ZELMANOV CHARACTERISTIC 0 THEOREM AND THE RESTRICTED BURNSIDE PROBLEM

CHRISTIAN D'ELBÉE

Description of the course

Those are the notes of a course given by the author at the *Nesin Matematik Köyü* in September 2025. This course is centered on the restricted Burnside Problem (RBP) for groups of prime exponents, which is the following question: *for fixed $r \in \mathbb{N}$ and prime p, are there only finitely many finite r-generated groups of exponent p?* Unlike the other classical Burnside problems, the answer to the RBP is positive. This course does not include a full proof of the positive answer to the RBP for groups of prime exponent, which is due to Kostrikin in 1958-1959, nor of the general solution of the RBP, which is due to Zelmanov in 1991 and for which he was awarded the Fields medal in 1994. In this course, we will present the method for reducing the RBP for groups of prime exponent to the question of nilpotency of finitely generated n-Engel Lie algebras over a field of characteristic p with $p > n$. We will then look at solutions for particular values of the pair (n, p) before diving into the proof of Zelmanov's *characterictic* 0 Theorem: every n-Engel Lie algebra over a field of characteristic 0 is nilpotent. This theorem yields an asymptotic solution to the RBP: for each n there exists N such that n-Engel Lie algebras over a field of characteristic $p > N$ are nilpotent. The prerequisites for the course are a fair level in group theory and basic knowledge in universal algebra, linear and multilinear algebra.

Contents

Date: January 24, 2026.

The author is supported by the Ramon y Cajal grant RYC2023-042677-I funded by MICIU/AEI/10.13039/501100011033 and by ESF+.

Introduction

In 1902, William Burnside raised his famous problems [4].

> **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 $A_1, A_2, \ldots, A_m$ 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?*

In 1964, Golod and Shafarevich [6] 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 [17, 18, 19]. For $r, n \in \mathbb{N}$, we let $B(r, n)$ denote the quotient of the free group on r generators by the normal closure of the set of n-th powers (see Subsection 1.3), $B(r, n)$ is known as the *(free) Burnside group of rank r and exponent n*. As every r-generated group of exponent n is a homomorphic image of $B(r, n)$, the Burnside problem is the question whether $B(r, n)$ is always finite. As there are values of (r, n) for which $B(r, n)$ is finite, the problem since then consist in describing those. For instance, $B(r, n)$ are known to be finite for $n \leq 6, n \neq 5$, whereas $B(r, n)$ are all infinite for $r \geq 2$ and $n \geq 8000$ [14, 2]. A recent result from Atkarskaya, Rips and Tent [3] is that $B(r, n)$ is infinite for n odd and greater than 557. Note that it is still open whether 2-generated group of exponent 5 can be infinite (!).

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?*

In other words, the restricted Burnside problem asks whether $B(r, n)$ has only finitely many *finite* quotients. Using a result of Hall and Higman [8] (and the classification of finite simple groups), the restricted Burnside problem reduces to the case of $B(r, p^k)$ which was solved positively in 1959 by Kostrikin [11, 12, 13] (announced in 1958 in [12]) for $B(r, p)$ and by Zelmanov [27, 28] in 1991 for the general case. 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.*

The basic idea is that, given a group G, we may consider the lower central series $G = \gamma_1 \trianglerighteq \gamma_2 \trianglerighteq \ldots$ whose consecutive quotients γ_i/γ_{i+1} are abelian groups. The direct sum

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

can be equipped with a bracket $[.,.]$ defined in terms of the group commutator and, as a result, $(L(G),+,[.,.])$ has the structure of a *Lie ring*. This will be studied in details in Subsection 1.2 for the case of groups of exponent p, for which the Lie ring is in fact a Lie algebra over a the prime field $\mathbb{F}_p$. Then the connection with the RBP will be made apparent in Subsection 1.4, where it is proved that the RBP holds for a finitely generated group of prime exponent if and only if its associated Lie algebra is nilpotent. Then the turning point is given by Theorem 1.21: the associated Lie algebra of a group of exponent p satisfies the $(p-1)$-Engel identity, which is the following:

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

Therefore, to solve the RBP for groups of exponent p, it suffices to prove that finitely generated $(p-1)$-Engel Lie algebras are nilpotent. We then 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 [9]. In fact, in those cases, the finitely generated assumption is superfluous.

We then turn to the main focus of this course, Zelmanov's *characteristic* 0 theorem [26].

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, see Section 3.

Before reaching the very technical proof of Zelmanov's Theorem (Subsection 3), we will dig a little deeper in the theory of n-Engel Lie algebras and introduce the main ingredients for the proof. Free Lie algebras will be defined in Subsection 2.2, then Subsection 2.3 will establish a classical result of Higgins [9]: a $(p-1)$-Engel Lie algebra is nilpotent if and only if it is solvable. Finally, as it constitute one of the main subtlety of Zelmanov's proof, the beautiful theory of representations of the symmetric group will be described (without proof) in Subsection 2.5, with an application to 4-Engel Lie algebras of characteristic $p > 5$ due to Traustason, see Subsection 2.6.

The last key ingredient for the proof of Zelmanov's Theorem will also be assumed, as it is extremely technical. It is an early result of Kostrikin [13].

Theorem (Kostrikin, 1959). *Let L be an n-Engel Lie algebra of characteristic 0 (or of characteristic $p > n$), then L contains an abelian ideal.*

The proof of Kostrikin Theorem, using his famous *method of sandwiches* might be even harder than the proof of Zelmanov's characteristic zero result and in fact implies the RBP for groups of prime exponents. There is another proof by Adian and Rasborov [1] but it is still vastly technical. Therefore, Kostrikin Theorem is a costly assumption and the philosophy of this course is somehow backwards: we assume a theorem which is stronger than the RBP to deduce a weak solution to the RBP, but Zelmanov *characteristic* 0 is a beautiful theorem of its own, with unique technics coming into play.

Acknowledgement. I am very grateful to Michael Vaughan-Lee for sharing with me an early draft (now available as a preprint [25]) of a very clean exposition of the proof of Zelmanov's Theorem that is the main inspiration for our presentation in Section 3. I am very grateful to the organisation of the *Nesin Matematic Köyü*, for maintaining such a special place. Finally, I want to give a special thanks to Bulut Uygun and Doğan Turhan for their very enjoyable company and for their immense bravery in staying until the end of the course.

1. Preparation

1.1. **Prerequisites on groups and 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 $\mathrm{End}(G)$, the group of automorphisms $\mathrm{Aut}(G)$. 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

$g_1, \ldots, g_r \in G$ such that $G = \langle g_1, \ldots, g_r \rangle$. The normal closure of a set $S \subseteq G$ is the smallest normal subgroup of G contining S, and it is the group generated by all conjugates of elements of S. 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 $\{1, \ldots, n\}$. We now turn to Lie algebras, which we define more formally.

