementary abelian group or an $\mathbb{F}_p$ -vector space, and we will use the additive notation for the group law. We may define the group $L(G)$ as the direct sum of those $\mathbb{F}_p$ -vector spaces:

$$L(G) := \bigoplus_{i \in \mathbb{N}} L_i = \bigoplus_{i \in \mathbb{N}} \gamma_i/\gamma_{i+1}.$$

The group law in $L(G)$ will also be denoted additively. In $L(G)$, we call an element of L_i a *homogeneous element*. Every element of $L(G)$ can be uniquely written as a sum of homogeneous elements. Given two homogeneous elements $g\gamma_{i+1} \in L_i$ and $h\gamma_{j+1} \in L_j$ we know by Lemma 5.4 that $[g, h] \in \gamma_{i+j}$ and we define the bracket on $g\gamma_{i+1}$ and $h\gamma_{j+1}$ as follows

$$[g\gamma_{i+1}, h\gamma_{j+1}] = [g, h]\gamma_{i+j+1}.$$

- The binary function $[\cdot, \cdot]$ is well-defined on homogeneous elements. Assume that $g\gamma_{i+1} = k\gamma_{i+1}$ so that $g = hu$ with $u \in \gamma_{i+1}$. Using (1) we have $[g, h] = [k, h]^*[u, h]^*$ and by Lemma 5.4 and Exercise 5.1 (1) we have $[u, h]^* \in \gamma_{i+j+1}$ so $[g, h]\gamma_{i+j+1} = [k, h]\gamma_{i+j+1}$. Similarly, if $h\gamma_{j+1} = k\gamma_{j+1}$ we have $[g, h]\gamma_{i+j+1} = [g, k]\gamma_{i+j+1}$, so the bracket is well-defined.
- The bracket is alternative and antisymmetric on homogeneous elements. By definition $[g\gamma_{i+1}, g\gamma_{i+1}] = [g, g]\gamma_{2i+1} = 0$. By Exercise 5.2, $[g\gamma_{i+1}, h\gamma_{j+1}] = [g, h]\gamma_{i+j+1} = [h, g]^{-1}\gamma_{i+j+1}$. In L_{i+j} , the element $[h, g]^{-1}\gamma_{i+j+1}$ is the inverse of $[h, g]\gamma_{i+j+1} = [h\gamma_{j+1}, g\gamma_{i+1}]$ so, according to our additive notation, we conclude $[g\gamma_{i+1}, h\gamma_{j+1}] = -[h\gamma_{j+1}, g\gamma_{i+1}]$.

- The bracket $[\cdot, \cdot]$ is bilinear on homogeneous elements. Assume that $g\gamma_{i+1}, h\gamma_{i+1} \in L_i$ and $k\gamma_{j+1} \in L_j$. The group law in L_i is denoted $+$ so that by definition $g\gamma_{i+1} + h\gamma_{i+1} = gh\gamma_{i+1}$. Now we prove that $[g\gamma_{i+1} + h\gamma_{i+1}, k\gamma_{j+1}] = [g\gamma_{i+1}, k\gamma_{j+1}] + [h\gamma_{i+1}, k\gamma_{j+1}]$, where the sum in the second term corresponds to the sum in L_{i+j} . By definition $[g\gamma_{i+1} + h\gamma_{i+1}, k\gamma_{j+1}] = [gh\gamma_{i+1}, k\gamma_{j+1}] = [gh, k]\gamma_{i+j+1}$. By (1), the latter term equals $[g, k]^* [h, k]^* \gamma_{i+j+1} = [g, k]^* \gamma_{i+j+1} + [h, k]^* \gamma_{i+j+1}$ and as γ_{i+j+1} is normal, the latter equals $[g, k]\gamma_{i+j+1} + [h, k]\gamma_{i+j+1}$, which by definition, is $[g\gamma_{i+1}, k\gamma_{j+1}] + [h\gamma_{i+1}, k\gamma_{j+1}]$. The argument on the right input of $[\cdot, \cdot]$ is similar.

The latter point implies that we may now extend the definition of the bracket to the whole of $L(G)$ by linearity, since $L(G)$ is spanned by the homogeneous elements. For each $\sum_i a_i, \sum_j b_j \in L(G)$ where a_i, b_j are homogeneous elements, we define

$$[\sum_i a_i, \sum_j b_j] = \sum_{i,j} [a_i, b_j].$$

Basic linear algebra gives that this yields a well-defined bilinear function on $L(G)$.

Lemma 5.6. *The structure $(L(G), +, [\cdot, \cdot])$ is a Lie algebra over $\mathbb{F}_p$.*

Proof. We already know that $[\cdot, \cdot]$ is a bilinear map on $L(G)$, so it remains to check antisymmetry and the Jacobi identity. Let $a = \sum_i a_i \in L(G)$ where a_i are homogeneous elements. Then by definition $[a, a] = \sum_{i,j} [a_i, a_j] = \sum_i [a_i, a_i] + \sum_{i < j} ([a_i, a_j] + [a_j, a_i])$. This vanishes by the second point above. We now turn to the Jacobi identity. The function

$$(x, y, z) \mapsto [x, y, z] + [y, z, x] + [z, x, y]$$

is multilinear, hence to check the Jacobi identity for $L(G)$ it is enough to check it on homogeneous elements. Let $a = f\gamma_{i+1} \in L_i, b = g\gamma_{j+1} \in L_j, c = h\gamma_{k+1} \in L_k$. Note that by Lemma 5.4, we have $[f, g, h][g, h, f][h, f, g] \in \gamma_{i+j+k}$ hence, it is enough to prove that

$$[f, g, h][g, h, f][h, f, g] \in \gamma_{i+j+k+1}.$$

By Exercise 5.2 (5) we have $[f, g^{-1}, h]^g [g, h^{-1}, f]^h [h, f^{-1}, g]^f = 1$. We consider $[f, g^{-1}, h]^g$. By (4) we have $[f, g^{-1}] = [f, g]^{-s}$, for some $s \in G$. By Lemma 5.4, we have $[f, g] \in \gamma_{i+j}$ and hence $[f, g]^{-1} \in \gamma_{i+j}$. For $u = [[f, g]^{-1}, s] \in \gamma_{i+j+1}$, we have $[f, g]^{-s} = [f, g]^{-1}u$. It follows that $[f, g^{-1}, h]^g = [[f, g]^{-1}u, h]^g$ which, by (1) equals $[[f, g]^{-1}, h]^* [u, h]^*$. By (3), we have $[[f, g]^{-1}, h]^* = [f, g, h]^{-*}$ and as above $[f, g, h]^{-*} = [f, g, h]^{-1}$ modulo $\gamma_{i+j+k+1}$. As $[u, h]^* \in \gamma_{i+j+k+1}$ we conclude that

$$[f, g^{-1}, h]^g = [f, g, h]^{-1} \text{ mod } \gamma_{i+j+k+1}.$$

Similarly, $[g, h^{-1}, f]^h = [g, h, f]^{-1} \text{ mod } \gamma_{i+j+k+1}$ and $[h, f^{-1}, g]^f = [h, f, g]^{-1} \text{ mod } \gamma_{i+j+k+1}$, so we conclude using the equality above. $\square$

5.2. Varieties of groups, free Burnside group. Given variables $x_1, \dots, x_n$ a word in $x_1, \dots, x_n$ (or simply a word) a string of characters $x_{i_1}^{\varepsilon_1} \dots x_{i_k}^{\varepsilon_k}$ with exponents $\varepsilon_i \in \{1, -1\}$, $1 \leq i_1, \dots, i_k \leq n$, $k \in \mathbb{N}$, which is subjects to the reduction

$$\dots x_{i_{n-1}}^{\varepsilon_{i_{n-1}}} x_{i_n}^{\varepsilon_{i_n}} x_{i_{n+1}}^{\varepsilon_{i_{n+1}}} \dots = \dots x_{i_{n-1}}^{\varepsilon_{i_{n-1}}} x_{i_{n+1}}^{\varepsilon_{i_{n+1}}} \dots$$

Equivalently it is an equivalence class of strings of characters $x_{i_1}^{\varepsilon_1} \dots x_{i_k}^{\varepsilon_k}$ for the relations that relates two strings if one can be obtained from the other by removing or adding strings of the form $x_i^1 x_i^{-1}$ in between the letters. We denote the empty word by 1. The set $W(n) = W(x_1, \dots, x_n)$ of words in $x_1, \dots, x_n$ can be given the structure of a group via concatenation, i.e. by setting

$$(x_{i_1}^{\varepsilon_1} \dots x_{i_k}^{\varepsilon_k}) \cdot (x_{j_1}^{l_1} \dots x_{j_l}^{l_l}) = x_{i_1}^{\varepsilon_1} \dots x_{i_k}^{\varepsilon_k} x_{j_1}^{l_1} \dots x_{j_l}^{l_l}$$

and $x_i x_i^{-1} = 1$. This group is called the *free group in n generators*. Similarly, by taking words in variables $(x_i)_{i \in \mathbb{N}}$, one constructs the group $W(\omega)$, which can also be seen as the group $\bigcup_{n \in \mathbb{N}} W(x_1, \dots, x_n)$ for the natural group embedding of $W(x_1, \dots, x_n)$ in $W(x_1, \dots, x_{n+1})$ given by mapping $x_i \in W(x_1, \dots, x_n)$ to $x_i \in W(x_1, \dots, x_{n+1})$.

Evaluation of words in a group. For any element $w \in W(\omega)$, we write $w(x_1, \dots, x_n)$ if the variables appearing in the word w are among $x_1, \dots, x_n$. We may simply write $w(x_1, \dots, x_n)$. Given $g_1, \dots, g_n$ in a group G and a word $w(x_1, \dots, x_n)$ we can consider $w(g_1, \dots, g_n)$ the element of G given by setting $x_i \mapsto g_i$ and interpreting the word in G . If $w = x_1^{-1} x_2^{-1} x_1 x_2$, then for $g, h \in G$ we have $w(g, h) = g^{-1} h^{-1} g h = [g, h] \in G$.

Definition 5.7 (Variety of groups). Let V be a set of words in $(x_i)_{i \in \mathbb{N}}$. The *variety of groups generated by V* is the collection of all groups G such that for all $w(x_1, \dots, x_n) \in V$ and for all $g_1, \dots, g_n \in G$ we have $w(g_1, \dots, g_n) = 1$. A *variety of group* is a variety defined by some set of words.

Exercise 5.8. Check that the following classes of groups are varieties: all groups, abelian groups, nilpotent groups. Note that the set V in the previous definition might be infinite: prove that the class of group of exponent dividing n is a variety, for all n .

Definition 5.9 (Free groups). Given a variety of groups $\mathcal{V}$ and a set S of variable. A *free group for $\mathcal{V}$ generated by S* (or a group *freely generated by S*) is an element $F(S)$ of $\mathcal{V}$ generated by S and such that for any group $G \in \mathcal{V}$, any mapping $S \rightarrow G$ extends uniquely to a group homomorphism $F(S) \rightarrow G$.

$$\begin{array}{ccc} S & \longrightarrow & G \\ & \searrow & \uparrow \\ & \cap & \\ & F(S) & \end{array}$$

Free objects always exists in a variety.

Theorem 5.10. Every variety of groups admits a free group in any given set of generators.

Exercise 5.11. Let $\mathcal{V}$ be a variety of p -groups. prove that there exists n such that every groups in $\mathcal{V}$ has exponent p^n .

It does not take long to see that the group of words $W(S)$ is a free group in S for the variety of all groups. Then free groups in other varieties can be constructed from $W(S)$, as we will see in the next lemma. It is generally admitted that only the free groups $W(S)$ for the variety of all groups are called free groups, whereas for any strictly smaller variety, free groups are usually called *relatively* free group. In fact, those are given by quotient of $W(S)$.

A subgroup H of $W(\omega)$ is called *fully invariant* if for all endomorphism θ of W we have $H\theta \subseteq H$.

Lemma 5.12. Let H be a fully invariant ideal of the free group $W = W(\omega)$. Then W/H is a free group for some variety $\mathcal{V}$.

Proof. We set $\mathcal{V}$ be the variety of groups defined by the following subset of $W(\omega)$:

$$\{w \in W(\omega) \mid w(x_{i_1}, \dots, x_{i_k}) \in H \text{ for some distinct } i_1, \dots, i_k \in \mathbb{N}\}.$$

Note that because H is fully invariant, W/H is an element of $\mathcal{V}$. We prove that W/H is a free group relatively to $\mathcal{V}$, freely generated by $(x_i H)_{i \in \mathbb{N}}$. Let G be any group in $\mathcal{V}$, and assume we are given a mapping $f: \{x_i H \mid i \in \mathbb{N}\} \rightarrow G$ say $x_i H \mapsto g_i \in G$. By freeness of W , we know that the mapping $x_i \mapsto g_i$ extends to a homomorphism $\theta: W \rightarrow G$. As $G \in \mathcal{V}$, we have that $H \subseteq \ker \theta$ hence $\theta_H: W/H \rightarrow G$ is well defined and is a group homomorphism extending f , as required. $\square$

As was mentioned above (see Exercise 5.8), given a fixed $n \in \mathbb{N}$, the groups of exponent dividing n form a variety.

Definition 5.13. The *free Burnside group* $B(r, n)$ is the free group in r generators for the variety of groups of exponent dividing n .

In fact, by Lemma 5.12, the Burnside group $B(r, n)$ is precisely the quotient of the free group $F(r)$ in r -generators by the normal closure of the set $\{w^n \mid w \in F(r)\}$.

5.3. Connection with the restricted Burnside problem. For this subsection, we draw mainly from [21] and [14] for Theorem 5.17.

Theorem 5.14. Let $r \in \mathbb{N}$ and p be a prime number. Let $B(r, p)$ be the Burnside group and $(\gamma_i)_{i \in \mathbb{N}}$ the LCS. The following are equivalent.

- (1) There exists a finite number of finite r -generated groups of exponent p (up to isomorphism).
- (2) There exists a bound on the order of finite r -generated groups of exponent p .
- (3) There exists $i_0 \in \mathbb{N}$ such that $\gamma_{i_0} = \gamma_j$ for all $j \geq i_0$.
- (4) The associated Lie algebra $L(B(r, p))$ is finite.
- (5) There exists i_0 such that $B(r, p)/\gamma_{i_0}$ is the largest finite r -generator group of exponent p and every other finite r -generated group of exponent p is isomorphic to a quotient of $B(r, p)/\gamma_{i_0}$.

Proof. Set $B = B(r, p)$ and consider the associated Lie algebra $L(B) = \bigoplus_{i \in \mathbb{N}} \gamma_i / \gamma_{i+1}$. By Lemma 5.5, each quotient γ_i / γ_{i+1} is again finitely generated, hence as they are abelian, each quotient is finite. Using iteratively the third isomorphism theorem, we have that $|B/\gamma_i| = |B/\gamma_2| |\gamma_2/\gamma_3| \dots |\gamma_{i-1}/\gamma_i|$ so G/γ_i is finite. (1) $\implies$ (2) is trivial. For (2) $\implies$ (3) assume that (3) does not hold. Then $\gamma_{i+1} \subsetneq \gamma_i$ for each i , hence $B/\gamma_i \subsetneq B/\gamma_{i+1} \subsetneq \dots$ is an infinite increasing chain of finite r -generated groups of exponent p , contradicting (2). (3) $\implies$ (4) is immediate since γ_i / γ_{i+1} is finite for $j < i_0$ and $\gamma_j / \gamma_{j+1} = 0$ for $j \geq i_0$. (4) $\implies$ (3) If there were no such i_0 , then by the

above, we would have an infinite decreasing chain $\gamma_1 \supsetneq \gamma_2 \supsetneq \dots$, hence the sum $\bigoplus_{i \in \mathbb{N}} \gamma_i / \gamma_{i+1}$ would be infinite. (3) $\implies$ (5). Let G be any finite r -generated group of exponent p . Then there exists a normal subgroup $N \trianglelefteq B$ such that $B/N \cong G$. As finite groups of exponent p are nilpotent, B/N is nilpotent, hence there exists i such that $\gamma_i \subseteq N$, so also $\gamma_{i_0} \subseteq N$ and hence B/N is a subgroup of B/γ_{i_0} . As G was arbitrary, we have the first part of (5). The second part is merely $G \cong B/N \cong (B/\gamma_{i_0})/(N/\gamma_{i_0})$. (5) $\implies$ (1) is trivial since B/γ_{i_0} is finite, hence has only a finite number of subgroups. $\square$

We will use the following classical formula, see e.g. [15, Theorem 12.3.1]

Fact 5.15 (Hall-Petrescu formula). *In any group, we have modulo $\gamma_{p+1}(\langle a, b \rangle)$:*

$$a^p b^p = (ab)^p [a, b^{(p-1)}] \cdot C$$

where C is a product of commutators in a, b of weight p and where a appears twice.

