ON THE STRUCTURE OF LARGE POWERS AND RANDOM GENERATION IN THE NOTTINGHAM GROUP

by

Tuğba Aslan

Submitted to Central European University

In partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics and its Applications

Supervisor: Pál Hegedűs

Secondary advisor: László Pyber

Budapest, Hungary

2016

I, the undersigned [Tuğba Aslan], candidate for the degree of Doctor of Philosophy in Mathematics and its Applications at the Central European University, Mathematics and its Applications, declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of work of others, and no part the thesis infringes on any person's or institution's copyright. I also declare that no part the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Budapest, 28 July 2016

_____

Signature

Canım Babam'a ,,,

ACKNOWLEDGEMENT

First of all, I wish to express my deepest gratitude to my supervisor professor ***Pál Hegedűs*** *for his patience, wisdom and sincere support in several matters. During the past four years, he was not only a supervisor, but also an elder brother.*

I gratefully acknowledge professor ***László Pyber*** *for guiding me throughout the introductory part of the present thesis and for his fruitful comments and suggestions in general.*

I am greatly indebted to professor ***Rachel Camina*** *for hosting me in the University of Cambridge for three months, and for guiding my work in this research visit program.*

My warmest thanks go to professor ***Gülin Ercan*** *for making the first moves in building my knowledge in group theory and for watching me time after time.*

Last, but not least, I shall like to express my deepest love to Dr. ***Mohamed Khaled****, my fiancee, who was my colleague few months ago. He has given me love, companionship and unlimited support ever since I met him. Also, I express my deepest love and respect for my* ***parents****,* ***brother*** *and* ***sister****, and excuse me I would like to write few words for them in Turkish.*

"Sevgili ***Cafer Aslan****,* ***Yeter Aslan****,* ***Tuğrul Aslan*** *ve* ***Tülin Aslan****, geçen dört yıl boyunca sizin varlığınızı, desteğinizi ve sonsuz sevginizi her zaman yanımda hissettim. Siz olmadan bu dört yılı asla bitiremezdim. Tüm kabimle tesekkürler."*

ABSTRACT

The Nottingham group, $\mathcal{N}$, is the group of formal power series over a finite field where the group operation is substitution. In the theory of the Nottingham group, power and commutator structures play a significant role. We confirm a conjecture that was posed by K. Keating in [Kea05] for most of its cases, but we also show that this conjecture is not true for the remaining few cases. In more details, we give the sharp upper bounds for the distance between large p-th powers of any given two elements of $\mathcal{N}$ which share same depth and some leading coefficients.

The key idea of our work is to extend some matrix that was introduced by Keating in [Kea05]. With this extended matrix, we could prove the sharpness of the proposed upper bounds and, moreover, we could show that the depths of large powers of two elements which satisfy Keating's bound grow as slow as possible. Keating's matrix can provide a very useful tool in the study of the Nottingham group. In this connection, we use this matrix to tackle a conjecture that was posed by A. Shalev "Any two random elements of the Nottingham group generate an open subgroup with probability 1". We could confirm this conjecture, but for the elements obeying some extra restrictions. We also argue some possible improvements of our approach that might be helpful for confirming Shalev's conjecture completely.

CONTENTS

1

CHAPTER

Introduction and Basic Concepts

There has been a growing interest to probabilistic questions in group theory during the last 30 years. The question addressing the probability that k random elements, for a given number k, generate the whole group or a dense subgroup (in the case of, say profinite groups) is particularly crucial. In 1882, E. Netto in his book [Net82] conjectured that "The probability, that two randomly chosen elements from the symmetric group S_n generate either the alternating group A_n or S_n, tends to 1 as $n \longrightarrow \infty$". After almost a century, J. D. Dixon proved the conjecture of Netto in [Dix69]. The proof consists of two main steps: finding the probability that two random elements generate a primitive group in S_n and the probability that both elements do not contain a p-cycle for some prime $p < n - 2$. In the same paper, Dixon asked the same question for finite simple groups. In [KL90], W. M. Kantor and A. Lubotzky proved Dixon's conjecture for finite classical simple groups and for several families of exceptional simple groups by assuming classification of finite simple groups. The proof of Dixon's conjecture was finally completed by M. W. Liebeck and A. Shalev in [MA96].

The discrete topology on finite groups gives a natural compact topology on profinite groups due to the fact that a profinite group is an inverse limit of finite groups. Hence, any profinite group can be considered as a probability space with respect to the normalized Haar measure. Let G be a profinite group. Let $\mathcal{Q}(G, k)$ denote the probability that k randomly chosen elements of G topologically generate an open subgroup of G. Parallel to Dixon's work, M. D. Fried and M. Jarden, in their book [FJ86], proved that for a free procyclic group G, $\mathcal{Q}(G, k) = 1$ for some k. They also conjectured that the same holds for free profinite groups of rank greater than 1 and finitely generated free profinite abelian groups. Then Kantor and Lubotzky confirmed Fried and Jarden's conjecture for finitely generated free profinite abelian groups and they proved that

the conjecture is not true for a free profinite group of rank greater than 1, see [KL90].

A profinite group has polynomial subgroup growth if the number of subgroups of index n is at most some given power of n. A. Mann [Man96] showed that $\mathcal{Q}(G,k) = 1$ for some k if G is a profinite group with polynomial subgroup growth. A natural question arises here: Is the converse also true? Mann and Shalev [MS96] proved that this is not the case. Moreover, they proved that $\mathcal{Q}(G,k) = 1$ for some k if and only if G has polynomial maximal subgroup growth, that is for any n, the number of maximal subgroups of index n is at most a power of n.

The property of having polynomial subgroup growth for profinite groups is analogous to being p-adic analytic for pro-p groups. This yields a similar question to the one in the previous paragraph. In [Man96], Mann asked: If $\mathcal{Q}(G,k) = 1$ for some k for a pro-p group G then is G p-adic analytic? Shalev [Sha99] predicted that this is not the case. Shalev's prediction was verified by Y. Barnea and M. Larsen in [BL04]. They proved that if G is a simply connected, semisimple algebraic group over a non-archimedean local field of characteristic p, for a fixed prime number p then $\mathcal{Q}(G,k) = 1$. The first congruence subgroup of $SL_3(\mathbb{F}_p[[t]])$ is not a p-adic analytic pro-p group but it is among those groups which are simply connected, semisimple algebraic over a non-archimedean local field of prime characteristic. Barnea and Larsen, in their proof, used a characterization of compact groups linear over a local field that was given by R. Pink [Pin98].

Another important concept in pro-p group theory is lower rank. The *lower rank* of a pro-p group is the minimal integer r (possibly infinity) such that G has a base for the neighborhoods of identity consisting of r-generated subgroups. Shalev [Sha00] pointed out that if $\mathcal{Q}(G,k) = 1$ for some k then the lower rank of the pro-p group G is at most k. As no counterexample has been constructed yet, it is not known whether the converse is true or not.

In the present thesis, we deal with an important pro-p group called the Nottingham group, denoted by $\mathcal{N}$. It is the group of formal power series over a finite field where the group operation is substitution. Number theorists used to call it the group of wild automorphisms of the local field of prime characteristic. Then, it was introduced as a pro-p group by D. Johnson [Joh88] and his Ph.D. student I. York [Yor90a], [Yor90b]. It gained a keen interest after the result of R. Camina [Cam97], [Cam96]: The Nottingham group, over a field of characteristic

p, contains an isomorphic copy of every finitely generated pro-p group. In other words, the Nottingham group is "S-universal". On the other hand, M. Ershov proved that it is finitely presented [Ers05]. By those two results, the question talking about "the existence of a finitely presented S-universal pro-p group" which was posed in [dSSS00] has got a positive answer. The Nottingham group and some of its subgroups are just infinite, that is, any non-trivial closed normal subgroup is of finite index. Just infinite pro-p groups can be seen as the simple objects of the category of pro-p groups and, so, it brings the question of classification of just infinite pro-p groups, see [KLGP97], [Ers04].

The Nottingham group has lower rank 2 for $p \geq 3$ and it has lower rank at most 3 for $p = 2$, see [Cam00], [Heg01]. We discussed the connection between the concept of lower rank and the random generation in the previous paragraphs. Associated to this connection, Shalev [Sha00] asked the following question: Is $\mathcal{Q}(\mathcal{N}, k) = 1$? In the present thesis, we try to investigate his question. The motivation of our research relies on two papers. The first one is by B. Szegedy [Sze05] which provides a proof for the theorem " two random elements of the Nottingham group generate a free group with probability 1". We are inspired by some probability theorems which he used in his proof. The technical side of our research is inspired by the second paper which is by K. Keating [Kea05]. Keating gave a bound for the distance of p^m-th powers of two elements which are similar at the beginning. He also proved in the same paper that the bound is sharp for $m = 1$ and conjectured that it is sharp as well for $m > 1$. Our research gave us a positive answer for his conjecture except for some cases. In fact, we found the accurate sharp bounds for those exceptional cases.

Chapter 4 of the present thesis is devoted to discuss and to give a partial solution to Shalev's conjecture. The main idea is to change the tail of two random elements and to prove that these new modified elements generate an open subgroup with a positive probability. To test whether two elements generate an open subgroup, we use a simple criterion from [Cam00]: two elements generate an open subgroup if their depths satisfy some conditions.

To confirm Shalev's conjecture, we figured out that it might be useful to construct an Engel word from the modified elements. The n-th Engel word, $[x,_n y]$, is defined recursively as follows: $[x,_1 y] = [x, y]$ and, for $n \geq 2$, $[x,_n y] = [[x,_{n-1} y], y]$. Keating, in [Kea05], constructed a

matrix to compute the depth of the $(p-1)$-st Engel word. We extend this matrix in a way that allows us to determine some leading coefficients of the $s(p-1)$-st Engel word for $s \geq 1$. Then, we show that, given any two randomly chosen elements subjected to some certain restrictions, one can construct other two elements (one of them is an Engel word) from the modified ones such that these new elements satisfy the conditions to generate an open subgroup. This gives a partial solution for Shalev's conjecture.

In Chapter 2 (and in [AH16]), we give our generalization of Keating's matrix then we compute the depth and some first coefficients of the $s(p-1)$-th Engel word for $s \geq 1$. We provide examples, except for the exceptional case, showing that Keating's bound is sharp. Moreover, we prove that the sequences $D(f), D(f^p), D(f^{p^2}), \ldots$ and $D(g), D(g^p), D(g^{p^2}), \ldots$ increase as slow as possible for $f, g \in \mathcal{N}$ which satisfy the Keating's bound. Thus, by our technique, we get more detailed information about the commutator and p-th power structures of the Nottingham group.

In the study of the Nottingham group, it is already a commonplace that the even characteristic is problematic. In Chapter 3 (and in [Asl16]), we show that Keating's bound is not sharp for the case when $p = 2$ and the depths of the given elements are 1. The difficulty of this case comes from the fact that in even characteristic the sequence $D(g), D(g^2), D(g^{2^2}), \ldots$ increases very rapidly for an element $g \in \mathcal{N}$ of depth 1. The commutator and power structures of such an element needed to be investigated deeply. We provide different bounds and examples for the sharpness of the bounds depending on the similarity that exists between the two given elements.

Now, we will briefly recall some basic concepts and elementary results in the theory of pro-p groups and the theory of the Nottingham group.

§1.1 Pro-p groups

This section mainly based on [DdSMS99], [dSSS00].

Definition 1.1.1. Let G be a group which is also a topological space such that the maps

(i) $g \mapsto g^{-1} : G \longrightarrow G$

(ii) $(g, h) \mapsto gh : G \times G \longrightarrow G$

are both continuous. Then G is called a *topological group*.

Definition 1.1.2. A group G is called a *profinite group* if it is a compact Hausdorff topological group whose open subgroups form a base for the neighborhoods of the identity. A profinite group is called a *pro-p group* if every open normal subgroup of it has index equal to some power of a prime number p.

We can redefine profinite and pro-p groups in terms of inverse limit. Here, we recall the concept of inverse limit.

A *directed set* is a non-empty partially ordered set $(\Lambda, \leq)$ such that for every λ, μ there exists $\nu \in \Lambda$ with $\nu \geq \lambda, \mu$. An *inverse system* of sets (or groups) over Λ is a family of sets $(G_\lambda)_{\lambda \in \Lambda}$, with maps (or homomorphisms) $\pi_{\lambda\mu} : G_\lambda \longrightarrow G_\mu$ whenever $\lambda \geq \mu$, satisfying the natural compatibility conditions:

$$\pi_{\lambda\lambda} = id_{G_\lambda}, \quad \pi_{\lambda\mu}\pi_{\mu\nu} = \pi_{\lambda\nu}$$

for all $\lambda \geq \mu \geq \nu$. The *inverse limit*

$$\varprojlim G_\lambda = \varprojlim (G_\lambda)_{\lambda \in \Lambda} = \Big\{ (g_\lambda) \in \prod_{\lambda \in \Lambda} G_\lambda : g_\lambda \pi_{\lambda\mu} = g_\mu \text{ whenever } \lambda \geq \mu \Big\}.$$

Proposition 1.1.3. *A topological group G is a profinite (pro-p) group if and only if G is topologically isomorphic to an inverse limit of finite groups (finite p-groups).*

Haar Measure. Let G be a profinite group. Let G_i be a filtration of G, that is a descending chain of normal subgroups that form a base for the neighborhoods of identity. There is a G-invariant metric on G defined by $d(x, y) = \inf\{|G : G_i|^{-1} : xy^{-1} \in G_i\}$. Note that the balls in G with respect to this metric are the cosets of G_i, and the diameter of such balls is $|G : G_i|^{-1}$. Now one can define a profinite topology on G induced by this metric and so there is a probability measure on the Borel sets of G called normalized *Haar measure*.

Notation. Let G be a profinite group. Let $\mathcal{Q}(G, k)$ denote the probability that k random elements generate an open subgroup.

Lower rank. Let G be a profinite group. The minimal integer r (possibly infinity) is called the lower rank of G if it has a basis of neighborhoods of the identity consisting r-generated subgroups.

§1.2 The Nottingham group

This section is based on [Cam00]. Let R be a commutative ring with identity. Denote by $\mathcal{N}(R)$ the *Nottingham group* over R which is the group of formal power series of the form

$$f = x + \sum_{k=1}^{\infty} \alpha_k x^{k+1} \in R[[x]]$$

under the substitution: given $f, g \in \mathcal{N}(R)$ set $fg(x) = f(g(x))$.

From now on, assume that $R = \mathbb{F}_p$ where $\mathbb{F}_p$ is the finite field of p elements for a prime number p. Also from now on denote $\mathcal{N} = \mathcal{N}(\mathbb{F}_p)$. Consider the following chain of open subgroups of $\mathcal{N}$.

$$\mathcal{N}_k = \Big\{ g \in \mathcal{N} : g = x + \sum_{i=k}^{\infty} \alpha_i x^{i+1}, \alpha_i \in \mathbb{F}_p \Big\}, \text{ for each } k \in \mathbb{N}.$$

For all $k \in \mathbb{N}$ consider the finite groups $(\mathcal{N}/\mathcal{N}_k)$ which contain the polynomials of the form

$$\mathcal{N}/\mathcal{N}_k = \Big\{ f = x + \sum_{i=1}^{k-1} \alpha_i x^{i+1} : \alpha_i \in \mathbb{F}_p \Big\}.$$

It is clear that for each k, $\mathcal{N}/\mathcal{N}_k$ is a finite p-group of order p^{k-1}. Now $(\mathcal{N}/\mathcal{N}_k)_{k=1}^{\infty}$ is an inverse system with the homomorphisms $\pi_{nk} : \mathcal{N}/\mathcal{N}_n \longrightarrow \mathcal{N}/\mathcal{N}_k$, for $n \geq k$, such that

$$\pi_{nk} : x + \sum_{i=1}^{n-1} \alpha_i x^{i+1} \mapsto x + \sum_{i=1}^{k-1} \alpha_i x^{i+1}.$$

It is easy to see that the maps π_{nk} satisfy the required compatibility conditions. Now, $\mathcal{N} \cong \varprojlim(\mathcal{N}/\mathcal{N}_k)$ is a pro-p group and $\mathcal{N}_k$ form a base for the neighborhoods of the identity.

The following definitions are fundamental in the theory of the Nottingham group.

Definition 1.2.1. For $k \geq 1$ and $\alpha \in \mathbb{F}_p$, define $e_k[\alpha] \in \mathcal{N}_k$ where

$$e_k[\alpha] = x + \alpha x^{k+1}.$$

Definition 1.2.2. The depth of $f \in \mathcal{N}\backslash\{1\}$, $D(f)$, is the integer $k \geq 1$ such that $f \in \mathcal{N}_k \backslash \mathcal{N}_{k+1}$, while $D(1) = \infty$.

Notation. Let $f, g \in \mathcal{N}$. Let $[f, g] = f^{-1}g^{-1}fg$ denote the commutator of f and g.

Here are some important facts for the Nottingham group:

Proposition 1.2.3. *Let $f, g \in \mathcal{N}$. Then $D([f, g]) \geq D(f) + D(g)$ and equality holds only if $D(f) \not\equiv D(g) \pmod{p}$.*

Proposition 1.2.4. *Let $f \in \mathcal{N}$ with $D(f) = k$. Then $D(f^p) \geq pk + k_0$ where k_0 is the least non-negative residue of k modulo p. In fact, $D(f^p) \equiv k_0 \pmod{p}$, if $f^p \neq 1$.*

Proposition 1.2.3 and Proposition 1.2.4 imply the following theorems. For the proofs, see [Cam00, Theorem 2, Theorem 6, and Remark 3].

Theorem 1.2.5. *Suppose that $p \neq 2$. Then*

$$[\mathcal{N}_k, \mathcal{N}_n] = \overline{[\mathcal{N}_k, \mathcal{N}_n]} = \begin{cases} \mathcal{N}_{k+n} & \text{if } k \not\equiv n \pmod{p}, \\ \mathcal{N}_{k+n+1} & \text{if } k \equiv n \pmod{p}. \end{cases}$$

Theorem 1.2.6. *For all p, $\mathcal{N}_k^p \leq \mathcal{N}_{kp}$. If p is odd then $\mathcal{N}_k^p = \mathcal{N}_{kp+k_0} = \overline{\mathcal{N}_{kp+k_0}}$, where k_0 is the least non-negative residue of k modulo p.*

Theorem 1.2.7. *$\mathcal{N}$ is a 2-generator pro-p group and $\mathcal{N} = \overline{\langle e_1[\alpha], e_2[\alpha]\rangle}$ where $\alpha \in \mathbb{F}_p$.*

The following results concerning the lower rank of $\mathcal{N}$ are due to R. Camina [Cam00, Theorem 7] and P. Hegedűs [Heg01, Theorem 11].

Theorem 1.2.8. *For $p \geq 3$ the lower rank of $\mathcal{N}$ is 2.*

Theorem 1.2.9. *For $p = 2$ the lower rank of $\mathcal{N}$ is at most 3.*

2

CHAPTER

Maximal deviation of large powers

The distance of f and g is usually defined by $d(g,f) = p^{-D(fg^{-1})}$ which makes $\mathcal{N}$ an ultrametric space. It is clear that $d(f,g) = d(f^j, g^j)$ if $p \nmid j$. So in understanding the distance of powers the crucial case is when the exponent j is a power of p.

The following definition is from [Kea05].

Definition 2.0.1. $n \geq k \geq 1$ and let k_0 be the least nonnegative residue of k modulo p. Let

$$e(k,n) = \begin{cases} 0 & \text{if } p \mid k \text{ and } n = k, \\ 1 & \text{if } p \mid k,\ p \mid n, \text{ and } n > k, \\ 0 & \text{if } p \mid k \text{ and } p \nmid n, \\ i & \text{if } p \nmid k \text{ and } n \equiv 2k - i \pmod{p} \text{ for some } 0 \leq i \leq k_0, \\ k_0 & \text{if } p \nmid k \text{ and } n \not\equiv 2k - i \pmod{p} \text{ for all } 0 \leq i \leq k_0. \end{cases}$$

Keating defined this constant to prove the following theorem [Kea05, Corollary 2] and to show that for $m = 1$ the bound is sharp [Kea05, Theorem 1(b)].

Theorem 2.0.2. *Let p be a prime and $n \geq k \geq 1$. Suppose $f, g \in \mathcal{N}$ are such that $D(f) \geq k$ and $D(gf^{-1}) \geq n$. Then, for all $m \geq 1$, we have*

$$D(g^{p^m} f^{-p^m}) \geq n + (p^m - 1)k + \frac{p^m - p}{p-1} k_0 + e(k,n).$$

Keating conjectured that the bound is sharp for every $m > 1$ as well. In this chapter, we confirm his conjecture except for the case $p = 2$, $k = 1$, when it is not sharp for $m > 2$.

Theorem 2.0.3. *Let p be a prime and $n \geq k \geq 1$. If $p = 2$ then assume $k > 1$. Then there exist*

$f, g \in \mathcal{N}$ such that $D(f) \geq k$, $D(gf^{-1}) \geq n$ and for all $m \geq 1$ we have

$$D(g^{p^m} f^{-p^m}) = n + (p^m - 1)k + \frac{p^m - p}{p-1} k_0 + e(k, n).$$

For $p = 2$, $k = 1$ we have $D(g^{2^m} f^{-2^m}) \geq n + 2^{m+2} - 11 + e(k, n)$ which is strictly stronger than what is implied by Theorem 2.0.2 if $m > 2$. The results need more careful analysis on the power and commutator structure of an element whose depth is 1. Therefore we treated this case separately in Chapter 3.

Theorems 2.0.2 and 2.0.3 immediately imply

Theorem 2.0.4. *Let $d_1 = p^{-n} \leq d_2 = p^{-k}$. Suppose $f, g \in \mathcal{N}$ such that $d(f, g) \leq d_1$ and $d(f, 1) \leq d_2$. Let $j = p^m j'$ be an integer such that $p \nmid j'$. Then $d(f^j, g^j) \leq p^{-(n + (p^m - 1)k + \frac{p^m - p}{p-1} k_0 + e(k,n))}$ and for every choice of d_1, d_2, j (except for $p = 2$, $d_2 = p^{-1}$) there exist f, g as above such that equality holds.*

Also, we have the following Corollary of Theorem 2.0.2 and Theorem 2.0.3;

Corollary 2.0.5. *Let $p \geq 3$ and suppose that Theorem 2.0.3 holds for some $f, g \in \mathcal{N}$ with $D(g) = D(f) = k$ and $D(gf^{-1}) = n \geq k$. Then for all $m \geq 1$*

$$D(f^{p^m}) = p^m k + \frac{p^m - 1}{p-1} k_0 \quad \text{if } n \geq k,$$
$$D(g^{p^m}) = p^m k + \frac{p^m - 1}{p-1} k_0 \quad \text{if } n > k.$$

That is, the p-th powers of f and g have the slowest possible growth.

Proof. First note that by using Proposition 1.2.4 repeatedly we have $D(f^{p^m}) \geq p^m k + \frac{p^m - 1}{p-1} k_0$. Let $m \geq 1$, then by Theorem 2.0.3 we have

$$D(g^{p^m} f^{-p^m}) = n + (p^m - 1)k + \frac{p^m - p}{p-1} k_0 + e(k, n) := n_1$$

where $n_1 \equiv n \pmod p$ if $n \not\equiv 2k - i \pmod p$ for any $0 \leq i \leq k_0$, otherwise $n_1 \equiv k_0 \pmod p$. Therefore, $e(k, n_1) = k_0$. Apply Theorem 2.0.2 to g^{p^m}, f^{p^m}, i.e.,

$$\begin{aligned} D((g^{p^m})^p (f^{p^m})^{-p}) &\geq n_1 + (p-1) D(f^{p^m}) + k_0 \\ &\geq n + (p^m - 1)k + \frac{p^m - p}{p-1} k_0 + e(k, n) + (p-1) D(f^{p^m}) + k_0. \end{aligned}$$

On the other hand, since Theorem 2.0.3 holds for f and g,

$$D((g^{p^m})^p(f^{p^m})^{-p}) = D(g^{p^{m+1}}f^{-p^{m+1}}) = n + (p^{m+1}-1)k + \frac{p^{m+1}-p}{p-1}k_0 + e(k,n).$$

It follows that

$$\begin{aligned}(p-1)D(f^{p^m}) &\leq (p^{m+1}-1)k - (p^m-1)k + \frac{p^{m+1}-p}{p-1}k_0 - \frac{p^m-p}{p-1}k_0 - k_0 \\ \Rightarrow D(f^{p^m}) &\leq p^m k + \frac{p^m-1}{p-1}k_0.\end{aligned}$$

Thus,

$$D(f^{p^m}) = p^m k + \frac{p^m-1}{p-1}k_0.$$

Now suppose that $n > k$ then $n - k + e(k,n) - k_0 > 0$. If there exists $m \geq 1$ such that $D(g^{p^m}) > p^m k + \frac{p^m-1}{p-1}k_0$ then

$$D(g^{-p^m}f^{p^m}) = D(f^{p^m}) = p^m k + \frac{p^m-1}{p-1}k_0.$$

Since Theorem 2.0.3 holds for f, g, we have

$$\begin{aligned}D(g^{-p^m}f^{p^m}) &= n + (p^m-1)k + \frac{p^m-p}{p-1}k_0 + e(k,n) \\ &= n - k + e(k,n) - k_0 + p^m k + \frac{p^m-1}{p-1}k_0 > p^m k + \frac{p^m-1}{p-1}k_0\end{aligned}$$

which gives a contradiction. Therefore, $D(g^{p^m}) = p^m k + \frac{p^m-1}{p-1}k_0$ for all $m \geq 1$. □

§2.1 The Engel word

Let $a, b \in \mathcal{N}$. The m-th Engel word $[a,_m b]$ is defined inductively as follows $[a,_1 b] = [a, b]$, $[a,_m b] = [[a,_{m-1} b], b]$ for $m \geq 2$.

Lemma 2.1.1. *Let $g, h \in \mathcal{N}$ with $D(g) = n$, $D(h) = k$. Write $g = x + \gamma(x)$ and $h = \delta(x)$ where $\deg(\gamma(x)) = n+1$ and $\deg(\delta(x) - x) = k+1$. Then*

$$[g,h] \equiv x - \gamma(x) + \frac{\gamma(\delta(x))}{\delta'(x)} \pmod{x^{2n+k+1}},$$

where $\delta'(x)$ is the formal derivative of $\delta(x)$.