- A *Lie algebra* over a field $\mathbb{F}$ (or a Lie $\mathbb{F}$-algebra) is an $\mathbb{F}$-vector space equipped with a bilinear map $[.,.]$ which satisfies
  - (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 (in characteristic 2). 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 $[a + L, b + L] = [a, b] + L$ and the quotient map $\pi : L \to L/I$ is a Lie algebra homomorphism. The three isomorphism theorems hold for Lie algebras.
- A *monomial (or commutator) in* $x_1, \ldots, x_n$ is a term obtained using iteration of the bracket, for instance $[[x_1, [x_3, x_2]], x_4]$ or $[x_1, [x_2, [x_4, x_3]]]]$. A distinguished kind of monomials is given by *left-normed commutators* which are inductively defined as $[x_1, \ldots, x_n] = [[x_1, \ldots, x_{n-1}], x_n]$. We also use the notation:
$$[\ldots[x, \underbrace{y], y], \ldots, y}_{n \text{ times}}] = [x, y^{(n)}].$$
- 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 1.1, it is the span of all left-normed commutators evaluated in elements of S.
- For each a in a Lie $\mathbb{F}$-algebra L, we will denote by ad_a the map $x \mapsto [x, a]$, it is an element of $\mathrm{End}(L)$, the associative algebra of linear endomorphisms of L. **In this course, we will use the right-sided notation for the application of functions**, for instance $(a)\,\mathrm{ad}_b = [a, b]$. It might seems a bit odd at first but it is very consistent with our convention of left-normed bracketing. The composition $\mathrm{ad}_{b_1} \circ \ldots \circ \mathrm{ad}_{b_s}$ applied at an element $a \in L$ gives
$$(a)(\mathrm{ad}_{b_1} \circ \ldots \circ \mathrm{ad}_{b_s}) = ([a, b_1])\,\mathrm{ad}_{b_2} \circ \ldots \circ \mathrm{ad}_{b_s} = [a, b_1, \ldots, b_s]$$
  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 $\mathrm{ad}_b^n \in \mathrm{End}(L)$ at a. The Jacobi identity yields
$$[a, [b, c]] = [a, b, c] - [a, c, b]$$
  so that $\mathrm{ad}_{[b,c]} = \mathrm{ad}_b\,\mathrm{ad}_c - \mathrm{ad}_c\,\mathrm{ad}_b$.
- Any element of $\mathrm{End}(L)$ of the form ad_a is called a *Lie element* of $\mathrm{End}(L)$. Any algebraic relation true for Lie elements will yield a relation in the Lie algebra. For instance, $(x + y)^2 = x^2 + xy + yx + y^2$ is true in $\mathrm{End}(L)$, so in particular for Lie elements $x = \mathrm{ad}_b$, $y = \mathrm{ad}_c$ and hence in any Lie algebra:
$$[a, (b + c)^{(2)}] = [a, b^{(2)}] + [a, b, c] + [a, c, b] + [a, c^{(2)}].$$

Exercise 1.1. Any commutator in $x_1, \ldots, x_n$ can be written as a sum of left-normed commutators in $x_1, \ldots, x_n$.

Exercise 1.2. Prove that in a non-abelian Lie algebra L, there is no element 1 such that $[a, 1] = a$.

Definition 1.3. A Lie algebra is *n-Engel* if it satisfies the identity
$$0 = [x, \underbrace{y, y, \ldots, y}_{n \text{ times}}] = [x, y^{(n)}].$$

Lemma 1.4. *Let L be a Lie algebra over any field $\mathbb{F}$. If L is n-Engel then L satisfies the identity*
$$\sum_{\sigma \in \mathfrak{S}_n} [x, y_{1\sigma}, \ldots, y_{n\sigma}] = 0$$

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

Proof. The main point is the following claim.

Claim 1.

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

Proof of the claim. It suffices to develop the right hand side. To convince yourself, observe that in any associative algebra, we have for $n=3$:

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

hence

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

□

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 $[x,y^{(n)}]=0$. □

1.2. **Associated Lie $\mathbb{F}_p$-algebra of a group of exponent p.** The exposition draws mainly from [22]. 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

$$\gamma_1\trianglerighteq\gamma_2\trianglerighteq\dots$$

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 1.5). 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 1.5).

Exercise 1.5. 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 1.6. 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 1.7. 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 1.24.

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 1.6 (using that $gh=h^*g$), we have the following:

$$[gh,k]=[g,k]^*[h,k]^* \tag{1}$$
$$[g,hk]=[g,h]^*[g,k]^* \tag{2}$$
$$[g^{-1},h]=[g,h]^{-*} \tag{3}$$
$$[g,h^{-1}]=[g,h]^{-*} \tag{4}$$
$$[g,h,k]=[k,g^{-1},h^{-1}]^{-*}[h^{-1},k^{-1},g]^{-*} \tag{5}$$

Lemma 1.8. *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 1.6(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 1.9. *Let G be a group, let $S \subseteq G$ such that $G = \langle S\rangle$. For each $n \in \mathbb{N}$, define $S_n = \{[s_1, \ldots, s_n] \mid s_1, \ldots, s_n \in S\}$. Then*

(1) γ_n is the normal closure of S_n,
(2) If $\pi : \gamma_i \to \gamma_i/\gamma_{i+1}$, 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 $[s_1, \ldots, s_n]^k$ 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 $[[s_1, \ldots, s_n]^k, g]$. By Exercise 1.6, we have $[[s_1, \ldots, s_n]^k, g] = [[s_1, \ldots, s_n], g^{k^{-1}}]^k$. As $G = \langle S\rangle$, $g^{k^{-1}}$ is a product of elements of S and their inverses, hence by (2) and (4), $[[s_1, \ldots, s_n], g^{k^{-1}}]^k$ is in the normal closure of elements of the form $[[s_1, \ldots, s_n], s]$, 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 $a = [s_1, \ldots, s_{n+1}]$ to conclude that $[s_1, \ldots, s_{n+1}]^g = [s_1, \ldots, s_{n+1}]$ 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 1.8 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 $[., .]$ 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 1.8 and Exercise 1.5 (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 1.6, $[g\gamma_{i+1}, h\gamma_{j+1}] = [g, h]\gamma_{i+j+1} = [h, g]^{-1}\gamma_{i+j+1}$. 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 $[., .]$ 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 $g\gamma_{i+1} + h\gamma_{i+1} = gh\gamma_{i+j}$. Now we prove that $[g\gamma_{i+1} + h\gamma_{i+1}, k\gamma_{j+1}] = [g\gamma_{i+1}, k\gamma_{j+1}] + [k\gamma_{i+1}, k\gamma_{j+1}]$, 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 $[., .]$ 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 1.10. *The structure $(L(G), +, [.,.])$ is a Lie algebra over $\mathbb{F}_p$.*