Lemma 5.16. *Let $G = \langle x, y_1, \dots, y_s \rangle$ be a group of exponent p . Then*

$$[x, (y_1 \cdots y_s)^{p-1}]$$

is equal modulo γ_{p+1} to a product of commutators in $x, y_1, \dots, y_s$ of length p and where x appears twice.

Proof. Using the Hall-Petrescu formula on $a = x$ and $b = y_1 \cdots y_s$, one obtains, since G has exponent p and using (3) above

$$[x, (y_1 \cdots y_s)^p] = C$$

modulo $\gamma_{p+1}(\langle a, b \rangle)$, where C is a product of commutators in $x, y_1 \cdots y_s$ of weight p and where x appears twice. Note that the inverse of a commutator is again a commutator since $[g, h]^{-1} = [h, g]$. Let c be such a commutator. Using $[x, yz] = [x, z][x, y][x, y, z]$ we immediately see that c is equal modulo γ_{p+1} to a product of commutators of length p in $x, y_1, \dots, y_s$ where x still appears twice. $\square$

Theorem 5.17. *Let $r \in \mathbb{N}$ and p be a prime number. Then $L(B(r, p))$ satisfies the $(p-1)$ -Engel identity:*

$$0 = [x, y^{(p-1)}].$$

Proof. We work in $B = B(p, p)$, the relatively free Burnside group of exponent p generated by $x, y_1, \dots, y_{p-1}$. Let $a = x\gamma_2, b_i = y_i\gamma_2$. From Claim 1, we have the following equality in $L(B)$

$$\begin{aligned} \sum_{\sigma \in \mathfrak{S}_{p-1}} [a, b_{1\sigma}, \dots, b_{(p-1)\sigma}] &= \sum_{\emptyset \subsetneq S \subsetneq \{1, \dots, p-1\}} (-1)^{p-1-|S|} [a, (\sum_{i \in S} b_i)^{(p-1)}] \\ &= [a, (b_1 + \dots + b_{p-1})^{(p-1)}] + \sum_{\emptyset \subsetneq S \subsetneq \{1, \dots, p-1\}} (-1)^{p-1-|S|} [a, (\sum_{i \in S} b_i)^{(p-1)}] \\ &= [a, (b_1 + \dots + b_{p-1})^{(p-1)}] \\ &\quad + \sum_{\substack{\emptyset \subsetneq S \subsetneq \{1, \dots, p-1\}, \\ i_1, \dots, i_{p-1} \in S}} (-1)^{p-1-|S|} [a, b_{i_1}, \dots, b_{i_{p-1}}] \end{aligned}$$

As y_i are in γ_1 , b_i belong to $L_1 = \gamma_1/\gamma_2$ hence each sum of the form $\sum_{i \in S} b_i$ translate in B as the product $\prod_{i \in S} y_i$. Similarly, by Lemma 5.4, commutators of weight p in x, y_i belongs to γ_p hence they commute modulo γ_{p+1} , hence the equality above translates readily to the following equation in B , modulo γ_{p+1} :

$$\prod_{\sigma \in \mathfrak{S}_{p-1}} [x, y_{1\sigma}, \dots, y_{(p-1)\sigma}] = [x, (y_1 \cdots y_{p-1})^{(p-1)}] \cdot \prod_{\substack{\emptyset \subsetneq S \subsetneq \{1, \dots, p-1\}, \\ i_1, \dots, i_{p-1} \in S}} [x, y_{i_1}, \dots, y_{i_{p-1}}]^{(-1)^{p-1-|S|}}$$

Using Lemma 5.16, we know that the term $[x, (y_1 \cdots y_{p-1})^{(p-1)}]$ of the right hand side can be written modulo γ_{p+1} as product of commutators in x, y_i of weight p where x appears twice. In turn we have

$$\prod_{\sigma \in \mathfrak{S}_{p-1}} [x, y_{1\sigma}, \dots, y_{(p-1)\sigma}] = c_1 \cdots c_s \pmod{\gamma_{p+1}} \quad (\dagger)$$

where each c_i is a commutator of weight p involving either x or some y_i twice. Note that all of them involve x at least once.

Let θ_1 be an endomorphism of B mapping y_1 to 1 and fixing x, y_i pointwise for each $i > 1$. Applying θ_1 on $(\dagger)$, the left hand side yields 1 while on the right hand side only commutators not involving y_1 remain. As γ_{p+1} is stable under θ_1 , we conclude that the product of all commutators among $c_1, \dots, c_s$ which do not involve y_1 are zero modulo γ_{p+1} . Each of the c_i are in γ_p and commute modulo γ_{p+1} hence we may regroup them and as any of their product is in γ_{p+1} , we may assume that all c_i in $(\dagger)$ involve at least one y_1 . Using an endomorphism θ_2 of B mapping y_2 to 1 and fixing $x, y_1, y_3, \dots, y_{p+1}$, we may similarly assume that all c_i involve y_2 . Iterating the argument,

we see that each c_i involve all of $x, y_1, \dots, y_{p-1}$ and either x or y_i appears twice in c_j for each j . This contradicts that the c_j 's are of weight p , unless they are all 1, which proves that $\prod_{\sigma \in \mathfrak{S}_{p-1}} [x, y_1 \sigma, \dots, y_{(p-1) \sigma}] \in \gamma_{p+1}$. In particular, there exists a word $f(x, y_1, \dots, y_{p-1})$ consisting of product of commutators of weight at least $p+1$ such that we have the following equality in B :

$$\prod_{\sigma \in \mathfrak{S}_{p-1}} [x, y_1 \sigma, \dots, y_{(p-1) \sigma}] = f(x, y_1, \dots, y_{p+1})$$

Let $u \in \gamma_{i_0}, v_1 \in \gamma_{i_1}, \dots, v_{p-1} \in \gamma_{i_{p-1}}$. Then, using an endomorphism of B sending x to u and y_i to v_i , we have

$$\prod_{\sigma \in \mathfrak{S}_{p-1}} [u, v_1 \sigma, \dots, v_{(p-1) \sigma}] = f(u, v_1, \dots, v_{p+1}).$$

By Lemma 5.4, we know that $\prod_{\sigma \in \mathfrak{S}_{p-1}} [u, v_1 \sigma, \dots, v_{(p-1) \sigma}] \in \gamma_{i_0 + \dots + i_{p-1}}$. As $f(u, v_1, \dots, v_{p-1})$ is a product of elements each belonging to $\gamma_{i_0 + \dots + 2i_k + \dots + i_{p-1} + 1}$ for some k (again by Lemma 5.4), in particular $f(u, v_1, \dots, v_{p-1}) \in \gamma_{i_0 + \dots + i_{p-1} + 1}$ and so $\prod_{\sigma \in \mathfrak{S}_{p-1}} [u, v_1 \sigma, \dots, v_{(p-1) \sigma}] \in \gamma_{i_0 + i_1 + \dots + i_{p-1} + 1}$. In $L(B)$, this translates as

$$\sum_{\sigma \in \mathfrak{S}_{p-1}} [c, d_1 \sigma, \dots, d_{(p-1) \sigma}] = 0 \quad (\ddagger)$$

for all homogeneous elements $c, d_1, \dots, d_{p-1}$. Because the map

$$(x, y_1, \dots, y_{p-1}) \mapsto \sum_{\sigma \in \mathfrak{S}_{p-1}} [x, y_1 \sigma, \dots, y_{(p-1) \sigma}]$$

is multilinear, the equality $(\ddagger)$ holds for all elements of $L(B)$. By Lemma 4.2 we conclude the theorem for $r = p$. To conclude for all $r \in \mathbb{N}$, if $r < p$, embed $B(r, p)$ in $B(p, p)$ and if $r > p$, this is immediate. $\square$

Remark 5.18. It is a longstanding open question whether the Lie algebra $L(B(r, p))$ is itself a free Lie algebra. The more general question whether $L(F)$ is itself a relatively free Lie algebra when F is a relatively free group was open for a long time and solved negatively by Daniel Groves in the 1999 [6].

Corollary 5.19. *The restricted Burnside problem for groups of exponent p holds if every finitely generated $(p-1)$ -Engel Lie $\mathbb{F}_p$ -algebras are nilpotent.*

Proof. Using Theorem 5.14, the restricted Burnside problem for groups of exponent p holds provided for each $r \in \mathbb{N}$ and prime p , $L(B(r, p))$ is finite. By definition, as $B(r, p)$ is finitely generated as a group, $L(B(r, p))$ is finitely generated as a Lie algebra (see Exercise 5.21). By Theorem 5.17, $L(B(r, p))$ is $(p-1)$ -Engel hence under our assumption, $L(B(r, p))$ is nilpotent. It is immediate that a finitely generated nilpotent Lie algebra over $\mathbb{F}_p$ is finite, since it is spanned by all monomials of length at most the nilpotency class evaluated in the generators. $\square$

Exercise 5.20. In this exercise, we prove that finite p -groups are nilpotent. Let G be a finite p -group.

- (1) The class equation gives that the action of G on itself has non-central representant $x_1, \dots, x_s$ and that

$$|G| = |Z(G)| + \sum_i |G/C_G(x_i)|.$$

Prove that $Z(G) \neq 1$.

- (2) Prove that if $G/Z(G)$ is nilpotent then G is nilpotent.
- (3) Prove by induction on $|G|$ that a finite p -group is nilpotent.

Exercise 5.21. Prove that if G is generated as a group by $g_1, \dots, g_r$ then $L(G)$ is generated as a Lie algebra by $g_1, g_2, \dots, g_r$.

Exercise 5.22. (1) Prove that for any exponent p group G , we have $L_{i+1} = [L_i, L_1]$ in $L(G)$.

- (2) Prove that $\gamma_i(L) = \bigoplus_{j \geq i} L_j$.

- (3) Prove that one can also add the following statement as an equivalent condition in Theorem 5.14:

The associated Lie algebra $L(B(r, p))$ is nilpotent.

6. PROBLEM

Exercise 6.1. The goal of this exercise is to prove Higgins' Theorem:

A n -Engel Lie algebra over $\mathbb{F}_p$ with $p > n$ is solvable if and only if it is nilpotent.

Given a Lie algebra L , the *derived series* $(L^{(k)})_{k \in \mathbb{N}}$ is defined as follows: $L^{(0)} = L^1 = L$, and $L^{(k+1)} = [L^{(k)}, L^{(k)}]$. L is called *solvable* if there exists $k \in \mathbb{N}$ such that $L^{(k)} = 0$. Recall that the lower central series $(L^k)_{k \geq 1}$ is defined as $L^1 = L$ and $L^{k+1} = [L^k, L]$.

- (1) Prove that $L^1 = L^{(0)}$ and that $L^2 = L^{(1)}$.
- (2) Prove that $L^{(k)} \subseteq L^{k+1}$.
- (3) Deduce that if L is nilpotent, then it is solvable.
- (4) Let $M = L^2 = L^{(1)} = [L, L]$. We prove by induction that $L^{ni+2} \subseteq M^{i+1}$ for all $i \in \mathbb{N}$.
 - (a) Check the initialization $i = 0$.
 - (b) Prove that for each $k \in \mathbb{N}$, M^k is an ideal of L .
 - (c) Let $a, b_1, \dots, b_n \in L$. Prove that for each $\sigma \in \mathfrak{S}_n$, a term $[a, b_{1\sigma}, \dots, b_{n\sigma}]$ can be written as the sum of $[a, b_1, \dots, b_n]$ and a sum of commutators of the form $Z(f, k) = [a, b_{f_1}, \dots, [b_{f_k}, b_{f_{k+1}}], b_{f_{k+2}}, \dots, b_{f_n}]$ for $f = (f_1, \dots, f_n)$ and $\{f_1, \dots, f_n\} = \{1, \dots, n\}$.
 - (d) Deduce that $n![a, b_1, \dots, b_n]$ can be written as a sum of elements of the form $Z_{f,k}$ for some f, k as in the previous question.
 - (e) Deduce that $[a, b_1, \dots, b_n]$ can be written as a sum of elements of the form $Z_{f,k}$ for some f, k as in the previous question.
 - (f) Using (b) and (e), prove that $[M^i, b_1, \dots, b_n] \subseteq M^{i+1}$.
 - (g) Conclude, using induction and the previous question, that we have $L^{ni+2} \subseteq M^{i+1}$.
- (5) Assume by induction hypothesis, that for all $k < d$ and all n -Engel Lie algebra S , we have $S^{(k)} = 0$ implies that S is nilpotent. Assume that L is an n -Engel Lie algebra which satisfies $L^{(d)} = 0$.
 - (a) Prove that $M^{(d-1)} = 0$.
 - (b) Apply the induction hypothesis and the result of Question (4) to conclude Higgins' Theorem.

Exercise 6.2. Let L be a Lie algebra satisfying the identity $[x, y^{(2)}, z^{(2)}] = 0$ and $[x, y, z^2] = [x, z^2, y]$.

- (1) Let I be the vector span of $\{[a, b^2] \mid a, b \in L\}$. Prove that I is an ideal of L .
- (2) Prove that L/I is nilpotent.
- (3) Prove that I is nilpotent.
- (4) Conclude.
- (5) Prove that a Lie algebra satisfying $[x, y^{(2)}, z^{(2)}, t^{(2)}] = 0$ and $[x, y, z^2] = [x, z^2, y]$ is also nilpotent.

Exercise 6.3. Let A be an associative algebra of nilindex 3.

- (1) Prove that for all $a, b, c \in A$ we have $abc + acb + bac + bca + cab + cba = 0$.
- (2) Assume further that A satisfies: for every $a, b \in A$, $ab = 0$ or $ba = 0$. Prove that A is nilpotent of class at most 2.

Exercise 6.4. Let G be a group of exponent 3.

- (1) Prove that the enveloping algebra $L(G)$ is 2-Engel.
 - (2) Deduce that G is nilpotent.
-

APPENDIX A. MORE ON n -ENGEL LIE ALGEBRAS

A.1. Varieties of Lie algebras, free Lie algebras. We fix a field $\mathbb{F}$, and we consider the set $\text{Lie}(x_1, \dots, x_n)$ consisting of $\mathbb{F}$ -linear combination of (formal) Lie monomials in $x_1, \dots, x_n$. Those are called *(formal) Lie polynomials over $\mathbb{F}$* in $x_1, \dots, x_n$. For a Lie algebra L and $a_1, \dots, a_n \in L$, we may again evaluate any element $p(x_1, \dots, x_n)$ from $\text{Lie}(x_1, \dots, x_n)$ by interpreting the expression $p(a_1, \dots, a_n)$ in L . For instance if $p(x_1, x_2, x_3) = [x_1, x_2, x_3] + [x_2, x_3, x_1] + [x_3, x_1, x_2]$, then $p(a_1, a_2, a_3) = 0$. Note that there is no structure on $\text{Lie}(x_1, \dots, x_n)$, for instance, $[x_1, x_1]$ and $[x_2, x_2]$ are formally two different objects (both evaluated as 0 in any Lie algebra).

Definition A.1 (Varieties of Lie $\mathbb{F}$ -algebras). Let V be a set of Lie polynomials in $\text{Lie}(x_i \mid i \in \mathbb{N})$. The *variety generated by V* is the collection of all Lie algebras L such that for all $p(x_1, \dots, x_n) \in V$ and for all $a_1, \dots, a_n \in L$ we have $p(a_1, \dots, a_n) = 0$. A *variety of Lie algebras* is a variety defined by some set of words.

Exercise A.2. Check that the class of n -Engel Lie $\mathbb{F}$ -algebras is a variety.

Definition A.3 (Free Lie algebras). Given a variety of Lie $\mathbb{F}$ -algebras $\mathcal{V}$ and a set S of variable, a *free Lie $\mathbb{F}$ -algebra for $\mathcal{V}$ generated by S* is an element $F(S)$ of $\mathcal{V}$ containing S and such that for any Lie $\mathbb{F}$ -algebra $L \in \mathcal{V}$, any mapping $S \rightarrow L$ extends uniquely to a Lie $\mathbb{F}$ -algebra homomorphism $F(S) \rightarrow L$.

$$\begin{array}{ccc} S & \xrightarrow{\quad} & L \\ \cap & \nearrow & \\ F(S) & & \end{array}$$

Theorem A.4. Every variety of Lie $\mathbb{F}$ -algebra admits a free Lie algebra in any given set of generators.