Proof. We know that $hg = \delta(x + \gamma(x))$ and $gh = \delta(x) + \gamma(\delta(x))$. Let us denote $\varepsilon(x) = -\gamma(x) + \frac{\gamma(\delta(x))}{\delta'(x)}$. To prove the claim we have to verify

$$\delta(x + \varepsilon(x) + \gamma(x + \varepsilon(x))) \equiv \delta(x) + \gamma(\delta(x)) \pmod{x^{2n+k+1}}.$$

First, note that

$$\gamma(\delta(x)) \equiv \gamma(x) + \gamma'(x)(\delta(x) - x) \pmod{x^{n+2k+1}},$$

so

$$\varepsilon(x) = \frac{\gamma(\delta(x)) - \gamma(x)\delta'(x)}{\delta'(x)} \equiv \frac{\gamma'(x)(\delta(x) - x) - \gamma(x)(\delta'(x) - 1)}{\delta'(x)} \pmod{x^{n+2k+1}}.$$

Here the numerator has degree at least $n + k + 1$ (with equality if and only if $n \equiv k \pmod{p}$), so $\varepsilon(x)$ has degree at least $n + k + 1$. Consequently, $\gamma(x + \varepsilon(x)) \equiv \gamma(x) \pmod{x^{2n+k+1}}$. The required approximation follows:

$$\begin{aligned} \delta(x + \varepsilon(x) + \gamma(x + \varepsilon(x))) &\equiv \delta(x + \varepsilon(x) + \gamma(x)) = \\ &= \delta(x + \tfrac{\gamma(\delta(x))}{\delta'(x)}) \equiv \delta(x) + \gamma(\delta(x)) \pmod{x^{2n+k+1}}. \end{aligned}$$

□

The next two results will be used in the proof of the main theorem but they are also interesting in their own. The first is a generalisation of the previous lemma. Note that it also holds when $b = 1$, $k = \infty$.

Lemma 2.1.2. *Let $a, b \in \mathcal{N}$ with $D(a) = n$ and $D(b) = k$. Suppose that $a(x) = x + \alpha(x)$ and $b(x) = \beta(x)$. Set $\beta_0(x) = x$, $\beta_1(x) = \beta(x)$, $\beta_i(x) = \beta_{i-1}(\beta(x))$ for $i \geq 1$. Then*

$$[a,_m b] \equiv x + \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i} \frac{\alpha(\beta_i(x))}{\beta_i'(x)} \pmod{x^{2n+mk+1}}. \tag{2.1}$$

Moreover, if $p \nmid k$ and either $m \geq p + 1$ or $m = p$, $p \nmid 2n - k$ then (2.1) *is valid modulo $x^{2n+mk+2}$.*

Proof. The proof is by induction on m. The $m = 1$ case is exactly Lemma 2.1.1. Suppose that we have the result for $m \geq 1$, that is, we can write $[a,_m b] = cd$, where

$$c = x + \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i} \frac{\alpha(\beta_i(x))}{\beta_i'(x)}$$

and $D(d) \geq 2n + mk$. Now $[a,_{m+1} b] = [cd, b] = [c, b][c, b, d][d, b]$, and here $D([c, b, d]) \geq D(c) + D(d) + D(b) > 2n + (m + 1)k$. On the other hand, $D([d, b]) \geq D(b) + D(d)$ with equality if and only if $D(b) \not\equiv D(d) \pmod{p}$, see Proposition 1.2.3. So

$$[a,_{m+1} b] \equiv [c, b] \pmod{x^{2n+(m+1)k+1}}$$

holds unconditionally. But even

$$[a,_{m+1} b] \equiv [c, b] \pmod{x^{2n+(m+1)k+2}}$$

holds if

$$D(d) > 2n + mk, \text{ or}$$
$$k = D(b) \equiv 2n + mk \pmod{p}.$$

If $p \nmid k$ then there is a unique $1 \leq m_0 \leq p$ which is a solution of $k \equiv 2n + m_0 k \pmod{p}$. Thus, $m_0 = p$ if only if $p \mid 2n - k$. For $m = m_0$ the second condition holds, while for $m > m_0$ the first does.

For the commutator $[c, b]$, we apply Lemma 2.1.1. As $D(c) \geq D(a) + mD(b) = n + mk$ by Proposition 1.2.3, the congruence is valid modulo $x^{2n+(2m+1)k+1}$

$$\begin{aligned}
[c, b] &\equiv x - \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i} \frac{\alpha(\beta_i(x))}{\beta_i'(x)} + \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i} \frac{\alpha(\beta_i(\beta(x)))}{\beta_i'(\beta(x))\beta'(x)} \\
&\equiv x + \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i+1} \frac{\alpha(\beta_i(x))}{\beta_i'(x)} + \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i} \frac{\alpha(\beta_{i+1}(x))}{\beta_{i+1}'(x)} \\
&\equiv x + (-1)^{m+1}\alpha(x) + \frac{\alpha(\beta_{m+1}(x))}{\beta_{m+1}'(x)} + \sum_{i=1}^{m} \left(\binom{m}{i} + \binom{m}{i-1} \right) (-1)^{m+i-1} \frac{\alpha(\beta_i(x))}{\beta_i'(x)} \\
&\equiv x + \sum_{i=0}^{m+1} \binom{m+1}{i} (-1)^{m+1+i} \frac{\alpha(\beta_i(x))}{\beta_i'(x)} \pmod{x^{2n+(2m+1)k+1}}
\end{aligned}$$

Thus, we are done by induction. □

Lemma 2.1.3. *Let $a, b \in \mathcal{N}$ with $D(a) = n$ and $D(b) = k$. Then*

$$[a,_m b^{p^l}] \equiv [a,_{mp^l} b] \pmod{x^{2n+mp^l k+1}} \text{ for } m, l \geq 1. \tag{2.2}$$

Moreover, if $p \nmid k$ then (2.2) *is valid modulo $x^{2n+mp^l k+2}$, unless $p > 2$ and $m = l = 1$.*

Proof. We use Lemma 2.1.2 for both sides of the congruence. Note that b^{p^l} is of depth at least

$p^l D(b)$ by Proposition 1.2.4. Note also that $\binom{mp^l}{j} = 0$ if $p^l \nmid j$ and $\binom{mp^l}{ip^l} = \binom{m}{i}$ in $\mathbb{F}_p$, a field of characteristic p. We also clearly have $(-1)^{m+j} = (-1)^{p^l(m+j)}$. So over $\mathbb{F}_p$

$$x + \sum_{i=0}^{m} \binom{m}{i} (-1)^{m+i} \frac{\alpha(\beta_{ip^l}(x))}{\beta'_{ip^l}(x)} = x + \sum_{j=0}^{mp^l} \binom{mp^l}{j} (-1)^{mp^l+j} \frac{\alpha(\beta_j(x))}{\beta'_j(x)}.$$

By Lemma 2.1.2, the first expression is $[a_{,m}\, b^{p^l}]$ modulo $x^{2n+mD(b^{p^l})+1}$, while the second is $[a_{,mp^l}\, b]$ modulo x^{2n+mp^lk+1}, as required.

If $p \nmid k$ then $D(b^{p^l}) > p^l k$, by Proposition 1.2.3, so we get the refinement from Lemma 2.1.2. (Note that for $p = 2$, $k \equiv 2n \pmod p$ implies k even.) □

Now we will introduce the matrix of Keating [Kea05] which is very useful to determine the coefficients of an Engel word. Let $b, u \in \mathcal{N}$ be such that $b = x + x^{k+1}\beta(x)$, where $\beta(x) = r_k + r_{k+1}x + r_{k+2}x^2 + \ldots$, and $D(u) = n \geq k + k_0$.

Assume that $p \nmid k$. Define $u_0 = u$ and $u_{h+1} = [u_h, b]$ for $h \geq 0$. Following [Kea05], we determine the coefficients of u_{h+1} from those of $u_h = x + x^{(h-1)k+n+1}\theta(x)$ where $\theta(x) = t_0 + t_1x + t_2x^2 + \ldots$:

$$[u_h, b] \equiv x + ((h-2)k+n)x^{hk+n+1}\beta(x)\theta(x) + x^{hk+n+2}(\beta(x)\theta'(x) - \beta'(x)\theta(x)) \pmod{x^{(h+1)k+n+1}}. \tag{2.3}$$

It follows from the above expansion that, for $0 \leq j \leq k-1$, the coefficient of $x^{hk+n+j+1}$ in u_{h+1} is

$$\sum_{i=1}^{j} ((h-2)k + n + 2i - j) r_{k+j-i} t_i. \tag{2.4}$$

For $s \leq k$ and $h \geq 1$, consider the following matrix $A_h = A_{h,n,s} \in M_{s\times s}(\mathbb{F}_p)$, where $n_h = (h-2)k + n$,

$$A_h = \begin{pmatrix} n_h r_k & (n_h-1)r_{k+1} & (n_h-2)r_{k+2} & \ldots & (n_h-s+2)r_{k+s-1} \\ 0 & (n_h+1)r_k & n_h r_{k+1} & \ldots & (n_h-s+4)r_{k+s-2} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \ldots & (n_h+s-1)r_k \end{pmatrix}. \tag{2.5}$$

Clearly the matrix depends only on s, b and the remainder of n modulo p.

Let $\vec{\nu}_h \in \mathbb{F}_p^s$ be the row vector whose entries are the coefficients of $x^{(h-1)k+n+j+1}$ in u_h for $0 \leq j \leq s-1$. It follows from (2.4) that $\vec{\nu}_{h+1} = \vec{\nu}_h A_{h,n,s}$. For $h \geq 1$, define a matrix $\Pi_h = \Pi_{h,n,s} \in M_{s\times s}(\mathbb{F}_p)$ by setting $\Pi_h = A_1 A_2 ... A_h$. Then we have $\vec{\nu}_p = \vec{\nu}_1 \Pi_{p-1}$. Keating used these matrices for $s = e+1$ to make these observations, we follow his notations. He went on to show [Kea05, Cases 3,4] that

Lemma 2.1.4. *For $p \nmid k$ and $e = e(k,n)$ we have $\Pi_{p-1,n,e} = 0$.*

By this we instantly get

Corollary 2.1.5. *Let $p \nmid k > k_0$, so $2k_0 < k$ and write $e = e(k,n)$. For a fixed $s < 2k_0 + 1$, denote by T the top right hand corner $s-e \times s-e$ block of $\Pi_{p-1,n,s}$ and let $\vec{\nu}'$ be the first $s-e$ coefficients of $\vec{\nu}_1$. The first e entries of $\vec{\nu}_p$ are 0. The remaining $s-e$ entries are $\vec{\nu}'T$.*

Let $p \nmid k$. Set $n_1 := n + (p-1)k + e(k,n)$. Now if $n \not\equiv 2k - i$, for any $0 \leq i \leq k_0$, then $n_1 \equiv n \pmod p$, otherwise $n_1 \equiv k_0 \pmod p$. Now $D(u_p) \geq n_1$ and $e(k, n_1) = k_0$. Then by Theorem 2.0.2, we have

$$n_2 := D([u_{,p-1}\, b]) = D([u_{,2(p-1)}\, b]) \geq n_1 + (p-1)k + k_0 = n + 2(p-1)k + e(k,n) + k_0.$$

Now $n_2 \equiv n \pmod p$ or $n \equiv k_0 \pmod p$ again and thus $e(k, n_2) = k_0$. Hence, following this way inductively we have the following lemma

Lemma 2.1.6. *Let $u, b \in \mathcal{N}$ such that $D(u) \geq n \geq D(b) = k$ and $k \equiv k_0 \pmod p$, where $0 < k_0 \leq p-1$. Then*

$$D([u_{,s(p-1)}\, b] \geq n + s(p-1)k + e(k,n) + (s-1)k_0$$

for $s \geq 1$.

§2.2 Maximal deviation of large powers

The proof of [Kea05, Lemma 4] can be mimicked to prove the following stronger statement.

Lemma 2.2.1. *Suppose that $n' > n \geq k$ are such that Theorem* 2.0.2 *holds for (k, n), and Theorem* 2.0.3 *holds for (k, n'). If*

$$n + e(k, n) = n' + e(k, n')$$

then Theorem 2.0.3 *also holds for (k, n).*

By the definition of $e(k, n)$ the sum $n + e(k, n)$ is constant in the intervals $k + pt \leq n \leq k + tp + k_0$. The lemma allows us to assume that $k + k_0 + tp \leq n < k + (t + 1)p$. If $n = k + k_0 + pt$ then $e(k, n) = 0$, otherwise, $e(k, n) = k_0$. We are going to assume this in the proof of Theorem 2.0.3, below.

Let $p \nmid k$ and $k + k_0 + pt \leq n \leq k + (t + 1)p$ for some nonnegative integer t. We define

$$e'(k, n) = \begin{cases} 0 & \text{if } k + k_0 + pt < n < k + p(t + 1), \\ k_0 & \text{if } n = k + k_0 + pt \text{ or } n = k + p(t + 1). \end{cases}$$

This describes the number of interesting digits for the following definition. For $u \in \mathcal{N}_n$, let $\overrightarrow{\mu_n(u)} \in \mathbb{F}_p^{e'(k,n)+1}$ denote the vector consisting of the first $e'(k, n) + 1$ coefficients of u.

For given n, k as above and $\lambda \in \mathbb{F}_p$, we define an $e'(k, n) + 1 \times e'(k, n) + 1$ matrix $T_n[\lambda]$ as

$$\begin{pmatrix} -\lambda & 0 & \dots & 0 & 2\lambda^2 \\ 0 & \lambda & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 & \dots \\ 0 & 0 & \dots & \lambda & 0 \\ -1 & 0 & \dots & 0 & 2\lambda \end{pmatrix} \text{ if } n \equiv k \pmod p, \quad \begin{pmatrix} -1 & 0 & \cdots & 0 & 2\lambda \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \text{ if } n \equiv 2k \pmod p,$$

and $T_n[\lambda] = (\lambda)$ if $k + k_0 + pt < n < k + p(t + 1)$.

Lemma 2.2.2. *Let $k > p$, $p \nmid k$ and $k + k_0 + pt \leq n \leq k + (t + 1)p$ for some nonnegative integer t. Define $b = x + x^k + x^{k+k_0}$. For every $u \in \mathcal{N}_n$, we have $[u,_{p-1} b] \in \mathcal{N}_{n+(p-1)k+e(k,n)}$ and $\overrightarrow{\mu_{n+(p-1)k+e(k,n)}([u,_{p-1} b])} = \overrightarrow{\mu_n(u)} T_n[-1]$.*

Proof. By Corollary 2.1.5, we have to show that the top right hand $e'(k, n) + 1 \times e'(k, n) + 1$ block of $\Pi_{p-1,n,e(k,n)+e'(k,n)+1}$ is $T_n[-1]$.

By (2.5) and the choice of our b, we see that in the matrices A_h the entries (i, j) are 0 unless

$j - i \in \{0, k_0\}$. Consequently, in the products $\Pi_h = A_1 \cdots A_h$ the entries (i, j) are 0 unless $j - i$ is a nonnegative integer multiple of k_0.

First, we deal with the cases $n \equiv 2k \pmod{p}$ and $m \equiv k \pmod{p}$. Note $e(k, n) = 0$, $e(k, m) = k_0$. By (2.5), we also see the following block decomposition of the matrices $A_{h,m,2k_0+1}, \Pi_{h,m,2k_0+1} \in M_{(2k_0+1)\times(2k_0+1)}(\mathbb{F}_p)$ for $h \geq 1$:

$$A_{h,m,2k_0+1} = \left(\begin{array}{c|c} A_{h,m,k_0} & * \\ \hline 0 & A_{h,n,k_0+1} \end{array}\right), \quad \Pi_{h,m,2k_0+1} = \left(\begin{array}{c|c} \Pi_{h,m,k_0} & * \\ \hline 0 & \Pi_{h,n,k_0+1} \end{array}\right). \quad (2.6)$$

So, it is enough to compute the larger matrices $A_{h,m,2k_0+1} \in M_{(2k_0+1)\times(2k_0+1)}(\mathbb{F}_p)$, and $\Pi_{p-1,m,2k_0+1} \in M_{(2k_0+1)\times(2k_0+1)}(\mathbb{F}_p)$. The (i, j) entry of A_h for $0 \leq i \leq j \leq 2k_0$ is

$$a_{hij} = \begin{cases} m_h + i & \text{if } i = j, \\ m_h - k_0 + i & \text{if } j = k_0 + i, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $m_h + k_0 = (h-2)k + m + k_0 \equiv m_{h+1} \pmod{p}$. Thus A_h has the form

$$\begin{array}{c} \\ 0 \\ \\ \\ \\ k_0 \\ \\ \\ \\ 2k_0 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 0 & & & & k_0 & & & & 2k_0 \end{array} \\ \left(\begin{array}{ccccccccc} m_h & 0 & \dots & 0 & m_{h-1} & 0 & \dots & 0 & 0 \\ 0 & m_h+1 & 0 & \dots & 0 & m_{h-1}+1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & 0 & \dots & 0 & \ddots & 0 & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \dots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & m_{h+1} & 0 & \dots & 0 & m_h \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & m_{h+2} \end{array}\right) \end{array}.$$

By the above observation, the (i, j) entry of Π_h is zero if $j - i \notin \{0, k_0, 2k_0\}$. By induction on $h \geq 1$, we get that the $(i, i + k_0)$ entry of Π_h is

$$\left(\sum_{t=1}^{h-1}(m_1+i)(m_2+i)...(m_t+i)^2(m_{t+3}+i)(m_{t+4}+i)...(m_{h+1}+i)\right) \\ +(m_0+i)(m_3+i)...(m_{h+1}+i). \tag{2.7}$$

Finally, for $h = p-1$ we get that the $(i, i+k_0)$ entry of Π_{p-1} is

$$\pi_{i,i+k_0} = \left(\sum_{t=1}^{p-2}(m_1+i)(m_2+i)...(m_t+i)^2(m_{t+3}+i)(m_{t+4}+i)...(m_p+i)\right) \\ +(m_0+i)(m_3+i)...(m_p+i) = \sum_{t=0}^{p-2}(m_t+i)\prod_{j=3}^{p}(m_{t+j}+i). \tag{2.8}$$

Note that each summand in (2.8) is a product of $p-2$ consecutive numbers from $\{m_t+i\}_{t=1}^{p}$, the last with multiplicity 2. Pick the unique $l \in \{1, ..., p\}$ such that $m_l+i = 0$. So, $\prod_{j\neq l}(m_j+i) = -1$ and, consequently, it follows that

$$(m_{l-1}+i)\prod_{j=3}^{p}(m_{l-1+j}+i) = -\frac{-k_0}{k_0} = 1, \qquad (m_{l-2}+i)\prod_{j=3}^{p}(m_{l-2+j}+i) = -\frac{-2k_0}{-k_0} = -2.$$

$$\pi_{i,i+k_0} = \begin{cases} -2+1=-1 & \text{if } l \neq 1, p; \\ -2 & \text{if } l = p; \\ 1 & \text{if } l = 1. \end{cases} \tag{2.9}$$

By Lemma 2.1.4, we know that the first k_0 columns and the last k_0 rows of Π_{p-1} are zero. The remaining diagonal entry is

$$\pi_{k_0,k_0} = m_2m_3m_4...m_p = (k_0)(2k_0)(3k_0)...((p-1)k_0) = -1.$$

To compute $\pi_{0,2k_0}$, note that $a_{p-1,2k_0,2k_0} = m_{p+1} = 0$ and $a_{p-1,k_0,2k_0} = m_{p-1}$ in A_{p-1}. So, $\pi_{0,2k_0}$ is m_{p-1} times the $(0, k_0)$ entry of Π_{p-2}. Using (2.7), we get

$$\pi_{0,2k_0} = m_0m_3m_4...m_{p-1}m_{p-1} = (-k_0)(2k_0)(3k_0)...((p-2)k_0)^2 = 2.$$

Finally, by (2.9) the top right hand corner $k_0+1 \times k_0+1$ matrix of $\Pi_{p-1,m,2k_0+1}$ corresponding to $m \equiv k \pmod p$, is indeed

$$T_m[-1] = \begin{pmatrix} 1 & 0 & \dots & 0 & 2 \\ 0 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 & \dots \\ 0 & 0 & \dots & -1 & 0 \\ -1 & 0 & \dots & 0 & -2 \end{pmatrix}. \tag{2.10}$$

From (2.6) and (2.10), the matrix $\Pi_{p-1,n,k_0+1} \in M_{(k_0+1)\times(k_0+1)}(\mathbb{F}_p)$ corresponding to $n \equiv 2k \pmod p$, is indeed

$$T_n[-1] = \begin{pmatrix} -1 & 0 & \cdots & 0 & -2 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Now, assume that $k + k_0 + pt < n < k + p(t+1)$, i.e., $n \not\equiv 2k - i \pmod p$ for any $0 \le i \le k_0$, thus $e(k,n) = k_0$. Then the corresponding matrix A_{h,n,k_0+1} is

$$\begin{array}{c} \\ 0 \\ \\ \\ \\ k_0 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & & & & k_0 \end{array} \\ \begin{pmatrix} n_h & 0 & \dots & 0 & n_{h-1} \\ 0 & n_h+1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & n_{h+1} \end{pmatrix} \end{array}.$$

As above, we determine the (i,j) entries of $\Pi_{p-1} = \Pi_{p-1,n,k_0+1}$.

$$\pi_{0,k_0} = \left(\sum_{t=1}^{p-2} n_1 n_2 ... n_t^2 n_{t+3} n_{t+4} ... n_p \right) + n_0 n_3 ... n_p = \sum_{t=0}^{p-2} n_t \prod_{j=3}^{p} n_{t+j}.$$

Note that we get $n_1 = n - k \neq 0$, $n_p = (p-2)k + n \neq 0$ by our assumptions, so $\pi_{0,k_0} = -1$ as in (2.9). Again, for the diagonal entries we have

$$\pi_{i,i} = \prod_{t=1}^{p-1} (n_t + i) \in \{0, -1\},$$

with $\pi_{i,i} = -1$ if and only if $0 \equiv n_0 + i = n - 2k + i \pmod p$. In our case $n \not\equiv 2k - i \pmod p$ for any $0 \le i \le k_0$, so $\pi_{i,i} = 0$. We cut the first k_0 columns and last k_0 rows and we indeed get

the matrix

$$T_n[-1] = (-1).$$

□

Lemma 2.2.3. *Let $p > 2$. Suppose $k = k_0 < p$ and $k + k_0 + pt \le n \le k + (t+1)p$ for some nonnegative integer t. If $k < p-1$ then let $b = x + x^{k+1}$, otherwise, let $b = x + x^{k+1} + x^{2k+1}$. For every $u \in \mathcal{N}_n$ we have $[u,_{p-1} b] \in \mathcal{N}_{n+(p-1)k+e(k,n)}$ and $\overrightarrow{\mu_{n+(p-1)k+e(k,n)}([u,_{p-1} b])} = \overrightarrow{\mu_n(u)} T_n[\lambda]$, where $\lambda = \frac{k-1}{2}$ if $k = p-1$ and $\lambda = \frac{k+1}{2}$ otherwise.*

Proof. The proof follows from the argument of Lemma 2.2.2. However, as k is small we have to be concerned with more terms than in (2.3) and in (2.4). We extend the dimension of the matrices to $e(k,n) + e'(k,n) + 1 \le 2k+1$. Then we can use Lemma 2.1.4 and it remains to show that the top right hand $e'(k,n)+1 \times e'(k,n)+1$ block of $\Pi_{p-1,n,e(k,n)+e'(k,n)+1}$ is $T_n[\lambda]$.

Here is the refinement of (2.3) (using $n \ge 2k$) for u_{h+1},

$$\begin{aligned}
[u_h, b] \;\equiv\;& x + ((h-2)k+n)\beta(x)\theta(x)x^{hk+n+1} + (\beta(x)\theta'(x) - \beta'(x)\theta(x))x^{hk+n+2} + \\
+\;& \left(\tbinom{(h-1)k+n+1}{2} - (k+1)((h-2)k+n)\right)\beta(x)^2\theta(x)x^{(h+1)k+n+1} + \\
+\;& (((h-2)k+n)\beta(x)^2\theta'(x) + ((3-h)k-n+1)\beta(x)\beta'(x)\theta(x))\,x^{(h+1)k+n+2} + \\
+\;& \left(\tfrac{1}{2}\beta(x)^2\theta''(x) + \beta'(x)^2\theta(x) - \beta(x)\beta'(x)\theta'(x)\right)x^{(h+1)k+n+3} + E_{h,n} \pmod{x^{(h+2)k+n+2}}.
\end{aligned} \tag{2.11}$$

Here, for some $q_{h,n}, q'_{h,n} \in \mathbb{F}_p$,

$$E_{h,n} = \begin{cases} q_{h,n}x^{(h+2)k+n+1}\beta(x)^3\theta(x) & \text{if } (h-1)k+n > 2k, \\ q_{h,n}x^{(h+2)k+n+1}\beta(x)^3\theta(x) + q'_{h,n}\beta(x)\theta(x)^2 & \text{if } (h-1)k+n = 2k. \end{cases}$$

The coefficient of $x^{hk+n+j+1}$ in u_{h+1} for $0 \le j < k$ is the same as in (2.4), and for $k \le j < 2k$ it is

$$\begin{aligned}
\sum_{i=1}^{k+a}((h-2)k+n+2i-j)r_{k+j-i}t_i + \left(\binom{n_h+a}{2} + \binom{k+1}{2}\right)r_k^2 t_a + \sum_{i=0}^{1} c_{1i}r_{k+i}r_{k+1-i})t_{a-1} \\
+(\sum_{i=0}^{2} c_{2i}r_{k+i}r_{k+2-i})t_{a-2} + \cdots + (\sum_{i=0}^{a} c_{ai}r_{k+i}r_{k+a-i})t_0
\end{aligned} \tag{2.12}$$

where $a = j-k$ and $c_{sw} \in \mathbb{F}_p$ for $1 \le s \le a$, $0 \le w \le a$, and $w \le s$. For $j = 2k$, there will

be an extra term in the coefficient of t_0 and the rest will satisfy the above formula. But its value is irrelevant for our proof.

We get again that the (i,j) entry of Π_h is 0 unless $j-i \in \{0,k,2k\}$ as in Lemma 2.2.2.