Proof. We already know that $[.,.]$ 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 1.8, 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 1.6 (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 1.8, 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

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

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

1.3. **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 $x_{i_1}^{\epsilon_1} \dots x_{i_k}^{\epsilon_k}$ with exponents $\epsilon_i \in \{1, -1\}$, $1 \leq i_1, \dots, i_k \leq n$, $k \in \mathbb{N}$, which is subjects to the reduction

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

Equivalently it is an equivalence class of strings of characters $x_{i_1}^{\epsilon_1} \dots x_{i_k}^{\epsilon_k}$ 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

$$(x_{i_1}^{\epsilon_1} \dots x_{i_k}^{\epsilon_k}) \cdot (x_{j_1}^{\iota_1} \dots x_{j_k}^{\iota_l}) = x_{i_1}^{\epsilon_1} \dots x_{i_k}^{\epsilon_k} x_{j_1}^{\iota_1} \dots x_{j_k}^{\iota_l}$$

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 $w(x_{i_1}, \dots, x_{i_n})$ if the variables appearing in the word w are among $x_{i_1}, \dots, x_{i_n}$. 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 $w(g, h) = g^{-1}h^{-1}gh = [g, h] \in G$.

Definition 1.11 (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 1.12. 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 1.13 (Free groups)**.** Given a variety of groups $\mathscr{V}$ and a set S of variable. A *free group for $\mathscr{V}$ generated by S* (or a group *freely generated by S*) is an element $F(S)$ of $\mathscr{V}$ generated by S and such that for any group $G \in \mathscr{V}$, any mapping $S \to G$ extends uniquely to a group homomorphism $F(S) \to G$.

$$\begin{tikzcd} S \arrow[r] \arrow[d, phantom, "\cap"] & G \\ F(S) \arrow[ur, dashed] & \end{tikzcd}$$

Free objects always exists in a variety.

Theorem 1.14. *Every variety of groups admits a free group in any given set of generators.*

Exercise 1.15. Let $\mathscr{V}$ be a variety of p-groups. prove that there exists n such that every groups in $\mathscr{V}$ 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 1.16. *Let H be a fully invariant ideal of the free group $W = W(\omega)$. Then W/H is a free group for some variety $\mathscr{V}$.*

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

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

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

As was mentioned above (see Exercise 1.12), given a fixed $n \in \mathbb{N}$, the groups of exponent dividing n form a variety.

Definition 1.17. 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 1.16, 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 $\{w^n \mid w \in F(r)\}$.

1.4. **The restricted Burnside problem.** For this subsection, we draw mainly from [22] and [15] for Theorem 1.21.

Theorem 1.18. *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 1.9, 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. [16, Theorem 12.3.1]

Fact 1.19 (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 1.20. *Let $G = \langle x, y_1, \ldots, y_s\rangle$ 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 $x, y_1, \ldots, y_s$ 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 $x, y_1, \ldots, y_s$ where x still apears twice. $\square$

Theorem 1.21. *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 $x, y_1, \ldots, y_{p-1}$. Let $a = x\gamma_2, b_i = y_i\gamma_2$. From Claim 1, we have the following equality in $L(B)$

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

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 1.8, 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}:

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

Using Lemma 1.20, 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

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

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 (†), 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 $c_1, \ldots, c_s$ 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 (†) involve at least one y_1. Using an endomorphism θ_2 of B mapping y_2 to 1 and fixing $x, y_1, y_3, \ldots, y_{p+1}$, we may similarly assume that all c_i involve y_2. Iterating the argument, we see that each c_i involve all of $x, y_1, \ldots, y_{p-1}$

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 1.8, 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 1.8), 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

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

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 (‡) holds for all elements of $L(B)$. By Lemma 1.4 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. □

Remark 1.22. 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 [7].

Corollary 1.23. *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 1.18, 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 1.25). By Theorem 1.21, $L(B(r, p))$ 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. □

Exercise 1.24. 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 representent $x_1, \dots, x_s$ 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 1.25. 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 $g_1\gamma_2, \dots, g_r\gamma_2$.

Exercise 1.26. (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 1.18:

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

2. A first study of n-Engel Lie algebra

2.1. **n-Engel Lie algebras for $n = 2, 3$.** We now look at particular case of nilpotency of n-Engel Lie algebras. This section draws from Higgins [9] and Traustason [20].

Example 2.1 (2-Engel Lie $\mathbb{F}_p$-algebras are nilpotent)**.** Let L be a 2-Engel Lie $\mathbb{F}_p$-algebra. For convenience, we will write $A, B, C, \ldots$ for $\mathrm{ad}_a, \mathrm{ad}_b, \mathrm{ad}_c, \ldots$. As L is 2-Engel, we have $[c, a^{(2)}] = [c, b^{(2)}] = 0$, so that, in $\mathrm{End}(L)$ we have $A^2 = B^2 = 0$. Because $A + B$ is also a Lie element (it is ad_{a+b}) we also have:

$$0 = (A + B)^2 = AB + BA. \tag{$\star$}$$

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

$$BC - 2CB = 0 \tag{$\star\star$}$$

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 $L^4 = 0$.

Example 2.2 (3-Engel Lie $\mathbb{F}_p$-algebras are nilpotent for $p \neq 2, 5$)**.** Again, we write $A, B, \ldots$ for $\mathrm{ad}_a, \mathrm{ad}_b, \ldots$.
Linearization. We developp the 3-Engel equation $0 = (A + \lambda B)^3$ for $a, b \in L$ and $\lambda \in \mathbb{F}_p$:

$$\lambda(A^2B + ABA + AB^2) + \lambda^2(AB^2 + BAB + B^2A) = 0 \tag{👻(λ)}$$

This equation holds for all values of λ. As $p \neq 2$, there exists $\lambda \in \mathbb{F}_p$ such that $\lambda^2 \neq \lambda$. By substracting λ times 👻(1) from 👻(λ) and dividing the result by $\lambda - \lambda^2$, we obtain

$$AB^2 + BAB + B^2A = 0 \tag{👻}$$

Applying (👻) 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

$$3CB^2 - 3BCB + B^2C = 0 \tag{👻👻}$$

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

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

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 (👻) and (👻👻) we deduce

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

In particular, we have

$$3AB^2 = -2B^2A. \tag{🎃}$$

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

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

From (☺), 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 2.2 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 (☺) 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 $L^5 = 0$.