The existence of a free Lie $\mathbb{F}$ -algebra (for the class of all Lie $\mathbb{F}$ -algebras) can be defined via the free associative $\mathbb{F}$ -algebra $A = A(x_i \mid i \in \mathbb{N})$. Before defining A , observe that the subset $W^+(\omega)$ of $W(\omega)$ of positive words (i.e. that do not involve x_i^{-1}) is a monoid for the multiplication. Now let $A(\omega)$ be the associative algebra of $\mathbb{F}$ -linear combination of words in $W^+(\omega)$ (which is an $\mathbb{F}$ -vector space) and where multiplication in $A(\omega)$ is defined as such

$$\left(\sum_i \lambda_i v_i \right) \cdot \left(\sum_i \mu_i w_i \right) = \sum_{i,j} \lambda_i \mu_j v_i w_j$$

for $\lambda_i, \mu_j \in \mathbb{F}$, $v_i, w_j \in W^+(\omega)$. Then the bracket $[a, b] = ab - ba$ in $A(\omega)$ turns $(A, +, [, \cdot])$ into a Lie algebra. Then the Lie subalgebra $L(\omega)$ of $A(\omega)$ generated by $(x_i)_{i \in \mathbb{N}}$ is a free Lie $\mathbb{F}$ -algebra. One should think of $L(\omega)$ as the Lie algebra of Lie polynomials in $(x_i)_{i \in \mathbb{N}}$, just as the usual polynomials $\mathbb{F}[(X_i)_{i \in \mathbb{N}}]$ are in fact the free object in the variety of all commutative $\mathbb{F}$ -associative algebras. We conflate monomials in $L(\omega)$ with formal Lie monomials in variables $(x_i)_{i \in \mathbb{N}}$. The free Lie algebra $L(\omega)$ is *multigraded*, in the sense that it comes equipped with a well-defined notion of *multiweight* which associated to each nonzero monomial in variables $(x_1, \dots, x_n)$ the tuple $(i_1, \dots, i_n)$ where each i_j is the number of times x_i appears in the monomial. For a given $(i_1, \dots, i_n)$ we define $V_{(i_1, \dots, i_n)}$ to be the vector space spanned by elements of multiweight $(i_1, \dots, i_n)$. One key fact about multiweight is that for two different multiweight $(i_1, \dots, i_n), (j_1, \dots, j_n)$ we have

$$V_{(i_1, \dots, i_n)} \cap V_{(j_1, \dots, j_n)} = \{0\}.$$

We also have that

$$[V_{(i_1, \dots, i_n)}, V_{(j_1, \dots, j_n)}] \subseteq V_{(i_1+j_1, \dots, i_n+j_n)}.$$

Note that by allowing the multiweight to The *weight* of the monomial is the sum $i_1 + \dots + i_n$. For instance, $[x, y^{(3)}]$ has multiweight $(1, 3)$ and weight 4. If $[x, y^{(3)}]$ can be written as a linear combination $\lambda_1 m_1 + \dots + \lambda_k m_k$ then by the above, we may assume that each m_i is again of multiweight $(1, 3)$ in (x, y) . Another important fact that will be used later is the following: assume that $J = \phi(L(\omega))$ is an ideal which is given by the image of a multilinear map $\phi : L^k \rightarrow L$, then $L(\omega)/J$ is again multigraded.

Again, we might sometimes call *relatively free* a free Lie algebra for a variety strictly smaller than the variety of all Lie $\mathbb{F}$ -algebras. An ideal I of $F = L(\omega)$ is called *fully invariant* if for all endomorphism θ of F we have $I\theta \subseteq I$.

Lemma A.5. Let I be a fully invariant ideal of the free Lie $\mathbb{F}$ -algebra $F = L(\omega)$. Then F/I is a free Lie $\mathbb{F}$ -algebra for some variety $\mathcal{V}$.

Proof. We set $\mathcal{V}$ be the variety of Lie $\mathbb{F}$ -algebras defined by the following subset of $\text{Lie}((x_i)_{i \in \mathbb{N}})$:

$$\{q \in \text{Lie}((x_i)_{i \in \mathbb{N}}) \mid q(x_{i_1}, \dots, x_{i_k}) \in I \text{ for some distinct } i_1, \dots, i_k \in \mathbb{N}\}.$$

We prove that L/I is a free Lie algebra relatively to $\mathcal{V}$, freely generated by $(x_i + I)_{i \in \mathbb{N}}$. Note that L/I belongs to $\mathcal{V}$ since I is fully invariant. Let L be any Lie algebra in $\mathcal{V}$, and assume we are given a mapping $f : \{x_i + I \mid i \in \mathbb{N}\} \rightarrow L$ say $x_i + I \mapsto a_i \in L$. By freeness of F , we know that the mapping $x_i \mapsto a_i$ extends to a homomorphism $\theta : F \rightarrow L$.

As $L \in \mathcal{V}$, we have that $I \subseteq \ker \theta$ hence $\theta_I : F/I \rightarrow L$ is well defined and is a Lie algebra homomorphism extending f , as required. $\square$

A relatively free Lie algebra of characteristic 0 is always multigraded.

Exercise A.6. Let $\mathcal{V}$ be a variety of Lie $\mathbb{F}$ -algebra with free group $F = F(x_1, \dots, x_n)$. Assume that $[x_1, \dots, x_n] = 1$.

- (1) Prove that F is nilpotent of class $< n$.
- (2) Prove that every element of $\mathcal{V}$ is nilpotent.

Exercise A.7. Let $\mathcal{V}$ be a variety of Lie $\mathbb{F}$ -algebras such that every element of $\mathcal{V}$ is n -Engel for some $n \in \mathbb{N}$. Prove that there exists $N \in \mathbb{N}$ such that every element of $\mathcal{V}$ is N -Engel.

Exercise A.8. Prove that if the free n -Engel Lie $\mathbb{Q}$ -algebra in countably generators F is nilpotent, then so are any n -Engel Lie algebra over a field of characteristic 0.

Exercise A.9. (1) Prove that for each $n \in \mathbb{N}$, the vector space

$$\text{Span}_{\mathbb{F}} \left(\sum_{\sigma \in \mathfrak{S}_n} [a, b_{1\sigma}, b_{2\sigma}, \dots, b_{n\sigma}] \mid a, b_i \in L \right)$$

is an ideal.

- (2) Deduce that if $\mathbb{F}$ is of characteristic p with $p > n$, then the free n -Engel Lie algebra is multigraded.

A.2. Higgins Theorem. Given a Lie algebra L , the *derived series* is defined as follows: $L^{(0)} = L^1 = L$, and $L^{(i+1)} = [L^{(i)}, L^{(i)}]$. A Lie algebra is *solvable* if there exists $k \in \mathbb{N}$ such that $L^{(k)} = 0$. Every nilpotent Lie algebra is solvable, but the converse does not hold in general. However, for $(p-1)$ -Engel Lie algebras, it holds. This is a classical result of Higgins and the proof here is the one of the original paper [8].

Theorem A.10. Let L be a n -Engel Lie algebra over a field of characteristic 0 or over $\mathbb{F}_p$ with $n < p$. If $L^{(d)} = 0$ then $L^k = 0$ where $k = \frac{n^d-1}{n-1} + 1$.

Proof. Let $M = L^2 = [L, L]$. We first establish by induction that for all $i \in \mathbb{N}$ we have

$$L^{ni+2} \subseteq M^{i+1}. \quad (\star)$$

It is trivial for $i = 0$, so assume $i > 0$. By Lemma 4.2, we have that $\sum_{\sigma \in \mathfrak{S}_n} [a, b_{1\sigma}, \dots, b_{n\sigma}] = 0$. Using the Jacobi identity, we have

$$\begin{aligned} [c, d_1, \dots, d_{i-1}, d_i, d_{i+1}, d_{i+2}, \dots, d_n] &= [c, d_1, \dots, d_{i-1}, d_{i+1}, d_i, d_{i+2}, \dots, d_n] \\ &\quad + [c, d_1, \dots, d_{i-1}, [d_i, d_{i+1}], d_{i+2}, \dots, d_n] \end{aligned}$$

Using this, each term $[a, b_{1\sigma}, \dots, b_{n\sigma}]$ can be written as the sum of $[a, b_1, \dots, b_n]$ and a sum of commutators of the form $Z(f, k) = [a, b_{f_1}, \dots, [b_{f_k}, b_{f_{k+1}}], b_{f_{k+2}}, \dots, b_{f_n}]$. Therefor, we have

$$0 = \sum_{\sigma \in \mathfrak{S}_n} [a, b_{1\sigma}, \dots, b_{n\sigma}] = n![a, b_1, \dots, b_n] + Z$$

where Z is a sum of $Z(f, k)$. It follows that

$$n![M^i, b_1, \dots, b_n] = \sum_{f, k} [M_i, b_{f_1}, \dots, [b_{f_k}, b_{f_{k+1}}], b_{f_{k+2}}, \dots, b_{f_n}]$$

As M^i is an ideal, we have

$$[M_i, b_{f_1}, \dots, [b_{f_k}, b_{f_{k+1}}], b_{f_{k+2}}, \dots, b_{f_n}] \subseteq [M_i, [b_{f_k}, b_{f_{k+1}}], b_{f_{k+2}}, \dots, b_{f_n}]$$

Now as $[b_{f_k}, b_{f_{k+1}}] \in M$ and M^{i+1} is an ideal we conclude

$$[M_i, [b_{f_k}, b_{f_{k+1}}], b_{f_{k+2}}, \dots, b_{f_n}] \subseteq M^{i+1}.$$

In turn, we have proved that $[M^i, b_1, \dots, b_n] \subseteq M^{i+1}$. Now by induction hypothesis, we have that

$$L^{n(i-1)+2} \subseteq M^i.$$

A generic element of $L^{ni+2} = [L^{n(i-1)+2}, \underbrace{L, \dots, L}_{n \text{ times}}]$ will belong to $[M^i, b_1, \dots, b_n] \subseteq M^{i+1}$, so we conclude that $(\star)$ holds.

Suppose that $L^{(d)} = 0$, we prove the lemma by induction on d . If $d = 1$, this is clear as $L^{(1)} = L^2 = 0$. Assume that $d > 1$ and $L^{(d)} = 0$. Then, for $M = L^2$, we easily check that $M^{(d-1)} = 0$ and that M is n -Engel, so that the induction hypothesis applies and $M^k = 0$ for $k = \frac{n^{d-1}-1}{n-1} + 1$. By $(\star)$, $L^{n(k-1)+2} \subseteq M^k = 0$ and $n(k-1) + 2 = \frac{n^d-1}{n-1} + 1$, so we conclude. $\square$

We immediately get:

Corollary A.11. *Let L be a $(p-1)$ -Engel Lie algebra over $\mathbb{F}_p$. Then L is nilpotent if and only if L is solvable.*

A.3. $\mathbb{Z}_2$ -gradings on Engel Lie algebras.

Lemma A.12. *Let L be an n -Engel Lie algebra over a field of characteristic $p > n$. Then for all $m \geq n$ and $x, y_1, \dots, y_m$, the term $[x, y_1, \dots, y_m]$ can be written as a linear combination of monomials of the form $[x, u_1, \dots, u_{n-1}]$ where each u_i is a monomial in $y_1, \dots, y_m$.*

Proof. Using $y_i y_j = y_j y_i + [y_i, y_j]$, we obtain that for each $\sigma \in \mathfrak{S}(n)$, the monomial $[x, y_{1\sigma}, \dots, y_{n\sigma}]$ can be written as the sum of $[x, y_1, \dots, y_n]$ and a sum of monomials of the form

$$[x, y_{i_1}, \dots, y_{i_{s-1}}, [y_{i_s}, y_{i_{s+1}}], y_{i_{s+2}}, \dots, y_{i_n}].$$

Doing this for all $\sigma \in \mathfrak{S}(n)$ and summing up, we obtain:

$$n![x, y_1, \dots, y_n] = \sum_{\sigma \in \mathfrak{S}(n)} [x, y_{1\sigma}, \dots, y_{n\sigma}] + P$$

where P is a sum of terms of the form $[x, y_{i_1}, \dots, y_{i_{s-1}}, [y_{i_s}, y_{i_{s+1}}], y_{i_{s+2}}, \dots, y_{i_n}]$, hence monomials of the form $[x, u_1, \dots, u_{n-1}]$ where each u_i is a monomial in $y_1, \dots, y_n$. By Lemma 4.2, the term $\sum_{\sigma \in \mathfrak{S}(n)} [x, y_{1\sigma}, \dots, y_{n\sigma}]$ vanishes and as $p > n$ we may divide by $n!$ so that the statement of the lemma holds for $m = n$. By induction, assume that the statement holds for $m \geq n$ and consider $[x, y_1, \dots, y_m, y_{m+1}]$. As $[x, y_1, \dots, y_m]$ can be written as a linear combination of terms $[x, u_1, \dots, u_{n-1}]$ where the u_i 's are monomials in $y_1, \dots, y_m$, the term $[x, y_1, \dots, y_{m+1}]$ can be written as a linear combination of monomials of the form $[x, u_1, \dots, u_{n-1}, y_{m+1}]$. By previously, each $[x, u_1, \dots, u_{n-1}, y_{m+1}]$ can be written as the sum of terms of the form $[x, v_1, \dots, v_{n-1}]$, where each v_i are monomials in $u_1, \dots, u_{n-1}, y_{m+1}$ hence in fact monomials in $y_1, \dots, y_{m+1}$, so we conclude the lemma. $\square$

Definition A.13. A $\mathbb{Z}_2$ -grading on a Lie algebra L is a decomposition $L = L_0 \oplus L_1$ into vector spaces such that $[L_0, L_0] \subseteq L_0$, $[L_0, L_1] \subseteq L_1$, $[L_1, L_1] \subseteq L_0$. In order words: $[L_i, L_j] \subseteq L_{i+j}$ where the addition happens in $\mathbb{Z}_2$.

Example A.14. Let L be a relatively free Lie algebra with free generators $x_1, \dots, x_K$. Let $S \subseteq \{1, \dots, K\}$, we define a $\mathbb{Z}_2$ -grading $L = L_0 + L_1$ with

$$\begin{aligned} x_i \in L_0 &\iff i \in S \\ x_i \in L_1 &\iff i \notin S. \end{aligned}$$

To do so, consider the set U of all left-normed commutators in $x_1, \dots, x_K$ and split U in two subsets $U = U_0 \cup U_1$ where U_0 is the set of commutators where an even number of elements from $\{x_j \mid j \notin S\}$ occurs and $U_1 = U \setminus U_0$. Define $L_0 = \text{Span}(U_0)$ and $L_1 = \text{Span}(U_1)$. As every monomial is a linear combination of left-normed commutators, we have $L = L_0 \oplus L_1$. The number of such gradings on L is 2^K .