Again, we assume first that $n \equiv 2k \pmod{p}$ and $m \equiv k \pmod{p}$ and hence $e(k,n) = 0$, $e(k,m) = k$. As a subcase, assume that $k < p-1$ so $b = x + x^{k+1}$. Now from (2.12) we compute the entries of the matrix $A_{h,m,2k+1}$

$$a_{hij} = \begin{cases} m_h + i & \text{if } i = j, \\ \binom{m_h+i}{2} + \binom{k+1}{2} & \text{if } j = i+k, \\ C_{h,m} & \text{if } (i,j) = (0,2k), \\ 0 & \text{otherwise.} \end{cases}$$

The value of $C_{h,m} \in \mathbb{F}_p$ is going to be irrelevant. By induction on $h \geq 1$, we get that the $(i, i+k)$ entry of Π_h is

$$\left(\sum_{t=1}^{h} (m_1+i)(m_2+i)\cdots(m_t+i)(m_{t+2}+i)(m_{t+3}+i)\cdots(m_{h+1}+i)\frac{m_t+i-1}{2} \right) + \binom{k+1}{2} \sum_{t=2}^{h+1} (m_1+i)(m_2+i)\cdots(m_{t-2}+i)(m_{t+1}+i)\cdots(m_{h+1}+i) \tag{2.13}$$

Consider $h = p-1$. For $i = 0$, the first sum is 0, because $m_1 = 0$ is a factor in each product. For $i > 0$, pick the unique $l \in \{2, \ldots, p\}$ such that $m_l + i = 0$ and observe that the first sum has a single nonzero term, that for $t = l-1$. Its value is $\frac{k+1}{2}$. For $i = 0, k$, the second sum has a single nonzero summand, two otherwise. For $i = 0$ the value is $\frac{-k-1}{2}$, for $i = k$ it is $\frac{k+1}{2}$ and for $i \neq 0, k$ the two terms add up to 0. So, finally

$$\pi_{i,i+k} = \begin{cases} \frac{-k-1}{2}, \text{ if i=0;} \\ k+1, \text{ if i=k;} \\ \frac{k+1}{2}, \text{ otherwise.} \end{cases} \tag{2.14}$$

Note that Π_{p-2} has $(0,0)$ entry 0 and $(0,k)$ entry $\frac{k+1}{2k}$ and A_{p-1} has $(k,2k)$ entry $k(k+1)$ and

$(2k, 2k)$ entry 0. So, $\pi_{0,2k} = \frac{(k+1)^2}{2}$. For the diagonal entries we again have

$$\pi_{i,i} = \prod_{h=1}^{p-1}(m_h + i) \in \{0, -1\},$$

with $\pi_{i,i} = -1$ if and only if $0 \equiv m_0 + i = m - 2k + i \equiv -k + i \pmod{p}$. Put $\lambda = \frac{k+1}{2} \neq 0$ as $k < p - 1$. Now we have the claimed matrix

$$T_m[\lambda] = \begin{pmatrix} -\lambda & 0 & \dots & 0 & 2\lambda^2 \\ 0 & \lambda & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 & \dots \\ 0 & 0 & \dots & \lambda & 0 \\ -1 & 0 & \dots & 0 & 2\lambda \end{pmatrix}.$$

As (2.6) is still valid, we also have

$$T_n[\lambda] = \begin{pmatrix} -1 & 0 & \cdots & 0 & 2\lambda \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Now suppose that $k = p - 1$ so take $b = x + x^p + x^{2p-1}$. We continue to assume that $n \equiv 2k \pmod{p}$ and $m \equiv k \pmod{p}$, hence $e(k, n) = 0$, $e(k, m) = k$. By using (2.11), (2.12), we have the corresponding matrix $A_{h,m,2k+1}$ whose (i, j) entries for $0 \leq i, j \leq 2k$ are

$$a_{hij} = \begin{cases} m_h + i & \text{if } i = j, \\ m_h + i - k + \binom{m_h+i}{2} + \binom{k+1}{2} & \text{if } j = k + i, \\ D_{h,m} & \text{if } (i, j) = (0, 2k) \\ 0 & \text{otherwise.} \end{cases}$$

The value of $D_{h,m} \in \mathbb{F}_p$ is uninteresting, as before. Again the (i, j) entry of $\Pi_{h,m,2k+1}$ is zero if $j - i \notin \{0, k, 2k\}$. By the value of $a_{h,i,i+k}$, we get that $(i, i + k)$ entry of Π_h is the sum of

(2.13) and (2.7). Therefore,

$$\pi_{i,i+k} = \begin{cases} \frac{-k-1}{2} + 1 = -\frac{k-1}{2}, \text{ if } i = 0; \\ k + 1 - 2 = k - 1, \text{ if } i = k; \\ \frac{k+1}{2} - 1 = \frac{k-1}{2}, \text{ otherwise.} \end{cases}$$

Note that Π_{p-2} has $(0,0)$ entry 0 and $(0,k)$ entry $\frac{k+1}{2k} + \frac{-1}{k} = \frac{k-1}{2k}$. On the other hand, A_{p-1} has $(k,2k)$ entry $-2k + k(k+1) = k(k-1)$ and $(2k,2k)$ entry 0. Thus, $\pi_{0,2k} = \frac{(k-1)^2}{2}$. For the diagonal entries, we again have

$$\pi_{i,i} = \prod_{h=1}^{p-1}(m_h + i) \in \{0, -1\},$$

with $\pi_{i,i} = -1$ if and only if $0 \equiv m_0 + i = m - 2k + i \equiv -k + i \pmod{p}$. Put $\lambda = \frac{k-1}{2}$. (We use the assumption $p > 2$ to claim $\lambda \neq 0$). We obtained the claimed matrix

$$T_m[\lambda] = \begin{pmatrix} -\lambda & 0 & \dots & 0 & 2\lambda^2 \\ 0 & \lambda & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 & \dots \\ 0 & 0 & \dots & \lambda & 0 \\ -1 & 0 & \dots & 0 & 2\lambda \end{pmatrix}.$$

We get $T_n[\lambda]$ in a similar way exactly as above.

Assume finally that $k + k_0 + pt < n < k + p(t+1)$, that is, $n \not\equiv 2k - i \pmod{p}$ for any $0 \leq i \leq k_0$, then $e(k,n) = k_0$. Of course $k < p - 1$. So we take the element $b = x + x^{k+1}$ and obtain the corresponding matrix $T_n[\lambda] = (\frac{k+1}{2})$ (from (2.14)) exactly as in Lemma 2.2.2. $\square$

Proof of Theorem 2.0.3. Suppose that $p \nmid k$. By the remark before Lemma 2.2.1 we may assume that $k + k_0 + pt \leq n < k + (t+1)p$, for some integer $t \geq 0$. If $n = k + k_0 + pt$ then $e(k,n) = 0$, otherwise, $e(k,n) = k_0$.

Let $g^{-1} = x + x^{k+1} + x^{k+k_0+1}$, or $x + x^{k+1}$ as in Lemma 2.2.2 and Lemma 2.2.3. We pick $u = u_0 \in \mathcal{N}_n$, but its leading coefficient, α, will be chosen appropriately later. Set $f^{-1} = g^{-1}u$ and then $u_m = g^{p^m} f^{-p^m}$. Let $n_m = n + (p^m - 1)k + \frac{p^m - p}{p-1}k_0 + e(k,n)$.

We prove that $D(u_m) = n_m$ by induction on m. We assume $m \geq 1$ and that for every

smaller index the theorem holds. Note that $n_m = n_{m-1} + (p-1)p^{m-1}k + p^{m-1}k_0$.

By [LGM02, Proposition 1.1.32(i)], we have

$$u_m \equiv u_{m-1}^p [u_{m-1}, g^{-p^{m-1}}]^{\binom{p}{2}} [u_{m-1,2}\, g^{-p^{m-1}}]^{\binom{p}{3}} \dots [u_{m-1,p-1}\, g^{-p^{m-1}}] \bmod K(g^{-p^{m-1}}, u_{m-1})$$

where $K(g^{-p^{m-1}}, u_{m-1})$ is the normal closure in $\mathcal{N}$ of the set of all formal basic commutators in $\{g^{-p^{m-1}}, u_{m-1}\}$ of weight at least p and of weight at least 2 in u_{m-1} and also the p-th powers of all basic commutators of weight at least 3 and of weight at least 2 in u_{m-1}. The only weight p commutator with weight 2 in u_{m-1} is $w = [[u_{m-1,p-2}\, g^{-p^{m-1}}], u_{m-1}]$. We denote its multiplicity by δ_0. (The exact value $\delta_0 = p^2 - p - 1$ is irrelevant for our proof.)

If $m = 1$ and $n_{m-1} = n = k + k_0$ then $e(k,n) = 0$ and $e'(k,n) = k_0$ so $D([u_{,p-1}\, g^{-1}]) = n + (p-1)k = n + (p-1)k + e(k,n)$. The element w in $K(g^{-1}, u)$ has depth exactly $2n + (p-2)k = n + (p-1)k + e(k,n) + e'(k,n)$, every other element in $K(g^{-1}, u)$ has larger depth. The leading coefficient of w is

$$(n-k)\alpha n(n+k)\cdots(n+(p-4)k)\alpha(p-2)k = -2\alpha^2,$$

where α is the leading coefficient of u. If $n_{m-1} > k + k_0$ then every element in $K(g^{-p^{m-1}}, u)$ has depth at least $2n_{m-1} + (p-2)k > n_{m-1} + (p-1)k + e(k, n_{m-1}) + e'(k, n_{m-1})$.

Also note that $D([u_{m-1,i}\, g^{-p^{m-1}}]^{\binom{p}{i+1}}) \geq p(n_{m-1} + k) > n_{m-1} + (p-1)k + e(k, n_{m-1}) + e'(k, n_{m-1})$ for $1 \leq i \leq p-2$ for every $n_{m-1} \geq k + k_0$. Let $\overline{n_{m-1}}$ be the remainder of n_{m-1} modulo p. Then $D(u_{m-1}^p) \geq pn_{m-1} + \overline{n_{m-1}} = n_{m-1} + (p-1)k + k_0 + (p-2)k_0 + (p-1)(n_{m-1} - k - k_0) + \overline{n_{m-1}} > n_{m-1} + (p-1)k + e(k, n_{m-1}) + e'(k, n_{m-1})$.

We conclude that

$$\overrightarrow{\mu_{n_m}(u_m)} = \begin{cases} \overrightarrow{\mu_{n_m}([u_{m-1,p-1}\, g^{-p^{m-1}}])} & \text{if } n_{m-1} > k + k_0, \\ \overrightarrow{\mu_{n_1}([u_{,p-1}\, g^{-1}])} + (0, \dots, 0, -2\delta_0\alpha^2) & \text{if } m = 1,\ n = k + k_0. \end{cases} \tag{2.15}$$

Suppose that $m > 1$. By the last line of Lemma 2.1.3 (which holds in our case, as $p - 1 = 1$ would imply $p = 2$),

$$[u_{m-1,p-1}\, g^{-p^{m-1}}] \equiv [u_{m-1,(p-1)p^{m-1}}\, g^{-1}] \pmod{x^{2n_{m-1} + (p-1)p^{m-1}k + 2}}.$$

Now

$$2n_{m-1} + (p-1)p^{m-1}k + 2 = n_m + n_{m-1} - p^{m-1}k_0 + 2 =$$

$$n_m + n + (p^{m-1}-1)k + (-p^{m-1} + \frac{p^{m-1}-p}{p-1})k_0 + e(k,n) + 2 > n_m + e'(k,n_m) + 1.$$

So, we obtain $\overrightarrow{\mu_{n_m}(u_m)} = \overrightarrow{\mu_{n_m}([u_{m-1,(p-1)p^{m-1}}\, g^{-1}])}$ for $m > 1$. By induction and by (2.15), we have

$$\overrightarrow{\mu_{n_m}(u_m)} = \begin{cases} \overrightarrow{\mu_{n_m}([u_{1,p^m-p}\, g^{-1}])} & \text{for } m > 1, \\ \overrightarrow{\mu_{n_m}([u_{,p^m-1}\, g^{-1}])} & \text{for } m \geq 1 \text{ if } n > k + k_0. \end{cases} \tag{2.16}$$

After these preliminary observations we turn to the proof.

First, let $k + k_0 + pt < n < k + (t+1)p$. Then $e'(k,n) = 0$ and $e(k,n) = k_0$. Lemma 2.2.2 or Lemma 2.2.3 shows that for any $u \in \mathcal{N}_n$ of depth n we have $D([u_{,s(p-1)}\, g^{-1}]) = n + s(p-1)k + sk_0$ for every $s \geq 1$. Thus, (2.16) implies that

$$D(u_m) = D([u_{m-1,p-1}\, g^{-p^{m-1}}]) = D([u_{,p^m-1}\, g^{-1}]) = n_m,$$

as required.

Now suppose $n = k + k_0 + tp$. Then $e'(k,n) = k_0$, $e(k,n) = 0$. Recall that $\alpha \neq 0$ denotes the leading coefficient of $u \in \mathcal{N}_n$. If $k > p$ then $\lambda = -1$, if $k = p-1$ then $\lambda = \frac{k-1}{2}$ and if $k < p-1$ then $\lambda = \frac{k+1}{2}$. Then Lemma 2.2.2 or Lemma 2.2.3 shows that for the first k_0 coefficients of $v = [u_{,p-1}\, b]$ we have

$$\overrightarrow{\mu_{n_1}(v)} = (-\alpha, 0, \ldots, 0, 2\lambda\alpha),$$

where $\lambda \in \mathbb{F}_p$ nonzero. By (2.15),

$$\overrightarrow{\mu_{n_1}(g^p f^{-p})} = (-\alpha, \ldots, 0, 2\lambda\alpha - 2\delta\alpha^2),$$

where $\delta = \delta_0$ if $n = k + k_0$ and $\delta = 0$ otherwise. The leading coefficient is $-\alpha \neq 0$, so $D(u_1) = D(g^p f^{-p}) = n + (p-1)k + e(k,n) = n_1$ holds. Now $n_1 \equiv k \pmod p$, so $e'(k,n_1) = e(k,n_1) = k_0$.

We apply Lemma 2.2.2 or Lemma 2.2.3 again to get

$$D(u_2) = D([u_{1,p-1}\, g^{-p}]) = D([u_{1,(p-1)p}\, g^{-1}]) = n_2.$$

We also obtain

$$\overrightarrow{\mu_{n_2}(g^{p^2} f^{-p^2})} = (-\lambda\alpha + 2\delta\alpha^2, 0, \ldots, 0, 2\lambda^2\alpha - 4\lambda\delta\alpha^2). \tag{2.17}$$

The leading coefficient is nonzero if $\alpha \neq 0$ and $2\alpha\delta \neq \lambda$. Such an α exists because $\lambda \neq 0$ and if $\delta \neq 0$ then $p > 2$. Fix this α. The vector at (2.17) is an eigenvector of $T_{n_m}[\lambda]$ with eigenvalue $\lambda \neq 0$ for all $m \geq 1$.

Now we can use Lemma 2.2.2 or Lemma 2.2.3 and (2.16) to obtain that

$$D(u_m) = D([u_{m-1,p-1}\, g^{-p^{m-1}}]) = D([u_{1,p^m-1}\, g^{-1}] = n_m,$$

as required.

Now it remains to consider the case of $p \mid k$. Our arguments are similar to those of [Kea05]. Suppose that $p \nmid n$. Let g, u be arbitrary and set $f^{-1} := g^{-1}u$. Then $D(g^p f^{-p}) = D([u_{,p-1}\, g^{-1}]) = n + (p-1)k \equiv n \pmod{p}$ since $2n + (p-2)k > n + (p-1)k$. We can proceed with induction to get $D(g^{p^m} f^{-p^m}) = D([u_{,p^{m-1}(p-1)}\, g^{-1}]) = n + (p^m - 1)k$ for $m \geq 1$. Suppose finally that $p \mid n$. If $n = k$, let g be arbitrary of depth k and $f = g^p$, then $g^{p^m} f^{p^{-m}} = g^{p^m(1-p)}$ so $D(g^{p^m} f^{p^{-m}}) = p^m k$ by Proposition 1.2.4. If $n > k$ then pick g and f to work for the pair $k, n+1$. Then $e(k,n) = 1 = e(k,n+1) + 1$. Therefore, the required follows by Lemma 2.2.1. □

3

CHAPTER

Characteristic 2

In the previous chapter we showed that the depth of $g^{p^m} f^{-p^m}$ can be determined by finding the depth of $[u,_{s(p-1)}\, g^{-1}]$ where $s = \frac{p^m-1}{p-1}$. This argument fails for $p = 2$ and $k = 1$. The main reason behind it is that if $D(g^{-1}) = 1$ then the sequence $D(g^{-1}), D(g^{-2}), D(g^{-2^2}), \ldots$ increases more rapidly. However, when $p = 2$ the 2^m-th powers of an element can be computed more easily than when $p \geq 3$. Therefore, the matrix of Keating can be used to compute $[u, g^{-2^m}]$ for $m \geq 1$, and the problem can be reduced to determine $D([u, g^{-1}, g^{-2}, g^{-2^2}, g^{-2^3}, \ldots, g^{2^{m-1}}])$. It turns out that the bound is almost the double of Keating's bound for $p = 2, k = 1$. Recall that by Definition 2.0.1, $e(1, n) = 1$ if n is odd, otherwise, zero . Denote $D_m := D(g^{2^m} f^{-2^m})$, our main result is the following

Theorem 3.0.1. *Let $p = 2$, and $n \geq 2$. Suppose $f, g \in \mathcal{N}$ such that $D(f) = D(g) = 1$ and $D(gf^{-1}) \geq n$. The table below describes the lower bounds on D_m for the various values of n, m. Furthermore, the bounds are sharp, that is, in each case there exist $f, g \in \mathcal{N}$ for which equality holds.*

	$m = 1$	$m = 2$	$m = 3$	$m \geq 4$
$D_m \geq n + e(1,n) + 2^{m+2} - 7$	$n \geq 2$	$n = 2$	$n = 2, 3, 4$ $n \equiv 7, 8 \pmod 8$	$n = 2, 3, 4$ $n \equiv 3, 4, 7, 8 \pmod 8$ for $n > 6$
$D_m \geq n + e(1,n) + 2^{m+2} - 9$		$n = 3, 4$ $n \equiv 3, 4, 7, 8 \pmod 8$ for $n > 6$	$n = 5, 6$ $n \equiv 1, 2, 3, 4 \pmod 8$ for $n > 6$	$n = 5, 6$ $n \equiv 1, 2, 5, 6 \pmod 8$ for $n > 6$
$D_m \geq n + e(1,n) + 2^{m+2} - 11$		$n = 5, 6$ $n \equiv 1, 2, 5, 6 \pmod 8$ for $n > 6$	$n \equiv 5, 6 \pmod 8$ for $n > 6$	

Table 3.1: Theorem 1

§3.1 Commutator and power structure for $p = 2, k = 1$

From now on, suppose that $p = 2$ and $k = 1$. Note that if $D(f) = D(g) = 1$ then $D(gf^{-1}) \geq 2$ since $p = 2$. So, we can assume that $n \geq 2$. Since $k = 1$, the size of the matrix A_h at (2.5) is too small for our purposes. Hence we redefine the matrix A_h, for $h = 1$, by extending its size. This needs more careful computations. Let $b = x + x^2 + r_2x^3 + r_3x^4 + r_4x^5 + \ldots$, and $u = x + s_0x^3 + s_1x^4 + s_2x^5 + s_3x^6 + \ldots$. Note that $r_1 = 1$ by our assumption on k. Let $[u, b] \equiv x + \theta_0x^4 + \theta_1x^5 + \theta_2x^6 + \theta_3x^7 + \theta_4x^8 + \theta_5x^9 + \theta_6x^{10} + \theta_7x^{11} \pmod{x^{12}}$. Now we have

$$\begin{aligned}
ub \equiv\ & x + x^2 + (r_2 + s_0)x^3 + (r_3 + s_1 + s_0)x^4 + (r_4 + s_2 + r_2s_0 + s_0)x^5 \\
& + (r_5 + s_2 + s_3 + r_3s_0 + s_0)x^6 + (r_6 + s_2r_2 + s_4 + r_4s_0)x^7 \\
& + (r_7 + s_1 + s_2r_3 + s_3 + s_4 + s_5 + r_5s_0 + r_3s_0 + r_2s_0)x^8 \\
& + (r_8 + s_2r_4 + s_2 + s_4r_2 + s_4 + s_6 + r_6s_0 + r_3s_0 + r_4s_0 + r_2s_0)x^9 \\
& + (r_9 + s_2r_5 + s_2 + s_3r_2 + s_3 + s_4r_3 + s_4 + s_6 + s_7 + r_7s_0 + r_5s_0 + r_3s_0 + r_2r_3s_0)x^{10} \\
& + (r_{10} + s_2r_6 + s_2r_2 + s_4r_4 + s_4r_2 + s_4r_2 + s_4 + s_6r_2 \\
& + s_8 + r_8s_0 + r_4s_0 + r_6s_0 + r_2r_4s_0 + r_2r_3s_0)x^{11} \pmod{x^{12}}
\end{aligned}$$

$$\begin{aligned}
bu \equiv\ & x + x^2 + (r_2 + s_0)x^3 + (s_1 + r_3)x^4 + (s_2 + r_4 + r_2s_0)x^5 + (s_3 + r_5 + r_2s_1 + s_0)x^6 \\
& + (s_4 + r_6 + r_2s_2 + r_2s_0 + r_4s_0)x^7 + (s_5 + r_7 + r_2s_3 + s_1 + r_4s_1)x^8 \\
& + (s_6 + r_8 + r_2s_4 + r_2s_1 + r_4s_2 + r_2s_0 + r_6s_0)x^9 \\
& + (s_7 + r_9 + s_2 + r_2s_5 + r_4s_3 + r_6s_1 + r_2s_1s_0 + r_5s_0)x^{10} \\
& + (s_8 + r_{10} + r_2s_6 + r_2s_2 + r_4s_4 + r_6s_2 + r_2s_2s_0 + r_2s_1s_0 + r_6s_0 + r_8s_0)x^{11} \pmod{x^{12}}
\end{aligned}$$

$$\begin{aligned}
bu[u, b] \equiv\ & x + x^2 + (r_2 + s_0)x^3 + (r_3 + s_1 + \theta_0)x^4 + (s_2 + r_4 + r_2s_0 + \theta_1)x^5 \\
& + \Big(\theta_2 + (r_2 + s_0)\theta_0 + s_3 + r_5 + r_2s_1 + s_0\Big)x^6 \\
& + \Big(\theta_3 + (r_2 + s_0)\theta_1 + s_4 + r_6 + r_2s_2 + r_2s_0 + r_4s_0\Big)x^7 \\
& + \Big(\theta_4 + (r_2 + s_0)\theta_2 + \theta_0 + (s_2 + r_4 + r_2s_0)\theta_0 + s_5 + r_7 + r_2s_3 + s_1 + r_4s_1\Big)x^8
\end{aligned}$$

$$
\begin{aligned}
&+\Big(\theta_5+(r_2+s_0)\theta_3+(s_2+r_4+r_2s_0)\theta_1\\
&\quad+s_6+r_8+r_2s_4+r_2s_1+r_4s_2+r_2s_0+r_6s_0\Big)x^9\\
&+\Big(\theta_6+\theta_1+(r_2+s_0)\theta_4+(s_2+r_4+r_2s_0)\theta_2+(s_4+r_6+r_2s_2+r_2s_0+r_4s_0)\theta_0\\
&\quad+s_7+r_9+s_2+r_2s_5+r_4s_3+r_6s_1\Big)x^{10}\\
&+\Big(\theta_7+(r_2+s_0)\theta_5+(r_2+s_0)\theta_1+(s_4+r_2s_2+r_6+r_2s_0+r_4s_0)\theta_1\\
&\quad+(s_2+r_4+r_2s_0)\theta_3+s_8+r_{10}+r_2s_6+r_2s_2+r_4s_4+r_6s_2\\
&\quad+(r_2s_2+r_2s_1+r_6+r_8)s_0\Big)x^{11}\pmod{x^{12}}.
\end{aligned}
$$

By solving the equation $bu[u,b] \equiv ub \pmod{x^{12}}$, we can get the commutator $[u,b]$ mod $\pmod{x^{12}}$ which is

$$
\begin{aligned}
[u,b] \equiv\ & x+s_0x^4+s_0x^5+\Big[(r_2+r_3+1)s_0+r_2s_1+s_2\Big]x^6+s_0x^7\\
+\ & \Big[(1+r_2(r_3+1)+r_2s_1+r_4+r_5)s_0+(r_2+r_4)s_1+(r_2+r_3)s_2+(r_2+1)s_3+s_4\Big]x^8\\
+\ & \Big[(r_2+r_3+s_2)s_0+r_2s_1+s_2+s_4\Big]x^9+\Big[Bs_0+r_6s_1+(r_2(1+r_3)+r_4+r_5+1)s_2\\
+\ & (r_4+r_2+1)s_3+(r_2+r_3+1)s_4+r_2s_5+s_6\Big]x^{10}\\
+\ & \Big[Cs_0+r_2s_1+r_2s_2+(r_2+1)s_4\Big]x^{11}\pmod{x^{12}}
\end{aligned}
$$

where B, C are polynomials in the coefficients of u, and b whose explicit expressions are irrelevant. Let $\vec{\nu}=(s_0,s_1,s_2,\ldots,s_7)$ and $\vec{\nu}_1=(\theta_0,\theta_1,\theta_2,\ldots,\theta_7)$ be the coefficients vectors of u, $[u,b]$ for the terms x^{3+i} and x^{4+i}, respectively, for $i\geq 0$. Then the formula above gives us $\vec{\nu}A_{b,2}=\vec{\nu}_1$, where $A_{b,2}$ is

$$
\begin{pmatrix}
1 & 1 & r_2+r_3+1 & 1 & 1+r_2(r_3+1)+r_2s_1+r_4+r_5 & r_2+r_3+s_2 & * & *\\
0 & 0 & r_2 & 0 & r_2+r_4 & r_2 & r_6 & r_2\\
0 & 0 & 1 & 0 & r_2+r_3 & 1 & r_2+r_2r_3+r_4+r_5+1 & r_2\\
0 & 0 & 0 & 0 & r_2+1 & 0 & r_4+r_2+1 & 0\\
0 & 0 & 0 & 0 & 1 & 1 & r_2+r_3+1 & r_2+1\\
0 & 0 & 0 & 0 & 0 & 0 & r_2 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}. \tag{3.1}
$$