2.2. **Varieties of Lie algebras, free Lie algebras.** We fix a field $\mathbb{F}$, and we consider the set $\mathrm{Lie}(x_1, \ldots, x_n)$ consisting of $\mathbb{F}$-linear combination of (formal) Lie monomials in $x_1, \ldots, x_n$. Those are called *(formal) Lie polynomials over* $\mathbb{F}$ in $x_1, \ldots, x_n$. For a Lie algebra L and $a_1, \ldots, a_n \in L$, we may again evaluate any element $p(x_1, \ldots, x_n)$ from $\mathrm{Lie}(x_1, \ldots, x_n)$ by interpreting the expression $p(a_1, \ldots, a_n)$ 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 $\mathrm{Lie}(x_1, \ldots, x_n)$, 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 2.3 (Varieties of Lie $\mathbb{F}$-algebras)**.** Let V be a set of Lie polynomials in $\mathrm{Lie}(x_i \mid i \in \mathbb{N})$. The *variety generated by* V is the collection of all Lie algebras L such that for all $p(x_1, \ldots, x_n) \in V$ and for all $a_1, \ldots, a_n \in L$ we have $p(a_1, \ldots, a_n) = 0$. A *variety of Lie algebras* is a variety defined by some set of words.

Exercise 2.4. Check that the class of n-Engel Lie $\mathbb{F}$-algebras is a variety.

Definition 2.5 (Free Lie algebras)**.** Given a variety of Lie $\mathbb{F}$-algebras $\mathscr{V}$ and a set S of variable, a *free Lie $\mathbb{F}$-algebra for $\mathscr{V}$ generated by S* is an element $F(S)$ of $\mathscr{V}$ containing S and such that for any Lie $\mathbb{F}$-algebra $L \in \mathscr{V}$, any mapping $S \to L$ extends uniquely to a Lie $\mathbb{F}$-algebra homomorphism $F(S) \to L$.

$$\begin{tikzcd} S \arrow[r] \arrow[d, phantom, "\cap"] & L \\ F(S) \arrow[ur, dashed] & \end{tikzcd}$$

Theorem 2.6. *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 $(A, +, [.,.])$ 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 $(x_1, \ldots, x_n)$ the tuple $(i_1, \ldots, i_n)$ where each i_j is the number of times x_i appears in the monomial. For a given $(i_1, \ldots, i_n)$ we define $V_{(i_1, \ldots, i_n)}$ to be the vector space spanned by elements of multiweight $(i_1, \ldots, i_n)$. One key fact about multiweight is that for two different multiweight $(i_1, \ldots, i_n), (j_1, \ldots, j_n)$ we have

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

We also have that

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

Note that by allowing the multiweight to The *weight* of the monomial is the sum $i_1 + \ldots + i_n$. For instance, $[x, y^{(3)}]$ has multiweight $(1, 3)$ and weight 4. If $[x, y^{(3)}]$ can be written as a linear combination $\lambda_1 m_1 + \ldots + \lambda_k m_k$ 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 $\phi : L^k \to L$, 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 2.7. *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 $\mathscr{V}$.*

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

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

We prove that L/I is a free Lie algebra relatively to $\mathscr{V}$, freely generated by $(x_i + I)_{i\in\mathbb{N}}$. Note that L/I belongs to $\mathscr{V}$ since I is fully invariant. Let L be any Lie algebra in $\mathscr{V}$, and assume we are given a mapping $f : \{x_i + I \mid i \in \mathbb{N}\} \to L$ 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 $\theta : F \to L$. As $L \in \mathscr{V}$, we have that $I \subseteq \ker\theta$ hence $\theta_I : F/I \to L$ 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 2.8. Let $\mathscr{V}$ 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 $\mathscr{V}$ is nilpotent.

Exercise 2.9. Let $\mathscr{V}$ be a variety of Lie $\mathbb{F}$-algebras such that every element of $\mathscr{V}$ is n-Engel for some $n \in \mathbb{N}$. Prove that there exists $N \in \mathbb{N}$ such that every element of $\mathscr{V}$ is N-Engel.

Exercise 2.10. 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 2.11. (1) Prove that for each $n \in \mathbb{N}$, the vector space

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

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.

2.3. **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 [9].

Theorem 2.12. *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

$$L^{ni+2} \subseteq M^{i+1}. \tag{$\star$}$$

It is trivial for $i = 0$, so assume $i > 0$. By Lemma 1.4, we have that $\sum_{\sigma\in\mathfrak{S}_n} [a, b_{1\sigma}, \dots, b_{n\sigma}] = 0$. Using the Jacobi identity, we have

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

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. □

We immediately get:

Corollary 2.13. *Let L be a $(p-1)$-Engel Lie algebra over $\mathbb{F}_p$. Then L is nilpotent if and only if L is solvable.*

2.4. **$\mathbb{Z}_2$-gradings on Engel Lie algebras.**

Lemma 2.14. *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 1.4, 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. □

Definition 2.15. 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 2.16. 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

$$x_i \in L_0 \iff i \in S$$
$$x_i \in L_1 \iff i \notin S.$$

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 $L_0 = \mathrm{Span}(U_0)$ and $L_1 = \mathrm{Span}(U_1)$. 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 2.17. *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

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

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 2.14, $[b, a_1, \dots, a_k]$ 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 2.12), L is nilpotent of class $\frac{n^{k+m}-1}{n-1}$. □

2.5. **Group algebra, Young tableau and representation theory of the symmetric group.** References for this section are [5] and [10]. 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:

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

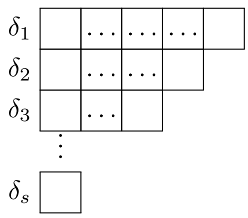
and multiplication is defined as follows:

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

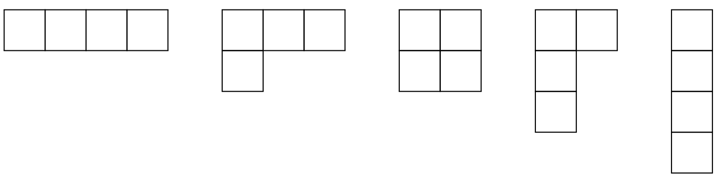
A *representation of G (over $\mathbb{F}$)* is the data of a right-action of G on an $\mathbb{F}$-vector space V which is linear, in the sense that $(\lambda v + \mu w)g = \lambda vg + \mu wg$. A representation V of G (with a given right-action of G on V) is the same thing as equipping V with the action 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 $V = \mathrm{Span}_{\mathbb{F}}(X)$ 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 $\delta = (\delta_1 \geq \ldots \geq \delta_s)$ of d (i.e. such that $d = \delta_1 + \ldots + \delta_s$), *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 $\{1,\ldots,d\}$. Canonical examples are given by numbering the boxes in the (descending) lexicographical order:

$$\begin{array}{|c|c|c|c|}\hline 1&2&3&4\\\hline\end{array}\qquad \begin{array}{|c|c|c|}\hline 1&2&3\\\hline \multicolumn{1}{|c|}{4}&\multicolumn{2}{c}{}\\\cline{1-1}\end{array}\qquad \begin{array}{|c|c|}\hline 1&2\\\hline 3&4\\\hline\end{array}\qquad \begin{array}{|c|c|}\hline 1&2\\\hline \multicolumn{1}{|c|}{3}&\multicolumn{1}{c}{}\\\cline{1-1} \multicolumn{1}{|c|}{4}&\multicolumn{1}{c}{}\\\cline{1-1}\end{array}\qquad \begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline 4\\\hline\end{array}$$