Lemma A.15. *Let $L = L_0 \oplus L_1$ be an n -Engel Lie algebra with a $\mathbb{Z}_2$ -grading and suppose that L_0 is nilpotent of class at most $m-1$. Then L is nilpotent of class bounded by*

$$\frac{n^{(n-1)(m-1)+m+1} - 1}{n - 1}.$$

Proof. Recall that the derived series of L is defined by $L^{(0)} = L$ and $L^{(i+1)} = [L^{(i)}, L^{(i)}]$. Then $L^{(1)} \subseteq L_0 + [L_1, L_0]$, $L^{(2)} \subseteq L_0 + [L_1, L_0, L_0]$ and more generally

$$L^{(k)} \subseteq L_0 + [L_1, \underbrace{L_0, \dots, L_0}_{k \text{ times}}]. \quad (\star)$$

First, we prove that $[L_1, L_0, \dots, L_0] = 0$ for $k = (m-1)(n-1) + 1$. For any $b, a_1, \dots, a_k$ with $k \geq n$, by Lemma A.12, $[b, a_1, \dots, a_k]$ is a linear combination of monomials of the form $[b, c_1, \dots, c_{n-1}]$ where $c_1, \dots, c_{n-1}$ are commutators in $a_1, \dots, a_k$ whose weight add up to k . Assume now that $b \in L_1$, $a_i \in L_0$ and that c is the nilpotency class of L_0 , then for $k = (m-1)(n-1) + 1$ one of the commutator c_i has weight m hence vanishes. Now by $(\star)$, we have $L^{(k)} \subseteq L_0$ hence $L^{(k+m)} = 0$. By Higgins theorem (Theorem A.10), L is nilpotent of class $\frac{n^{k+m-1}}{n-1}$. $\square$

A.4. Group algebra, Young tableau and representation theory of the symmetric group. References for this section are [4] and [9]. Given a field $\mathbb{F}$ and a finite group G , we may consider the *group algebra* $\mathbb{F}G$ which is the set of all formal linear combination $\sum_{g \in G} \lambda_g g$ with $\lambda \in \mathbb{F}$ where addition is defined naturally:

$$\left(\sum_g \lambda_g g \right) + \left(\sum_g \mu_g g \right) = \sum_g (\lambda_g + \mu_g) g$$

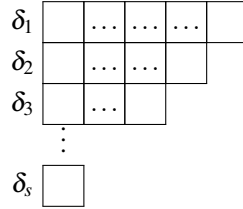
and multiplication is defined as follows:

$$\left(\sum_g \lambda_g g\right) \cdot \left(\sum_g \mu_g g\right) = \sum_{g,h} \lambda_g \mu_h gh.$$

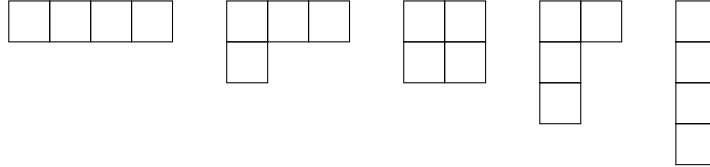
A *representation of G (over $\mathbb{F}$)* is the data of a right-action of G on an $\mathbb{F}$ -vector space V which is linear, in the sense that $(\lambda v + \mu w)g = \lambda vg + \mu wg$. A representation V of G (with a given right-action of G on V) is the same thing as equipping V with the action of a $\mathbb{F}G$ -module in the standard sense. Note that $\mathbb{F}G$ itself is already a $\mathbb{F}G$ -module (and of course a representation of G over $\mathbb{F}$), which is called the *standard representation*. Given any action of G on a subset X of a vector space over $\mathbb{F}$, this action extends linearly to $V = \text{Span}_{\mathbb{F}}(X)$ which gives a representation of G .

An $\mathbb{F}G$ -module V which does not admit nontrivial $\mathbb{F}G$ submodule is called *simple*. The equivalent notion for a representation is called *irreducible*, that is, a representation V of G such that no nontrivial vector subspace of V is closed under the action of G . The algebra $\mathbb{F}G$ is called *semisimple* if can be written as a direct sum of simple submodule, equivalently, the representation V can be written as a direct sum of irreducible representations. As G is finite, the condition for semisimplicity is given by the celebrated Maschke's theorem: $\mathbb{F}G$ is semisimple if and only if the characteristic of $\mathbb{F}$ does not divide $|G|$.

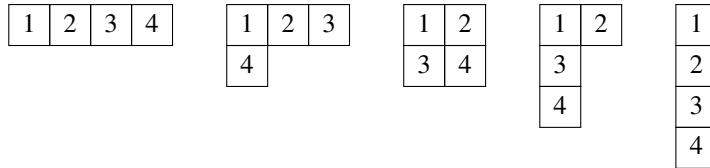
We are now interested in the representation theory of symmetric groups $\mathfrak{S}_d$. The irreducible representations of $\mathfrak{S}_d$ are classified and are given using the notion of *Young tableau*, which we define now. We fix $d \in \mathbb{N}$. Given a partition $\delta = (\delta_1 \geq \dots \geq \delta_s)$ of d (i.e. such that $d = \delta_1 + \dots + \delta_s$), a *Young diagram* is the following diagram



with δ_i boxes at the i -th row. For instance the Young diagrams associated to the five partitions $4 = 4$, $4 = 3 + 1$, $4 = 2 + 2$, $4 = 2 + 1 + 1$, $4 = 1 + 1 + 1 + 1$ are given as follows respectively



Given a Young diagram, one defines a *Young tableau* by filling in the boxes of a Young diagram with the numbers in $\{1, \dots, d\}$. Canonical examples are given by numbering the boxes in the (descending) lexicographical order:



One crucial property of a Young tableau for d is that it has either a first row of length at least (the smallest integers greater than or equal to) $\sqrt{d}$.

For each Young tableau T , let $V = V_T$ be the subgroup of $\mathfrak{S}_d$ which permutes elements within each column and $H = H_T$ the subgroup of $\mathfrak{S}_d$ which permutes elements within each row. The representation theory of the symmetric group is given by the following.

Theorem A.16. *Let $\mathbb{F}$ be either $\mathbb{Q}$ or $\mathbb{F}_p$, for $p > d$. For each Young tableau T there exists $s = s(T) \in \mathbb{N}$ such that s divides $d!$ and such that the element $e = e(T)$ defined by*

$$e = \frac{1}{s} \sum_{\pi \in V, \rho \in H} \text{sgn}(\pi) \pi \rho \in \mathbb{F}\mathfrak{S}_d$$

is idempotent, i.e. satisfies $e^2 = e$. We call it a primitive idempotent. Let $W = W(T)$ be the image of $\mathbb{F}\mathfrak{S}_d$ by right-multiplication by e , then W is an irreducible representation of $\mathfrak{S}_d$ and every irreducible representation of $\mathfrak{S}_d$ occurs this way. Let $T_1, \dots, T_k$ be all Young tableau (up to permutation of the numbering of the boxes). Let $e_1, \dots, e_k$ be the corresponding primitive idempotent and $W_i = \mathbb{F}\mathfrak{S}_d e_i$ for $i = 1, \dots, k$. Then the map $x \mapsto xe_1 + \dots + xe_k$ defines an isomorphism $\mathbb{F}\mathfrak{S}_d \cong W_1 \oplus \dots \oplus W_k$.

Of course, $\mathbb{F}\mathfrak{S}_d \cong V_1 \oplus \dots \oplus V_k$ provides the decomposition of the algebra $\mathbb{F}\mathfrak{S}_d$ as a sum of simple modules mentioned above. The main fact we will use from the above theorems is the following equivalent: $1 \in \mathbb{F}\mathfrak{S}_d$ can be written as a sum of primitive idempotent.

A.5. A theorem of Traustason on 4-Engel Lie algebras. In this subsection we will illustrate the use of representation theory of the symmetric group to give the strategy of the proof of Traustason's theorem below [19, Section 3.1, 3.2, 4.2]. We will not give a complete proof. We follow a presentation which draws more from Vaughan-Lee's approach in [23, 24].

Theorem A.17 (Traustason, 2011). *Let L be a 4-Engel Lie algebra of characteristic $p > 5$, the L is nilpotent of class at most 7.*

Let L be a free 4-Engel Lie algebra over $\mathbb{F}_p$ with $p > 5$, freely generated by $x_1, \dots, x_8$. By freeness, the theorem follows if we are able to prove that

$$[x_1, \dots, x_8] = 0. \quad (\dagger)$$

Goal. By putting together Claims 2, 3 and 4 below we will conclude the following statement.

The equality

$$[x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8] = 0$$

in free 8-generated 4-Engel Lie algebra of characteristic $p > 5$ follows from all the equalities

$$M(x_1, x_2, x_3, x_4, x_5, x_6) = 0$$

for all monomial M of multiweight $(3, 1, 1, 1, 1, 1)$ in a free 6-generated 4-Engel Lie algebra of characteristic $p > 5$ and in a free 6-generated Lie superalgebra of characteristic $p > 5$ which satisfies $\sum_{\sigma \in \mathfrak{S}(4)} [X, Y_{1\sigma}, Y_{2\sigma}, Y_{3\sigma}, Y_{4\sigma}] = 0$.

It might not be completely evident why this reduction is a reduction at all, in fact for some aspect it seems we have reduced the question to an even harder thing to prove! But the main point comes from the fact that the number of generators required has been reduced from 8 to 6. The number of generators is a huge problem when it comes to computational algebra and a reduction of the number of generators might be the only way to reduce the computing times drastically. In fact, Gunnar Traustason goes even further than that reduction above and goes down to 4 generators and the uses a computer program to check the identities. The strategy laid down in this section is a very good illustration of the methods that will be involved in the proof of Zelmanov's theorem.

Let M be the vector subspace of L generated by all multilinear monomials in $x_1, \dots, x_8$ of weight 8. Note that

$$M = \text{Span}_{\mathbb{F}_p}([x_{i_1}, \dots, x_{i_8}] \mid \{i_1, \dots, i_8\} = \{1, \dots, 8\}).$$

We let $\mathfrak{S}(8)$ act naturally on $x_1, \dots, x_8$ via $x_i\sigma = x_{i\sigma}$. This action extends to every monomial:

$$[x_{i_1}, \dots, x_{i_8}]\sigma = [x_{i_1\sigma}, \dots, x_{i_8\sigma}]$$

Then this action extends linearly by

$$\left(\sum_j \lambda_j [x_{i_1(j)}, \dots, x_{i_8(j)}] \right) \sigma = \sum_j \lambda_j [x_{i_1(j)\sigma}, \dots, x_{i_8(j)\sigma}] \sigma$$

so that M is given the structure of an $\mathbb{F}_p\mathfrak{S}(8)$ -module. We also see $\mathfrak{S}(5) \subseteq \mathfrak{S}(8)$ acting on M by the action on $x_1, \dots, x_5$ (with $x_6\sigma = x_6, x_7\sigma = x_7, x_8\sigma = x_8$) so that we may consider M as an $\mathbb{F}_p\mathfrak{S}(5)$ -module.

By Subsection A.4, the identity in $\mathbb{F}_p\mathfrak{S}(5)$ may be written as a sum of primitive idempotent $e_1 + \dots + e_s$. It follows that $(\dagger)$ holds if we can prove that for all primitive idempotent e of $\mathbb{F}_p\mathfrak{S}(5)$, we have

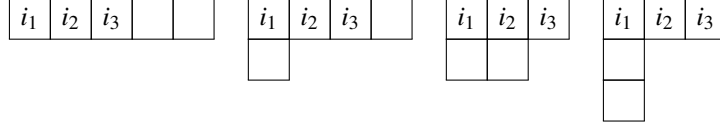
$$[x_1, \dots, x_8]e = 0 \quad (\ddagger)$$

We fix a primitive idempotent e corresponding to a Young tableau T and write

$$e = \frac{1}{s} \sum_{\pi \in V, \rho \in H} \text{sgn}(\pi) \pi \rho$$

for V the subgroup of $\mathfrak{S}(5)$ which permutes the columns of T and H the subgroup of $\mathfrak{S}(5)$ which permutes the rows of T . The tableau T has either the first row of length at least 3, or the first column of length at least 3, respectively type 1 or type 2.

Case 1: e has type 1. Assume first that T has type 1 and let i_1, i_2, i_3 be the first three elements of the first row. This corresponds to the following four Young diagrams.



Let S be the subgroup of H which permutes the entries i_1, i_2, i_3 in T , so that $S \cong \mathfrak{S}(3)$. Let C be a left transversal for S in H , i.e. such that $H = \bigsqcup_{c \in C} cS$ and let $f = \sum_{\sigma \in S} \sigma$. Then

$$\begin{aligned}
[x_1, \dots, x_8]e &= [x_1, \dots, x_8] \left(\frac{1}{s} \sum_{\pi \in V, \rho \in H} \text{sgn}(\pi) \pi \rho \right) \\
&= \frac{1}{s} \sum_{\pi \in V, c \in C, \sigma \in S} \text{sgn}(\pi) [x_1, \dots, x_8] \pi c \sigma \\
&= \frac{1}{s} \sum_{\pi \in V, c \in C} \text{sgn}(\pi) [x_1, \dots, x_8] \pi c \left(\sum_{\sigma \in S} \sigma \right) \\
&= \frac{1}{s} \sum_{\pi \in V, c \in C} \text{sgn}(\pi) [x_1, \dots, x_8] \pi c f
\end{aligned}$$

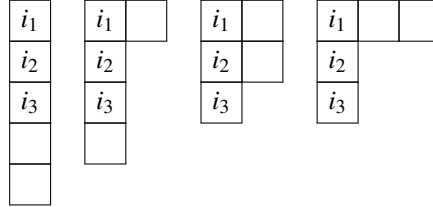
In order to establish $(\ddagger)$, it is enough to prove that each summand vanishes. Fix $\pi \in V, c \in C$ and Let $1 \leq j_1, j_2, j_3 \leq 5$ be such that $\{i_1, i_2, i_3\} = \{j_1 < j_2 < j_3\} \pi c$. It suffices to prove that

$$\sum_{\sigma \in S} [x_1 \pi c, \dots, (x_{j_1} \pi c) \sigma, \dots, (x_{j_2} \pi c) \sigma, \dots, (x_{j_3} \pi c) \sigma, \dots, x_5 \pi c, x_6, x_7, x_8] = 0$$

By renaming the variables, we have that for e of type 1, $(\ddagger)$ follows by establishing all the following equalities, for all $1 \leq i_1 < i_2 < i_3 \leq 5$

$$\sum_{\sigma \in \mathfrak{S}(3)} [x_1, \dots, x_{i_1 \sigma}, \dots, x_{i_2 \sigma}, \dots, x_{i_3 \sigma}, \dots, x_5, x_6, x_7, x_8]$$

Case 2: e has type 2. Now we assume that e has type 2, hence the first column has length at least 3 and let i_1, i_2, i_3 be the first three elements. This corresponds to the following four diagrams.



Let $S \cong \mathfrak{S}(3)$ be the subgroup of V which permutes $\{i_1, i_2, i_3\}$. Let C be a right transversal for S in V , i.e. such that $V = \bigsqcup_{c \in C} Sc$ and let $f = \sum_{\sigma \in S} \text{sgn}(\sigma) \sigma$. Then

$$\begin{aligned}
[x_1, \dots, x_8]e &= [x_1, \dots, x_8] \left(\frac{1}{s} \sum_{\pi \in V, \rho \in H} \text{sgn}(\pi) \pi \rho \right) \\
&= \frac{1}{s} \sum_{\sigma \in S, c \in C, \rho \in H} [x_1, \dots, x_8] \text{sgn}(\sigma c) \sigma c \rho \\
&= \frac{1}{s} \sum_{c \in C, \rho \in H} \text{sgn}(c) [x_1, \dots, x_8] \left(\sum_{\sigma \in S} \text{sgn}(\sigma) \sigma \right) c \rho \\
&= \frac{1}{s} \sum_{c \in C, \rho \in H} \text{sgn}(c) [x_1, \dots, x_8] f c \rho
\end{aligned}$$

Again, to establish $(\ddagger)$, it is enough to prove that each summand is zero and hence we obtain that it suffices to prove that for all $1 \leq i_1 < i_2 < i_3 \leq 5$, we have

$$\sum_{\sigma \in \mathfrak{S}(3)} \text{sgn}(\sigma) [x_1, \dots, x_{i_1 \sigma}, \dots, x_{i_2 \sigma}, \dots, x_{i_3 \sigma}, \dots, x_5, x_6, x_7, x_8]$$

We have established the following claim.

Claim 2. Equation $(\dagger)$ follows if we can establish the two following equations for all $1 \leq i_1 < i_2 < i_3 \leq 5$:

$$\sum_{\sigma \in \mathfrak{S}(3)} [x_1, \dots, x_{i_1\sigma}, \dots, x_{i_2\sigma}, \dots, x_{i_3\sigma}, \dots, x_5, x_6, x_7, x_8] \quad (E_{\text{sym}})$$

$$\sum_{\sigma \in \mathfrak{S}(3)} \text{sgn}(\sigma) [x_1, \dots, x_{i_1\sigma}, \dots, x_{i_2\sigma}, \dots, x_{i_3\sigma}, \dots, x_5, x_6, x_7, x_8] \quad (E_{\text{skew}})$$

Let E be the associative algebra generated by three elements e_1, e_2, e_3 and relations $e_i^2 = 0$ and $e_i e_j = e_j e_i$. Then consider the tensor product of algebras

$$\tilde{L} = L \otimes E.$$

Then L is a Lie algebra, and by Lemma 4.2, it is 4-Engel.

Exercise A.18. Prove that $\tilde{L}$ is a 4-Engel Lie algebra.