We use $*$ for the entries which are irrelevant for the computation. We can obtain the matrix $A_{b,n+1}$ from the matrix $A_{b,n}$ by deleting the first row and column of $A_{b,n}$ for $n = 2, 3, 4, 5$. We get

$$A_{b,3} = \begin{pmatrix} 0 & r_2 & 0 & r_2+r_4 & r_2 & r_6 & r_2 \\ 0 & 1 & 0 & r_2+r_3 & 1 & r_2+r_2r_3+r_4+r_5+1 & r_2 \\ 0 & 0 & 0 & r_2+1 & 0 & r_4+r_2+1 & 0 \\ 0 & 0 & 0 & 1 & 1 & r_2+r_3+1 & r_2+1 \\ 0 & 0 & 0 & 0 & 0 & r_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.2}$$

$$A_{b,4} = \begin{pmatrix} 1 & 0 & r_2+r_3 & 1 & r_2+r_2r_3+r_4+r_5+1 & r_2 \\ 0 & 0 & r_2+1 & 0 & r_4+r_2+1 & 0 \\ 0 & 0 & 1 & 1 & r_2+r_3+1 & r_2+1 \\ 0 & 0 & 0 & 0 & r_2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3.3}$$

$$A_{b,5} = \begin{pmatrix} 0 & r_2+1 & 0 & r_4+r_2+1 & 0 \\ 0 & 1 & 1 & r_2+r_3+1 & r_2+1 \\ 0 & 0 & 0 & r_2 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; A_{b,6} = \begin{pmatrix} 1 & 1 & r_2+r_3+1 & r_2+1 \\ 0 & 0 & r_2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.4}$$

Note that to determine the above matrices we need to consider the coefficients coming from the terms $x^{2n+2}, x^{2n+3}, x^{2n+4}$ in bu and $bu[u, b]$. Let $n \geq 7$, b as above, and $u = x + s_0x^{n+1} + s_1x^{n+2} + s_2x^{n+3} + \ldots$. Then $2n + 2 \geq n + 9$, so those terms in bu and $bu[u, b]$ do not affect the solution of $ub \equiv bu[u, b] \pmod{x^{n+9}}$. Hence, for $n \geq 7$, by using similar computations as in the case $n < 7$, we can write a general formula modulo x^{n+9} to determine $A_{b,n}$. Indeed,

$$\begin{aligned} ub \;\equiv\; & x + x^2 + r_2x^3 + \ldots + r_{n-1}x^n + (r_n + s_0)x^{n+1} + \Big[r_{n+1} + s_0(n+1) + s_1\Big]x^{n+2} \\ + \; & \Big[r_{n+2} + s_0r_2(n+1) + s_0\tbinom{n+1}{2} + s_1(n+2) + s_2\Big]x^{n+3} \\ + \; & \Big[r_{n+3} + s_0r_3(n+1) + s_0\tbinom{n+1}{3} + s_1r_2(n+2) + s_1\tbinom{n+2}{2} + s_2(n+3) + s_3\Big]x^{n+4} \end{aligned}$$

$$
\begin{aligned}
+ \quad & \Big[r_{n+4} + s_0 r_4 (n+1) s_0 r_2 \tbinom{n+1}{2} + s_0 r_2 \tbinom{n+1}{3} + s_0 \tbinom{n+1}{4} + s_1 r_3 (n+2) + s_1 \tbinom{n+2}{3} \\
& + s_2 r_2 (n+3) + s_2 \tbinom{n+3}{2} + s_3 (n+4) + s_4 \Big] x^{n+5} \\
+ \quad & \Big[r_{n+5} + s_0 r_5 (n+1) + s_0 r_3 \tbinom{n+1}{2} + s_0 r_2 \tbinom{n+1}{3} + s_0 \tbinom{n+1}{5} + s_1 r_4 (n+2) \\
& + s_1 r_2 \tbinom{n+2}{2} + s_1 r_2 \tbinom{n+2}{3} + s_1 \tbinom{n+2}{4} + s_2 r_3 (n+3) + s_2 \tbinom{n+3}{3} + s_3 r_2 (n+4) \\
& + s_3 \tbinom{n+4}{2} + s_4 (n+5) + s_5 \Big] x^{n+6} \\
+ \quad & \Big[r_{n+6} + s_0 r_6 (n+1) + s_0 r_3 \tbinom{n+1}{2} + s_0 r_4 \tbinom{n+1}{3} + s_0 r_2 \tbinom{n+1}{3} + s_0 r_2 \tbinom{n+1}{5} \\
& + s_0 \tbinom{n+1}{6} + s_1 r_5 (n+2) + s_1 r_3 \tbinom{n+2}{3} + s_1 r_2 \tbinom{n+2}{3} + s_1 \tbinom{n+2}{5} + s_2 r_4 (n+3) \\
& + s_2 r_2 \tbinom{n+3}{2} + s_2 r_2 \tbinom{n+3}{3} + s_2 \tbinom{n+3}{4} + s_3 r_3 (n+4) + s_3 \tbinom{n+4}{3} + s_4 r_2 (n+5) \\
& + s_4 \tbinom{n+5}{2} + s_5 (n+6) + s_6 \Big] x^{n+7} \\
+ \quad & \Big[r_{n+7} + s_0 r_7 (n+1) + s_0 r_5 \tbinom{n+1}{3} + s_0 r_3 \tbinom{n+1}{3} + s_0 r_2 r_3 \tbinom{n+1}{3} + s_0 r_3 \tbinom{n+1}{5} \\
& + s_0 \tbinom{n+1}{7} + s_1 r_6 (n+2) + s_1 r_3 \tbinom{n+2}{2} + s_1 r_4 \tbinom{n+2}{3} + s_1 r_2 \tbinom{n+2}{3} + s_1 r_2 \tbinom{n+2}{5} \\
& + s_1 \tbinom{n+2}{6} + s_2 r_5 (n+3) + s_2 r_3 \tbinom{n+3}{3} + s_2 r_2 \tbinom{n+3}{3} + s_2 \tbinom{n+3}{5} + s_3 r_4 (n+4) \\
& + s_3 r_2 \tbinom{n+4}{2} + s_3 r_2 \tbinom{n+4}{3} + s_3 \tbinom{n+4}{4} + s_4 r_3 (n+5) + s_4 \tbinom{n+5}{3} + s_5 r_2 (n+6) \\
& + s_5 \tbinom{n+6}{2} + s_6 (n+7) + s_7 \Big] x^{n+8} \pmod{x^{n+9}},
\end{aligned}
$$

$$
\begin{aligned}
bu \equiv\ & x + x^2 + r_2 x^3 + \ldots + r_{n-1} x^n + (r_n + s_0) x^{n+1} + (r_{n+1} + s_1) x^{n+2} + (r_2 s_0 + s_2) x^{n+3} \\
& + (r_2 s_1 + s_3) x^{n+4} + (r_2 s_2 + r_4 s_0 + s_4) x^{n+5} + (r_2 s_3 + r_4 s_1 + s_5) x^{n+6} \\
& + (r_2 s_4 + r_4 s_2 + r_6 s_0 + s_6) x^{n+7} + (r_2 s_5 + r_4 s_3 + r_6 s_1 + s_7) x^{n+8} \pmod{x^{n+9}}.
\end{aligned}
$$

Take $[u, b] \equiv x + \theta_0 x^{n+2} + \theta_1 x^{n+3} + \theta_2 x^{n+4} + \theta_3 x^{n+5} + \theta_4 x^{n+6} + \theta_5 x^{n+7} + \theta_6 x^{n+8} \pmod{x^{n+9}}$.

Then

$$
\begin{aligned}
bu[u, b] \equiv\ & x + x^2 + r_2 x^3 + \ldots + r_{n-1} x^n + (r_n + s_0) x^{n+1} + (r_{n+1} + s_1 + \theta_0) x^{n+2} \\
& + (r_{n+2} + r_2 s_0 + s_2 + \theta_1) x^{n+3} + (r_{n+3} + r_2 s_1 + r_2 \theta_0 + s_3 + \theta_2) x^{n+4}
\end{aligned}
$$

$$
\begin{aligned}
&+ (r_{n+4} + r_2 s_2 + r_4 s_0 + r_2\theta_1 + s_4 + \theta_3)x^{n+5} \\
&+ (r_{n+5} + r_2 s_3 + r_4 s_1 + r_2\theta_2 + r_4\theta_0 + s_5 + \theta_4)x^{n+6} \\
&+ (r_{n+6} + r_2 s_4 + r_4 s_2 + r_6 s_0 + r_4\theta_1 + r_2\theta_3 + s_6 + \theta_5)x^{n+7} \\
&+ (r_{n+7} + r_2 s_5 + r_4 s_3 + r_6 s_1 + r_6\theta_0 + r_4\theta_2 + r_2\theta_4 + s_7 + \theta_6)x^{n+8} \pmod{x^{n+9}}.
\end{aligned}
$$

Now we have

$$
\begin{aligned}
[u,b] \;\equiv\; & x + \Big[(n+1)s_0\Big]x^{n+2} + \Big[\Big(nr_2 + \tbinom{n+1}{2}\Big)s_0 + ns_1\Big]x^{n+3} \\
+ \;& \Big[\Big(n+1)(r_2+r_3) + \tbinom{n+1}{3}\Big)s_0 + \Big((n+1)r_2 + \tbinom{n+2}{2}\Big)s_1 + (n+1)s_2\Big]x^{n+4} \\
+ \;& \Big[\Big(n(r_2+r_4) + \tbinom{n+1}{3}r_2 + \tbinom{n+1}{4}\Big)s_0 + \Big(n(r_2+r_3) + \tbinom{n+2}{3}\Big)s_1 \\
& + \Big(nr_2 + \tbinom{n+3}{2}\Big) + ns_3\Big]x^{n+5} \\
+ \;& \Big[\Big((n+1)(r_2 + r_2r_3 + r_4 + r_5) + \tbinom{n+1}{3}r_3 + \tbinom{n+1}{5}\Big)s_0 \\
& + \Big((n+1)(r_2+r_4) + \tbinom{n+2}{3}r_2 + \tbinom{n+2}{4}\Big)s_1 + \Big((n+1)(r_2+r_3) + \tbinom{n+3}{3}\Big)s_2 \\
& + \Big((n+1)r_2 + \tbinom{n+4}{2}\Big)s_3 + (n+1)s_4\Big]x^{n+6} \\
+ \;& \Big[\Big(n(r_2+r_6) + \tbinom{n+1}{4}r_2 + \tbinom{n+1}{2}(r_4+r_3) + \tbinom{n+1}{3}r_4 + \tbinom{n+1}{5}r_2 + \tbinom{n+1}{6}\Big)s_0 \\
+ \;& \Big(n(r_2 + r_2r_3 + r_4 + r_5) + \tbinom{n+2}{3}r_3 + \tbinom{n+2}{5}\Big)s_1 \\
& + \Big(n(r_2+r_4) + \tbinom{n+3}{3}r_2 + \tbinom{n+3}{4}\Big)s_2 + \Big(n(r_2+r_3) + \tbinom{n+4}{3}\Big)s_3 \\
& + \Big(nr_2 + \tbinom{n+5}{2}\Big)s_4 + ns_5\Big]x^{n+7} \\
+ \;& \Big[\Big((n+1)(r_6 + r_3r_4 + r_2(1 + r_3 + r_5) + r_7) \\
& + \tbinom{n+1}{3}(r_3 + r_4 + r_5) + \tbinom{n+1}{5}(r_2 + r_3) + \tbinom{n+1}{7}\Big)s_0 + \Big((n+1)(r_2 + r_6) \\
& + \tbinom{n+2}{2}(r_3 + r_4) + \tbinom{n+2}{3}r_4 + \tbinom{n+2}{4}r_2 + \tbinom{n+2}{5}r_2 + \tbinom{n+2}{6}\Big)s_1 \\
& + \Big((n+1)(r_2 + r_4 + r_5 + r_2r_3) + \tbinom{n+3}{3}r_3 + \tbinom{n+3}{5}\Big)s_2 \\
& + \Big((n+1)(r_4 + r_2) + \tbinom{n+4}{3}r_2 + \tbinom{n+4}{4}\Big)s_3 + \Big((n+1)(r_2 + r_3) + \tbinom{n+5}{3}\Big)s_4 \\
& + \Big((n+1)r_2 + \tbinom{n+6}{2}\Big)s_5 + \Big((n+1)\Big)s_6\Big]x^{n+8} \pmod{x^{n+9}}.
\end{aligned}
$$

Then $A_{b,n}$ is

$$\begin{pmatrix} 0 & r_2+1 & 0 & * & 0 \\ 0 & 1 & 1 & * & r_2+\binom{n+2}{4} \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & r_2 & 0 & * & 0 & * & 0 \\ 0 & 1 & 0 & * & \binom{n+2}{4} & * & 0 \\ 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 1 & * & r_2+\binom{n+4}{4} \\ 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.5}$$

if $n \equiv 1, 5 \pmod 8$ if $n \equiv 3, 7 \pmod 8$

For the even values of n, we can obtain the matrices by deleting the first row and column of $A_{b,n-1}$, for example,

$$A_{b,n} = \begin{pmatrix} 1 & 1 & * & r_2+\binom{(n-1)+2}{4} \\ 0 & 0 & * & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ if } n \equiv 2, 6 \pmod 8. \tag{3.6}$$

Before going further, we need some observations related to the powers of b. We have $b = x + x^2 + r_2x^3 + r_3x^4 + r_4x^5 + \ldots$ then by direct substitution we obtain

$$b^2 = x + \alpha_3x^4 + \alpha_5x^6 + \alpha_7x^8 + \alpha_8x^9 + \alpha_9x^{10} + \ldots \tag{3.7}$$

where $\alpha_3 = 1+r_2$, $\alpha_5 = r_2r_3+r_4$, $\alpha_7 = r_2(r_5+r_3+1)+r_3r_4+r_5+r_6$, $\alpha_8 = r_2(r_3+r_4+1)+r_6$, $\alpha_9 = r_2r_7 + r_4r_5 + r_5 + r_3r_6 + r_6 + r_8$, $\alpha_{10} = r_2(r_4 + r_3) + r_6$. From here,

$$b^{2^2} = x + b_{11}x^{12} + b_{13}x^{14} + b_{15}x^{16} + b_{16}x^{17} + b_{17}x^{18} + b_{18}x^{19} + \ldots \tag{3.8}$$

where $b_{11} = \alpha_3(\alpha_5+\alpha_8)$, $b_{13} = \alpha_5\alpha_8+\alpha_{10}\alpha_3$, $b_{15} = \alpha_5(1+\alpha_{10})+\alpha_3+\alpha_8\alpha_7+\alpha_9\alpha_3+\alpha_{12}\alpha_3$, $b_{16} = \alpha_8 + \alpha_{10}\alpha_3$, $b_{17} = \alpha_5(1 + \alpha_{12}) + \alpha_8\alpha_9 + \alpha_{10}\alpha_7 + \alpha_{14}$, and $b_{18} = 0$. The following formula from [Heg01, Formula (4)] will help us to compute larger powers of b.

Let $a = x + a_kx^{k+1} + a_{k+1}x^{k+2} + a_{k+3}x^{k+3} + \ldots \in \mathcal{N}$ with $k \geq 1$. Then

$$a^2 = x + \sum_{i,j\geq k} a_ia_j(i+1)x^{i+j+1} + \sum_{i,j\geq k} a_ia_j^2\frac{i(i+1)}{2}x^{i+2j+1} \pmod{x^{4k+1}}. \tag{3.9}$$

The following observation is a direct application of the formula (3.9).

Lemma 3.1.1. *Let* $a = x + \sum_{i=0}^{l} a_{k+2i}x^{k+2i+1} + \sum_{j=2l+1}^{\infty} a_{k+j}x^{k+j+1}$ *where* k *is odd,* $l \in \mathbb{N}$*, and* $k \geq 2l+1$ *then*

$$a^2 = x + \sum_{i=1}^{l+1} c_i x^{2k+2l+2i} + \sum_{j=1}^{\infty} c_j x^{2k+4l+2+j}.$$

In Lemma 3.1.1, the first odd power term in a is x^{k+2l+2}. This causes a larger drop of the depth of a^2 than expected, that is, $D(a^2) \geq 2k+2l+1$ and the first odd power term in a^2 is $x^{2k+4l+2+1}$.

Let us see an application of Lemma 3.1.1 which will give us the larger powers of our $b = x+x^2+r_2x^3+r_3x^4+r_4x^5+\ldots$. We have $b^{2^2} = x+b_{11}x^{12}+b_{13}x^{14}+b_{15}x^{16}+b_{16}x^{17}+b_{17}x^{18}+b_{18}x^{19}+\ldots$, here $k = 11 = 2^4-5$ and $l = 2$. Then $D(b^{2^3}) \geq 2(2^4-5)+2.2+1 = 2^5-5 = 27$ and $b^{2^3} = x + c_{27}x^{28} + c_{29}x^{30} + c_{31}x^{32} + c_{32}x^{33} + \ldots$. Following this way, inductively we get

$$b^{2^m} = x + \alpha x^{2^{m+2}-4} + \beta x^{2^{m+2}-2} + \gamma x^{2^{m+2}} \pmod{x^{2^{m+2}+1}} \tag{3.10}$$

for $m \geq 1$, and for $\alpha, \beta, \gamma \in \mathbb{F}_p$ (depending on m).

Now we have $b^2 = x + \alpha_3x^4 + \alpha_5x^6 + \alpha_7x^8 + \alpha_8x^9 + \alpha_9x^{10} + \ldots$. Let n be odd and $u_1 = x + s_0x^{n+1} + s_1x^{n+2} + s_2x^{n+3} + \ldots$. We need to find the solution of $b^2u_1[u_1, b^2] \equiv u_1b^2 \pmod{x^{n+13}}$ in order to find the coefficients of $[u_1, b^2]$ modulo $\pmod{x^{n+13}}$. Note that the first term in b^2u_1 or $b^2u_1[u_1, b^2]$, which may affect the solution of the equation above, is $2n+5$. If $n \geq 9$ then $2n+5 \geq n+13$. Therefore, for $n \geq 9$, we can write a general formula $\pmod{x^{n+13}}$ to determine $A_{b^2,n}$ of size 9. Indeed,

$$\begin{aligned}
u_1b^2 \equiv\ & x + \alpha_3x^4 + \alpha_5x^6 + \alpha_7x^8 + \alpha_8x^9 + \ldots + \alpha_{n-1}x^n + (s_0+\alpha_n)x^{n+1} \\
+\ & (s_1+\alpha_{n+1})x^{n+2} + (s_2+\alpha_{n+2})x^{n+3} + \Big[s_3 + \alpha_{n+3} + (n+1)s_0\alpha_3\Big]x^{n+4} \\
+\ & \Big[s_4 + \alpha_{n+4} + (n+2)s_1\alpha_3\Big]x^{n+5} + \Big[s_5 + \alpha_{n+4} + (n+1)s_0\alpha_5 + (n+3)s_2\alpha_3\Big]x^{n+6} \\
+\ & \Big[s_6 + \alpha_{n+6} + \tbinom{n+1}{2}s_0\alpha_3 + (n+2)s_1\alpha_5 + (n+4)s_3\alpha_3\Big]x^{n+7} \\
+\ & \Big[s_7 + \alpha_{n+7} + (n+1)s_0\alpha_7 + \tbinom{n+2}{2}s_1\alpha_3 + (n+3)s_2\alpha_5 + (n+5)s_4\alpha_3 + \Big]x^{n+8}
\end{aligned}$$

$$
\begin{aligned}
+ \quad & \Big[s_8 + \alpha_{n+8} + (n+1)s_0\alpha_8 + (n+2)s_1\alpha_7 + \tbinom{n+3}{2}s_2\alpha_3 + (n+4)s_3\alpha_5 \\
& + (n+6)s_5\alpha_3 \Big] x^{n+9} \\
+ \quad & \Big[s_9 + \alpha_{n+9} + (n+1)s_0\alpha_9 + \tbinom{n+3}{3}s_0\alpha_3 + (n+2)s_1\alpha_8 + (n+3)s_2\alpha_7 \\
& + \tbinom{n+4}{2}s_3\alpha_3 + (n+5)s_4\alpha_5 + (n+7)s_6\alpha_3 \Big] x^{n+10} \\
+ \quad & \Big[s_{10} + \alpha_{10} + (n+1)s_0\alpha_{10} + \tbinom{n+1}{2}s_0\alpha_5 + (n+2)s_1\alpha_9 + \tbinom{n+2}{3}s_1\alpha_3 + (n+3)s_2\alpha_8 \\
& + (n+4)s_3\alpha_7 + \tbinom{n+5}{2}s_4\alpha_3 + (n+6)s_5\alpha_5 + (n+8)s_7\alpha_3 \Big] x^{n+11} \\
+ \quad & \Big[s_{11} + \alpha_{n+11} + (n+1)s_0\alpha_{11} + \tbinom{n+1}{3}s_0\alpha_3\alpha_5 + (n+2)s_1\alpha_{10} + \tbinom{n+2}{2}s_1\alpha_5 \\
& + (n+3)s_2\alpha_9 + \tbinom{n+3}{3}s_2\alpha_3 + (n+4)s_3\alpha_8 + (n+5)s_4\alpha_7 + \tbinom{n+6}{2}s_5\alpha_3 \\
& + (n+7)s_6\alpha_5 + (n+9)s_8\alpha_3 \Big] x^{n+12} \pmod{x^{n+13}}
\end{aligned}
$$

$$
\begin{aligned}
b^2u_1 \equiv\ & x + \alpha_3x^4 + \alpha_5x^6 + \alpha_7x^8 + \alpha_8x^9 + \ldots + \alpha_{n-1}x^n + (\alpha_n + s_0)x^{n+1} \\
& + (\alpha_{n+1} + s_1)x^{n+2} + \ldots + (\alpha_{n+7} + s_7)x^{n+8} + (\alpha_{n+8} + s_8 + s_0\alpha_8)x^{n+9} \\
& + (\alpha_{n+9} + s_9 + s_1\alpha_8)x^{n+10} + (\alpha_{n+10} + s_{10} + s_2\alpha_8 + s_0\alpha_{10})x^{n+11} \\
& + (\alpha_{n+11} + s_{11} + s_3\alpha_8 + s_1\alpha_{10})x^{n+12} \pmod{x^{n+13}}
\end{aligned}
$$

Let $[u_1, b^2] \equiv x + \theta_0x^{n+4} + \theta_1x^{n+5} + \theta_2x^{n+6} + \ldots + \theta_8x^{n+12} \pmod{x^{n+13}}$. Then

$$
\begin{aligned}
b^2u_1[u_1, b^2] \equiv\ & x + \alpha_3x^4 + \alpha_5x^6 + \alpha_7x^8 + \alpha_8x^9 + \ldots + \alpha_{n-1}x^n + (\alpha_n + s_0)x^{n+1} \\
& + (\alpha_{n+1} + s_1)x^{n+2} + (\alpha_{n+2} + s_2)x^{n+3} + (\alpha_{n+3} + s_3 + \theta_0)x^{n+4} \\
& + (\alpha_{n+4} + s_4 + \theta_1)x^{n+5} + \ldots + (\alpha_{n+7} + s_7 + \theta_4)x^{n+8} \\
& + (\alpha_{n+8} + s_8 + s_0\alpha_8 + \theta_5)x^{n+9} + (\alpha_{n+9} + s_9 + s_1\alpha_8 + \theta_6)x^{n+10} \\
& + (\alpha_{n+10} + s_{10} + s_2\alpha_8 + s_0\alpha_{10} + \theta_7)x^{n+11} \\
& + (\alpha_{n+11} + s_{11} + s_3\alpha_8 + s_1\alpha_{10} + \alpha_8\theta_0 + \theta_8)x^{n+12} \pmod{x^{n+13}}.
\end{aligned}
$$

Now by solving the equation $b^2u_1[u_1, b^2] \equiv u_1b^2 \pmod{x^{n+13}}$, we get the following matrix for an odd number $n \geq 9$

$$A_{b^2,n} = \begin{pmatrix} 0 & 0 & 0 & \binom{n+1}{2}\alpha_3 & 0 & \alpha_8 & 0 & * & 0 \\ 0 & \alpha_3 & 0 & \alpha_5 & \binom{n+2}{2}\alpha_3 & \alpha_7 & 0 & * & \binom{n+2}{2}\alpha_5 \\ 0 & 0 & 0 & 0 & 0 & \binom{n+3}{2}\alpha_3 & 0 & * & 0 \\ 0 & 0 & 0 & \alpha_3 & 0 & \alpha_5 & \binom{n+4}{2}\alpha_3 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_3 & 0 & * & \binom{n+6}{2}\alpha_3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.11}$$

Let $n = 3$, then

$$\begin{aligned} u_1 b^2 \equiv\ & x + (\alpha_3 + s_0)x^4 + s_1x^5 + (\alpha_5 + s_2)x^6 + s_3x^7 + (\alpha_7 + \alpha_3 s_1 + s_4)x^8 + (\alpha_8 + s_5)x^9 \\ & + (\alpha_9 + \alpha_5 s_1 + \alpha_3 s_3 + s_6)x^{10} + (\alpha_{10} + s_7)x^{11} + (\alpha_{11} + \alpha_7 s_1 + \alpha_3 s_2 + \alpha_5 s_3 + \alpha_3 s_5 + s_8)x^{12} \\ & + (\alpha_{12} + \alpha_8 s_1 + \alpha_3 s_3 + s_9)x^{13} + (\alpha_{13} + \alpha_9 s_1 + \alpha_7 s_3 + \alpha_5 s_5 + \alpha_3 s_7 + s_{10})x^{14} \\ & + (\alpha_{14} + \alpha_{10} s_1 + \alpha_8 s_3 + s_{11})x^{15} \\ & + (\alpha_{15} + \alpha_3 s_0 + \alpha_{11} s_1 + \alpha_5 s_2 + \alpha_9 s_3 + \alpha_3 s_3 + \alpha_7 s_5 + \alpha_3 s_6 + \alpha_5 s_7 + \alpha_3 s_9 + s_{12})x^{16} \\ & + (\alpha_{16} + \alpha_{12} s_1 + \alpha_3 s_1 + \alpha_{10} s_3 + \alpha_5 s_3 + \alpha_8 s_5 + \alpha_3 s_7 + s_{13})x^{17} \\ & + (\alpha_{17} + \alpha_{13} s_1 + \alpha_3 s_2 + \alpha_{11} s_3 + \alpha_3\alpha_5 s_3 + \alpha_9 s_5 + \alpha_7 s_7 + \alpha_5 s_9 + \alpha_3 s_{11} + s_{14})x^{18} \\ & + (\alpha_{18} + \alpha_{14} s_1 + \alpha_{12} s_3 + \alpha_3 s_3 + \alpha_{10} s_5 + \alpha_8 s_7 + s_{15})x^{19} \\ & + (\alpha_{19} + \alpha_{15} s_1 + \alpha_3 s_1 + \alpha_7 s_2 + \alpha_{13} s_3 + \alpha_7\alpha_3 s_3 + \alpha_5\alpha_3 s_3 + \alpha_{11} s_5 \\ & \quad + \alpha_5 s_6 + \alpha_9 s_7 + \alpha_3 s_7 + \alpha_7 s_9 + \alpha_3 s_{10} + \alpha_5 s_{11} + \alpha_3 s_{13} + s_{16})x^{20} \pmod{x^{21}} \end{aligned}$$

$$\begin{aligned} b^2 u_1 \equiv\ & x + (\alpha_3 + s_0)x^4 + s_1x^5 + (\alpha_5 + s_2)x^6 + s_3x^7 + (\alpha_7 + s_4)x^8 + (\alpha_8 + s_5)x^9 \\ & + (\alpha_9 + s_6)x^{10} + (\alpha_{10} + s_7)x^{11} + (\alpha_{11} + \alpha_5 s_0 + \alpha_8 s_0 + s_8)x^{12} \\ & + (\alpha_{12} + \alpha_8 s_1 + s_9)x^{13} + (\alpha_{13} + \alpha_5 s_1 + \alpha_8 s_2 + \alpha_{10} s_0 + s_{10})x^{14} \\ & + (\alpha_{14} + \alpha_{10} s_1 + \alpha_8 s_3 + s_{11})x^{15} \\ & + (\alpha_{15} + \alpha_3 s_0 + \alpha_5 s_2 + \alpha_8 s_4 + \alpha_9 s_0 + \alpha_{10} s_2 + \alpha_{12} s_0 + s_{12})x^{16} \\ & + (\alpha_{16} + \alpha_{12} s_1 + \alpha_{10} s_3 + \alpha_8 s_5 + \alpha_{10} s_0 + s_{13})x^{17} \\ & + (\alpha_{17} + \alpha_5 s_3 + \alpha_5 s_0 + \alpha_8 s_6 + \alpha_9 s_1 + \alpha_{10} s_4 + \alpha_{12} s_2 + \alpha_{14} s_0 + s_{14})x^{18} \end{aligned}$$

$$+ (\alpha_{18} + \alpha_{14}s_1 + \alpha_{12}s_3 + \alpha_{10}s_5 + \alpha_8 s_7 + \alpha_{10}s_1 + s_{15})x^{19}$$
$$+ (\alpha_{19} + \alpha_3 s_1 + \alpha_5 s_4 + \alpha_8 s_8 + \alpha_9 s_2 + \alpha_{10}s_6 + \alpha_{10}s_0 + \alpha_{12}s_4$$
$$+\alpha_{13}s_0 + \alpha_{14}s_3 + \alpha_{16}s_0 + s_{16})x^{20} \pmod{x^{21}}.$$