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 2.18. *Let $\mathbb{F}$ be either $\mathbb{Q}$ of $\mathbb{F}_p$, 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*

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

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 $T_1,\ldots,T_k$ be all Young tableau (up to permutation of the numbering of the boxes). Let $e_1,\ldots,e_k$ be the corresponding primitive idempotent and $W_i = \mathbb{F}\mathfrak{S}_d e_i$ for $i = 1,\ldots,k$. Then the map $x \mapsto xe_1 + \ldots + xe_k$ defines an isomorphism $\mathbb{F}\mathfrak{S}_d \cong W_1 \oplus \ldots \oplus W_k$.

Of course, $\mathbb{F}\mathfrak{S}_d \cong V_1 \oplus \ldots \oplus V_k$ 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.

2.6. **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 [20, 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 [24, 25].

Theorem 2.19 (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 $x_1,\ldots,x_8$. By freeness, the theorem follows if we are able to prove that

$$[x_1,\ldots,x_8] = 0. \tag{†}$$

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.

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

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

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

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

Then this action extends linearly by

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

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 $x_1, \ldots, x_5$ (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 2.5, the identity in $\mathbb{F}_p\mathfrak{S}(5)$ may be written as a sum of primitive idempotent $e_1 + \ldots + e_s$. It follows that (†) holds if we can prove that for all primitive idempotent e of $\mathbb{F}_p\mathfrak{S}(5)$, we have

$$[x_1, \ldots, x_8]e = 0 \tag{‡}$$

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

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

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.

$$\begin{array}{|c|c|c|c|c|}\hline i_1 & i_2 & i_3 & & \\ \hline \end{array} \qquad \begin{array}{|c|c|c|c|}\hline i_1 & i_2 & i_3 & \\ \hline \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|c|c|}\hline i_1 & i_2 & i_3 \\ \hline & \\ \cline{1-2} \end{array} \qquad \begin{array}{|c|c|c|}\hline i_1 & i_2 & i_3 \\ \hline \\ \cline{1-1} \\ \cline{1-1} \end{array}$$

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

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

In order to establish (‡), 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

$$\sum_{\sigma \in S} [x_{1\pi c}, \ldots, (x_{j_1\pi c})\sigma, \ldots, (x_{j_2\pi c})\sigma, \ldots, (x_{j_3\pi c})\sigma, \ldots, x_{5\pi c}, x_6, x_7, x_8] = 0$$

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

$$\sum_{\sigma \in \mathfrak{S}(3)} [x_1, \ldots, x_{i_{1\sigma}}, \ldots, x_{i_{2\sigma}}, \ldots, x_{i_{3\sigma}}, \ldots, x_5, x_6, x_7, x_8]$$

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 $f = \sum_{\sigma \in S} \mathrm{sgn}(\sigma)\sigma$. Then

$$\begin{array}{|c|}\hline i_1 \\ \hline i_2 \\ \hline i_3 \\ \hline \\ \hline \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline i_1 & \\ \hline i_2 \\ \cline{1-1} i_3 \\ \cline{1-1} \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|c|}\hline i_1 & \\ \hline i_2 & \\ \hline i_3 \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|c|c|}\hline i_1 & & \\ \hline i_2 \\ \cline{1-1} i_3 \\ \cline{1-1} \end{array}$$

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

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

$$\sum_{\sigma\in\mathfrak{S}(3)}\operatorname{sgn}(\sigma)[x_1,\dots,x_{i_{1\sigma}},\dots,x_{i_{2\sigma}},\dots,x_{i_{3\sigma}},\dots,x_5,x_6,x_7,x_8]$$

We have established the following claim.

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

$$\sum_{\sigma\in\mathfrak{S}(3)}[x_1,\dots,x_{i_{1\sigma}},\dots,x_{i_{2\sigma}},\dots,x_{i_{3\sigma}},\dots,x_5,x_6,x_7,x_8] \tag{E_{sym}}$$

$$\sum_{\sigma\in\mathfrak{S}(3)}\operatorname{sgn}(\sigma)[x_1,\dots,x_{i_{1\sigma}},\dots,x_{i_{2\sigma}},\dots,x_{i_{3\sigma}},\dots,x_5,x_6,x_7,x_8] \tag{E_{skew}}$$

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

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

Then L is a Lie algebra, and by Lemma 1.4, it is 4-Engel.

Exercise 2.20. 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 $X_i=x_i\otimes 1$. On has that

$$[X_1,X_{234}^{(3)},X_5,X_6,X_7,X_8]=\left(\sum_{\sigma\in\mathfrak{S}(2,3,4)}[x_1,x_{2\sigma},x_{3\sigma},x_{4\sigma},x_5,x_6,x_7,x_8]\right)\otimes e_1e_2e_3$$

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

$$[X_{134},X_2,X_{134}^{(2)},X_5,X_6,X_7,X_8]=\left(\sum_{\sigma\in\mathfrak{S}(1,3,4)}[x_{1\sigma},x_2,x_{3\sigma},x_{4\sigma},x_5,x_6,x_7,x_8]\right)\otimes e_1e_2e_3.$$

By mimicking the above, we have that every sum in (E_{sym}) can be realized as a product in $\tilde{L}$ where $X_{i_1i_2i_3}$ 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 $u_i^2=0$, $u_iu_j=-u_ju_i$ and consider

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

Let $L_0 = L \otimes \mathrm{Span}(u_i u_j \mid 1 \leq i,j,k \leq 3)$ and $L_1 = L \otimes \mathrm{Span}(u_i, u_i u_j u_k \mid 1 \leq i,j,k \leq 3)$. Then one easily checks that L^{super} is a Lie superalgebra, and by Lemma 1.4, 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 2.21. 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

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

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,

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

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 $P(X_1, \ldots, X_6) = 0$ 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 stuctures. Once he reaches Step 2 and the reduction to equations of type:

$$\sum_{\sigma \in \mathfrak{S}(3)} [x_1, \ldots, x_{i_{1\sigma}}, \ldots, x_{i_{2\sigma}}, \ldots, x_{i_{3\sigma}}, \ldots, x_5, x_6, x_7, x_8]$$

$$\sum_{\sigma \in \mathfrak{S}(3)} \mathrm{sgn}(\sigma) [x_1, \ldots, x_{i_{1\sigma}}, \ldots, x_{i_{2\sigma}}, \ldots, x_{i_{3\sigma}}, \ldots, x_5, x_6, x_7, x_8]$$

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 $[x_1, \ldots, x_8] = 0$ 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.