Consider $X_{234} = x_2 \otimes e_1 + x_3 \otimes e_2 + x_4 \otimes e_3$ and set $X_i = x_i \otimes 1$. One has that

$$[X_1, X_{234}^{(3)}, X_5, X_6, X_7, X_8] = \left(\sum_{\sigma \in \mathfrak{S}(2,3,4)} [x_1, x_{2\sigma}, x_{3\sigma}, x_{4\sigma}, x_5, x_6, x_7, x_8] \right) \otimes e_1 e_2 e_3$$

Similarly, for $X_{134} = x_1 \otimes e_1 + x_3 \otimes e_2 + x_4 \otimes e_3$, we have that

$$[X_{134}, X_2, X_{134}^{(2)}, X_5, X_6, X_7, X_8] = \left(\sum_{\sigma \in \mathfrak{S}(1,3,4)} [x_{1\sigma}, x_2, x_{3\sigma}, x_{4\sigma}, x_5, x_6, x_7, x_8] \right) \otimes e_1 e_2 e_3.$$

By mimicking the above, we have that every sum in (E_{sym}) can be realized as a product in $\tilde{L}$ where $X_{i_1 i_2 i_3}$ appears 3 times and the X_i once, for each $i \notin \{i_1, i_2, i_3\}$. We have proved the following claim.

Claim 3. In order to establish (E_{sym}) , it is enough to prove that every 6-generated 4-Engel Lie $\mathbb{F}_p$ -algebra satisfy $P(X_1, \dots, X_6) = 0$ for every monomial P of multiweight $(3, 1, 1, 1, 1, 1)$.

We now turn to (E_{skew}) . Let U be the associative algebra generated by three elements u_1, u_2, u_3 and submitted to the relations $u_i^2 = 0$, $u_i u_j = -u_j u_i$ and consider

$$L^{\text{super}} = L \otimes U.$$

Let $L_0 = L \otimes \text{Span}(u_i u_j \mid 1 \leq i, j, k \leq 3)$ and $L_1 = L \otimes \text{Span}(u_i, u_i u_j u_k \mid 1 \leq i, j, k \leq 3)$. Then one easily checks that L^{super} is a Lie superalgebra, and by Lemma 4.2, it satisfies the identity $\sum_{\sigma \in \mathfrak{S}(4)} [X, Y_{1\sigma}, Y_{2\sigma}, Y_{3\sigma}, Y_{4\sigma}] = 0$,

Exercise A.19. Prove that L^{super} is a Lie superalgebra.

Let $\iota : 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 4$, observe that for any $\sigma \in \mathfrak{S}(2, 3, 4)$ we have

$$u_{1\sigma} u_{2\sigma} u_{3\sigma} = \text{sgn}(\sigma) u_1 u_2 u_3,$$

hence for $X_{234} = x_2 \otimes u_1 + x_3 \otimes u_2 + x_4 \otimes u_3$ and $X_i = x_i \otimes 1$, similarly as above we obtain,

$$\begin{aligned} [X_1, X_{234}^{(3)}, X_5, X_6, X_7, X_8] &= \sum_{\sigma \in \mathfrak{S}(2,3,4)} ([x_1, x_{2\sigma}, x_{3\sigma}, x_{4\sigma}, x_5, x_6, x_7, x_8] \otimes u_{1\sigma} u_{2\sigma} u_{3\sigma}) \\ &= \left(\sum_{\sigma \in \mathfrak{S}(2,3,4)} \text{sgn}(\sigma) [x_1, x_{2\sigma}, x_{3\sigma}, x_{4\sigma}, x_5, x_6, x_7, x_8] \right) \otimes u_1 u_2 u_3 \end{aligned}$$

and we recognise an instance of E_{skew} . It follows that, as above, we may conclude the following claim.

Claim 4. In order to establish (E_{skew}) , it is enough to prove that every 6-generated Lie superalgebra algebra over $\mathbb{F}_p$ which satisfies $\sum_{\sigma \in \mathfrak{S}(4)} [X, Y_{1\sigma}, Y_{2\sigma}, Y_{3\sigma}, Y_{4\sigma}] = 0$ also satisfies $P(X_1, \dots, X_6) = 0$ for every monomial P of multiweight $(3, 1, 1, 1, 1, 1)$.

This establishes our result above. As mentioned above, Traustason goes even further than this reduction and reduces the question to 4-generated structures. Once he reaches Step 2 and the reduction to equations of type:

$$\begin{aligned} &\sum_{\sigma \in \mathfrak{S}(3)} [x_1, \dots, x_{i_1\sigma}, \dots, x_{i_2\sigma}, \dots, x_{i_3\sigma}, \dots, x_5, x_6, x_7, x_8] \\ &\sum_{\sigma \in \mathfrak{S}(3)} \text{sgn}(\sigma) [x_1, \dots, x_{i_1\sigma}, \dots, x_{i_2\sigma}, \dots, x_{i_3\sigma}, \dots, x_5, x_6, x_7, x_8] \end{aligned}$$

for all $1 \leq i_1 < i_2 < i_3 \leq 5$, one immediately sees that for a fixed i_1, i_2, i_3 , say $(i_1, i_2, i_3) = (2, 3, 4)$, there are still 5 variables left unchanged in the sum

$$\sum_{\sigma \in \mathfrak{S}(2,3,4)} [x_1, x_{2\sigma}, x_{3\sigma}, x_{4\sigma}, x_5, x_6, x_7, x_8]$$

One may therefore consider the action of $\mathfrak{S}(\{1, 5, 6, 7, 8\})$ on M and use another round of the same argument layed down above. That way, Traustason were able to reduce the equality $[x_1, \dots, x_8] = 0$ to equations in **4-generated structures**, Lie algebras, Lie superalgebras, and *colour* algebras. Then he used computer algebra and nilpotent quotient algorithm in order to check those identities and prove his theorem.

APPENDIX B. ZELMANOV CHARACTERISTIC 0 THEOREM

The goal of this section is to prove Zelmanov's *characteristic 0* theorem.

Theorem. *Every n -Engel Lie algebra over a field of characteristic 0 is nilpotent.*

We now explain how to deduce the *asymptotic solution to the RPB* from Zelmanov's characteristic 0 Theorem. Let Δ be the two-sorted theory of n -Engel Lie algebras over a field of characteristic 0, i.e. the field sort is axiomatised by the theory of fields and all sentences

$$\theta_p := \underbrace{1 + 1 + \dots + 1}_{p \text{ times}} \neq 0$$

whereas the Lie algebra sort is axiomatised by the theory of Lie algebra with the extra sentence

$$\forall x, y [x, y^{(n)}] = 0.$$

Being nilpotent of class $\leq c$ is expressible by a single sentence ϕ_c , and Zelmanov theorem can be stated as follows:

$$\Delta \models \bigvee_{c \in \mathbb{N}} \phi_c$$

Compactness implies that a finite fragment of Δ implies a finite fragment of the right hand side. In particular, only finitely many of the θ_p 's are necessary to imply finitely many of the ϕ_c 's hence we may conclude the following¹.

Theorem (Asymptotic nilpotency for n -Engel Lie algebras). *For each $n \in \mathbb{N}$, there exists $N, c \in \mathbb{N}$ such that every n -Engel Lie algebra over a field of characteristic $p > N$ is nilpotent of class $\leq c$.*

B.1. Step 1: Establishing the induction scheme. In order to prove the theorem, it is enough to prove that the relatively free n -Engel Lie algebra over $\mathbb{Q}$ in countably generators is nilpotent.

We will use the following well-known result of Kostrikin mentioned in the introduction.

Theorem (Kostrikin, 1958). *Let L be an n -Engel Lie algebra of characteristic 0 (or of characteristic $p > n$), then L contains an abelian ideal.*

During the rest of this section, $n \in \mathbb{N}$ is fixed and let L be the relatively free n -Engel Lie algebra over $\mathbb{Q}$ in countably many generators $\{x_i \mid i \in \mathbb{N}\}$.

We construct a chain of ideals $(I_\alpha)_{\alpha < \gamma}$ in the following way:

- if $\alpha = 0$ start with $I_0 = \{0\}$,
- if α is a successor ordinal hence $I_{\alpha-1}$ has been constructed, consider $\pi : L \rightarrow L/I_{\alpha-1}$ and let $\overline{I_\alpha}$ to be the sum of all abelian ideals of $L/I_{\alpha-1}$ and define $I_\alpha := \pi^{-1}(\overline{I_\alpha})$,
- if α is a limit ordinal, set $I_\alpha = \bigcup_{\lambda < \alpha} I_\lambda$.

As the homomorphic image of any n -Engel Lie algebra stays n -Engel, Kostrikin Theorem ensure that $L/I_{\alpha-1}$ always contains a nontrivial ideal, as soon as $I_{\alpha-1} \neq L$. By a cardinality argument there is a (countable) ordinal γ such that $L = I_\gamma$.

Claim 5. For each $\alpha \leq \gamma$, I_α is fully invariant. In particular L/I_α is relatively free.

Proof. We prove it by induction. If $\alpha = 0$ or α is limit, the result is clear, so we assume that α is successor and that $I_{\alpha-1}$ is fully invariant. As $L/I_{\alpha-1}$ is relatively free, every relation that holds in $L/I_{\alpha-1}$ is an identical relation. Using the isomorphism: $(L/I_{\alpha-1})/(I_\alpha/I_{\alpha-1}) \cong L/I_\alpha$, it suffices to prove that $I := I_\alpha/I_{\alpha-1}$ is a fully invariant ideal of $M := L/I_{\alpha-1}$. Let θ be any endomorphism of M and let a be an element generating an abelian ideal. As I is the sum of all abelian ideals of M , it is enough to show that $a\theta$ generates an abelian ideal. Let $(y_i)_{i < \omega}$ be the free

¹See also [22] for a quick argument for the same result which does not use any logic. Traustason wrote a constructive account of Zelmanov's proof, yielding explicit bounds for N, c , depending on n [20].

generators of M . Assume that a lives in the Lie subalgebra of L generated by $y_1, \dots, y_r$. Because a generates an abelian ideal in M , we have

$$[a, y_{r+1}, \dots, y_{r+k}, a] = 0$$

for any $k \geq 1$. Now for arbitrary elements $a_1, \dots, a_k$ of M , the tuple $(y_1, \dots, y_r, y_{r+1}, \dots, y_{r+k})$ can be mapped to the tuple $(y_1\theta, \dots, y_r\theta, a_1, \dots, a_k)$ by an endomorphism ϕ of M . Observe now that $a\phi = a\theta$ so that

$$[a\theta, a_1, \dots, a_k, a\theta] = 0.$$

As this holds for all $a_1, \dots, a_k \in M$ and $k \geq 1$, the ideal generated by $a\theta$ is abelian. $\square$

We prove that for all $0 < \beta \leq \gamma$ there exists $\alpha < \beta$ such that L/I_α is nilpotent. By basic properties of ordinals, this implies that $L/I_0 = L$ is nilpotent.

Claim 6. If β is a limit ordinal and L/I_β is nilpotent, then there is $\alpha < \beta$ such that I_α is nilpotent.

Proof. Assume that $I_\beta = \bigcup_{\alpha < \beta} I_\alpha$. As L/I_β is nilpotent, then $[x_1, \dots, x_m] \in I_\alpha$ for some $m \in \mathbb{N}$ and $\alpha < \beta$. Then L/I_α is a relatively free Lie algebra (by Claim 5) and because $[x_1, \dots, x_m] = 0$, the latter is an identity, which means that L/I_α is nilpotent. $\square$

It remains to prove that if β is a successor ordinal and L/I_β is nilpotent then $L/I_{\beta-1}$ is also nilpotent. This is the core of the proof, and will occupy the rest of this section.

Remark B.1. Another equivalent way of stating the strategy is to consider β to be the minimal element of the set $\{\alpha \leq \gamma \mid L/I_\alpha \text{ is nilpotent}\}$, which is nonzero since it contains γ . Then Claim 6 yields that β must be either zero or a successor cardinal. Assuming that β is a successor ordinal, we prove that $L/I_{\beta-1}$ is also nilpotent which forces β to be 0 and L to be nilpotent.

B.2. Step 2: First footprint in the induction step. We let $M = L/I_{\beta-1}$ and $I = I_\beta/I_{\beta-1}$. Then $M/I \cong L/I_\beta$ hence M/I is nilpotent and we want to prove that M is nilpotent. Of course, I is the sum of all abelian ideals of M . Recall also that M is relatively free and of characteristic 0 hence it is multigraded. By abuse of notation, we reuse the variables $(x_i)_{i < \omega}$ to denote a set of generators of the relatively free Lie algebra M . We also consider another set of generators of M , this time indexed by $\mathbb{N} \times \mathbb{N}$, which we denote $(x_{(i,j)})_{(i,j) \in \mathbb{N} \times \mathbb{N}}$. We will prove that for some $K, N \in \mathbb{N}$ we have that $(x_{(i,j)})_{i,j}$ satisfy the following identity.

$$[[x_{(1,1)}, x_{(1,2)}, \dots, x_{(1,K)}], [x_{(2,1)}, x_{(2,2)}, \dots, x_{(2,K)}], \dots, [x_{(N,1)}, \dots, x_{(N,K)}]] = 0. \quad (1)$$

As M is relatively free, this implies that the above is an identity in M and hence that M is solvable. Then using Higgins Theorem (Corollary A.11), we conclude that M is nilpotent.

B.3. Step 3: Linearization. Let m be such that $(M/I)^m = 0$, so that $[x_1, \dots, x_m] \in I$ hence there exists $a_1, \dots, a_{k-1} \in I$ with

$$[x_1, \dots, x_m] = a_1 + \dots + a_{k-1}$$

where each a_i generates an abelian ideal. This is a general fact, if a Lie algebra is generated by elements which generates abelian ideals, then the product of two elements that generates an abelian ideal also generates an abelian ideal so that every element can be written as a sum of elements that generates an abelian ideal. In particular, in any monomial where $[x_1, \dots, x_m]$ appears k -times, by developing the sums, one of the a_i appears at least twice in each summand hence vanishes, so the ideal generated by $[x_1, \dots, x_m]$ is nilpotent of class $< k$. It follows that we have the identities:

$$[[x_1, \dots, x_m], \dots, [x_1, \dots, x_m], \dots, [x_1, \dots, x_m]] = 0$$

where any number and any element can appear in between the monomials $[x_1, \dots, x_m]$, which appear k times.

Claim 7 (Linearization). M satisfies the following identity:

$$\sum_{\sigma_1 \in \mathfrak{S}(k), \dots, \sigma_m \in \mathfrak{S}(k)} [[x_{(1\sigma_1,1)}, \dots, x_{(1\sigma_m,m)}], \dots, [x_{(2\sigma_1,1)}, \dots, x_{(2\sigma_m,m)}], \dots, [x_{(k\sigma_1,1)}, \dots, x_{(k\sigma_m,m)}]] = 0 \quad (\Lambda)$$

In the above, we omit that any number of elements can be put in between the commutators $[x_{(i\sigma_1,1)}, \dots, x_{(i\sigma_m,m)}]$.