Let $[u_1, b^2] \equiv x + x^7\theta_0 + x^8\theta_1 + x^9\theta_2 + \ldots + x^{20}\theta_{13} \pmod{x^{21}}$. Then

$$\begin{aligned}
b^2u_1[u_1, b^2] \equiv\ & x + (\alpha_3 + s_0)x^4 + s_1x^5 + (\alpha_5 + s_2)x^6 + (\theta_0 + s_3)x^7 + (\theta_1 + \alpha_7 + s_4)x^8 \\
& + (\theta_2 + \alpha_8 + s_5)x^9 + (\theta_3 + \alpha_9 + s_6)x^{10} + (\theta_4 + s_1\theta_0 + \alpha_{10} + s_7)x^{11} \\
& + (\theta_5 + s_1\theta_1 + \alpha_{11} + \alpha_5 s_0 + \alpha_8 s_0 + s_8)x^{12} + (\theta_6 + s_1\theta_2 + s_3\theta_0 + \alpha_{12} + \alpha_8 s_1 + s_9)x^{13} \\
& + (\theta_7 + s_1\theta_3 + s_3\theta_1 + \alpha_{13} + \alpha_5 s_1 + \alpha_8 s_2 + \alpha_{10}s_0 + s_{10})x^{14} \\
& + (\theta_8 + s_1\theta_4 + s_3\theta_2 + (\alpha_8 + s_5)\theta_0 + \alpha_{14} + \alpha_{10}s_1 + \alpha_8 s_3 + s_{11})x^{15} \\
& + \Big(\theta_9 + s_1\theta_5 + s_3\theta_3 + (\alpha_8 + s_5)\theta_1 + \alpha_{15} + \alpha_3 s_0 \\
& + \alpha_5 s_2 + \alpha_8 s_4 + \alpha_9 s_0 + \alpha_{10}s_2 + \alpha_{12}s_0 + s_{12}\Big)x^{16} \\
& + \Big(\theta_{10} + s_1\theta_6 + s_3\theta_4 + (\alpha_8 + s_5)\theta_2 + (\alpha_{10} + s_7)\theta_0 \\
& + \alpha_{16} + \alpha_{12}s_1 + \alpha_{10}s_3 + \alpha_8 s_5 + \alpha_{10}s_0 + s_{13}\Big)x^{17} \\
& + \Big(\theta_{11} + s_1\theta_7 + (\alpha_5 + s_2)\theta_0 + s_3\theta_5 + (\alpha_8 + s_5)\theta_3 + (\alpha_{10} + s_7)\theta_1 \\
& + \alpha_{17} + \alpha_5 s_3 + \alpha_5 s_0 + \alpha_8 s_6 + \alpha_9 s_1 + \alpha_{10}s_4 + \alpha_{12}s_2 + \alpha_{14}s_0 + s_{14}\Big)x^{18} \\
& + \Big(\theta_{12} + s_1\theta_8 + s_3\theta_6 + s_3\theta_0 + (\alpha_8 + s_5)\theta_4 + (\alpha_{10} + s_7)\theta_2 + (\alpha_{12} + \alpha_8 s_1 + s_9)\theta_0 \\
& + \alpha_{18} + \alpha_{14}s_1 + \alpha_{12}s_3 + \alpha_{10}s_5 + \alpha_8 s_7 + \alpha_{10}s_1 + s_{15}\Big)x^{19} \\
& + \Big(\theta_{13} + s_1\theta_9 + (\alpha_5 + s_2)\theta_1 + s_3\theta_7 + (\alpha_8 + s_5)\theta_5 \\
& + (\alpha_{10} + s_7)\theta_3 + (\alpha_{12} + \alpha_8 s_1 + s_3)\theta_1 + \alpha_{19} \\
& + \alpha_3 s_1 + \alpha_5 s_4 + \alpha_8 s_8 + \alpha_9 s_2 + \alpha_{10}s_6 + \alpha_{10}s_0 \\
& + \alpha_{12}s_4 + \alpha_{13}s_0 + \alpha_{14}s_3 + \alpha_{16}s_0 + s_{16}\Big)x^{20} \pmod{x^{21}}.
\end{aligned}$$

Now by solving the equation $b^2u_1[u_1, b^2] \equiv u_1b^2 \pmod{x^{21}}$, we can get the matrix $A_{b^2,3}$ of size 14. Then we can obtain the matrix $A_{b^2,n+2}$ from the matrix $A_{b^2,n}$ by deleting the first two rows and columns of $A_{b^2,n}$ for $n = 3, 5$. We need only the upper left 9×9 corners of $A_{b^2,3}, A_{b^2,5}, A_{b^2,7}$. The form of the matrices is almost similar as in the case $n \geq 9$. The

differences are only at the irrelevant entries denoted by $*$, and at the $(0,6),(1,6)$ entries of $A_{b^2,3}$ which are $\alpha_8+\alpha_5$, and $\alpha_7+\alpha_3$, respectively.

Now consider b^{2^m} for $m \geq 2$. First of all, $D(b^{2^m}) \geq 2^{m+2}-5 \geq 11$ by (3.10). Let $2^{m+2}-5=k$ and so $D(b^{2^m}) \geq k$. Again by the equation (3.10), $b^{2^m} = x+c_kx^{k+1}+c_{k+2}x^{k+3}+c_{k+4}x^{k+5}+c_{k+5}x^{k+6}+c_{k+6}x^{k+7}+\ldots$. Let $u_2 = x+s_0x^{n+1}+s_1x^{n+2}+s_2x^{n+3}+\ldots$ where n is odd. Now we can determine the matrix $A_{b^{2^m}}$ of size 11 by using (2.5) which is

$$A_{b^{2^m}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & c_{k+5} & 0 & * & 0 & * & 0 \\ 0 & c_k & 0 & c_{k+2} & 0 & c_{k+4} & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & c_k & 0 & c_{k+2} & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & c_k & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.12}$$

Note that we change the notation and use $A_{b^{2^m}}$ instead of $A_{b^{2^m},n}$ because for $m \geq 2$ the form of the matrix is same for all odd $n \in \mathbb{N}$.

By using (3.7), (3.8), (3.9), we have got the following examples which will be our main examples to show the sharpness of the bounds in Theorem 3.0.1

	b	b^2	$b^{2^m}, m \geq 2$
1.	$x+x^2+x^3$	$x+x^8+x^9 \pmod{x^{12}}$	$x+x^{2^{m+2}}+x^{2^{m+2}+1} \pmod{x^{2^{m+2}+4}}$
2.	$x+x^2+x^3+x^4+x^7$	$x+x^6+x^8+x^9 \pmod{x^{12}}$	$x+x^{2^{m+2}-2}+\alpha x^{2^{m+2}}+x^{2^{m+2}+1} \pmod{x^{2^{m+2}+2}}$
3.	$x+x^2+x^4$	$x+x^4+x^{16}$	$x+x^{2^{2^m}}+x^{2^{2^m+1}}$
4.	$x+x^2+x^4+x^5$	$x+x^4+x^6+x^8 \pmod{x^{12}}$	$x+x^{12}+\alpha x^{16} \pmod{x^{18}}$ for $m=2$

Table 3.2: Examples

where $\alpha \in \mathbb{F}_p$.

§3.2 Maximal deviation of large powers, $p=2, k=1$

The main aim of this section is to prove Theorem 3.0.1. Let $f,g \in \mathcal{N}$ with $D(f)=D(g)=1$ and $D(gf^{-1}) \geq n$ where $n \geq 2$. Set $u_m = g^{2^m}f^{-2^m}$ for $m \geq 0$. Recall the well known commutator identity: for any group G and for any $x,y \in G$ $xy = yx[x,y]$. Note that $f^{-2^m} =$

$g^{-2^m}u_m$. By taking square of both sides and using the commutator identity, we get

$$g^{2^{m+1}}f^{-2^{m+1}} = u_m^2[u_m, g^{-2^m}][[u_m, g^{-2^m}], u_m]. \tag{3.13}$$

Let $m \geq 0$, denote by $\vec{\nu}_m$ the coefficient vector of u_m. For $m = 0$, denote $\vec{\nu} = \vec{\nu}_0$, and $u = u_0$. Let $D(u_m^2) \geq B_m$, $D([u_m, g^{-2^m}]) \geq C_m$, and $D([[u_m, g^{-2^m}], u_m]) \geq E_m$. Let $M_m = min\{B_m, C_m, E_m\}$, and $a \in \{u_m^2, [u_m, g^{-2^m}], [[u_m, g^{-2^m}], u_m]\}$. Let $\vec{\nu}_{m,a}$ denote the vector consisting of the coefficients of the terms x^{M_m+i} in a for $i \geq 0$.

Remark 3.2.1. Let $m \geq 0$ and suppose that $D(u_m) \geq 7$. Then $D([[u_m, g^{-2^m}], u_m]) \geq D([u_m, g^{-2^m}]) + 7$ by Proposition 1.2.3. It follows that

$$g^{2^{m+1}}f^{-2^{m+1}} \equiv u_m^2[u_m, g^{-2^m}] \text{ (mod } x^{D([u_m, g^{-2^m}])+8}).$$

Remark 3.2.2. Suppose that $D([u, g^{-1}]) \geq n + e(1, n) + 1$ and $n \geq 7$. By Proposition 1.2.3, we have $D(u^2) \geq 2n + e(1, n) \geq n + e(1, n) + 1 + 6$ and $D([[u, g^{-1}], u]) \geq n + e(1, n) + 1 + 7$. Therefore, by (3.13), $D(g^2f^{-2}) \geq n + e(1, n) + 1$ and

$$D(g^2f^{-2}) \equiv [u, g^{-1}] \text{ (mod } x^{n+e(1,n)+7}).$$

Let $\vec{\nu} = (s_0, s_1, s_3, \ldots)$ and $g^{-1} = x + x^2 + r_2x^3 + r_3x^4 + r_4x^5 + \ldots$. Throughout the proof instead of over complicating the notation we will use $C*$ to denote an undetermined value which is 0 if $C = 0$, or we will use only $*$ when C is not important.

The following is a key observation of our proof:

Lemma 3.2.1. *Let $m \geq 2$, and $n \geq 2$ arbitrary. Suppose that $\vec{\nu}_m = (*, 0, *, 0, *, *, \ldots, *)$ and $D(u_m) \geq n + e(1, n) + 2^{m+2} - K$, where $K \in \{7, 9, 11\}$ as in Theorem 3.0.1 (e.g., for $n = 3, 4$ $K = 9$ if $m = 2$, $K = 7$ if $m \geq 3$ and so on). Then*

$$D(u_{m+1}) = D(g^{2^{m+1}}f^{-2^{m+1}}) \geq n + e(1, n) + 2^{m+3} - K,\ \vec{\nu}_{m+1} = (*, 0, *, 0, *, *, \ldots, *).$$

Equality holds if $D(u_m) = n + e(1, n) + 2^{m+2} - K$, and one of the followings is satisfied:

(a) $n = 2$, $m \geq 2$, $\vec{\nu}_m = (1, 0, *, 0, *, 0, *, *, \ldots, *)$, and $(0, 6)$ entry of $A_{g^{-2^m}}$ is 1,

(b) $n = 3, 4, 5, 6$, $\vec{\nu}_m = (1, 0, *, 0, *, 0, *, *, \ldots, *)$, $m \geq 3$, *the* $(0, 6)$ *entry of* $A_{g^{-2^m}}$ *is* 1 *for* $m \geq 2$, *and the* $(5, 6)$ *entry of* $A_{g^{-2^3}}$ *is* 0,

(c) $n \geq 7$, $\vec{\nu}_m = (1, 0, *, 0, *, 0, *, *, \ldots, *)$, $m \geq 2$, *the* $(0, 6)$ *entry of* $A_{g^{-2^m}}$ *is* 1 *for* $m \geq 2$, *and the* $(5, 6)$ *entry of* $A_{g^{-2^2}}$ *is* 0.

Proof. First, note that by (3.10) we have $D(g^{-2^m}) \geq 2^{m+2} - 5$. Then, by the matrix at (3.12),

$$\vec{\nu}_m A_{g^{-2^m}} = (0, 0, 0, 0, 0, *, 0, *, 0, *, 0) \Longrightarrow$$
$$D([u_m, g^{-2^m}]) \geq n + e(1, n) + 2^{m+2} - K + 2^{m+2} - 5 + 5$$
$$= n + e(1, n) + 2^{m+3} - K. \tag{3.14}$$

On the other hand, $D(u_m^2) \geq 2(n + e(1, n) + 2^{m+2} - K) + 5 \geq n + e(1, n) + 2^{m+3} - K$, and $\vec{\nu}_{m+1, u_m^2} = (*, 0, *, 0, *, *, \ldots, *)$ by Lemma 3.1.1. Also, by (3.14), we have $\vec{\nu}_{m+1, [u_m, g^{-2^m}]} = (*, 0, *, 0, *, 0, *, *, \ldots, *)$. Then by Remark 3.2.1,

$$\vec{\nu}_{m+1} = (*, 0, *, 0, *, *, \ldots, *),\ D(u_{m+1}) \geq n + e(1, n) + 2^{m+3} - K.$$

Essentially, a similar proof implies our claim about equality. Indeed, for all the cases $(a), (b), (c)$ above, we have $\vec{\nu}_{m+1, [u_m, g^{-2^m}]} = (1, 0, *, 0, *, 0, *, *, \ldots, *)$, and $\vec{\nu}_{m+1, u_m^2} = (0, 0, *, 0, *, 0, *, *, \ldots, *)$. So we have $\vec{\nu}_{m+1} = (1, 0, *, 0, *, 0, *, *, \ldots, *)$. □

For the small values of n we do not have a regular formulation for the corresponding matrices, so we have to investigate case by case. Hence, we first look at the general cases, that is, when $n \geq 7$. Then we look at the small values of n. We divide our proof into cases corresponding to the different values of n modulo 8 since we have different matrices corresponding to in these cases.

$\mathbf{n \geq 7}$:

First we determine $\vec{\nu}_2$ and $D(u_2)$ for each of $n \equiv 1, 3, 5, 7 \pmod 8$.

$\mathbf{n \equiv 1 \pmod 8}$: By using the first matrix at (3.5), and Remark 3.2.2,

$$\vec{\nu} A_{g^{-1}, n} = (0, *, *, *, r_2 *) \ \Rightarrow D([u, g^{-1}]) \geq n + D(g^{-1}) + 1 \geq n + 2 \equiv 3 \pmod 8,$$

$$\Rightarrow D(u_1) = D(g^2 f^{-2}) \geq n+2, \quad \vec{\nu_1} = (*,*,*,r_2*,*,\ldots,*).$$

Now by using (3.7) we have $g^{-2} = x + \alpha_3 x^4 + \alpha_5 x^6 + \alpha_7 x^8 + \alpha_8 x^9 + \alpha_9 x^{10} + \ldots$, and $\alpha_3 r_2 = 0$. By using the matrix at (3.11), Remark 3.2.1 and Proposition 1.2.3,

$$\vec{\nu_1} A_{g^{-2},n+2} = (0, \alpha_3*, 0, \alpha_5 * + \alpha_3*, 0, *, \alpha_3 r_2*, *, 0) = (0, \alpha_3*, 0, *, 0, *, 0, *, 0, *, \ldots, *)$$

$$\Rightarrow D([u_1, g^{-2}]) \geq n+2+D(g^{-2})+1 \geq n+2+3+1 = n+6 = n+e(1,n)+2^4-11.$$

$$D(u_1^2) \geq 2(n+2)+1 \geq n+6+n-1 \geq n+6+8.$$

$$\Rightarrow D(u_2) = D(g^{2^2} f^{-2^2}) \geq n+6 = n+e(1,n)+2^4-11, \quad \vec{\nu_2} = (\alpha_3*, 0, *, 0, *, 0, *, 0, *, \ldots, *).$$

$\mathbf{n \equiv 3}$ $(\text{mod } \mathbf{8})$: By using the second matrix at (3.5) and Remark 3.2.2,

$$\vec{\nu} A_{g^{-1},n} = (0, *, 0, *, *, *, *) \Rightarrow D([u, g^{-1}]) \geq n+1+1 \equiv 5 \text{ (mod } 8),$$

$$\Rightarrow D(u_1) = D(g^2 f^{-2}) \geq n+2, \ \vec{\nu_1} = (*, 0, *, *, *, *, .., *).$$

By using the matrix at (3.11), Lemma 3.1.1 and Remark 3.2.1,

$$\vec{\nu_1} A_{g^{-2},n+2} = (0, 0, 0, \alpha_3*, 0, *, 0, *, \alpha_3*)$$

$$\Rightarrow D([u_1, g^{-2}]) \geq n+2+3+3 = n+8 = n+e(1,n)+2^4-9.$$

$$D(u_1^2) \geq 2(n+2)+3 \geq n+8+10.$$

$$\Rightarrow D(u_2) = D(g^{2^2} f^{-2^2}) \geq n+8 = n+e(1,n)+2^4-9, \ \vec{\nu_2} = (\alpha_3*, 0, *, 0, *, \alpha_3*, *, \ldots, *).$$

$\mathbf{n \equiv 5}$ $(\text{mod } \mathbf{8})$: By using the first matrix at (3.5) and Remark 3.2.2,

$$\vec{\nu} A_{g^{-1},n} = (0, *, *, *, *) \Rightarrow D([u, g^{-1}]) \geq n+e(1,n)+1 \equiv 7 \text{ (mod } 8)$$

$$\Rightarrow D(u_1) = D(g^2 f^{-2}) \geq n+2, \ \vec{\nu_1} = (*, *, , \ldots, *).$$

By using the matrix at (3.11), Remark 3.2.1 and Proposition 1.2.3,

$$\vec{\nu_1} A_{g^{-2},n+2} = (0, \alpha_3*, 0, *, 0, *, \alpha_3*, *, 0)$$

$$\Rightarrow D([u_1, g^{-2}]) \geq n+2+3+1 = n+6 = n+e(1,n)+2^4-11.$$

$$D(u_1^2) \geq 2(n+2)+1 \geq n+6+12.$$

$$\Rightarrow D(u_2) = D(g^{2^2} f^{-2^2}) \geq n + e(1,n) + 2^4 - 11,\ \vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, \alpha_3 *, *, 0, *, *, \ldots, *).$$

$\mathbf{n \equiv 7}\ (\text{mod } \mathbf{8})$: By using the second matrix at (3.5) and Remark 3.2.2,

$$\vec{\nu} A_{g^{-1},n} = (0, *, 0, *, *, *, r_2 *) \Rightarrow D([u, g^{-1}]) \geq n+2 \equiv 1\ (\text{mod } 8).$$

$$\Rightarrow D(u_1) = D(g^2 f^{-2}) \geq n+2,\ \vec{\nu_1} = (*, 0, *, *, *, r_2 *, *, *, \ldots, *).$$

By using the matrix at (3.11), Lemma 3.1.1 and Remark 3.2.1,

$$\vec{\nu_1} A_{g^{-2},n+2} = (0,0,0,\alpha_3 *, 0, *, 0, *, \alpha_3 r_2 *) = (0,0,0,\alpha_3 *, 0, *, 0, *, 0)\ (\text{as } \alpha_3 r_2 = 0)$$

$$\Rightarrow D([u_1, g^{-2}]) \geq n+2+3+3 = n + e(1,n) + 2^4 - 9.$$

$$D(u_1^2) \geq 2(n+2)+3 \geq n+8+6.$$

$$\Rightarrow D(u_2) = D(g^2 f^{-2}) \geq n + e(1,n) + 2^4 - 9,\ \vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, 0, *, *, \ldots, *).$$

Now we have the vector $\vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, \delta\alpha_3 *, *, *, \ldots, *)$, where $\delta = 0$ if $n \equiv 1, 7\ (\text{mod } 8)$ and $\delta = 1$ if $n \equiv 3, 5\ (\text{mod } 8)$. Also $D(u_2) \geq n + e(1,n) + 2^4 - K$, where $K = 11$ if $n \equiv 1, 5\ (\text{mod } 8)$, $K = 9$ if $n \equiv 3, 7\ (\text{mod } 8)$.

Suppose that $\alpha_3 = 1$ then by using (3.7), (3.8) and (3.9), the coefficients of the terms x^{17} in b^{2^2} and x^{32} in b^{2^3} are both zero. So $(0,6)$ entry of the matrix at (3.12) is zero for $m = 2, 3$. Thus,

$$\vec{\nu_2} A_{g^{-2^2}} = (0,0,0,0,0,\delta *, 0, *, 0, *, 0) \Rightarrow D([u_2, g^{-2^2}]) \geq n + e(1,n) + 2^5 - K + \lambda$$

where $\lambda = 2$ if $\delta = 0$, $\lambda = 0$ otherwise . On the other hand, by Lemma 3.1.1, $D(u_2^2) \geq 2(n + e(1,n) + 2^4 - K) + 5 + \lambda \geq n + e(1,n) + 2^5 - K + 4 + \lambda$ and $\vec{\nu}_{3,u_2^2} = (0,0,0,0,*,0,*,*,\ldots,*)$. Now by Remark 3.2.1,

$$\vec{\nu_3} = (\delta *, 0, *, 0, *, 0, *, *, \ldots, *),\ D(u_3) \geq n + e(1,n) + 2^5 - K + \lambda.$$

Above we noted that $(0,6)$ entry of the matrix at (3.12) is zero for $m = 3$ so

$$\vec{\nu_3} A_{g^{-2^3}} = (0,0,0,0,0,0,0,*,0,*,0),\ D([u_3, g^{-2^3}]) \geq n + e(1,n) + 2^6 - K + \lambda + 2.$$

On the other hand, by Lemma 3.1.1, $D(u_3^2) \geq 2(n+e(1,n)+2^5-K+\lambda)+7 \geq n+e(1,n)+2^6-K+\lambda+2+6$. By Remark 3.2.1

$$\vec{\nu_4} = (*,0,*,0,*,*,\ldots,*),\ D(u_4) \geq n+e(1,n)+2^6-K+\lambda+2.$$

Here we can apply Lemma 3.2.1 with an inductive argument to get $D(g^{2^m}f^{-2^m}) \geq n+e(1,n)+2^{m+2}-K+\lambda+2$ for $m \geq 4$.

Now suppose that $\alpha_3 = 0$ then $\vec{\nu_2} = (*,0,*,0,*,*,\ldots,*)$ and $D(u_2) \geq n+e(1,n)+2^4-K+2$. Now we can apply Lemma 3.2.1 with an inductive argument to get $D(g^{2^m}f^{-2^m}) \geq n+e(1,n)+2^{m+2}-K+2$ for $m \geq 2$.

Now we have

$$D(g^{2^m}f^{-2^m}) \geq \begin{cases} n+e(1,n)+2^{m+2}-K & \text{if } m=2, \\ n+e(1,n)+2^{m+2}-K+2 & \text{if } m \geq 3. \end{cases} \tag{3.15}$$

$$D(g^{2^m}f^{-2^m}) \geq \begin{cases} n+e(1,n)+2^{m+2}-K' & \text{if } m=2,3, \\ n+e(1,n)+2^{m+2}-K'+2 & \text{if } m \geq 4. \end{cases} \tag{3.16}$$

where $K, K' = 11$ if $n \equiv 1 \pmod 8$, $n \equiv 5 \pmod 8$, respectively, and $K, K' = 9$ if $n \equiv 7 \pmod 8$, $n \equiv 3 \pmod 8$, respectively. This confirms the bound of Theorem 3.0.1 in this case.

$\mathbf{n \equiv 2, 4, 6, 8}$: The case of an even $n \geq 8$ is closely related to its odd counterpart $n-1$, above. We see that $\vec{\nu}A_{g^{-1}}$ is exactly the same vector as $\vec{\nu_1}$ obtained for $n-1$, above. On the other hand, $D(u_1) = n+e(1,n)+1 = n-1+e(1,n-1)+1$ as n is even. From now on the proof is the same as for $n-1$, above, so are the bounds. Now return to the sharpness of the bounds.