3. 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

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

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:

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

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 following[1].

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 $\leq c$.*

3.1. **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 $\pi : L \to L/I_{\alpha-1}$ 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 generators of M. Assume that a lives in the Lie subalgebra of L generated by $y_1, \ldots, y_r$. Because a generates an abelian ideal in M, we have

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

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

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

As this holds for all $a_1, \ldots, a_k \in M$ 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 $[x_1, \ldots, x_m] \in I_\alpha$ for some $m \in \mathbb{N}$ and $\alpha < \beta$. Then L/I_α is a relatively free Lie algebra (by Claim 5) and because $[x_1, \ldots, x_m] = 0$, 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 3.1. 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.

[1]See also [23] 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 [21].

3.2. **Step 2: First footstep 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 charactericstic 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.

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

As M is realively free, this implies that the above is an identity in M and hence that M is solvable. Then using Higgins Theorem (Corollary 2.13), we conclude that M is nilpotent.

3.3. **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

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

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:

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

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

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

then

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

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:

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

which is easily seens to equal

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

$\square$

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

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

is a consequence of the equations

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\epsilon_1} \sum_{\sigma_2 \in \mathfrak{S}_2}^{\epsilon_2} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\epsilon_K} [\ldots, [x^{\sigma_1}_{(1,1)}, \ldots, x^{\sigma_K}_{(1,K)}], \ldots, [x^{\sigma_1}_{(2,1)}, \ldots, x^{\sigma_K}_{(2,K)}] \ldots, [x^{\sigma_1}_{(R,1)}, \ldots, x^{\sigma_K}_{(R,K)}], \ldots] = 0 \tag{2}$$

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

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

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

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

The group $\mathfrak{S}(N)$ act on the set (on the right) $\{x_{(1,1)}, \ldots, x_{(N,1)}\}$ 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

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

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 2.5, 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

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

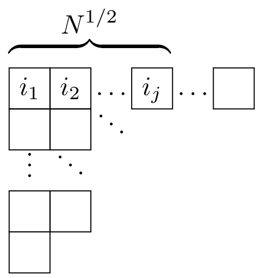
Indeed, the identity Id of $\mathbb{Q}\mathfrak{S}(N)$ can be written as a sum of primitive idempotents $\mathrm{Id} = e_1 + \ldots + e_s$ and for any linear combination u of monomials in $x_{(i,j)}$ we have $u = \mathrm{Id}(u) = \sum_{i=1}^{s} ue_i$.

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

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

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 $i_1, \ldots, i_{N^{1/2}}$ 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 $\{i_1, \ldots, i_{N^{1/2}}\}$ can be seen as the subgroup S of H which fixes elements of

$\{1\ldots,N\}\setminus\{i_1,\ldots,i_{N^{1/2}}\}$. 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

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

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

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

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

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

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

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

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

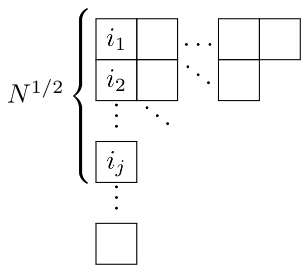
$$\sum_{\sigma\in S}[\ldots,t_{j_1}\pi c\sigma,\ldots,t_{j_{N^{1/2}}}\pi c\sigma,\ldots]=0$$

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:

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

Recall that extra entries are to be inserted between the entries $t_1^\sigma,\ldots,t_{N^{1/2}}^\sigma$, 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 $i_1,\ldots,i_{N^{1/2}}$ 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 $\{1,\ldots,N\}\setminus\{i_1,\ldots,i_{N^{1/2}}\}$ and let C be a right transversal for S in V, that is

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

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

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

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

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

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

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

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

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

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

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

We write

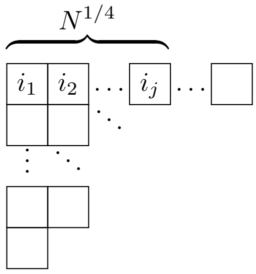
$$
e = \frac{1}{s}\sum_{\pi\in V,\rho\in H}\operatorname{sgn}(\pi)\pi\rho
$$

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

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



is a linear combination of terms of the form

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

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

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

and hence

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

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

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

Proof. First, the restriction $\sigma_{0\restriction\{j_1,\ldots,j_{N^{1/4}}\}}:\{j_1,\ldots,j_{N^{1/4}}\}\to X$ can be corrected by an element $a\in A$ so that $\sigma_0\restriction\{j_1,\ldots,j_{N^{1/4}}\}a=\sigma\restriction\{j_1,\ldots,j_{N^{1/4}}\}$. Similarly, there exists $b\in B$ such that

$$\sigma_{\restriction(\{1,\ldots,N^{1/2}\}\setminus\{j_1,\ldots,j_{N^{1/4}}\})}:(\{1,\ldots,N^{1/2}\}\setminus\{j_1,\ldots,j_{N^{1/4}}\})\to\{1,\ldots,N^{1/2}\}\setminus X$$

is equal to $\sigma_{0\restriction(\{1,\ldots,N^{1/2}\}\setminus\{j_1,\ldots,j_{N^{1/4}}\})}b$. Now because the domains and the ranges of $\sigma\restriction\{j_1,\ldots,j_{N^{1/4}}\}$ and $\sigma_{\restriction(\{1,\ldots,N^{1/2}\}\setminus\{j_1,\ldots,j_{N^{1/4}}\})}$ 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 $\sigma_0\restriction\{j_1,\ldots,j_{N^{1/4}}\}a$ and $\sigma_{0\restriction(\{1,\ldots,N^{1/2}\}\setminus\{j_1,\ldots,j_{N^{1/4}}\})}b$ gives $\sigma_0 ab$ (note that a and b commute since they have disjoint support). $\square$

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

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

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

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

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

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

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

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

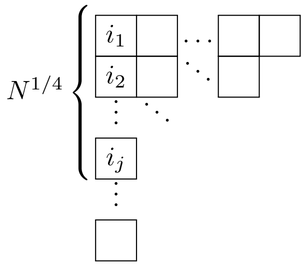
where

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

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

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



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

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

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

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

$$\sum_{\sigma_1\in\mathfrak{S}(R)}^{\epsilon_1}\cdots\sum_{\sigma_K\in\mathfrak{S}(R)}^{\epsilon_K} [\dots,t_1^{(\sigma_1,\dots,\sigma_K)},\dots,t_2^{(\sigma_1,\dots,\sigma_K)},\dots,t_R^{(\sigma_1,\dots,\sigma_K)},\dots] = 0$$

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

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

3.5. **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 $\epsilon_i = 1$ for all $i = 1,\dots,K$ we have the identity

$$\sum_{\sigma_1\in\mathfrak{S}(R)}^{\epsilon_1}\cdots\sum_{\sigma_K\in\mathfrak{S}(R)}^{\epsilon_K} [\dots,t_1^{(\sigma_1,\dots,\sigma_K)},\dots,t_2^{(\sigma_1,\dots,\sigma_K)},\dots,t_R^{(\sigma_1,\dots,\sigma_K)},\dots] = 0$$

Hence, assuming that $\mathrm{Id}\in\mathbb{Q}\mathfrak{S}(k)$, $\mathrm{Id}\in\mathbb{Q}\mathfrak{S}(k^2)$, $\mathrm{Id}\in\mathbb{Q}\mathfrak{S}(k^{2^2})$, …, $\mathrm{Id}\in\mathbb{Q}\mathfrak{S}(N)$ 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.