Proof. It is an easy exercise of linearization by setting x_i to be $\sum_{j=1}^k \lambda_{j,i} x_{(j,i)}$ for $\lambda_{1,i}, \dots, \lambda_{k,i} \in \mathbb{F}$. To convince yourself, do it with $k = 3$ and $m = 2$. Write

$$x_1 = \sum_{j=1}^3 \lambda_j x_{(j,1)}, \quad x_2 = \sum_{j=1}^3 \mu_j x_{(j,2)},$$

then

$$\begin{aligned} [[x_1, x_2], \dots, [x_1, x_2], \dots, [x_1, x_2]] &= \left[\sum_{1 \leq i, j \leq 3} [\lambda_i x_{(i,1)}, \mu_j x_{(j,2)}], \dots, \sum_{1 \leq i, j \leq 3} [\lambda_i x_{(i,1)}, \mu_j x_{(j,2)}], \dots, \sum_{1 \leq i, j \leq 3} [\lambda_i x_{(i,1)}, \mu_j x_{(j,2)}] \right] \\ &= \sum_{1 \leq i, j, k, l, s, t \leq 3} [[\lambda_i x_{(i,1)}, \mu_j x_{(j,2)}], \dots, [\lambda_k x_{(k,1)}, \mu_l x_{(l,2)}], \dots, [\lambda_s x_{(s,1)}, \mu_t x_{(t,2)}]] \end{aligned}$$

Now by evaluation $(\lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2, \mu_3) = (0, 1, 1, 1, 1, 1)$, one sees that in the sum above, the sum of all monomials where $x_{(1,1)}$ does not occur vanishes, hence we may assume that the sum only involves monomials where $x_{(1,1)}$ occurs. Iterating the argument above, we may assume that all monomials in the sum above involve each $x_{(i,j)}$ at least (and by the degree at most) once. In turn, the sum above equals:

$$\sum_{\substack{1 \leq i, j, k, l, s, t \leq 3 \\ i, j, k, l, s, t \text{ all distinct}}} [[x_{(i,1)}, x_{(j,2)}], \dots, [x_{(k,1)}, x_{(l,2)}], \dots, [x_{(s,1)}, x_{(t,2)}]]$$

which is easily seen to equal

$$\sum_{\sigma_1, \sigma_2 \in \mathfrak{S}(3)} [[x_{(1\sigma_1,1)}, x_{(1\sigma_2,2)}], \dots, [x_{(2\sigma_1,1)}, x_{(2\sigma_2,2)}], \dots, [x_{(3\sigma_1,1)}, x_{(3\sigma_2,2)}]].$$

□

B.4. Step 4: Reduction to symmetric/skew-symmetric sums. We prove that the equation

$$[[x_{(1,1)}, x_{(1,2)}, \dots, x_{(1,K)}], [x_{(2,1)}, x_{(2,2)}, \dots, x_{(2,K)}], \dots, [x_{(N,1)}, \dots, x_{(N,K)}]] = 0 \quad (1)$$

is a consequence of the equations

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\varepsilon_1} \sum_{\sigma_2 \in \mathfrak{S}_2}^{\varepsilon_2} \dots \sum_{\sigma_K \in \mathfrak{S}_K}^{\varepsilon_K} [\dots, [x_{(1,1)}^{\sigma_1}, \dots, x_{(1,K)}^{\sigma_K}], \dots, [x_{(2,1)}^{\sigma_1}, \dots, x_{(2,K)}^{\sigma_K}], \dots, [x_{(R,1)}^{\sigma_1}, \dots, x_{(R,K)}^{\sigma_K}], \dots] = 0 \quad (2)$$

where each $\mathfrak{S}_i$ is a copy of $\mathfrak{S}(R)$ which acts by permutation on the generators $\{x_{(1,i)}, \dots, x_{(R,i)}\}$ (and leaves the other generators invariant) and where $R^{2^K} = N$ and $\varepsilon_i \in \{+, -\}$, and by definition

$$\sum_{\sigma \in \mathfrak{S}(R)}^+ t^\sigma := \sum_{\sigma \in \mathfrak{S}(R)} t^\sigma \quad \text{and} \quad \sum_{\sigma \in \mathfrak{S}(R)}^- t^\sigma := \sum_{\sigma \in \mathfrak{S}(R)} \text{sgn}(\sigma) t^\sigma.$$

Action on the first coordinates. We start by acting on the “first coordinates” of the monomials

$$\begin{aligned} t_1 &:= [x_{(1,1)}, x_{(1,2)}, \dots, x_{(1,K)}] \\ &\vdots \\ t_N &:= [x_{(N,1)}, x_{(N,2)}, \dots, x_{(N,K)}] \end{aligned}$$

The group $\mathfrak{S}(N)$ act on the set (on the right) $\{x_{(1,1)}, \dots, x_{(N,1)}\}$ by $x_{(i,1)}\sigma = x_{(i\sigma,1)}$ and extends to all $x_{(i,j)}$ by $x_{(i,j)}\sigma = x_{(i,j)}$ if $j \neq 1$. Then this action extends to all monomials by

$$[x_{(i_1,j_1)}, x_{(i_2,j_2)}, \dots, x_{(i_s,j_s)}]\sigma = [x_{(i_1,j_1)}\sigma, x_{(i_2,j_2)}\sigma, \dots, x_{(i_s,j_s)}\sigma]$$

and finally the group algebra $\mathbb{Q}\mathfrak{S}(N)$ acts on the $\mathbb{Q}$ -vector span of those monomials (i.e. M) in the natural way.

By Subsection A.4, in order to obtain equation (1) it is enough to prove that for all primitive idempotent $e \in \mathbb{Q}\mathfrak{S}(N)$, we have

$$[[x_{(1,1)}, x_{(1,2)}, \dots, x_{(1,K)}], [x_{(2,1)}, x_{(2,2)}, \dots, x_{(2,K)}], \dots, [x_{(N,1)}, \dots, x_{(N,K)}]]e = 0. \quad (3)$$

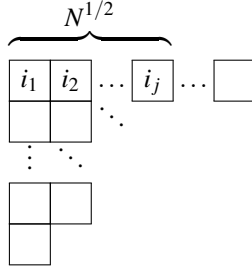
Indeed, the identity Id of $\mathbb{Q}\mathfrak{S}(N)$ can be written as a sum of primitive idempotents $\text{Id} = e_1 + \dots + e_s$ and for any linear combination u of monomials in $x_{(i,j)}$ we have $u = \text{Id}(u) = \sum_{i=1}^s u e_i$.

Any primitive idempotent corresponds to a Young diagram which has either a first row of length at least $N^{1/2}$ or the first column of length at least $N^{1/2}$. The first case will be called the “symmetric” case and the second case will be called the “skew-symmetric” case. The idempotent e can be written as

$$e = \frac{1}{s} \sum_{\pi \in V, \rho \in H} \text{sgn}(\pi) \pi \rho$$

for some s dividing $N!$ and where V is the subgroup of $\mathfrak{S}(N)$ that permutes the entries within each columns and H is the subgroup which permutes the entries of the rows.

Symmetric case. We assume here that the first row of the Young diagram associated to e has length at least $N^{1/2}$, and we denote $i_1, \dots, i_{N^{1/2}}$ the first $N^{1/2}$ elements of the associated Young tableau of e .



As H is the subgroup of $\mathfrak{S}(N)$ which permutes all the entries within each rows, the group of all permutations of the set $\{i_1, \dots, i_{N^{1/2}}\}$ can be seen as the subgroup S of H which fixes elements of $\{1, \dots, N\} \setminus \{i_1, \dots, i_{N^{1/2}}\}$. Let C be a left transversal for S in H , i.e. with $H = \bigsqcup_{c \in C} cS$. We set $f := \sum_{\sigma \in S} \sigma$. Then

$$\begin{aligned} [t_1, \dots, t_N]e &= [t_1, \dots, t_N] \frac{1}{s} \sum_{\pi \in V, c \in C, \sigma \in S} \text{sgn}(\pi) \pi c \sigma \\ &= \frac{1}{s} \sum_{\pi \in V, c \in C} \text{sgn}(\pi) [t_1, \dots, t_N] \pi c \sum_{\sigma \in S} \sigma \\ &= \frac{1}{s} \sum_{\pi \in V, c \in C} \text{sgn}(\pi) [t_1, \dots, t_N] \pi c f \end{aligned}$$

It is enough to show that for each $\pi \in V, c \in C$, $[t_1, \dots, t_N] \pi c f = 0$, so fix $\pi \in V, c \in C$ and let $j_1 < \dots < j_{N^{1/2}}$ be such that

$$\{j_1 \pi c, \dots, j_{N^{1/2}} \pi c\} = \{i_1, \dots, i_{N^{1/2}}\}.$$

Then, because S only acts on $\{i_1, \dots, i_{N^{1/2}}\}$,

$$\begin{aligned} [t_1, \dots, t_N] \pi c f &= \sum_{\sigma \in S} [t_1 \pi c \sigma, \dots, t_{j_1} \pi c \sigma, \dots, t_{j_{N^{1/2}}} \pi c \sigma, \dots, t_N \pi c \sigma] \\ &= \sum_{\sigma \in S} [t_1 \pi c, \dots, t_{j_1} \pi c \sigma, \dots, t_{j_{N^{1/2}}} \pi c \sigma, \dots, t_N \pi c] \end{aligned}$$

For $j \in \{j_1, \dots, j_{N^{1/2}}\}$ we have

$$t_j \pi c \sigma = [x_{(j \pi c \sigma, 1)}, x_{(j, 2)}, \dots, x_{(j, K)}]$$

and as σ runs over S , $(j_1 \pi c \sigma, \dots, j_{N^{1/2}} \pi c \sigma)$ runs over all permutations of $(i_1, \dots, i_{N^{1/2}})$. So in order to prove equation (3) for that idempotent e , it is enough to establish each equation of the form

$$\sum_{\sigma \in S} [\dots, t_{j_1} \pi c \sigma, \dots, t_{j_{N^{1/2}}} \pi c \sigma, \dots] = 0$$

for each $\pi \in V, c \in C$. Note that the left hand side of the previous equation is a commutator of multiweight NK which is multilinear because we started with the equation (3) which is multilinear, and we are applying permutations of the generators. We may now re-label the generators so that proving equation (3) reduces to proving:

$$\sum_{\sigma \in \mathfrak{S}(N^{1/2})} [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] = 0 \quad (4)$$

Recall that extra entries are to be inserted between the entries $t_1^\sigma, \dots, t_{N^{1/2}}^\sigma$, and those are fixed throughout the sum (and the expression is still multilinear). But those extra entries also varies when considering two different sums of the above form.

Skew-symmetric case. In this case, the sum will be twisted because the permutation subgroup of $N^{1/2}$ elements will be a subgroup of V instead of H .

Although the argument is similar, we will repeat it in details. Let $i_1, \dots, i_{N^{1/2}}$ be the first $N^{1/2}$ entries in the first column of the Young tableau, and let S be the subgroup of V which fixes $\{1, \dots, N\} \setminus \{i_1, \dots, i_{N^{1/2}}\}$ and let C be a right transversal for S in V , that is

$$V = \bigsqcup_{c \in C} Sc.$$

Let $f = \sum_{\sigma \in S} \text{sgn}(\sigma) \sigma$. For $c \in C$ and $\sigma \in S$ we have $\text{sgn}(\sigma c) = \text{sgn}(\sigma) \text{sgn}(c)$ hence

$$N^{1/2} \left\{ \begin{array}{c} \begin{array}{|c|c|} \hline i_1 & \\ \hline i_2 & \\ \hline \vdots & \\ \hline i_j & \\ \hline \vdots & \\ \hline \end{array} \cdots \begin{array}{|c|c|} \hline & \\ \hline & \\ \hline & \\ \hline & \\ \hline \end{array} \end{array} \right.$$

$$\begin{aligned} [t_1, \dots, t_N]e &= [t_1, \dots, t_N] \frac{1}{s} \sum_{\sigma \in S, c \in C, \rho \in H} \text{sgn}(\sigma c) \sigma c \rho \\ &= \frac{1}{s} \sum_{c \in C, \rho \in H} \text{sgn}(c) [t_1, \dots, t_N] \left(\sum_{\sigma \in S} \text{sgn}(\sigma) \sigma \right) c \rho \\ &= \frac{1}{s} \sum_{c \in C, \rho \in H} \text{sgn}(c) [t_1, \dots, t_N] f c \rho \end{aligned}$$

Now we may re-label i_j so that $i_1 < \dots < i_{N^{1/2}}$ and for each $c \in C, \rho \in H$:

$$\begin{aligned} [t_1, \dots, t_N] f c \rho &= \sum_{\sigma \in S} \text{sgn}(\sigma) [t_1 \sigma, \dots, t_{i_1} \sigma, \dots, t_{i_{N^{1/2}}} \sigma, \dots, t_N \sigma] c \rho \\ &= \sum_{\sigma \in S} \text{sgn}(\sigma) [t_1, \dots, t_{i_1} \sigma, \dots, t_{i_{N^{1/2}}} \sigma, \dots, t_N] c \rho \end{aligned}$$

since $t_i \sigma = t_i$ if $i \notin \{i_1, \dots, i_{N^{1/2}}\}$ and $t_i \sigma = [x_{(i\sigma,1)}, x_{(i\sigma,2)}, \dots, x_{(i\sigma,K)}]$ for $i \in \{i_1, \dots, i_{N^{1/2}}\}$. As above, in order to obtain equation (3) it is enough to prove that

$$\sum_{\sigma \in \mathfrak{S}(N^{1/2})} \text{sgn}(\sigma) [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] = 0 \quad (5)$$

Putting together (4) and (5), we conclude that in order to establish equation (1), it is enough to prove that for $\varepsilon \in \{+, -\}$ we have:

$$\sum_{\sigma \in \mathfrak{S}(N^{1/2})}^\varepsilon [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] = 0. \quad (6)$$

Action on the second coordinates. According to equation (6), we now no longer need to consider monomials $t_1, \dots, t_N$ but only subsets of size $N^{1/2}$ of those. We fixed one such subset and re-labeled it $t_1, \dots, t_{N^{1/2}}$. We continue by acting on the second coordinates of the monomials

$$\begin{aligned} t_1 &= [x_{(1,1)}, x_{(1,2)}, \dots, x_{(1,K)}] \\ &\vdots \\ t_{N^{1/2}} &= [x_{(N^{1/2},1)}, x_{(N^{1/2},2)}, \dots, x_{(N^{1/2},K)}]. \end{aligned}$$

We let $\mathfrak{S}(N^{1/2})$ acts on $\{x_{(1,2)}, \dots, x_{(N^{1/2},2)}\}$, leaving all other generators invariant, we extend the action on all monomials coordinatewise. Then $\mathbb{Q}\mathfrak{S}(N^{1/2})$ acts on the span of all those monomials, i.e. on all M . In order to establish (6), it is enough to prove that for every primitive idempotent $e \in \mathbb{Q}\mathfrak{S}(N^{1/2})$ we have

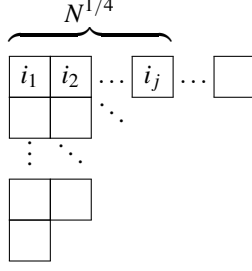
$$\left(\sum_{\sigma \in \mathfrak{S}(N^{1/2})}^\varepsilon [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] \right) e = 0. \quad (7)$$

We write

$$e = \frac{1}{s} \sum_{\pi \in V, \rho \in H} \text{sgn}(\pi) \pi \rho$$

where V is the group that permutes the values in the column of the Young tableau and H is the group that permutes the values inside each row. The Young tableau associated to e will again have either a row of length at least $N^{1/4}$ or a column of length at least $N^{1/4}$, which leads us again to two cases.

Symmetric case. If e corresponds to a Young tableau with the first row of length at least $N^{1/4}$ let $i_1, \dots, i_{N^{1/4}}$ be the first entries and let S be the subgroup of H that fixes $\{1, \dots, N^{1/2}\} \setminus \{i_1, \dots, i_{N^{1/4}}\}$.



Let C be a left transversal of S in H , so that $H = \bigsqcup_{c \in C} cS$ and $f = \sum_{\sigma \in S} \sigma$. Then

$$\sum_{\sigma \in \mathfrak{S}(N^{1/2})} [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] e$$

is a linear combination of terms of the form

$$\sum_{\sigma \in \mathfrak{S}(N^{1/2})} [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] \pi c f$$

for $\pi \in V, c \in C$. For a fixed $\pi \in V, c \in C$, there exists $j_1 < \dots < j_{N^{1/4}}$ with

$$\{j_1 \pi c, \dots, j_{N^{1/4}} \pi c\} = \{i_1, \dots, i_{N^{1/4}}\}$$

and hence

$$\begin{aligned} \sum_{\sigma \in \mathfrak{S}(N^{1/2})} [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] \pi c f &= \sum_{\sigma \in \mathfrak{S}(N^{1/2})} \sum_{\tau \in S} [\dots, t_1^\sigma, \dots, t_2^\sigma, \dots, t_{N^{1/2}}^\sigma, \dots] \pi c \tau \\ &= \sum_{\sigma \in \mathfrak{S}(N^{1/2})} \sum_{\tau \in S} [\dots, t_1^\sigma \pi c, \dots, t_{j_1}^\sigma \pi c \tau, \dots, t_{j_{N^{1/4}}}^\sigma \pi c \tau, \dots, t_{N^{1/2}}^\sigma \pi c, \dots] \end{aligned}$$

since $\tau \in S$ fixes $t_j \pi c$ unless $j \in \{j_1, \dots, j_{N^{1/4}}\}$, (for which it acts on the second coordinate). As τ runs over S , $(j_1 \pi c \tau, \dots, j_{N^{1/4}} \pi c \tau)$ runs over all permutations of $(i_1, \dots, i_{N^{1/4}})$. As σ ranges over $\mathfrak{S}(N^{1/2})$, $(j_1 \sigma, \dots, j_{N^{1/4}} \sigma)$ ranges over all possible permutations of $N^{1/4}$ element subsets X of $\{1, \dots, N^{1/2}\}$. We fix one particular $N^{1/4}$ element subset X of $\{1, \dots, N^{1/2}\}$. Let $A = \mathfrak{S}(X)$ and $B = \mathfrak{S}(\{1, \dots, N^{1/2}\} \setminus X)$. Both A and B are here seen as subgroups of $\mathfrak{S}(N^{1/2}) = \mathfrak{S}(\{1, \dots, N^{1/2}\})$, which fix pointwise the set $\{1, \dots, N^{1/2}\} \setminus X$ (respectively the set X).