Example for $\mathbf{n \equiv 1, 2, 5, 6} \pmod 8$:

$\mathbf{m \leq 3}$: Let $n \equiv 1, 5 \pmod 8$ and $g^{-1} = x + x^2 + x^4 + x^5$. We have the following matrices by using the matrices at (3.5), (3.11), (3.12), and example 4 in Table 3.2

$$A_{g^{-1},n} = \begin{pmatrix} 0 & 1 & 0 & * & 0 \\ 0 & 1 & 1 & * & \binom{n+2}{4} \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_{g^{-2^m}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 1 & 0 & * & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & \epsilon & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.17}$$

where $m = 1, 2$. Also $\epsilon = 1$ if $m = 1$, zero otherwise. Let $\vec{\nu} = (1, 1, *, *, \ldots, *)$ then $\vec{\nu} A_{g^{-1},n} = (0, 0, 1, *, \binom{n+2}{4})$. Now $\vec{\nu_1} = (0, 1, *, \binom{n+2}{4}, *, *, \ldots, *)$. It follows that $\vec{\nu_1} A_{g^{-2},n+2} = (0, 1, 0, *, 0, *, \binom{n+2}{4}, *, 0)$ and so $D(g^{2^2} f^{-2^2}) = n + e(1, n) + 2^4 - 11$. Hence, for $m = 2$ and $n \equiv 1, 5 \pmod 8$ we have the desired bound. Now suppose $n \equiv 5 \pmod 8$ then $\binom{n+2}{4} = 1$. So we have $\vec{\nu_2} = (1, 0, *, 0, *, 1, *, 0, *, *, \ldots, *)$. Thus, $\vec{\nu_2} A_{g^{-2^2}} = (0, 0, 0, 0, 0, 1, 0, *, 0)$ and so we have $D(g^{2^3} f^{-2^3}) = n + e(1, n) + 2^5 - 11$, which is the desired bound for $n \equiv 5 \pmod 8$. For the even values $n \equiv 2, 6 \pmod 8$, it can be directly seen that the same example produces the desired bound which is same as for $n \equiv 1, 5 \pmod 8$. Indeed, from $\vec{\nu} A_{g^{-1},n}$ we get $\vec{\nu_1} = (1, 1, *, \binom{n+1}{4})$ which produces same $\vec{\nu_2}$ and $D(u_2)$ as in the odd case.

$\mathbf{m \geq 3}$:Let $n \equiv 1, 5 \pmod 8$ and $g^{-1} = x + x^2 + x^3 + x^4 + x^7$. By using the matrices at (3.5), (3.11), (3.12), and example 2 in Table 3.2,

$$A_{g^{-1},n} = \begin{pmatrix} 0 & 0 & 0 & * & 0 \\ 0 & 1 & 1 & * & \binom{n+2}{4} \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_{g^{-2^m}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.18}$$

Note that for $m \geq 2$ we can extend the size of the matrix $A_{g^{-2^m}}$ to 11 with the 11-th column being zero by using the matrix at (3.12). Let $\vec{\nu} = (1, 1, *, *, \ldots, *)$. Then $\vec{\nu}A_{g^{-1},n} = (0, 1, 1, *, \binom{n+2}{4})$ and so $\vec{\nu_1} = (1, 1, *, \binom{n+2}{4}, *, *, \ldots, *)$. Now $\vec{\nu_1}A_{g^{-2},n+2} = (0, 0, 0, 1, 0, *, 0, *, 0)$, it follows that $D(u_2) = n + e(1, n) + 2^4 - 9$, and $\vec{\nu_2} = (1, 0, *, 0, *, 0, *, *, \ldots, *)$. The rest follows from Lemma 3.2.1. For the even values $n \equiv 2, 6 \pmod 8$ the same example works. Similarly to the previous case, after the first step we get $\vec{\nu_1} = (1, 1, *, \binom{n+1}{4})$ which produces same $\vec{\nu_2}$ and same $D(u_2)$ as in the odd case. So the rest is the same as in the odd case.

Example for $\mathbf{n \equiv 3, 4, 7, 8} \pmod 8$:

$\mathbf{m \leq 3}$: Let $n \equiv 3, 7 \pmod 8$, and $g^{-1} = x + x^2 + x^4 + x^5$. By using the matrices at (3.5), (3.11), (3.12) and example 4 in Table 3.2 we have the following matrices,

$$A_{g^{-1},n} = \begin{pmatrix} 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 1 & 0 & * & \binom{n+2}{4} & * & 0 \\ 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 1 & * & \binom{n+4}{4} \\ 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; A_{g^{-2},n+2} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & * & 0 & * & 0 \\ 0 & 1 & 0 & 1 & 1 & * & 0 & * & 1 \\ 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.19}$$

$$A_{g^{-2^2}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 1 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.20}$$

Let $\vec{\nu} = (1, 1, 1, \binom{n+2}{4}, *, *, \ldots, *)$. It follows that $\vec{\nu}A_{g^{-1},n} = (0, 1, 0, *, 0, *, \binom{n+2}{4}\binom{n+4}{4})$ and so $\vec{\nu_1} = (1, 0, *, 0, *, \binom{n+2}{4}\binom{n+4}{4}, *, *, \ldots, *)$. Now $\vec{\nu_1}A_{g^{-2},n+2} = (0, 0, 0, 1, 0, *, 0, *, \binom{n+2}{4}\binom{n+4}{4})$, so $D(u_2) = n + e(1, n) + 2^4 - 9$, and $\vec{\nu_2} = (1, 0, *, 0, *, \binom{n+2}{4}\binom{n+4}{4}, *, *, \ldots, *)$. Hence, we have the desired result that $D(g^{2^2}f^{-2^2}) =$

$n+e(1,n)+2^4-9$ for $n \equiv 3,7 \pmod 8$. Now suppose that $n \equiv 3 \pmod 8$ then $\binom{n+2}{4}\binom{n+4}{4} = 1$. Hence, $\vec{\nu_2} = (1,0,*,0,*,1,*,*,\ldots,*)$ and so $\vec{\nu_2}A_{g^{-2^2}} = (0,0,0,0,0,1,0,*,0)$ which implies $D(g^{2^3}f^{-2^3}) = n + e(1,n) + 2^5 - 9$ as desired. For the even values $n \equiv 4,8 \pmod 8$, we use a similar argument as in the previous cases so the same example works.

$\mathbf{m \geq 3}$:Let $n \equiv 3,7 \pmod 8$ and $g^{-1} = x + x^2 + x^3$. Then we have the following matrices by using the matrices (3.5), (3.11), (3.12), and example 5 in Table 3.2,

$$A_{g^{-1},n} = \begin{pmatrix} 0&1&0&*&0&*&0\\ 0&1&0&*&*&*&0\\ 0&0&0&*&0&*&0\\ 0&0&0&1&*&*&*\\ 0&0&0&0&0&*&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0 \end{pmatrix},\ A_{g^{-2^m}} = \begin{pmatrix} 0&0&0&0&0&1&0&*&0\\ 0&0&0&0&0&1&0&*&0\\ 0&0&0&0&0&0&0&*&0\\ 0&0&0&0&0&0&0&*&0\\ 0&0&0&0&0&0&0&*&0\\ 0&0&0&0&0&0&0&*&0\\ 0&0&0&0&0&0&0&*&0\\ 0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0 \end{pmatrix}. \tag{3.21}$$

for $m \geq 1$. Let $\vec{\nu} = (1,0,*,*,\ldots,*)$. Then $\vec{\nu}A_{g^{-1},n} = (0,1,0,*,*,*,*)$ and so $\vec{\nu_1} = (1,0,*,*,*,*)$. Now $\vec{\nu_1}A_{g^{-2},n+2} = (0,0,0,0,0,1,0,*,0)$ which implies $D(u_2) = n + e(1,n) + 2^4 - 7$ and $\vec{\nu_2} = (1,0,*,0,*,*,\ldots,*)$. From now we can apply Lemma 3.2.1 to get $D(g^{2^m}f^{-2^m}) = n + e(1,n) + 2^{m+2} - 7$ for $m \geq 2$. For the even values $n \equiv 4,6 \pmod 8$, the same example works because of the same argument as in the previous cases.

$\mathbf{n = 2}$:

Recall $\vec{\nu} = (s_0, s_1, s_2, \ldots)$. Without loss of generality, we may assume that $s_0 = 1$ because if $s_0 = 0$ then we are in the case $n = 3$. By using the matrix at (3.1),

$$\vec{\nu}A_{g^{-1},2} = (1,1,*,1,*,*,*,*) \Rightarrow D([u,g^{-1}]) \geq n+1 = n + e(1,n) + 1 = 3.$$

Let $u = x + x^3 + s_1x^4 + s_2x^5 + \cdots$. Then $u^2 = x + x^5 + s_1x^6 + x^7 + \cdots$ and so $\vec{\nu}_{1,u^2} = (0,1,s_1,1,*,*,\ldots,*)$. By the same technique as before, we can write the matrix $A_{u,3}$ to determine $[[u,g^{-1}],u]$

$$A_{u,3} = \begin{pmatrix} 1&0&1+s_2&0\\ 0&0&s_1&0\\ 0&0&1&0\\ 0&0&0&0 \end{pmatrix}. \tag{3.22}$$

Hence, $(1,1,*,1)A_{u,3} = (1,0,*,0)$, consequently we have $D([[u,g^{-1}],u]) \geq 5$ and $\vec{\nu}_{1,[[u,g^{-1}],u]} = (0,0,1,0,*,*,\ldots,*)$. Then by (3.13)

$$\vec{\nu_1} = (1,0,*,0,*,\ldots,*),\ D(u_1) = D(g^2f^{-2}) \geq n+1.$$

Now recall the matrix at (3.11). Then

$$\vec{\nu_1}A_{g^{-2},3} = (0,0,0,0,0,*,0,*,0) \Rightarrow D([u_1,g^{-2}]) \geq n+1+3+5 = n+e(1,n)+2^4-7.$$

By Lemma 3.1.1, $D(u_1^2) \geq 2(n+1)+5 = n+e(1,n)+2^4-7$ and $\vec{\nu}_{2,u_1^2} = (*,0,*,0,*,*,\ldots,*)$. On the other hand, by Proposition 1.2.3, $D([[u_1,g^{-2}],u_1]) \geq n+e(1,n)+2^4-7+4$. It follows that

$$\vec{\nu_2} = (*,0,*,0,*,*,\ldots,*),\ D(u_2) = D(g^{2^2}f^{-2^2}) \geq n+e(1,n)+2^4-7.$$

Now for $m \geq 2$, by Lemma 3.2.1,

$$D(g^{2^m}f^{-2^m}) \geq n+e(1,n)+2^{m+2}-7.$$

This is the required bound of Theorem 3.0.1.

Example for $\mathbf{n=2}$:

Let $g^{-1} = x+x^2+x^3$, use the matrix at (3.1) to get

$$A_{g^{-1},2} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1+s_2 & * & * \\ 0 & 0 & 1 & 0 & 1 & 1 & * & * \\ 0 & 0 & 1 & 0 & 1 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.23}$$

Let $\vec{\nu} = (1,0,1,0,1,0,*,*,\ldots,*)$, then $s_2 = 1$ and so

$$\vec{\nu}A_{g^{-1},2} = (1,1,1,1,0,0,*,*) \Rightarrow D([u,g^{-1}]) = n+1 = 3.$$

Then $[u,g^{-1}] = x+x^4+x^5+x^6+x^7+\gamma x^{10}+\ldots$. On the other hand, $u^2 = x+x^5+x^7+\alpha x^{10}+\ldots$ where $\alpha,\gamma \in \mathbb{F}_p$. Recall the matrix $A_{u,3}$ at (3.22). Since $(1,1,1,1)A_{u,3} = (1,0,1,0)$,

$[[u, g^{-1}], u] = x + x^6 + x^8 + \beta x^{10} + \dots$ where $\beta \in \mathbb{F}_p$. Direct substitution for the product $u^2[u, g^{-1}][[u, g^{-1}], u]$ gives

$$\vec{\nu_1} = (1, 0, 0, 0, 0, 0, *, *, \dots, *),\ D(u_1) = n + 1 = 3.$$

Recalling the second matrix at (3.21), we obtain

$$\vec{\nu_1} A_{g^{-2},3} = (0, 0, 0, 0, 0, 1, 0, *, 0, *, 0) \Rightarrow D([u_1, g^{-2}]) = n{+}1{+}3{+}5 = n{+}e(1, n){+}2^4{-}7 = 11$$

By Lemma 3.1.1, $D(u_1^2) \geq 2.3 + 7 = 13$ and $\vec{\nu}_{2,u_1^2} = (0, 0, *, 0, *, 0, *, *, \dots, *)$. On the other hand, $D([[u_1, g^{-2}], u_1]) \geq 11 + 3 + 5$ (we use the matrix at (3.11) as $D(u_1) = 3$ and it is in the same form of b^2). Thus

$$\vec{\nu_2} = (1, 0, *, 0, *, 0, *, \dots, *),\ \text{and } D(u_2) = 11 = n + e(1, n) + 2^4 - 7.$$

Now for $m \geq 2$, we can apply Lemma 3.2.1 to get

$$D(g^{2^m} f^{-2^m}) = n + e(1, n) + 2^{m+2} - 7.$$

So the bound is sharp in Theorem 3.0.1.

$\mathbf{n = 3, 4, 5, 6}$:

First we will find $\vec{\nu_2}$ and $D(u_2)$ for each of $n = 3, 4, 5, 6$.

$\mathbf{n = 3}$: Again, without loss of generality, let us assume that $s_0 = 1$ so $\vec{\nu} = (1, s_1, s_2, s_3, s_4, \dots)$. Recall the matrix at (3.2),

$$\vec{\nu} A_{g^{-1},3} = (0, r_2 + s_1, 0, *, r_2 + s_1 + s_3, *, r_2(1 + s_1 + s_3) + s_3) \Rightarrow D([u, g^{-1}]) \geq n{+}1{+}1 = 5.$$

By computing u^2 in a direct way, $\vec{\nu}_{1,u^2} = (0, 0, s_1, s_1, s_2 + s_3, 0, s_1 + s_2 + s_3 s_2 + s_5, *, *, \dots, *)$. We can write the matrix of u to compute $[[u, g^{-1}], u]$ using similar methods as before. We obtain

$$A_{u,5} = \begin{pmatrix} 0 & s_1 & 0 & s_3 + 1 \\ 0 & 1 & 0 & s_2 \\ 0 & 0 & 0 & s_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.24}$$

Then $\vec{\nu}_{1,[[u,g^{-1}],u]} = (0,0,0,0,s_1(r_2+s_1),0,*,*,\ldots,*)$. Thus, by (3.13),

$$\vec{\nu_1} = (r_2+s_1,0,*,r_2+s_3,*,r_2(1+s_1+s_3)+s_3,*,*,\ldots,*),\ D(u_1) = D(g^2f^{-2}) \geq n+2.$$

Recall the matrix at (3.11),

$$\vec{\nu_1}A_{g^{-2},5} = (0,0,0,\alpha_3(s_1+s_3),0,*,0,*,\alpha_3(r_2(1+s_1+s_3)+s_3))$$

$$\Rightarrow D([u_1,g^{-2}]) \geq n+2+3+3 = n+e(1,n)+2^4-9.$$

A direct computation gives the coefficient vector $\vec{\nu}_{2,u_1^2} = (0,0,(r_2+s_3)(r_2+s_1),0,*,r_2+s_3,*,*,\ldots,*)$. On the other hand, by Proposition 1.2.3, $D([[u_1,g^{-2}],u_1]) \geq n+e(1,n)+2^4-9+6$. Thus, by (3.13),

$$\vec{\nu_2} = (\alpha_3(s_1+s_3),0,*,0,*,r_2\alpha_3(1+s_1+s_3)+r_2+(\alpha_3+1)s_3,*,*,\ldots,*),$$

$$D(u_2) = D(g^{2^2}f^{-2^2}) \geq n+e(1,n)+2^4-9.$$

Note that $\alpha_3 r_2 = 0$ by (3.7). So we have $\vec{\nu_2} = (\alpha_3*,0,*,0,*,r_2*,*,*,\ldots,*)$.

$\mathbf{n=4}$: Recall the matrix at (3.3), then

$$\vec{\nu}A_{g^{-1},4} = (1,0,*,1+s_2,*,r_2(1+s_2)+s_2) \Rightarrow D([u,g^{-1}]) \geq n+1.$$

A direct computation gives $\vec{\nu}_{1,u^2} = (0,0,0,1,s_1,0,*,*,\ldots,*)$. On the other hand, $\vec{\nu}_{1,[[u,g^{-1}],u]} = (0,0,0,0,1,0,*,*,\ldots,*)$ by using the matrix at (2.5). Now by (3.13),

$$\vec{\nu_1} = (1,0,*,s_2,*,r_2(1+s_2)+s_2,*,*,\ldots,*),\ D(u_1) = D(g^2f^{-2}) \geq n+1.$$

Recall the matrix at (3.11)

$$\vec{\nu_1}A_{g^{-2},5} = (0,0,0,\alpha_3(1+s_2),0,*,0,*,\alpha_3(r_2(1+s_2)+s_2))$$

$$\Rightarrow D([u_1,g^{-2}]) \geq n+7 = n+e(1,k)+2^4-9.$$

On the other hand, $\vec{\nu}_{2,u_1^2} = (0,0,s_2,0,*,s_2,*,*,\ldots,*)$ by a direct computation, and $D([[u_1,g^{-2}],u_1]) \geq n+7+6$ by Proposition 1.2.3. Then by (3.13),

$$\vec{\nu_2} = (\alpha_3(1+s_2),0,*,0,*,\alpha_3(r_2(1+s_2)+s_2)+s_2,*,*,\ldots,*),\ D(g^{2^2}f^{-2^2}) \geq n+e(1,n)+2^4-9.$$

Again, since $r_2\alpha_3 = 0$ by (3.7), we have $\vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, r_2 *, *, *, \ldots, *)$.

$\mathbf{n = 5}$: Let $\vec{\nu} = (1, s_1, s_2, s_3, \ldots)$. Recall the first matrix at (3.4),

$$\vec{\nu} A_{g^{-2},5} = (0, r_2 + 1 + s_1, s_1, *, (r_2 + 1)s_1) \Rightarrow D([u, g^{-1}]) \geq n + 2 = 7.$$

On the other hand, by Proposition 1.2.3, $D(u^2) \geq 2.5 + 1 = 11$ and so $\vec{\nu}_{1,u^2} = (0, 0, 0, 0, *, *)$. Also $D([[u, g^{-1}], u]) \geq 7 + 6$ by Proposition 1.2.3. So by (3.13)

$$\vec{\nu_1} = (r_2 + s_1 + 1, s_1, *, (r_2 + 1)s_1, *, *, \ldots, *),\ D(g^2 f^{-2}) \geq n + 2.$$

Recall the matrix at (3.11), also recall $\alpha_3 = 1 + r_2$ from (3.7), then

$$\vec{\nu_1} A_{g^{-2},7} = (0, \alpha_3 s_1, 0, *, 0, *, \alpha_3 s_1, *, 0),\ D([u_1, g^{-2}]) \geq n+2+1+3 = 11 = n+e(1,n)+2^4-11.$$

$D(u_1^2) \geq 2.7 + 1 = 15$ and $\vec{\nu}_{2,u_1^2} = (0, 0, 0, 0, s_1 r_2, s_1, *, *, \ldots, *)$ by Proposition 1.2.3. So by Remark 3.2.1, we have

$$\vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, (\alpha_3 + 1)*, *, *, \ldots, *),\ D(g^{2^2} f^{-2^2}) \geq n + e(1, n) + 2^4 - 11.$$

By (3.7), $\alpha_3 = 1 + r_2$ so $\vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, r_2 *, *, *, \ldots, *)$.

$\mathbf{n = 6}$: Recall the matrix at (3.4), delete its first row and column to get the matrix $A_{g^{-1},6}$. Then

$$\vec{\nu} A_{g^{-1},6} = (1, 1, *, r_2 + 1, *, *, \ldots, *),\ D([u, g^{-1}]) \geq n + 1 = 7.$$

On the other hand, by using Proposition 1.2.3, $D(u^2) \geq 2.6 = 7+5$ and $D([[u, g^{-1}], u]) \geq 7+6$, so we have

$$\vec{\nu_1} = (1, 1, *, r_2 + 1, *, *, \ldots, *),\ D(g^2 f^{-2}) \geq n + 1 = 7.$$

Now recall the matrix at (3.11). Also recall that $\alpha_3 = 1 + r_2$, by (3.7), then

$$\vec{\nu_1} A_{g^{-2},7} = (0, \alpha_3, 0, *, 0, *, \alpha_3, *, 0),\ D([u_1, g^{-2}]) \geq n+1+3+1 = 11 = n+e(1,n)+2^4-11.$$

On the other hand, by Proposition 1.2.3 and a direct computation, $D(u_1^2) \geq 2.7 + 1$ and $\vec{\nu}_{2,u_1^2} = (0, 0, 0, 0, 1, 1)$. So by Remark 3.2.1,

$$\vec{\nu_2} = (\alpha_3, 0, *, 0, *, r_2, *, *, \ldots, *),\ D(g^{2^2} f^{-2^2}) \geq n + e(1, n) + 2^4 - 11.$$

We established that $\vec{\nu_2} = (\alpha_3 *, 0, *, 0, *, r_2 *, *, *, \ldots, *)$ for $n = 3, 4, 5, 6$. We also established that $D(u_2) \geq n + e(1, n) + 2^4 - K$ where $K = 9$ if $n = 3, 4$, and $K = 11$ if $n = 5, 6$.

Suppose that $\alpha_3 = 1$, then $\vec{\nu_2} = (*, 0, *, 0, *, 0, *, *, \ldots, *)$. By (3.7) and (3.8), the coefficient of the term x^{17} in b^{2^2} is zero. So the $(0, 6)$ entry of the matrix at (3.12) is zero for $m = 2$. So

$$\vec{\nu_2} A_{g^{-2^2}} = (0, 0, 0, 0, 0, 0, 0, *, 0, *, 0,),\ D([u_2, g^{-2^2}]) \geq n + e(1, n) + 2^5 - K + 2.$$

By Lemma 3.1.1, $D(u_2^2) \geq 2(n + e(1, n) + 2^4 - K) + 7 \geq n + e(1, n) + 2^5 - K + 2$ and $\vec{\nu}_{3,u_2^2} = (*, 0, *, 0, *, 0, *, *, \ldots, *)$. By Remark 3.2.1,

$$\vec{\nu_3} = (*, 0, *, 0, *, *, \ldots, *),\ D(g^{2^3} f^{-2^3}) \geq 29 = n + e(1, n) + 2^5 - K + 2.$$

Then for $m \geq 3$, we can apply Lemma 3.2.1 to get

$$D(g^{2^m} f^{-2^m}) \geq n + e(1, n) + 2^{m+2} - K + 2.$$

Now suppose that $\alpha_3 = 0$, then $\vec{\nu_2} = (*, 0, *, *, \ldots, *)$ and $D(u_2) \geq 13 = n + e(1, n) + 2^4 - K + 2$. On the other hand, by (3.8) and Lemma 3.1.1, the coefficients of the terms $x^{2^m - 4}$ in g^{-2^m} are always zero for $m \geq 2$. So the diagonal entries of $A_{g^{-2^m}}$ are zero. Thus,

$$\vec{\nu_2} A_{g^{-2^2}} = (0, 0, 0, 0, 0, *, 0, *, 0, *, 0),\ D([u_2, g^{-2^2}]) \geq 13 + 5 + 11 = n + e(1, n) + 2^5 - K + 2.$$

$D(u_2^2) \geq n + e(1, n) + 2^5 - K + 2$ and $\vec{\nu}_{3,u_2^2} = (*, 0, *, *, \ldots, *)$ by Lemma 3.1.1. Thus,

$$\vec{\nu_3} = (*, 0, *, *, \ldots, *),\ D(g^{2^3} f^{-2^3}) \geq n + e(1, n) + 2^5 - K + 2.$$

Here an inductive argument, very similar to Lemma 3.2.1, gives the bound $D(g^{2^m} f^{-2^m}) \geq n + e(1, n) + 2^{m+2} - K + 2$.

Hence, in general we have

$$D(g^{2^m} f^{-2^m}) \geq \begin{cases} n + e(1, n) + 2^{m+2} - K & \text{if } m = 2 \\ n + e(1, n) + 2^{m+2} - K + 2 & \text{if } m \geq 3 \end{cases} \tag{3.25}$$

where $K = 9$ if $n = 3, 4$, and $K = 11$ if $n = 5, 6$. This confirms the bound in Theorem 3.0.1.

Now we return to the sharpness of the bounds.

Examples for $\mathbf{n = 3, 4}$:

$\mathbf{m = 2}$: Let $n = 3$, $g^{-1} = x + x^2 + x^4$, and $\vec{\nu} = (1, 0, 0, 1, *, *, \ldots, *)$. By using the matrices at (3.2), (3.11), and example 3 in Table 3.2,

$$A_{g^{-1},3} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; A_{g^{-2},5} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & * & 0 & * & 0 \\ 0 & 1 & 0 & 0 & 1 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.26}$$

$\vec{\nu}A_{g^{-1},3} = (0, 0, 0, 1, 1, *, 1) \Rightarrow D([u, g^{-1}]) = n + 2 + 2$. By a direct calculation, $\vec{\nu}_{1,u^2} = (0, 0, 0, 0, 1, 0, *, *, \ldots, *)$. As $D([[u, g^{-1}], u]) \geq n + 4 + 4$ by Proposition 1.2.3, we have $\vec{\nu_1} = (0, 0, 1, 1, *, 1, *, *, \ldots, *)$ and so $D(u_1) = n + 4$. Now $\vec{\nu_1}A_{g^{-2},5} = (0, 0, 0, 1, 0, *, 0, *, 1)$, $D([u_1, g^{-1}]) = n+8 = n+e(1, n)+2^4-9$. As $D(u_1^2) \geq 2(n+4)+1 \geq n + 8 + 4$ and $D([[u_1, g^{-2}], u_1]) \geq n + 8 + 8$ by Proposition 1.2.3, we have

$$\vec{\nu_2} = (1, 0, *, 0, *, *, *, *, \ldots, *), \; D(u_2) = D(g^{2^2} f^{-2^2}) = n + 8 = n + e(1, n) + 2^4 - 9$$

as desired.

For $n = 4$, by using the matrix at (3.3), $\vec{\nu}A_{g^{-1},4} = (1, 0, 1, 1, *, 0, *, *, \ldots, *)$ and so $D([u, g^{-1}]) = n + 1$. By a direct computation we get the vector $\vec{\nu}_{1,u^2} = (0, 0, 0, 1, *, *, *, *, \ldots, *)$. On the other hand, $D([[u, g^{-1}], u]) \geq n+1+4$ by Proposition 1.2.3. Thus $\vec{\nu_1} = (1, 0, 1, 0, *, *, \ldots, *)$ and $D(u_1) = n + 1$. Now $\vec{\nu_1}A_{g^{-2},5} = (0, 0, 0, 1, 0, *, 0, *, *)$ and $D([u_1, g^{-2}]) = n + 7 = n + e(1, n) + 2^4 - 9$. On the other hand, $D(u_1^2) \geq n + 7 + 5$ by Lemma 3.1.1, and $D([[u_1, g^{-2}], u_1]) \geq n + 7 + 6$ by Proposition 1.2.3. Thus,

$$\vec{\nu_2} = (1, 0, *, 0, *, *, \ldots, *), \; D(u_2) = D(g^{2^2} f^{-2^2}) = n + e(1, n) + 2^4 - 9$$

as desired.