3.6. **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

$$\sum_{\sigma_1 \in \mathfrak{S}(R)}^{\epsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}(R)}^{\epsilon_K} [\ldots, t_1^{(\sigma_1,\ldots,\sigma_K)}, \ldots, t_2^{(\sigma_1,\ldots,\sigma_K)}, \ldots, t_R^{(\sigma_1,\ldots,\sigma_K)}, \ldots] = 0.$$

We now define the constants R and K. Recall that k is the number from Step 3 such that the ideal generated by $[x_1, \ldots x_m]$ 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 $x_1, \ldots, x_K$. It is a relatively free n-Engel Lie algebra with free generators $x_1, \ldots, x_K$. Assume that we are given a $\mathbb{Z}_2$-grading $F = F_0 \oplus F_1$ according to Example 2.16 with some $x_i \in F_0$ and the others in F_1. We have U the set of left-normed commutators in the $x_1, \ldots, x_K$ 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 $[c_1, \ldots, c_m]$ with $c_i \in U_0$. Then F/I satisfies the hypotheses of Lemma 2.17, and by the choice of K as defined above, we have $[x_1/I, \ldots, x_K/I] = 0$ in F/I hence $[x_1, \ldots, x_K] \in I$. This means that $[x_1, \ldots, x_K]$ is a finite linear combination of elements of the form $[c_1, \ldots, c_m, a_1, \ldots, a_t]$, for $c_i \in U_0$ and $a_i \in \{x_1, \ldots, x_K\}$. Note that because $[x_1, \ldots, x_K]$ is multilinear and F is relatively free and multigraded, we may assume that each element $[c_1, \ldots, c_m, a_1, \ldots, a_t]$ is multilinear and of weight K.

For any of those 2^K gradings on $F = \langle x_1, \ldots, x_K \rangle$ 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 $L = \langle x_1, \ldots, x_K \rangle$, the element $[x_1, \ldots, x_K]$ is a linear combination of at most T many monomials of the form $[c_1, \ldots, c_m, a_1, \ldots, a_t]$ with $c_i \in U_0$.

We set

$$R := kT.$$

We want to prove

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\epsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\epsilon_K} [\ldots, t_1, \ldots, t_2, \ldots, t_R, \ldots]\sigma_1 \ldots \sigma_K = 0 \tag{8}$$

where $\mathfrak{S}_j$ is a copy of $\mathfrak{S}(R)$ which acts on the set $\{x_{(1,j)}, \ldots, x_{(R,j)}\}$ 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 $\mathfrak{S}_1 \times \ldots \times \mathfrak{S}_K$.

We now fix a choice of $\epsilon_1, \ldots, \epsilon_K$, and we define a $\mathbb{Z}_2$-grading on $F := \langle x_1, \ldots, x_K \rangle = F_0 \oplus F_1$ by setting

$$x_i \in F_0 \iff \epsilon_i = +,$$
$$x_i \in F_1 \iff \epsilon_i = -.$$

By the above argument, using Lemma 2.17, there exists a number $s \leq T$ such that

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

where u_i are multilinear terms of weight K of the form

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

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

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

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

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

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\epsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\epsilon_K} [\ldots, u_{r_1}\theta_1, \ldots, u_{r_2}\theta_2, \ldots, u_{r_R}\theta_R, \ldots]\sigma_1 \ldots \sigma_K$$

for all $1 \le r_1, \dots, r_R \le s$. As $R = kT \ge ks$ 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

$$\begin{aligned}&\sum_{\sigma_1 \in \mathfrak{S}_1}^{\epsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\epsilon_K} [\dots, u_{r_1}\theta_1, \dots, u_{r_2}\theta_i, \dots, u_{r_R}\theta_R, \dots]\sigma_1 \dots \sigma_K \\ &= \sum_{\sigma_1 \in \mathfrak{S}_1}^{\epsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\epsilon_K} [\dots, u_r\theta_{i_1}, \dots, u_r\theta_{i_2}, \dots, u_r\theta_{i_k}, \dots]\sigma_1 \dots \sigma_K\end{aligned}$$

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

$$\{c_1\theta_i, \dots, c_m\theta_i, a_1\theta_i, \dots, a_q\theta_i\} \cap \{c_1\theta_j, \dots, c_m\theta_j, a_1\theta_j, \dots, a_q\theta_j\} = \emptyset$$

so that

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

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

$$\sum_{\sigma_1 \in \mathfrak{S}_1}^{\epsilon_1} \cdots \sum_{\sigma_K \in \mathfrak{S}_K}^{\epsilon_K} [\dots, [c_1, \dots, c_m]\theta_{i_1}, \dots, [c_1, \dots, c_m]\theta_{i_2}, \dots, [c_1, \dots, c_m]\theta_{i_k}, \dots]\sigma_1 \dots \sigma_K = 0 \tag{9}$$

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

$$\sum_{\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K} \tilde{\operatorname{sgn}}(\sigma)[\dots, [c_1, \dots, c_m]\theta_{i_1}, \dots, [c_1, \dots, c_m]\theta_{i_2}, \dots, [c_1, \dots, c_m]\theta_{i_k}, \dots]\sigma = 0$$

where if $\sigma = \sigma_1 \cdots \sigma_K$ we define $\tilde{\operatorname{sgn}}(\sigma)$ to be the product of the signatures $\operatorname{sgn}(\sigma_i)$ for $\epsilon_i = -$, in other words $\tilde{\operatorname{sgn}}(\sigma) = \prod_{\{i | \epsilon_i = -\}} \operatorname{sgn}(\sigma_i)$. 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 "$\dots]$" and without the extra hidden terms on the left "$[\dots,$" (for the latter, observe that any commutator $[\dots, x]$ is a linear combination of commutators of the form $[x, \dots]$). So, proving the above identity is equivalent to proving the following:

$$\sum_{\sigma \in \mathfrak{S}_1 \times \dots \times \mathfrak{S}_K} \tilde{\operatorname{sgn}}(\sigma)[[c_1, \dots, c_m]\theta_{i_1}, \dots, [c_1, \dots, c_m]\theta_{i_2}, \dots, [c_1, \dots, c_m]\theta_{i_k}]\sigma = 0 \tag{10}$$