Claim 8. Assume that $\sigma_0 \in \mathfrak{S}(N^{1/2}) = \mathfrak{S}(\{1, \dots, N^{1/2}\})$ is such that $\{j_1 \sigma_0, \dots, j_{N^{1/4}} \sigma_0\} = X$. Then any permutation $\sigma \in \mathfrak{S}(N^{1/2})$ with $\{j_1 \sigma, \dots, j_{N^{1/4}} \sigma\} = X$ can be written uniquely in the form $\sigma = \sigma_0 a b$ for $a \in A$ and $b \in B$.

Proof. First, the restriction $\sigma_0|_{\{j_1, \dots, j_{N^{1/4}}\}} : \{j_1, \dots, j_{N^{1/4}}\} \rightarrow X$ can be corrected by an element $a \in A$ so that $\sigma_0|_{\{j_1, \dots, j_{N^{1/4}}\}} a = \sigma|_{\{j_1, \dots, j_{N^{1/4}}\}}$. Similarly, there exists $b \in B$ such that

$$\sigma|_{(\{1, \dots, N^{1/2}\} \setminus \{j_1, \dots, j_{N^{1/4}}\})} : (\{1, \dots, N^{1/2}\} \setminus \{j_1, \dots, j_{N^{1/4}}\}) \rightarrow \{1, \dots, N^{1/2}\} \setminus X$$

is equal to $\sigma_0|_{(\{1, \dots, N^{1/2}\} \setminus \{j_1, \dots, j_{N^{1/4}}\})} b$. Now because the domains and the ranges of $\sigma|_{\{j_1, \dots, j_{N^{1/4}}\}}$ and $\sigma_0|_{(\{1, \dots, N^{1/2}\} \setminus \{j_1, \dots, j_{N^{1/4}}\})}$ are disjoint, the union of the graphs gives an element of $\mathfrak{S}(N^{1/2})$ which equals σ . For the same reason the union of the graphs of $\sigma_0|_{\{j_1, \dots, j_{N^{1/4}}\}} a$ and $\sigma_0|_{(\{1, \dots, N^{1/2}\} \setminus \{j_1, \dots, j_{N^{1/4}}\})} b$ gives $\sigma_0 a b$ (note that a and b commute since they have disjoint support). $\square$

Now, for the fixed X as above, in the sum

$$\sum_{\sigma \in \mathfrak{S}(N^{1/2})} \sum_{\tau \in S} [\dots, t_1^\sigma \pi c, \dots, t_{i_1}^\sigma \pi c \tau, \dots, t_{i_{N^{1/4}}}^\sigma \pi c \tau, \dots, t_{N^{1/2}}^\sigma \pi c, \dots]$$

we pick out the terms where σ satisfies $\{j_1 \sigma, \dots, j_{N^{1/4}} \sigma\} = X$, and by the claim this is equal to

$$\pm \sum_{a \in A} \sum_{b \in B} \sum_{\tau \in S} [\dots, t_1^{\sigma_0 b} \pi c, \dots, t_{j_1}^{\sigma_0 a} \pi c \tau, \dots, t_{j_{N^{1/4}}}^{\sigma_0 a} \pi c \tau, \dots, t_{N^{1/2}}^{\sigma_0 b} \pi c, \dots]$$

since $t_i^{\sigma_0 ab}$ either equals $t_i^{\sigma_0 a}$ if $i \in \{j_1, \dots, j_{N^{1/4}}\}$ or equals $t_i^{\sigma_0 b}$ otherwise. For a fixed $b \in B$, the latter is a sum of elements of the form

$$\pm \sum_{a \in A} \sum_{\tau \in S} [\dots, t_1^{\sigma_0 b} \pi_c \tau, \dots, t_{j_1}^{\sigma_0 a} \pi_c \tau, \dots, t_{j_{N^{1/4}}}^{\sigma_0 a} \pi_c \tau, \dots, t_{N^{1/2}}^{\sigma_0 b} \pi_c \tau, \dots].$$

In order to show that (7) holds, it is enough to prove that all the sums of the latter form vanish. By re-writting the above, re-labelling the generators $x_{(i,j)}$, we need to show that sums of the following form vanish

$$\sum_{\sigma \in \mathfrak{S}(N^{1/4})} \sum_{\tau \in \mathfrak{S}(N^{1/4})} [\dots, t_1^{(\sigma, \tau)} \dots, t_2^{(\sigma, \tau)} \dots, t_{N^{1/4}}^{(\sigma, \tau)} \dots]$$

where

$$t_i^{(\sigma, \tau)} = [x_{(i\sigma, 1)}, x_{(i\tau, 2)}, x_{(i, 3)}, \dots, x_{(i, K)}]$$

for $i = 1, \dots, N^{1/4}$.

Skew-symmetric case. If e is an idempotent corresponding to a Young tableau with first column of length at least $N^{1/4}$, then the situation can be carried out in a very similar manner, (with the same changes from left-transversal to right transversal, etc) so that (7) follows from proving the equality

$$\sum_{\sigma \in \mathfrak{S}(N^{1/4})} \sum_{\tau \in \mathfrak{S}(N^{1/4})} [\dots, t_1^{(\sigma, \tau)} \dots, t_2^{(\sigma, \tau)} \dots, t_{N^{1/4}}^{(\sigma, \tau)} \dots] = 0.$$

$$N^{1/4} \left\{ \begin{array}{c} \begin{array}{|c|c|c|} \hline i_1 & & \\ \hline i_2 & & \\ \hline \vdots & \ddots & \\ \hline i_j & & \\ \hline \vdots & & \\ \hline \end{array} \end{array} \right.$$

Iteration and conclusion. In turn, (7) follows from proving equalities of the form

$$\sum_{\sigma \in \mathfrak{S}(N^{1/4})} \sum_{\tau \in \mathfrak{S}(N^{1/4})} [\dots, t_1^{(\sigma, \tau)} \dots, t_2^{(\sigma, \tau)} \dots, t_{N^{1/4}}^{(\sigma, \tau)} \dots] = 0$$

with $t_i^{(\sigma, \tau)} = [x_{(i\sigma, 1)}, x_{(i\tau, 2)}, x_{(i, 3)}, \dots, x_{(i, K)}]$, and $\varepsilon, \nu \in \{+, -\}$.

Now we let $\mathbb{Q}\mathfrak{S}(N^{1/4})$ act on M by permuting the free generators $\{x_{(1,3)}, x_{(2,3)}, \dots, x_{(N^{1/4}, 3)}\}$ and we can carry out similar reductions. We iterate this process for K steps so that for R satisfying $R^{2^K} = N$ we establish that equation (1) follows from proving that all the following equations hold

$$\sum_{\sigma_1 \in \mathfrak{S}(R)} \dots \sum_{\sigma_K \in \mathfrak{S}(R)} [\dots, t_1^{(\sigma_1, \dots, \sigma_K)} \dots, t_2^{(\sigma_1, \dots, \sigma_K)} \dots, t_R^{(\sigma_1, \dots, \sigma_K)} \dots] = 0$$

with for each $i = 1, \dots, R$:

$$t_i^{(\sigma_1, \dots, \sigma_K)} = [x_{(1\sigma_1, i)}, \dots, x_{(K\sigma_K, i)}].$$

B.5. Remark: An escape route. We can already conclude the theorem in a very particular case. By Step 3, there exists $k, m \in \mathbb{N}$ such that the ideal generated by $[x_1, \dots, x_m]$ in M is k -nilpotent, and by Claim 7, M satisfies the identity (Λ) :

$$\sum_{\sigma_1 \in \mathfrak{S}(k), \dots, \sigma_m \in \mathfrak{S}(k)} [[x_{(1\sigma_1, 1)}, \dots, x_{(1\sigma_m, m)}], \dots, [x_{(2\sigma_1, 1)}, \dots, x_{(2\sigma_m, m)}], \dots, [x_{(k\sigma_1, 1)}, \dots, x_{(k\sigma_m, m)}]] = 0$$

where any number of elements can be put in between terms of the form $[x_{(i\sigma_1, 1)}, \dots, x_{(i\sigma_m, m)}]$. Let $N = k^{2^m}$, and $K = m$ then, assuming $\varepsilon_i = 1$ for all $i = 1, \dots, K$ we have the identity

$$\sum_{\sigma_1 \in \mathfrak{S}(R)} \dots \sum_{\sigma_K \in \mathfrak{S}(R)} [\dots, t_1^{(\sigma_1, \dots, \sigma_K)} \dots, t_2^{(\sigma_1, \dots, \sigma_K)} \dots, t_R^{(\sigma_1, \dots, \sigma_K)} \dots] = 0$$

Hence, assuming that $\text{Id} \in \mathbb{Q}\mathfrak{S}(k)$, $\text{Id} \in \mathbb{Q}\mathfrak{S}(k^2)$, $\text{Id} \in \mathbb{Q}\mathfrak{S}(k^{2^2})$, ..., $\text{Id} \in \mathbb{Q}\mathfrak{S}(N)$ can all be written as a linear combination of primitive idempotent only with first row of length at least $()^{1/2}$, then we would conclude the theorem.

B.6. Step 5: Symmetrization using $\mathbb{Z}_2$ -gradings. Recall from Step 4 that we want to prove that for some $R, K \in \mathbb{N}$ we have

$$\sum_{\sigma_1 \in \mathfrak{S}(R)}^{\varepsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}(R)}^{\varepsilon_K} [\dots, t_1^{(\sigma_1, \dots, \sigma_K)}, \dots, t_2^{(\sigma_1, \dots, \sigma_K)}, \dots, t_R^{(\sigma_1, \dots, \sigma_K)}, \dots] = 0.$$

We now define the constants R and K . Recall that k is the number from Step 3 such that the ideal generated by $[x_1, \dots, x_m]$ is $< k$ -nilpotent. We are already given n and m , and we define

$$K := \frac{n^{(n-1)(m-1)+m+1} - 1}{n-1} + 1.$$

Let F be the subalgebra of M generated by $x_1, \dots, x_K$. It is a relatively free n -Engel Lie algebra with free generators $x_1, \dots, x_K$. Assume that we are given a $\mathbb{Z}_2$ -grading $F = F_0 \oplus F_1$ according to Example A.14 with some $x_i \in F_0$ and the others in F_1 . We have U the set of left-normed commutators in the $x_1, \dots, x_K$ and $U = U_0 \cup U_1$ with $U_0 \subseteq F_0$ and $U_1 \subseteq F_1$. Recall that we are given m as in Step 1. Define I to be the ideal generated by elements of the form $[c_1, \dots, c_m]$ with $c_i \in U_0$. Then F/I satisfies the hypotheses of Lemma A.15, and by the choice of K as defined above, we have $[x_1/I, \dots, x_K/I] = 0$ in F/I hence $[x_1, \dots, x_K] \in I$. This means that $[x_1, \dots, x_K]$ is a finite linear combination of elements of the form $[c_1, \dots, c_m, a_1, \dots, a_t]$, for $c_i \in U_0$ and $a_i \in \{x_1, \dots, x_K\}$. Note that because $[x_1, \dots, x_K]$ is multilinear and F is relatively free and multigraded, we may assume that each element $[c_1, \dots, c_m, a_1, \dots, a_t]$ is multilinear and of weight K .

For any of those 2^K gradings on $F = \langle x_1, \dots, x_K \rangle$ we obtain such a linear combination, hence there is a maximal number T such that for a given such $\mathbb{Z}_2$ -grading U_0, U_1 on $L = \langle x_1, \dots, x_K \rangle$, the element $[x_1, \dots, x_K]$ is a linear combination of at most T many monomials of the form $[c_1, \dots, c_m, a_1, \dots, a_t]$ with $c_i \in U_0$.

We set

$$R := kT.$$

We want to prove

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\varepsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\varepsilon_K} [\dots, t_1, \dots, t_2, \dots, t_R, \dots] \sigma_1 \dots \sigma_K = 0 \quad (8)$$

where $\mathfrak{S}_j$ is a copy of $\mathfrak{S}(R)$ which acts on the set $\{x_{(1,j)}, \dots, x_{(R,j)}\}$ by permutation of the first index, and leaves the other generators invariant. Note that each groups is acting with support on disjoint sets of generators, hence we are really considering the action of the direct product $\mathfrak{S}_1 \times \dots \times \mathfrak{S}_K$.

We now fix a choice of $\varepsilon_1, \dots, \varepsilon_K$, and we define a $\mathbb{Z}_2$ -grading on $F := \langle x_1, \dots, x_K \rangle = F_0 \oplus F_1$ by setting

$$\begin{aligned} x_i \in F_0 &\iff \varepsilon_i = +, \\ x_i \in F_1 &\iff \varepsilon_i = -. \end{aligned}$$

By the above argument, using Lemma A.15, there exists a number $s \leq T$ such that

$$[x_1, \dots, x_K] = \sum_{i=1}^s \lambda_i u_i$$

where u_i are multilinear terms of weight K of the form

$$u_i = [c_1, \dots, c_m, a_1, \dots, a_q]$$

with $c_i \in U_0$ and $a_i \in \{x_1, \dots, x_K\}$.

Let $\theta_1, \dots, \theta_R$ be endomorphisms of M with $x_1 \theta_i = x_{(i,1)}, \dots, x_K \theta_i = x_{(i,K)}$ for $i = 1, \dots, R$. Then

$$t_i = [x_{(i,1)}, \dots, x_{(i,K)}] = [x_1, \dots, x_K] \theta_i = \sum_{r=1}^s \lambda_r \theta_i u_r.$$

We substitute t_i in (8) and expand to get a linear combination of terms of the form

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\varepsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\varepsilon_K} [\dots, u_{r_1} \theta_1, \dots, u_{r_2} \theta_2, \dots, u_{r_R} \theta_R, \dots] \sigma_1 \dots \sigma_K$$

for all $1 \leq r_1, \dots, r_R \leq s$. As $R = kT \geq ks$ there is some index r which appears at least k times among $r_1, \dots, r_R$, say $r = r_i$ for $i = i_1, \dots, i_k$. Then

$$\begin{aligned} & \sum_{\sigma_1 \in \mathfrak{S}_1}^{\varepsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\varepsilon_K} [\dots, u_{r_1} \theta_1, \dots, u_{r_2} \theta_i, \dots, u_{r_R} \theta_R, \dots] \sigma_1 \dots \sigma_K \\ &= \sum_{\sigma_1 \in \mathfrak{S}_1}^{\varepsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\varepsilon_K} [\dots, u_r \theta_{i_1}, \dots, u_r \theta_{i_2}, \dots, u_r \theta_{i_k}, \dots] \sigma_1 \dots \sigma_K \end{aligned}$$

Now $u_r = [c_1, \dots, c_m, a_1, \dots, a_q] = [[c_1, \dots, c_m], a_1, \dots, a_q]$ and for $i \neq j$ we have that

$$\{c_1 \theta_{i_1}, \dots, c_m \theta_{i_1}, a_1 \theta_{i_1}, \dots, a_q \theta_{i_1}\} \cap \{c_1 \theta_{i_2}, \dots, c_m \theta_{i_2}, a_1 \theta_{i_2}, \dots, a_q \theta_{i_2}\} = \emptyset$$

so that

$$\begin{aligned} & [\dots, u_r \theta_{i_1}, \dots, u_r \theta_{i_2}, \dots, u_r \theta_{i_k}, \dots] \\ &= [\dots, [c_1, \dots, c_m] \theta_{i_1}, \dots, [c_1, \dots, c_m] \theta_{i_2}, \dots, [c_1, \dots, c_m] \theta_{i_k}, \dots] \end{aligned}$$

is a multilinear commutator. It follows that in order to prove (8), it is enough to prove that

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\varepsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\varepsilon_K} [\dots, [c_1, \dots, c_m] \theta_{i_1}, \dots, [c_1, \dots, c_m] \theta_{i_2}, \dots, [c_1, \dots, c_m] \theta_{i_k}, \dots] \sigma_1 \dots \sigma_K = 0 \quad (9)$$

where $\mathfrak{S}_j$ is a copy of $\mathfrak{S}(R)$ which acts on the set $\{x_{(1,j)}, \dots, x_{(R,j)}\}$ and leaves the other generators invariant. Again, those groups are acting with support on disjoint sets hence we are really considering the action of the direct product $\mathfrak{S}_1 \times \dots \times \mathfrak{S}_K$, so we may re-write (9) as

$$\sum_{\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K} \text{s\grave{g}n}(\sigma) [\dots, [c_1, \dots, c_m] \theta_{i_1}, \dots, [c_1, \dots, c_m] \theta_{i_2}, \dots, [c_1, \dots, c_m] \theta_{i_k}, \dots] \sigma = 0$$

where if $\sigma = \sigma_1 \dots \sigma_K$ we define $\text{s\grave{g}n}(\sigma)$ to be the product of the signatures $\text{sgn}(\sigma_i)$ for $\varepsilon_i = -$, in other words $\text{s\grave{g}n}(\sigma) = \prod_{\{i | \varepsilon_i = -\}} \text{sgn}(\sigma_i)$. It should be clear that in order to establish the above identities, it is sufficient to prove the same identity without the extra hidden terms on the right “...” and without the extra hidden terms on the left “[...]” (for the latter, observe that any commutator $[...]x$ is a linear combination of commutators of the form $[x, \dots]$). So, proving the above identity is equivalent to proving the following:

$$\sum_{\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K} \text{s\grave{g}n}(\sigma) [[c_1, \dots, c_m] \theta_{i_1}, \dots, [c_1, \dots, c_m] \theta_{i_2}, \dots, [c_1, \dots, c_m] \theta_{i_k}] \sigma = 0 \quad (10)$$

B.7. Step 6: Conclusion. To conclude, it is enough to prove that the left hand side of (10) consists of instances of the linearized identity (Λ) from Claim 7.