$\mathbf{m \geq 3}$: Let $n = 3$ and $g^{-1} = x + x^2 + x^3$. By using the matrix at (3.2) and example 1 in

Table 3.2,

$$A_{g^{-1},3} = \begin{pmatrix} 0 & 1 & 0 & * & 0 & * & 1 \\ 0 & 1 & 0 & * & * & * & 1 \\ 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The form of the matrix $A_{g^{-2^m}}$ for $m \geq 1$ is same as the second matrix at (3.21). Let $\vec{\nu} = (1, 0, 0, 0, *, *, \ldots, *)$ then $\vec{\nu}A_{g^{-1},3} = (0, 1, 0, *, 0, *, 1)$, and $D([u, g^{-1}]) = n + 2$. By Lemma 3.1.1, $D(u^2) \geq 2n + 5 \geq n + 2 + 6$. Also $D([[u, g^{-1}], u]) \geq n + 2 + 6$ (we use the matrix at (3.11) since u and b^2 are in same form). Thus, $\vec{\nu_1} = (1, 0, *, 0, *, 1, *, *, \ldots, *)$, $D(u_1) = n + 2 = 5$. Now $\vec{\nu_1}A_{g^{-2},5} = (0, 0, 0, 0, 0, 1, 0, *, 0)$, and $D([u_1, g^{-2}]) \geq n + 10 = n + e(1, k) + 2^4 - 7$. On the other hand, by Lemma 3.1.1, $D(u_1^2) \geq 2.(n + 2) + 5 = n + 10 + 2$ and $\vec{\nu}_{2,u_1^2} = (0, 0, *, 0, *, 0, *, *, \ldots, *)$. By Proposition 1.2.3, $D([[u_{,1}, g^{-2}], u_1]) \geq 13 + 6$. Thus

$$\vec{\nu_2} = (1, 0, *, 0, *, *, \ldots, *),\ D(u_2) = n + e(1, n) + 2^4 - 7.$$

Now by Lemma 3.2.1 $D(g^{2^m} f^{-2^m}) \geq n + e(1, n) + 2^{m+2} - 7$ for $m \geq 3$.

Let $n = 4$ and same g^{-1} and u as above. By using the first matrix at (3.3), we have $\vec{\nu}A_{g^{-1},4} = (1, 0, *, 1, *, 1)$. Also by a direct computation, $\vec{\nu}_{1,u^2} = (0, 0, 0, 1, 0, 0, *, *, \ldots, *)$. By using the matrix at (2.5) $\vec{\nu}_{1,[[u,g^{-1}],u]} = (0, 0, 0, 0, 1, 0, *, *, \ldots, *)$. All together imply $\vec{\nu_1} = (1, 0, *, 0, *, 1, *, *, \ldots, *)$ and $D(u_1) = 5$. From now everything is same as in the above example for the case $n = 3$, so we have the desired result.

Examples for $\mathbf{n = 5, 6}$:

$\mathbf{m = 2}$: Let $n = 5$ and $g^{-1} = x + x^2 + x^4 + x^5$ then the matrices $A_{g^{-1},5}$, $A_{g^{-2},7}$ are same as the matrices at (3.17). Let $\vec{\nu} = (1, 1, 0, 0, *, *, \ldots, *)$. Then $\vec{\nu}A_{g^{-1},5} = (0, 0, 1, 0, 1, *, *, \ldots, *)$ and $D([u, g^{-1}]) \geq n + 2 = 7$. On the other hand, $D(u^2) \geq 2n + 1 = n + 2 + 4$ and $D([[u, g^{-1}], u]) \geq n + 2 + 6$ by Proposition 1.2.3. Then $\vec{\nu_1} = (0, 1, 0, 1, *, *, \ldots, *)$. As $\vec{\nu_1}A_{g^{-2},7} = (0, 1, 0, 0, 0, *, 1, *, 0)$, $D([u_1, g^{-2}]) = n + 6 = n + e(1, n) + 2^4 - 11$. On the other hand, $D(u_1^2) \geq 2.(n + 3) = n + 6 + 5$ by Proposition 1.2.3. So we have $D(g^{2^2} f^{-2^2}) =$

$n + e(1,n) + 2^4 - 11$. For $n = 6$, very similar argument as in previous cases works for the same example of the case $n = 5$.

$\mathbf{m \geq 3}$: Let $n = 5$ and $g^{-1} = x + x^2 + x^3 + x^4 + x^7$. Note that the matrices $A_{g^{-1},5}$, $A_{g^{-2^m}}$ for $m \geq 2$ are same as the matrices at (3.18). Let $\vec{\nu} = (1, 1, *, *, \ldots, *)$. Then $\vec{\nu}A_{g^{-1},5} = (0, 1, 1, *, 0, *, *, \ldots, *)$ and so $D([u, g^{-1}]) = n + 2 = 7$. On the other hand, $D(u^2) \geq 2.n + 1 = n + 2 + 4$ and $D([[u, g^{-1}], u]) \geq n + 2 + 6$ by Proposition 1.2.3. So we have $\vec{\nu_1} = (1, 1, *, 0, *, *, \ldots, *)$ and $D(u_1) = n + 2 = 7$. Now $\vec{\nu_1}A_{g^{-2},7} = (0, 0, 0, 1, 0, *, 0, *, 0)$ so $D([u_1, g^{-2}]) = n + 8 = n + e(1,n) + 2^4 - 9$. On the other hand, by a direct computation $D(u_1^2) \geq 2.(n + 2) + 1 = n + 8 + 2$ and $\vec{\nu}_{2,u_1^2} = (0, 0, 1, 1, *, 0, *, *, \ldots, *)$. So we have $\vec{\nu_2} = (1, 0, *, 1, *, 0, *, *, \ldots, *)$ and $D(u_2) = 13 = n + e(1,n) + 2^4 - 9$ by Remark 3.2.1. As $\vec{\nu_2}A_{g^{-2^2}} = (0, 0, 0, 0, 0, 0, 0, *, 0)$, we have $D([u_2, g^{-2^2}]) \geq n + e(1,n) + 2^5 - 9 + 2 \geq 31$ and $\vec{\nu}_{3,[u_2,g^{-2^2}]} = (0, 0, *, 0, *, *, \ldots, *)$. By direct computation $D(u_2^2) = 29$ and $\vec{\nu}_{3,u_2^2} = (1, 0, *, 1, *, *, \ldots, *)$. Hence, we have $\vec{\nu_3} = (1, 0, *, 1, *, *, \ldots, *)$ and $D(u_3) = 29 = n + e(1,n) + 2^5 - 9$ by Remark 3.2.1. Let $\vec{\nu_m} = (1, 0, *, 1, *, *, \ldots, *)$ and $D(u_m) = n + e(1,n) + 2^{m+2} - 9$ for $m \geq 3$. Then $(1, 0, *, 1, *, *, \ldots, *)A_{g^{-2^m}} = (0, 0, 0, 0, 0, 0, 0, *, 0)$. It follows that $D([u_m, g^{-2^m}]) \geq n + e(1,n) + 2^{m+3} - 9 + 2$ and $\vec{\nu}_{m+1,[u_m,g^{-2^m}]} = (0, 0, *, 0, *, *, \ldots, *)$. On the other hand, it can be easily seen that $D(u_m^2) = n + e(1,n) + 2^{m+3} - 9$ and $\vec{\nu}_{m+1,u_m^2} = (1, 0, *, 1, *, *, \ldots, *)$. Hence, $\vec{\nu_{m+1}} = (1, 0, *, 1, *, *, \ldots, *)$ and $D(u_{m+1}) = n + e(1,n) + 2^{m+3} - 9$ by Remark 3.2.1. Therefore, an inductive argument gives $D(g^{2^m} f^{-2^m}) = D(u_{m-1}^2) = n + e(1,n) + 2^{m+2} - 9$ for $m \geq 3$. For $n = 6$, because of a same argument as in the previous cases, same example works to get the desired bound.

4

CHAPTER

On random generation in the Nottingham group

Being a pro-p group, the Nottingham group can be seen as a probabilistic space with respect to the normalized Haar Measure. Recall that $\mathcal{Q}(G,k)$ is the probability that k random elements generate an open subgroup of G for a profinite group G. A. Shalev [Sha00] pointed out that if $\mathcal{Q}(G,k)=1$ then the lower rank of G is at most k. It is still not known whether the converse is true or not. For $p \geq 3$, the lower rank of $\mathcal{N}$ is 2, see [Cam00]. Following the relation between the lower rank and the random generation, Shalev [Sha00] conjectured the following:

Conjecture 1. *Is* $\mathcal{Q}(\mathcal{N},2)=1$*?*

The aim of this chapter is to investigate Conjecture 1 above. We will use a probabilistic theorem which Szegedy [Sze05] used to prove that two random elements of the Nottingham group generate a free subgroup. By invoking the following two propositions, we follow his footsteps. The first is a corollary of a measure theoretical theorem from [Sak64, Chapter 4, (15.7)Theorem].

Proposition 4.0.1. *Let X be a metric space, and let P_n be a sequence of its partitions such that P_n consists of Borel sets, P_n is a refinement of P_{n-1}, and the diameter of the sets in P_n is at most d_n, where d_n tends to zero. Then for every Borel set $E \subset X$ and finite Borel measure μ, we have, for almost all $x \in E$*

$$\lim_{n\to\inf} \frac{\mu(E \cap H_n)}{\mu(H_n)} = 1$$

where $x \in H_n \subset P_n$.

Let G be a pro-p group, G_i be a filtration of G and μ be the Haar measure induced by the G-invariant metric $d(x,y) = inf\{|G : G_i|^{-1} : xy^{-1} \in G_i\}$. Recall that the balls in G with respect to this metric are the cosets of G_i, and the diameter of such balls is $|G : G_i|^{-1}$. Applying the proposition above to G, we get the following

Proposition 4.0.2. *Let E be a Borel set of G. Then for almost all $x \in E$,*

$$\lim_{n\to\infty} \frac{\mu(E \cap xG_n)}{\mu(xG_n)} = 1.$$

By using Proposition 4.0.2, the following conjecture implies Conjecture 1.

Conjecture 2. *Let $p \geq 5$. Then $\exists \epsilon > 0$, $\forall a, b \in \mathcal{N}$, $\exists L_0 \in \mathbb{N}$, $\forall L_1 \geq L_0$ the conditional probability $P(\langle au, bv\rangle \text{ open} | u, v \in \mathcal{N}_{L_1}) \geq \epsilon$.*

To be more precise, take two random elements of $\mathcal{N}$ which generate a possibly non-open subgroup. Then Conjecture 2 states that: changing the tales of those two elements, the new two ones will generate an open subgroup with positive probability, depending only on p. To test whether two elements generate an open subgroup we have the following result which follows from the proof of a theorem in [Cam00, Theorem 7].

Theorem 4.0.3. *Let $p \geq 5$. Let $a, b \in \mathcal{N}$ with $D(a) = q, D(b) = r$. Then $\langle a, b\rangle$ is an open subgroup of $\mathcal{N}$ whenever*

(a) $q \not\equiv \mp r \pmod p$,

(b) $q, r \not\equiv 0 \pmod p$, *and*

(c) q *and* r *are coprime.*

From now on assume that $p \geq 5$. The m-th Engel word (see Section 2.1) plays a significant role in our approach. Related to this, we will state a lemma that will be one of our main tools. The following definition is a generalization of a definition in [LGM02, Proposition 1.1.32].

Definition 4.0.4. Let G be a group, $x_1, x_2, \ldots, x_k \in G$, and $n, i_1, i_2, \ldots, i_k \in \mathbb{N}$. Define $\{x_1, x_2, \ldots, x_k\}^n_{i_1,i_2,\ldots,i_k}$ to be the set of all formal basic commutators in $\{x_1, x_2, \ldots, x_k\}$ of weight at least n, and of weight at least i_j in x_j where $i_j < n$ for $1 \leq j \leq k$.

Here we have another definition which can be seen as a generalized m-th Engel word.

Definition 4.0.5. Let $x, y, z \in G$ and $m \in \mathbb{N}$. We define the set

$$[x,_{\{m\}} \{y, z\}] = \Big\{ [x,_{i_1} y,_{j_1} z, \ldots ,_{i_k} y,_{j_k} z] \mid m = i_1 + j_1 + \ldots i_k + j_k,\ i_1, j_1, \ldots, i_k, j_k, k \in \mathbb{N} \Big\}.$$

Lemma 4.0.6. *Let G be a group and $a, b, u, v \in G$. Then*

$$[au,_m bv] = [a,_m b][u,_m b]X$$

where $X = \prod x_t$, the factors x_t's are different from $[a,_m b], [u,_m b]$ and one of the following types of commutators.

(T1) $x_t \in [X^{(i)},_{\{m-i\}} \{b, v\}]$, *where* $1 \leq i \leq m$ *and we have one of the following:*

- $X^{(i)} \in \{a^{(i)}, u, v, a, b\}^{i+1}_{1,0,0,0,0}$ *involving at least one* u *or* v *and* $a^{(i)} \in [a,_{\{i\}} \{b, v\}]$,
- $X^{(i)} \in \{u^{(i)}, u, v, a, b\}^{i+1}_{1,0,1,0,0}$ *where* $u^{(i)} \in [u,_{\{i\}} \{b, v\}]$.

(T2) $x_t \in [Y^{(i)},_{\{m-i\}} \{b, v\}]$ *where* $Y^{(i)} \in \{u, v, a, b\}^{2i+1}_{0,0,0,0}$ *for some* $1 \leq i \leq m$ *such that the sum of weights of* u *and* v *is at least* 2 *and the sum of the weights of* b *and* v *is at least* $2i - 1$.

Proof. Recall the following well known commutator identities, see [LGM02, Proposition 1.1.6]. Let $x, y, z \in G$. The conjugate of x by y is denoted by $x^y = y^{-1}xy$. Then

(i) $[xy, z] = [x, z][x, z, y][y, z] = [x, z]^y[y, z]$,

(ii) $[x, yz] = [x, z][x, y][x, y, z] = [x, z][x, y]^z$,

(iii) $xy = yx[x, y] = yx^y$.

By using the identities above, we can get the following generalization. Let $x_1, x_2, \cdots, x_n, y \in G$. Then

$$\begin{aligned}
[x_1 x_2 \cdots x_n, y] &= [x_1, y]^{x_2 \cdots x_n}[x_2 \cdots x_n, y] \\
&= [x_1, y]^{x_2 \cdots x_n}[x_2, y]^{x_3 \cdots x_n}[x_3 \cdots x_n, y] \\
&\vdots \\
&= [x_1, y]^{x_2 \cdots x_n}[x_2, y]^{x_3 \cdots x_n} \cdots [x_n, y].
\end{aligned}$$

On the other hand,

$$\begin{aligned}
[x_i, y]^{x_{i+1}\cdots x_n} &= [x_i, y][x_i, y, x_{i+1}\cdots x_n] \\
&= [x_i, y][x_i, y, x_{i+2}\cdots x_n][x_i, y, x_{i+1}]^{x_{i+2}\cdots x_n} \\
&\vdots \\
&= [x_i, y][x_i, y, x_n][x_i, y, x_{n-1}]^{x_n}[x_i, y, x_{n-2}]^{x_{n-1}x_n}\ldots[x_i, y, x_{i+1}]^{x_{i+2}\cdots x_n}.
\end{aligned}$$

Note that, by a similar technique above, any conjugate of $[x_i, y, x_j]$ is a product of $[x_i, y, x_j]$ and commutators involving $[x_i, y, x_j]$. Hence, we have the following identity

(iv) The commutator $[x_1x_2\cdots x_n, y]$ can be written as a product of $[x_i, y]$ for $i = 1, 2, \cdots n$ and product of commutators involving $[x_i, y, x_j]$ for all $1 \leq i < j \leq n$.

Back to the proof, we use induction on $m \geq 1$. Let $m = 1$. By using the identities (i), (ii) and (iii), we have

$$\begin{aligned}
[au, bv] &= [au, v][au, b][au, b, v] \\
&= [a, v][a, v, u][u, v][a, b][a, b, u][u, b]\big[[a, b][a, b, u][u, b], v\big]
\end{aligned}$$

Now we can bring forward $[a, b]$ and $[u, b]$ by using the commutator identity (iii). Then the rest, but the last factor $\big[[a, b][a, b, u][u, b], v\big]$, is a product of $[a, v], [a, v, u], [u, v], [a, b, u]$, which belong to $[X^{(i)}, _{\{m-i\}}\{b, v\}]$ for $i = 1, m = 1$ that involves at least one u or v, and product of commutators involving $[a, b], [u, b]$ which are again in $[X^{(i)}, _{\{m-i\}}\{b, v\}]$ for $i = 1, m = 1$ that involves at least one u or v. Now by using identity (iv), the last factor can be written as a product of commutators $[a, b, v], [a, b, u, v], [u, b, v]$ and commutators involving $[x_i, v, x_j]$ where $1 \leq i < j \leq 3$ and $x_1 = [a, b], x_2 = [a, b, u], x_3 = [u, b]$. So, they belong to $[X^{(i)}, _{\{m-i\}}\{b, v\}]$ for $i = 1, m = 1$ that involves at least one u or v. Let $m \geq 1$ and suppose that we have

$$[au, _m b] = [a, _m b][u, _m b]X$$

where X is a product of commutators of the types in $(T1)$ or $(T2)$. By using the identity (ii),

we have

$$\begin{aligned}[au_{,m+1}\, b] &= \big[[a_{,m}\, b][u_{,m}\, b]X, bv\big] \\ &= \big[[a_{,m}\, b][u_{,m}\, b]X, v\big]\big[[a_{,m}\, b][u_{,m}\, b]X, b\big]\big[[a_{,m}\, b][u_{,m}\, b]X, b, v\big].\end{aligned}$$

Let c be either b or v. Now by using identity (iv), similar to the case when $m = 1$, the above equality can be written as product of $[[a_{,m}\, b], v]$, $[[u_{,m}\, b], v]$, $[[a_{,m}\, b], b]$, $[[u_{,m}\, b], b]$, $[X, v]$, $[X, b]$, $[[a_{,m+1}\, b], v]$, $[[u_{,m+1}\, b], v]$, $[X, b, v]$ and commutators involving $[x_i, c, x_j]$ for $1 \leq i < j \leq 3$ where $x_1 = [a_{,m}\, b], x_2 = [u_{,m}\, b], x_3 = X$. Now we can bring $[a_{,m+1}\, b]$ and $[u_{,m+1}\, , b]$ to front by using the identity (iii) and the rest is product of commutators $[[a_{,m}\, b], v]$, $[[u_{,m}\, b], v]$, $[a_{,m+1}\, b, v]$, $[u_{,m+1}\, b, v]$, $[X, v]$, $[X, b]$, $[X, b, v]$, commutators involving $[a_{,m+1}\, b]$ or $[u_{,m+1}\, b]$ and commutators involving $[x_i, c, x_j]$ for $1 \leq i < j \leq 3$, where $x_1 = [a_{,m}\, b], x_2 = [u_{,m}\, b], x_3 = X$. It can be easily seen that all, but $[X, v]$, $[X, b]$ and $[X, b, v]$, are in $[X^{(i)}{}_{,\{m+1-i\}}\, \{b, v\}]$ where $i = m + 1$. Now by induction, $[X, c] = [\prod x_t, c]$, where x_t's are commutators from $[X^{(i)}{}_{,\{m-i\}}\, \{b, v\}]$ or $[Y^{(i)}{}_{,\{m-i\}}\, \{b, v\}]$ for some $1 \leq i \leq m$. By using the identity (iv), we can rewrite $[X, b]$ as product of commutators $[x_t, c]$ and commutators involving $[x_t, c, x_s]$ for $t < s$. It can be directly seen that the commutators $[x_t, c]$'s are from $[X^{(i)}{}_{,\{m+1-i\}}\, \{b, v\}]$ or $[Y^{(i)}{}_{,\{m+1-i\}}\, \{b, v\}]$. Since any x_t is of weight at least $m + 1$ and the sum of weights of b and v is at least m and involves at least one u or v, then the commutators $[x_t, c, x_s] \in \{u, v, a, b\}^{2(m+1)+1}_{0,0,0,0}$, the sum of weights of u and v is at least 2 and the sum of weights of b and v is at least $2(m + 1) - 1$. Hence, they belong to $[Y^{(i)}{}_{,\{m+1-i\}}\, \{b, v\}]$ for $i = m + 1$. It remains only to consider $[X, b, v]$. Again, by using identity (iv), similar arguments which we applied for $[X, c]$ will give the desired. □

§4.1 Random generation

Fix $a, b \in \mathcal{N}$. Theorem 4.0.3 allows us to divide our investigation into four cases.

Case 0. Assume that $\langle a, b\rangle$ is open.

Case 1. Assume that all elements of $\langle a, b\rangle$ has p-divisible depth.

Case 2. Assume that every element in $\langle a, b\rangle$ has depth congruent to k_0, $-k_0$, or 0 modulo p for some $k_0 \not\equiv 0 \pmod{p}$.

Case 3. Assume that the depths of all elements of $\langle a, b\rangle$ have a common divisor $r \not\equiv 0 \pmod{p}$, where $r > 1$.

Case 0:

Assume that $\langle a, b\rangle$ is open. Then $\exists L \not\equiv 0, p-1 \pmod{p}$ such that $\langle a, b\rangle \supseteq \mathcal{N}_L$. Now, $\exists w_1, w_2 \in \langle a, b\rangle$ such that $D(w_1) = L \neq 0$ and $D(w_2) = L + 1$. Let $L_0 = L + 2$ then for any $u, v \in \mathcal{N}_{L_0}$, $w_1' := w_1(au, bv) \equiv w_1(a, b) \pmod{x^{L_0+1}}$ and $w_2' := w_2(au, bv) \equiv w_2(a, b) \pmod{x^{L_0+1}}$. Now, by Theorem 4.0.3, $\langle w_1', w_2'\rangle$ is open with probability 1. Hence, for all $L_1 \geq L_0$, $P(\langle au, bv\rangle \textit{ open} \mid u, v \in \mathcal{N}_{L_1}) = 1$.

Case 1:

Assume that all elements of $\langle a, b\rangle$ has p-divisible depth.

Proposition 4.1.1. *Let $L_0 \geq D(a), D(b)$. Then for all $L_1 \geq L_0$ the probability that $\langle au, bv\rangle$, for $u, v \in \mathcal{N}_{L_1}$, has an element whose depth is not divisible by p is $\geq \frac{1}{p^3}$.*

Proof. Pick $n, m \in \mathbb{N}$ such that:

1. $n \not\equiv 0 \pmod{p}$,
2. $pm > n - D(a)$.

Consider the commutator $[a,_m b]$. Since all elements in $\langle a, b\rangle$ has p-divisible depths, we have

$$D([a,_m b]) \geq D(a) + mD(b) + pm \tag{4.1}$$

for all $m \geq 0$. Indeed, note that $D([a, b]) = D(a) + D(b) + d$ for some $d \geq 1$. Since $D([a, b]), D(a), D(b) \equiv 0 \pmod{p}$, $d \equiv 0 \pmod{p}$. On the other hand, we have $d \geq 1$, so $d \geq p$. Hence, we have $D([a, b]) \geq D(a) + D(b) + p$. Now applying an inductive proof will give

$$D([a,_m b]) \geq D(a) + mD(b) + mp.$$

Let $u, v \in \mathcal{N}_{L_0}$ such that $D(u) = n$ and $D(v) > n$. Since $D(u) \not\equiv 0 \pmod{p}$, we have

$$D([u,_m b]) = D(u) + mD(b).$$

Consider the commutator $[au,_m bv]$. By Lemma 4.0.6, we have

$$[au,_m bv] = [a,_m b][u,_m b]X,$$

where X is the product of commutators of the types in $(T1)$ or $(T2)$. It can be seen easily that all factors of X has depth $> D(u) + mD(b) = D([u,_m b])$. Indeed, let us first consider the factors in X involving at least one u, but no v. Then the weight of such a factor is at least $m+2$ and the weight of b is at least m. Consequently, any such factor has depth strictly larger than $D(u) + mD(b)$. Now, consider the factors which do not involve any u, but at least one v. The weight of such a factor is at least $m + 1$ and the sum of weights of b and v is at least m. It follows that any such factor has depth strictly larger than $D(u) + mD(b)$ as $D(v) > n = D(u)$. Since $pm > n - D(a)$, we have

$$D([a,_m b]) \geq D(a) + mD(b) + pm > D(u) + mD(b) = D([u,_m b]).$$

Thus, we have

$$D([au,_m bv]) = D([u,_m b]) = D(u) + mD(b) \not\equiv 0 \pmod{p}.$$

Since $n \not\equiv 0 \pmod{p}$, $L_0 \leq D(u) \leq L_0 + 1$. Also $D(v) > D(u)$. So, for all $L_1 \geq L_0$, the probability that $\langle au, bv \rangle$, for $u, v \in \mathcal{N}_{L_1}$, has an element whose depth is not divisible by p is $\geq \frac{1}{p^3}$. □

We have the following corollary of Proposition 4.1.1.

Corollary 4.1.2. *The probability that two randomly chosen elements from $\mathcal{N}$ generates a subgroup where every element has depth divisible by p is zero.*

By Corollary 4.1.2, we need to consider Case 0, Case 2 or Case 3 only.

Case 2 and Case 3:

The following observation is important for the investigations of Case 2 and Case 3.

Remark 4.1.1. Let $w \in \mathcal{N}$ with $D(w) = k \equiv k_0 \pmod p$ for $0 < k_0 < p-1$. Let $n \geq k$ be such that $n \not\equiv 2k - i \pmod p$ for any $0 \leq i \leq k_0$. So, $n+1 \not\equiv 2k - j \pmod p$ for any $0 \leq j < k_0$ and, therefore, $e(k, n+1) = k_0$ and $e(k, n) = k_0$, see Definition 2.0.1. By Lemma 2.1.4, we have $\Pi_{p-1,n+1,k_0} = 0$ and $\Pi_{p-1,n,k_0} = 0$. On the other hand, recall that we obtain $\Pi_{p-1,n+1,k_0}$ by deleting the first column and row of Π_{p-1,n,k_0+1}. Recall also the definition of the corresponding matrix $T_n[\lambda]$ from Section 2.2. Now we have $T_n[\lambda] = (\lambda)$ (note that λ is the entry at the top right corner of Π_{p-1,n,k_0+1} which may or may not be zero) for the element w with respect to the integer n.