3.7. **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 $\tau = \tau_1 \cdots \tau_q$. Now recall that the term

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

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 $[c_1,\ldots,c_m]\theta_i$ and $[c_1,\ldots,c_m]\theta_j$ within (Δ) we obtain

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

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

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

Since $c_1=[x_{k_1},\ldots,x_{k_q}]\in F_0$, the number of elements of $\{x_{k_1},\ldots,x_{k_q}\}$ which lie in F_1 is even. We have that $q\leq K$ and τ can be seen as an element of $\mathfrak{S}_1\times\ldots\times\mathfrak{S}_K$ (putting identity elements in between) hence appears as a permutation involved in the sum (10). We argue that $\tilde{\operatorname{sgn}}(\tau)=1$. Indeed, the transposition τ_l will appear with $\operatorname{sgn}(\tau_l)$ in the sum if and only if $\epsilon_{k_l}=-$. 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 $\{x_{(1,k_l)},\ldots,x_{(R,k_l)}\}$ and hence are acted on by permutations from $\mathfrak{S}_{k_l}$, which corresponds to a sum twisted by ϵ_{k_l}. Hence $\operatorname{sgn}(\tau_l)$ 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 $\tilde{\operatorname{sgn}}(\tau)=1$. This implies that $\tilde{\operatorname{sgn}}(\sigma\tau)=\tilde{\operatorname{sgn}}(\sigma)$, i.e.

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

and

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

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

$$\sum_{\sigma\in\mathfrak{S}_1\times\ldots\times\mathfrak{S}_K}\tilde{\operatorname{sgn}}(\sigma)[[c_1\theta_{i_1}\sigma,\ldots,c_m\theta_{i_1}\sigma],\ldots,[c_1\theta_{i_2}\sigma,\ldots,c_m\theta_{i_2}\sigma],\ldots,[c_1\theta_{i_k}\sigma,\ldots,c_m\theta_{i_k}\sigma]] \tag{11}$$

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

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

equals

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

By the above argument,

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

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

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

all appear in (11) with the same sign. Similarly for $\nu_1,\ldots,\nu_m\in\mathfrak{S}(k)$ the following terms all share the same sign in (11)

$$\begin{aligned}&[[y_{(1\nu_1,1)},y_{(1,2)},\ldots,y_{(1,m)}],\ldots,[y_{(2\nu_1,1)},y_{(2,2)},\ldots,y_{(2,m)}],\ldots,[y_{(k\nu_1,1)},y_{(k,2)}\cdots,y_{(k,m)}]]\\ &[[y_{(1\nu_1,1)},y_{(1\nu_2,2)}\cdots,y_{(1,m)}],\ldots,[y_{(2\nu_1,1)},y_{(2\nu_2,2)}\cdots,y_{(2,m)}],\ldots,[y_{(k\nu_1,1)},y_{(k\nu_2,2)}\cdots,y_{(k,m)}]]\\ &\vdots\\ &[[y_{(1\nu_1,1)},\ldots,y_{(1\nu_m,m)}],\ldots,[y_{(2\nu_1,1)},\ldots,y_{(2\nu_m,m)}],\ldots,[y_{(k\nu_1,1)},\ldots,y_{(k\nu_m,m)}]]\end{aligned}$$

hence, gathering the sub-sum of (11) consisting of all the latter terms as $\nu_1,\ldots,\nu_m$ range in $\mathfrak{S}(k)$, we find an instance of identity (Λ)

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

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

References

[1] S. Adian and A. Razborov. Periodic groups and lie algebras. *Russian Math. Surveys*, 42(2):1–81, 1987.

[2] 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.

[3] Agatha Atkarskaya, Eliyahu Rips, and Katrin Tent. The Burnside problem for odd exponents (arxiv:2303.15997), 2024.

[4] William Burnside. On groups in which every two conjugate operations are permutable. *Proc. Lond. Math. Soc.*, s1-35, 1902.

[5] W. Fulton and J. Harris. *Representation theory: a first course*. Springer-Verlag., 1991.

[6] E. S. Golod and I. R. Shafarevitch. On the class field tower. *Izv. Akad. Nauk SSSR Ser. Mat.*, 28:261–272, 1964.

[7] Daniel Groves. A note on nonidentical Lie relators. *J. Algebra*, 211(1):15–25, 1999.

[8] 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.

[9] P. J. Higgins. Lie rings satisfying the Engel condition. *Proc. Cambridge Philos. Soc.*, 50:8–15, 1954.

[10] G. D. James. *The Representation Theory of the Symmetric Groups*. Springer-Verlag., 1978.

[11] A. I. Kostrikin. Lie rings satisfying the Engel condition. *Izv. Akad. Nauk SSSR Ser. Mat.*, 21:515–540, 1957.

[12] A. I. Kostrikin. On Burnside's problem. *Dokl. Akad. Nauk SSSR*, 119:1081–1084, 1958.

[13] A. I. Kostrikin. The Burnside problem. *Izv. Akad. Nauk SSSR Ser. Mat.*, 23:3–34, 1959.

[14] I. G. Lysënok. Infinite Burnside groups of even period. *Izv. Ross. Akad. Nauk Ser. Mat.*, 60(3):3–224, 1996.

[15] Avinoam Mann and Consuelo Martinez. Groups nearly of prime exponent and nearly engel lie algebras. *Archiv der Mathematik*, 71(1):5–11, 1998.

[16] Jr. Marshall Hall. *The Theory of Groups*. The Macmillan Company. The Macmillan Company, 7th printing 1965 edition, 1965.

[17] P. S. Novikov and S. I. Adjan. Infinite periodic groups. I. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:212–244, 1968.

[18] P. S. Novikov and S. I. Adjan. Infinite periodic groups. II. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:251–524, 1968.

[19] P. S. Novikov and S. I. Adjan. Infinite periodic groups. III. *Izv. Akad. Nauk SSSR Ser. Mat.*, 32:709–731, 1968.

[20] Gunnar Traustason. Engel Lie-algebras. *Quart. J. Math. Oxford Ser. (2)*, 44(175):355–384, 1993.

[21] 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.

[22] 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.

[23] 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.

[24] Michael Vaughan-Lee. 5-Engel lie algebras, 2024.

[25] Michael Vaughan-Lee. On zel'manov's global nilpotence theorem for engel lie algebras, 2025.

[26] E. I. Zelmanov. On Engel Lie algebras. *Sibirsk. Mat. Zh.*, 29(5):112–117, 238, 1988.

[27] 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.

[28] 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/