Fix $\sigma_1 \in \mathfrak{S}_1, \dots, \sigma_K \in \mathfrak{S}_K$ and let $\sigma = \sigma_1 \dots \sigma_K$. We take a closer look to a summand from (10), which is of the form

$$[[c_1, \dots, c_m] \theta_{i_1}, \dots, [c_1, \dots, c_m] \theta_{i_2}, \dots, [c_1, \dots, c_m] \theta_{i_k}] \sigma.$$

For $i, j \in \{i_1, \dots, i_k\}$ with $i < j$, the action of σ on $[c_1, \dots, c_m] \theta_i$ and $[c_1, \dots, c_m] \theta_j$ is given by

$$\begin{aligned} [c_1, \dots, c_m] \theta_i \sigma &= [c_1 \theta_i \sigma, \dots, c_m \theta_i \sigma], \\ [c_1, \dots, c_m] \theta_j \sigma &= [c_1 \theta_j \sigma, \dots, c_m \theta_j \sigma] \end{aligned}$$

and if $c_1 = [x_{k_1}, \dots, x_{k_q}]$, then

$$\begin{aligned} c_1 \theta_i \sigma &= [x_{(i,k_1)}, \dots, x_{(i,k_q)}] \sigma = [x_{(i,k_1)} \sigma_{k_1}, \dots, x_{(i,k_q)} \sigma_{k_q}] \\ c_1 \theta_j \sigma &= [x_{(j,k_1)}, \dots, x_{(j,k_q)}] \sigma = [x_{(j,k_1)} \sigma_{k_1}, \dots, x_{(j,k_q)} \sigma_{k_q}] \end{aligned}$$

since σ_k only permutes the set $\{x_{(1,k)}, \dots, x_{(R,k)}\}$ and leaves the other $x_{(i,j)}$ invariant.

Let τ_1 be the transposition which swaps $x_{(i,k_1)} \sigma_{k_1}$ and $x_{(j,k_1)} \sigma_{k_1}$, let τ_2 be the transposition which swaps $x_{(i,k_2)} \sigma_{k_2}$ and $x_{(j,k_2)} \sigma_{k_2}$ and so on until τ_q which permutes $x_{(i,k_q)} \sigma_{k_q}$ and $x_{(j,k_q)} \sigma_{k_q}$. Let $\tau = \tau_1 \dots \tau_q$. Now recall that the term

$$[[c_1, \dots, c_m] \theta_{i_1}, \dots, [c_1, \dots, c_m] \theta_{i_2}, \dots, [c_1, \dots, c_m] \theta_{i_k}] \quad (\Delta)$$

is multilinear in $(x_{(i,j)})$, i.e. each variable $x_{(i,j)}$ appears at most once. So the action by a permutation of generators gives another multilinear term. By looking on how $\sigma \tau$ acts on $[c_1, \dots, c_m] \theta_i$ and $[c_1, \dots, c_m] \theta_j$ within (Δ) we obtain

$$\begin{aligned} & [\dots, [c_1, \dots, c_m] \theta_i, \dots, [c_1, \dots, c_m] \theta_j, \dots] \sigma \tau \\ &= [\dots, [c_1 \theta_i \sigma \tau, \dots, c_m \theta_i \sigma \tau], \dots, [c_1 \theta_j \sigma \tau, \dots, c_m \theta_j \sigma \tau], \dots] \\ &= [\dots, [c_1 \theta_i \sigma \tau, c_2 \theta_i \sigma, \dots, c_m \theta_i \sigma], \dots, [c_1 \theta_j \sigma \tau, c_2 \theta_j \sigma, \dots, c_m \theta_j \sigma], \dots], \end{aligned}$$

the last line being justified by the fact that, as the term is multilinear, among the whole (Δ) , the generators $x_{(i,j)}$ which are in the support of τ only appear in $c_1 \theta_i \sigma$ and $c_1 \theta_j \sigma$, so that τ leaves invariant the other terms of the form $c_i \theta_j \sigma$. By choice of τ , we have, of course, $c_1 \theta_i \sigma \tau = c_1 \theta_j \sigma$ and $c_1 \theta_j \sigma \tau = c_1 \theta_i \sigma$, so

$$\begin{aligned} & [\dots, [c_1, \dots, c_m] \theta_i, \dots, [c_1, \dots, c_m] \theta_j, \dots] \sigma \tau \\ &= [\dots, [c_1 \theta_j \sigma, c_2 \theta_i \sigma, \dots, c_m \theta_i \sigma], \dots, [c_1 \theta_i \sigma, c_2 \theta_j \sigma, \dots, c_m \theta_j \sigma], \dots]. \end{aligned}$$

Since $c_1 = [x_{k_1}, \dots, x_{k_q}] \in F_0$, the number of elements of $\{x_{k_1}, \dots, x_{k_q}\}$ which lie in F_1 is even. We have that $q \leq K$ and τ can be seen as an element of $\mathfrak{S}_1 \times \dots \times \mathfrak{S}_K$ (putting identity elements in between) hence appears as a permutation involved in the sum (10). We argue that $\text{sgn}(\tau) = 1$. Indeed, the transposition τ_l will appear with $\text{sgn}(\tau_l)$ in the sum if and only if $\varepsilon_{k_l} = -$. This holds because, τ_l permutes $x_{(i,k_l)} \sigma_{k_l}$ and $x_{(j,k_l)} \sigma_{k_l}$ which are elements of $\{x_{(1,k_l)}, \dots, x_{(R,k_l)}\}$ and hence are acted on by permutations from $\mathfrak{S}_{k_l}$, which corresponds to a sum twisted by ε_{k_l} . Hence $\text{sgn}(\tau_l)$ appears in the sum (9) if and only if $x_l \in F_1$. The number of times this happens when considering τ is even, hence, according to this choice of ε_i , we have $\text{sgn}(\tau) = 1$. This implies that $\text{sgn}(\sigma \tau) = \text{sgn}(\sigma)$, i.e.

$$[\dots, [c_1 \theta_i \sigma, \dots, c_m \theta_i \sigma], \dots, [c_1 \theta_j \sigma, \dots, c_m \theta_j \sigma], \dots]$$

and

$$[\dots, [c_1 \theta_j, c_2 \theta_i, \dots, c_m \theta_i], \dots, [c_1 \theta_i \sigma, c_2 \theta_j \sigma, \dots, c_m \theta_j \sigma], \dots]$$

are two summands from (10) with the same sign. We can re-write the left hand side of (10) as

$$\sum_{\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K} \text{sgn}(\sigma) [[c_1 \theta_{i_1} \sigma, \dots, c_m \theta_{i_1} \sigma], \dots, [c_1 \theta_{i_2} \sigma, \dots, c_m \theta_{i_2} \sigma], \dots, [c_1 \theta_{i_k} \sigma, \dots, c_m \theta_{i_k} \sigma]] \quad (11)$$

Fix $\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K$, and set $y_{(u,v)} = c_v \theta_{i_u} \sigma$. Then

$$[[c_1 \theta_{i_1} \sigma, \dots, c_m \theta_{i_1} \sigma], \dots, [c_1 \theta_{i_2} \sigma, \dots, c_m \theta_{i_2} \sigma], \dots, [c_1 \theta_{i_k} \sigma, \dots, c_m \theta_{i_k} \sigma]]$$

equals

$$[[y_{(1,1)}, y_{(1,2)}, \dots, y_{(1,m)}], \dots, [y_{(2,1)}, y_{(2,2)}, \dots, y_{(2,m)}], \dots, [y_{(k,1)}, \dots, y_{(k,m)}]]$$

By the above argument,

$$[[y_{(2,1)}, y_{(1,2)}, \dots, y_{(1,m)}], \dots, [y_{(1,1)}, y_{(2,2)}, \dots, y_{(2,m)}], \dots, [y_{(k,1)}, \dots, y_{(k,m)}]]$$

appear as a summand in (11) with the same sign as the previously display term. Iterating the argument, for all permutation $v \in \mathfrak{S}(k)$, the terms

$$[[y_{(1v_1,1)}, \dots, y_{(1,m)}], \dots, [y_{(2v_1,1)}, \dots, y_{(2,m)}], \dots, [y_{(kv_1,1)}, \dots, y_{(k,m)}]]$$

all appear in (11) with the same sign. Similarly for $v_1, \dots, v_m \in \mathfrak{S}(k)$ the following terms all share the same sign in (11)

$$\begin{aligned} & [[y_{(1v_1,1)}, y_{(1,2)}, \dots, y_{(1,m)}], \dots, [y_{(2v_1,1)}, y_{(2,2)}, \dots, y_{(2,m)}], \dots, [y_{(kv_1,1)}, y_{(k,2)}, \dots, y_{(k,m)}]] \\ & [[y_{(1v_1,1)}, y_{(1v_2,2)}, \dots, y_{(1,m)}], \dots, [y_{(2v_1,1)}, y_{(2v_2,2)}, \dots, y_{(2,m)}], \dots, [y_{(kv_1,1)}, y_{(kv_2,2)}, \dots, y_{(k,m)}]] \\ & \vdots \\ & [[y_{(1v_1,1)}, \dots, y_{(1v_m,m)}], \dots, [y_{(2v_1,1)}, \dots, y_{(2v_m,m)}], \dots, [y_{(kv_1,1)}, \dots, y_{(kv_m,m)}]] \end{aligned}$$

hence, gathering the sub-sum of (11) consisting of all the latter terms as $v_1, \dots, v_m$ range in $\mathfrak{S}(k)$, we find an instance of identity (A)

$$\sum_{\sigma_1 \in \mathfrak{S}(k), \dots, \sigma_m \in \mathfrak{S}(k)} [[x_{(1\sigma_1,1)}, \dots, x_{(1\sigma_m,m)}], \dots, [x_{(2\sigma_1,1)}, \dots, x_{(2\sigma_m,m)}], \dots, [x_{(k\sigma_1,1)}, \dots, x_{(k\sigma_m,m)}]] = 0.$$

Doing this for each $\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K$, we establish that the expression in (11) vanishes, which proves the theorem.

REFERENCES

- [1] S. I. Adian. *The Burnside problem and identities in groups*, volume 95 of *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]*. Springer-Verlag, Berlin-New York, 1979. Translated from the Russian by John Lennox and James Wiegold.
- [2] Agatha Atkarskaya, Eliyahu Rips, and Katrin Tent. The Burnside problem for odd exponents (arxiv:2303.15997), 2024.
- [3] William Burnside. On groups in which every two conjugate operations are permutable. *Proc. Lond. Math. Soc.*, s1-35, 1902.
- [4] W. Fulton and J. Harris. *Representation theory: a first course*. Springer-Verlag., 1991.
- [5] E. S. Golod and I. R. Shafarevitch. On the class field tower. *Izv. Akad. Nauk SSSR Ser. Mat.*, 28:261–272, 1964.
- [6] Daniel Groves. A note on nonidentical Lie relators. *J. Algebra*, 211(1):15–25, 1999.
- [7] P. Hall and Graham Higman. On the p -length of p -soluble groups and reduction theorems for Burnside's problem. *Proc. London Math. Soc. (3)*, 6:1–42, 1956.
- [8] P. J. Higgins. Lie rings satisfying the Engel condition. *Proc. Cambridge Philos. Soc.*, 50:8–15, 1954.

- [9] G. D. James. *The Representation Theory of the Symmetric Groups*. Springer-Verlag., 1978.
- [10] A. I. Kostrikin. Lie rings satisfying the Engel condition. *Izv. Akad. Nauk SSSR Ser. Mat.*, 21:515–540, 1957.
- [11] A. I. Kostrikin. On Burnside’s problem. *Dokl. Akad. Nauk SSSR*, 119:1081–1084, 1958.
- [12] A. I. Kostrikin. The Burnside problem. *Izv. Akad. Nauk SSSR Ser. Mat.*, 23:3–34, 1959.
- [13] I. G. Lysénok. Infinite Burnside groups of even period. *Izv. Ross. Akad. Nauk Ser. Mat.*, 60(3):3–224, 1996.
- [14] Avinoam Mann and Consuelo Martinez. Groups nearly of prime exponent and nearly engel lie algebras. *Archiv der Mathematik*, 71(1):5–11, 1998.
- [15] Jr. Marshall Hall. *The Theory of Groups*. The Macmillan Company. The Macmillan Company, 7th printing 1965 edition, 1965.
- [16] P. S. Novikov and S. I. Adjan. Infinite periodic groups. I. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:212–244, 1968.
- [17] P. S. Novikov and S. I. Adjan. Infinite periodic groups. II. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:251–524, 1968.
- [18] P. S. Novikov and S. I. Adjan. Infinite periodic groups. III. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:709–731, 1968.
- [19] Gunnar Traustason. Engel Lie-algebras. *Quart. J. Math. Oxford Ser. (2)*, 44(175):355–384, 1993.
- [20] Gunnar Traustason. A constructive approach to Zelmanov’s global nilpotency theorem for n -Engel Lie algebras over a field of characteristic zero. *Internat. J. Algebra Comput.*, 8(3):317–326, 1998.
- [21] Michael Vaughan-Lee. *The restricted Burnside problem*, volume 8 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, second edition, 1993.
- [22] Michael Vaughan-Lee. Lie methods in Engel groups. In *Groups St Andrews 2009 in Bath. Volume 2*, volume 388 of *London Math. Soc. Lecture Note Ser.*, pages 551–559. Cambridge Univ. Press, Cambridge, 2011.
- [23] Michael Vaughan-Lee. 5-Engel lie algebras, 2024.
- [24] Michael Vaughan-Lee. On zel’manov’s global nilpotence theorem for engel lie algebras, 2025.
- [25] E. I. Zelmanov. On Engel Lie algebras. *Sibirsk. Mat. Zh.*, 29(5):112–117, 238, 1988.
- [26] E. I. Zelmanov. Solution of the restricted Burnside problem for groups of odd exponent. *Izv. Akad. Nauk SSSR Ser. Mat.*, 54(1):42–59, 221, 1990.
- [27] E. I. Zelmanov. Solution of the restricted Burnside problem for 2-groups. *Mat. Sb.*, 182(4):568–592, 1991.

UNIVERSITY OF THE BASQUE COUNTRY, DEPARTMENT OF MATHEMATICS (LEIOA) / INSTITUTE FOR LOGIC, COGNITION, LANGUAGE AND INFORMATION (DONOSTIA-SAN SEBASTIÁN), SPAIN

Email address: christian.delbee@ehu.eus

URL: <http://choum.net/~chris/page.perso/>