Based on the above, we make the following definition. Let $w \in \mathcal{N}$ with $D(w) = k \equiv k_0 \pmod p$ where $0 < k_0 < p-1$. We define the following property for w:

$(P*)$ $\exists n_0 \not\equiv 0, -k_0, 2k-i \pmod p$ for $i = 0, 1, \ldots, k_0$ such that $T_n[\lambda] = (\lambda) \neq 0 \; \forall n \equiv n_0 \pmod p$.

Remark 4.1.2. Let $c, d \in \mathcal{N}$ be such that $D(c) \geq D(d)$, $D(c) \equiv D(d) \equiv k_0 \pmod p$ for some $0 < k_0 \leq p-1$ and the first $k_0 + 1$ coefficients of c and d are same. Then they have the same matrix $\Pi_{p-1,n,e(k,n)+1}$ for any $n \geq D(c), D(d)$. So, for any $D(u) \geq n \geq D(c), D(d)$, by Lemma 2.1.6, we have

$$D(u^{(s(p-1))}) \geq n + s(p-1)D(d) + (s-1)k_0$$

where $u^{(s(p-1))} \in [u, _{\{s(p-1)\}} \{c, d\}]$.

Case 2

Assume that every element in $\langle a, b \rangle$ has depth congruent to k_0, $-k_0$, or 0 modulo p for some $0 < k_0 < p-1$.

Proposition 4.1.3. *Let $L_0 \geq D(a), D(b)$. Suppose that $D(a) > D(b) + k_0$ where $D(b) = k \equiv k_0 \pmod p$, $k_0 \not\equiv (p-3)^{-1}, 2(p-3)^{-1}, (p-2)^{-1}, 2(p-2)^{-1}, 2(p-1) \pmod p$ and b has*

property $(P*)$. *Then for all* $L_1 \geq L_0$

$$P(\langle au, bv\rangle \text{ open} \mid u, v \in \mathcal{N}_{L_1}) \geq \frac{1}{p^{2p^2+9p}}$$

Before proving Proposition 4.1.3, we need the following result.

Lemma 4.1.4. *Let* $c, d \in \langle a, b\rangle$ *be such that* $D(d) \equiv k_0 \pmod p$ *and* $k_0 \not\equiv (p-3)^{-1}, 2(p-3)^{-1}, (p-2)^{-1}, 2(p-2)^{-1}, 2(p-1) \pmod p$. *Then*

$$D([c,_{3m}\, d]) \geq D(c) + 3mD(d) + 3m$$

for all $m \geq 0$.

Proof. We use induction on $m \geq 0$. For $m = 0$ the claim is trivial. Now suppose that $D([c,_{3m}\, d]) \geq D(c)+3mD(d)+3m$ for $m > 0$. We have $D([c,_{3m}\, d]) \equiv k_0, -k_0, \text{ or } 0 \pmod p$.

If $D([c,_{3m+1}\, d]) \equiv k_0 \pmod p$ then $D([c,_{3m+2}\, d]) = D([c,_{3m+1}\, d]) + D(d) + l \equiv k_0, -k_0, 0 \pmod p$, where $l \geq 1$. Then $l \equiv -k_0, -3k_0, -2k_0 \pmod p$. Since $k_0 \not\equiv (p-3)^{-1}, 2(p-3)^{-1}, (p-2)^{-1}, 2(p-2)^{-1}, 2(p-1), (p-1), 0 \pmod p$, we have $l \geq 3$. Hence, by induction hypothesis, $D([c,_{3m+3}\, d]) \geq D([c,_{3m+2}\, d])+D(d) \geq D([c,_{3m+1}\, d])+2D(d)+l \geq D(c) + (3m+3)D(d) + 3m + 3$.

If $D([c,_{3m+1}\, d]) \equiv -k_0 \pmod p$ then $D([c,_{3m}\, d]) \equiv k_0 \pmod p$. Hence, $D([c,_{3m+1}\, d]) \geq D([c,_{3m}\, d]) + D(d) + 3$ because of our assumptions on k_0 as in the previous case. Therefore, $D([c,_{3m+3}\, d]) \geq D([c,_{3m+1}\, d]) + 2D(d) \geq D([c,_{3m}\, d]) + 3D(d) + l \geq D(c) + (3m+3)D(d) + 3m + 3$ by induction.

If $D([c,_{3m+1}\, d]) \equiv 0 \pmod p$ then $D([c,_{3m+2}\, d]) = D([c,_{3m+1}\, d]) + D(d) \equiv k_0 \pmod p$. So, $D([c,_{3m+3}\, d]) = D([c,_{3m+2}\, d]) + D(d) + 3$ because of the same reason in the above cases. Hence, by induction, $D([c,_{3m+3}\, d]) \geq D([c,_{3m+2}\, d]) + D(d) + 3 \geq D([c,_{3m}\, d]) + 3D(d) + 3 \geq D(c) + (3m+3)D(d) + 3m + 3$. □

Note that for $m' \geq 1$ the proof above shows that

$$D([a,_{m'}\, b]) \geq D(a) + m'D(b) + m' - l',$$

where $m' \equiv l' \pmod 3$ and $0 \leq l' \leq 2$. Also if we don't have the conditions $k_0 \not\equiv (p-3)^{-1}, 2(p-3)^{-1}, (p-2)^{-1}, 2(p-2)^{-1}, (p-1), 2(p-1), (p-1), 0 \pmod p$ then the proof above ensures that for $m \geq 0$ we still have

$$D([a,_{3m}\, b]) \geq D(a) + 3mD(b) + m.$$

Proof of Proposition 4.1.3. By assumption, b has property $(P*)$. So $\exists n_0$ such that $n_0 \not\equiv 0, -k_0, 2k-i \pmod p$ for any $0 \leq i \leq k_0$ and $T_n[\lambda] = (\lambda) \neq 0\ \forall n \equiv n_0 \pmod p$. Let $\gcd(k, k_0) = k_1$. Since p and k_1 are coprime then $\exists n \equiv n_0 \pmod p$, $n \equiv 1 \pmod{k_1}$ and $L_0 + k_0 + 1 \leq n < L_0 + k_0 + 1 + pk_1$ by using the Chinese Remainder Theorem. Note that

$$n \equiv 1 \pmod{k_1} \Rightarrow k_1 \mid 1-n \Rightarrow \exists\ s' \text{ such that } s'k_0 \equiv 1-n \pmod k$$

$$\Rightarrow n + s'k_0 \equiv 1 \pmod k \Rightarrow \gcd(n+s'k_0, k) = 1 \Rightarrow \gcd(n + s'(p-1)k + s'k_0, k) = 1$$

$$\Rightarrow \gcd(n + s(p-1)k + sk_0, k) = 1\ \forall s \equiv s' \pmod k.$$

Now pick $s \equiv s' \pmod k$ such that $n - D(a) < s(p-1) - sk_0 - l$ where $s(p-1) \equiv l \pmod 3$ and $0 \leq l \leq 2$. Let $u, v \in \mathcal{N}$ be such that $D(u) = n$, $D(v) > n + 5k_0 + p$, $D(v) \equiv k_0 \pmod p$ and first $k_0 + 1$ coefficients of v are same as first $k_0 + 1$ coefficients of b. Since b has property $(P*)$, $D([u,_{s(p-1)}\, b]) = n + s(p-1)k + sk_0$ by the proof of Theorem 2.0.3 in Chapter 2. Now consider the commutator $[au,_{s(p-1)}\, bv]$. Set $m = s(p-1)$. By Lemma 4.0.6 we have

$$[au,_m\, bv] = [a,_m\, b][u,_m\, b]X,$$

where X is the product of commutators of type $(T1), (T2)$. Let X' be the factor which has the minimum depth. Now X' is from $[X^{(i)},_{\{m-i\}}\, \{b, v\}]$, or $[Y^{(i)},_{\{m-i\}}\, \{b, v\}]$ for some $1 \leq i \leq m$. Recall that we have one of the following

- $X^{(i)} \in \{a^{(i)}, u, v, a, b\}^{i+1}_{1,0,0,0}$ which involves at least one u or v and $a^{(i)} \in [a,_{\{i\}}\, \{b, v\}]$,
- $X^{(i)} \in \{u^{(i)}, u, v, a, b\}^{i+1}_{1,0,1,0}$ where $u^{(i)} \in [u,_{\{i\}}\, \{b, v\}]$.

Recall also, $Y^{(i)} \in \{u, v, a, b\}^{2i+1}_{0,0,0,0}$ is such that the sum of weights of u and v is at least 2 and the sum of weights of b and v is at least $2i-1$. Suppose that X' is in $[X^{(i)},_{\{m-i\}}\, \{b, v\}]$ for some $1 \leq i \leq m$ and $X^{(i)} \in \{a^{(i)}, u, v, a, b\}_{i+1,1,0,0,0}$ which involves at least one u or v and

$a^{(i)} \in [a,_{\{i\}} \{b, v\}]$. Suppose that $D(X^{(i)})$ involves at least one v, but does not involve any u. Set $m = s(p-1)$. Now we have either $a^{(i)} \in [a,_{j-1} b, v,_{\{i-j\}} \{b, v\}]$ for some $1 \leq j \leq i$ or $a^{(i)} = [a,_i b]$. Suppose that we have the former. Let $j-1 = s_1(p-1)+i_1$, $i-j = s_2(p-1)+i_2$ and $m - i = s_3(p-1) + i_3$. Then $s \geq s_1 + s_2 + s_3 \geq s - 2$. Now, by Remark 4.1.2 and Lemma 4.1.4,

$$\begin{aligned} D(X^{(i)}) &\geq D([a,_{j-1} b]) + D(v) + (i-j)D(b) + (s_2 - 1)k_0 \\ &\geq D(a) + (j-1)D(b) + j - 3 + D(v) + (i-j)D(b) + (s_2 - 1)k_0 \\ &\geq D(a) + (i-1)D(b) + D(v) + j - 3 + (s_2 - 1)k_0. \end{aligned}$$

Then

$$\begin{aligned} D(X) \geq D(X') &\geq D(a) + (j-1)D(b) + j - 3 + D(v) + (i-j)D(b) \\ &\quad +(s_2 - 1)k_0 + (m-i)D(b) + (s_3 - 1)k_0 \\ &> n + mD(b) + (s_1 + s_2 + s_3 - 2 + 5)k_0 \\ &\geq n + s(p-1)k + sk_0 \end{aligned}$$

by Remark 4.1.2, Lemma 4.1.4, $D(a) > D(b) + k_0$, $D(v) > n + 5k_0 + p$ and $k_0 < p - 1$. Now suppose that we have the latter or $X^{(i)}$ involves at least one u, but does not involve any v. Then in both cases we have $a^{(i)} = [a,_i b]$ and $X^{(i)}$ involves at least one u or v. So

$$D(X^{(i)}) \geq D(a) + iD(b) + i - 2 + D(u)$$

by using Lemma 4.1.4. Let $i = t_1(p-1)+l_1$ and $m-i = t_2(p-1)+l_2$ for some $l_1, l_2 < p-1$. Then $s \geq t_1 + t_2 \geq s - 1$. Now similar to above, by Remark 4.1.2, $D(a) > D(b) + k_0$ and $k_0 < p - 1$, we have

$$\begin{aligned} D(X) &\geq D(a) + iD(b) + i - 2 + D(u) + (m-i)D(b) + (t_2 - 1)k_0 \\ &> n + mD(b) + (t_0 + t_2 + 1)k_0 \geq n + s(p-1)k + sk_0. \end{aligned}$$

Now for the other possible forms of X', by using very similar techniques above, we can show that $D(X') > n + s(p-1)k + sk_0$ and, so, $D(X) > n + s(p-1)k + sk_0 = D([u,_{s(p-1)} b])$.

Thus

$$[au,_m bv] \equiv [a,_m b][u,_m b] \pmod{x^{D([u,_m b])+2}}.$$

Recall that we picked s such that $n - D(a) > s(p-1) - sk_0 - l$. Then by Lemma 4.1.4,

$$\begin{aligned} D([a,_{s(p-1)} b]) &\geq D(a) + s(p-1)D(b) + s(p-1) - l \\ &> n + s(p-1)k + sk_0 = D([u,_{s(p-1)} b]). \end{aligned}$$

Finally,

$$D([au,_{s(p-1)} bv]) = D([u,_{s(p-1)} b]) = n + s(p-1)k + sk_0.$$

Now $D(bv) = k$ and $D([au,_s (p-1)bv]) = n + s(p-1)k + sk_0$ satisfy the conditions of Theorem 4.0.3. It follows that the subgroup $\langle [au,_{s(p-1)} bv], bv\rangle$ is open with probability $\geq \frac{1}{p^{p^2+p^2+p+5k_0+p+p+k_0+1}} \geq \frac{1}{p^{2p^2+9p}}$ as $L_0 \leq D(u) \leq L_0+k_0+pk_1+1 \leq L_0+p^2$, $D(v) > n+5k_0+p$, $D(v) \equiv k_0 \pmod{p}$ and its first k_0+1 coefficients are same as the first k_0+1 coefficients of b. Hence, $\langle au, bv\rangle$ is open with probability $\geq \frac{1}{p^{2p^2+9p}}$. □

Case 3:

Assume that the depths of all elements of $\langle a, b\rangle$ have a common divisor $r \not\equiv 0 \pmod{p}$ where $r > 1$ and there exist elements whose depths are not divisible by p (to distinguish Case 1 and Case 3).

Proposition 4.1.5. *Let $L_0 \geq D(a), D(b)$. Suppose that $D(a) > D(b) + k_0$ where $D(b) = k \equiv k_0 \pmod{p}$, $0 < k_0 < p-1$, $k_0 < r$ and b has the property $(P*)$. Then, for all $L_1 \geq L_0$,*

$$P(\langle au, bv\rangle \text{ open} \mid u, v \in \mathcal{N}_{L_1}) \geq \frac{1}{2p^2+9p}.$$

To prove the above proposition we need the following lemma.

Lemma 4.1.6. *Let $c, d \in \langle a, b\rangle$ be such that $D(d) \not\equiv 0 \pmod{p}$. Then*

$$D([c,_{s(p-1)+1} d]) \geq D(c) + (s(p-1)+1)D(d) + sr$$

for all $s \geq 0$.

Proof. We will prove by induction on $s \geq 0$. For $s = 0$, we have the claim. Now suppose that we have

$$D([c,_{s(p-1)+1}\, d]) \geq D(c) + (s(p-1)+1)D(d) + sr$$

for $s \geq 0$. Note that there exists $1 \leq i \leq p-1$ such that $D([c,_{s(p-1)+i}\, d]) \equiv D(d),$ or $2D(d) \pmod p$. If $D([c,_{s(p-1)+i}\, d]) \equiv 2D(d) \pmod p$ then $D([c,_{s(p-1)+i-1}\, d]) \equiv D(d) \pmod p$. It follows that

$$D([c,_{s(p-1)+i}\, d]) \geq D([c,_{s(p-1)+i-1}\, d]) + D(d) + pr$$

$$\Rightarrow\ D([c,_{(s+1)(p-1)+1}\, d]) \geq D(c) + ((s+1)(p-1)+1)D(d) + (s+1)r$$

by induction. Now if $D([c,_{s(p-1)+i}\, d]) \equiv D(d) \pmod p$ then

$$D([c,_{s(p-1)+i+1}\, d]) \geq D([c,_{s(p-1)+i}\, d]) + D(d) + r$$

$$\Rightarrow\ D([c,_{(s+1)(p-1)+1}\, d]) \geq D(c) + ((s+1)(p-1)+1)D(d) + (s+1)r$$

by induction. □

Let us turn to the proof of Proposition 4.1.5.

Proof of Proposition 4.1.5. The proof is very similar to the proof of Proposition 4.1.3. Indeed, since b has the property $(P*)$, $\exists n_0$ such that $n_0 \not\equiv 2k-i$ for any $0 \leq i \leq k_0$ and $T_n(\lambda) = (\lambda) \neq 0\ \forall n \equiv n_0 \pmod p$. Let $k_1 = \gcd(k, k_0)$. Now, by the Chinese Remainder Theorem, we can pick n such that $n \equiv n_0 \pmod p$, $n \equiv 1 \pmod{k_1}$, and $L_0 + k_0 + 1 \leq n < L_0 + k_0 + pk_1 + 1$. Also by using a very similar argument to the proof of Proposition 4.1.3, we can pick s such that $\gcd(n + s(p-1)k + sk_0, k) = 1$ and $n - D(a) < s(r - k_0) - r$ (note that we could guarantee the second condition on s because $k_0 < r$). Then, by applying the same argument as in the proof of Proposition 4.1.3, we get

$$[au,_{s(p-1)}\, bv] \equiv [a,_{s(p-1)}\, b][u,_{s(p-1)}\, b] \pmod{x^{[u,_{s(p-1)}b]+2}},$$

where $u, v \in \mathcal{N}_{L_0}$ such that $D(u) = n$, $D(v) > n + 5k_0 + p$, $D(v) \equiv k_0 \pmod p$ and the first

$k_0 + 1$ coefficients of b and v are same. On the other hand, by Lemma 4.1.6,

$$D([a,_{s(p-1)}\, b]) \geq D(a) + s(p-1)D(b) + (s-1)r$$

since $s(p-1) = (s-1)(p-1) + 1 + p - 2$. Then

$$D([a,_{s(p-1)}\, b]) > D([u,_{s(p-1)}\, b]) = D(u) + s(p-1)D(b) + sk_0,$$

since $n - D(a) < s(r - k_0) - r$. Hence, we have the commutator $[au,_{s(p-1)}\, bv]$ whose depth is as desired. So, $[au,_{s(p-1)}\, bv]$ and bv generate an open subgroup with probability $\geq \frac{1}{p^{2p^2+9p}}$ by Theorem 4.0.3 . Therefore, $\langle au, bv\rangle$ is open with probability $\geq \frac{1}{p^{2p^2+9p}}$. □

§4.2 Discussions

Here, we argue some possible developments of some techniques that seem promising for a complete confirmation of Shalev's conjecture. It is worth to mention that we have tried to accomplish these ideas, but we could not be successful for the time being. This is due to several difficulties that we will also discuss below.

First idea: Suppose that neither a nor b has the property $(P*)$. We argue a possible extension of our work for the case when there exists a word in a, b having property $(P*)$. So, suppose that there exists $w_2(a,b) \in \langle a, b\rangle$ that has the property $(P*)$. Let $w_1(a,b) \in \langle a, b\rangle$ be such that $D(w_1) \geq D(w_2)$ and $L_0 \geq D(w_1), D(w_2)$. Then for any $u, v \in \mathcal{N}_{L_0+1}$, we have

$$w_1(au, bv) \equiv w_1(a,b)\bar{u} \quad \text{for some } \bar{u} \in \mathcal{N}_{L_0+1}$$
$$w_2(au, bv) \equiv w_2(a,b)\bar{v} \quad \text{for some } \bar{v} \in \mathcal{N}_{L_0+1}$$

Now, the idea is to apply the arguments used in the previous section for the commutator $[w_1\bar{u},_{s(p-1)}\, w_2\bar{v}]$. For that, we need to guarantee that these new $\bar{u}$ and $\bar{v}$ satisfy the desired properties. Here is the difficulty. For instance, it is not easy to guarantee that $D(\bar{u})$ is co-prime to $D(w_2)$ and, $D(\bar{v})$ is congruent to $D(w_2)$ and both of them share the first $k_0 + 1$ coefficients (here $D(w_2) \equiv k_0 \pmod p$ and $0 < k_0 < p - 1$). Although, we are free only to prescribe the first few coefficients of u and v, we do not know yet how to choose these first few coefficients

in a way that $\bar{u}$ and $\bar{v}$ have the desired properties. A careful analysis on how the coefficients of u and v affect the new word $w(au, bv)$ might help us to overcome the difficulties above. But certainly, this is intricate.

To accomplish this idea and to give a complete answer of Shalev's conjecture, we need also to consider the following question. Why would there exist a $w(a, b) \in \langle a, b\rangle$ with property $(P*)$?

Second idea: This idea depends on the role of the growth rate of certain Engel words. First, we introduce the concept of the speed of an element, which is interesting in its own. From now on, k_0 denotes the smallest non-negative residue (modulo p) of the depth of a given element b.

Definition 4.2.1. Let $a, b \in \mathcal{N}$. Define

$$s(a, b) = \liminf_{m\to\infty} \frac{D([a,_m\, b]) - D(a) - mD(b)}{m}$$

the relative speed of b with respect to a. Now define

$$s(b) = \min_{a\in\mathcal{N}} s(a, b)$$

the speed of b.

Now, we consider the relative speed of b in a subgroup of $\mathcal{N}$ that contains b. Let $H \leq \mathcal{N}$ be such that $b \in H$. Then the relative speed of b in H is given by

$$s_H(b) = \min_{a\in H} s(a, b).$$

Lemma 2.1.6 is essentially saying that $s(b) \geq \frac{k_0}{p-1}$ for any $b \in \mathcal{N}$. Moreover, in Lemma 2.2.2 and Lemma 2.2.3, we could construct b such that $s(b) = \frac{k_0}{p-1}$. In fact, we can tell more. Recall the constant $e'(k, n)$ (see Section 2.2). By Corollary 2.1.5 if the top right hand $e'(k, n) + 1 \times e'(k, n) + 1$ block of $\Pi_{p-1,n,e(k,n)+e'(k,n)+1}$ (for $k + k_0 + pt < n < k + p(t + 1)$ where $k = D(b)$ and t is some nonnegative integer) is a non-nilpotent matrix then $s(b) = \frac{k_0}{p-1}$. So, our condition on b, having the property $(P*)$, in Case 2 and Case 3 implies that $s(b) = \frac{k_0}{p-1}$.

Now let us compare $s_H(b)$ with $s(b)$, where H is a non-open subgroup that contains b. For example, in Case 1, $s_{\langle a,b\rangle}(b) \geq p$ while $s(b) = 0$. In Case 2 and Case 3, $s_{\langle a,b\rangle}(b) \geq 1$ and $\frac{r}{p-1}$,

respectively, while $s(b) = \frac{k_0}{p-1}$. As we see in all cases $s(b)$ is small compared with $s_{\langle a,b\rangle}(b)$. In fact, $s(b)$ is the minimum possible. However, there might be many elements whose speeds are strictly larger than $\frac{k_0}{p-1}$. For example, if $b \equiv x + x^{k+1} \pmod{x^{k+2k_0+2}}$, where $k_0 \not\equiv 0 \pmod{p}$, then the associated matrix mentioned above is nilpotent. So, very possibly, $s(b) > \frac{k_0}{p-1}$. Here, we may try to find the exact speed of b by extending the matrix at (2.5) more carefully. But, even if we find the exact speed, we may face another problem in comparing $s(b)$ with $s_{\langle a,b\rangle}(b)$. To overcome this problem, we may try to determine the best possible minimum relative speed of an element, whose depth is not divisible by p, in a non-open subgroup. Essentially, Lemma 4.1.4 and Lemma 4.1.6 imply that the relative speed of an element in a non-open subgroup must be $\geq \frac{2}{p-1}$ if $p \geq 7$ (remember that without the conditions on k_0 in the statement of Lemma 4.1.4, we have $s_{\langle a,b\rangle}(d) \geq \frac{1}{3}$). In fact, this gives us the following proposition.

Proposition 4.2.2. *Let $a, b \in \mathcal{N}$. If there exists $c \in \langle a, b\rangle$ such that $D(c) \not\equiv 0 \pmod{p}$ and $s_{\langle a,b\rangle}(c) < \frac{2}{p-1}$ then $\langle a, b\rangle$ is open.*

Here, we have to mention that even if we can overcome the above problems, the real challenge is how to use this to confirm the conjecture under the question. In more details, even if we have $s(b) < s_{\langle a,b\rangle}(b)$, can we find a $u \in \mathcal{N}$ such that $s(u, b) = s(b)$ with $D([u,_m b])$ and $D(b)$ satisfying the criterion given in Theorem 4.0.3?

We also note that the concepts of speed and relative speed are interesting in their own, and worth investigating. We are hoping that a future work on this topic would give better understanding and would result in solving Shalev's conjecture completely.

BIBLIOGRAPHY

[AH16] T. Aslan and P. Hegedűs. Maximal deviation of large powers in the Nottingham group. Submitted, 2016.

[Asl16] T. Aslan. The distance of large powers in the Nottingham group for $p = 2$. Submitted, 2016.

[BL04] Y. Barnea and M. Larsen. Random generation in semisimple algebraic groups over local fields. *J. Algebra*, 271(1):1–10, 2004.

[Cam96] R. Camina. *Subgroups of the Nottingham Group*. PhD thesis, University of London, UK, 1996.

[Cam97] R. Camina. Subgroups of the Nottingham group. *J. Algebra*, 196(1):101–113, 1997.

[Cam00] R. D. Camina. The Nottingham group. In M. du Sautoy, D. Segal, and A. Shalev, editors, *New Horizons in Pro-p Groups*, volume 184 of *Progress in Mathematics*, pages 205–221. Birkhäuser, Basel, 2000.

[DdSMS99] J. D. Dixon, M.P.F. du Sautoy, A. Mann, and D. Segal. *Analytic pro-p Groups*, volume 61 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 1999.

[Dix69] J. D. Dixon. The probability of generating the symmetric group. *Math. Z.*, 110:199–205, 1969.

[dSSS00] M. du Sautoy, D. Segal, and A. Shalev. *New Horizons in Pro-p Groups*, volume 184 of *Progress in Mathematics*. Birkhäuser, Basel, 2000.

[Ers04] M. Ershov. New just-infinite pro-p groups of finite width and subgroups of the Nottingham group. *J. Algebra*, 275(1):419–449, 2004.

[Ers05] M. Ershov. The Nottingham group is finitely presented. *J. London Math. Soc.*, 71(2):362–378, 2005.

[FJ86] M. D. Fried and M. Jarden. *Field Arithmetic*. Springer, Berlin; Heidelberg, 1986.

[Heg01] P. Hegedűs. The Nottingham group for $p = 2$. *J. Algebra*, 246(1):55–69, 2001.

[Joh88] D. L. Johnson. The group of formal power series under substitution. *J. Austral. Math. Soc.*, 45(3):296–302, 1988.

[Kea05] K. Keating. How close are p-th powers in the Nottingham group? *J. Algebra*, 287(2):294–309, 2005.

[KL90] W. M. Kantor and A. Lubotzky. The probability of generating a finite classical group. *Geom. Dedicata*, 36(1):67–87, 1990.

[KLGP97] G. Klaas, C. R. Leedham-Green, and W. Plesken. *Linear Pro-p-Groups of Finite Width*, volume 1674 of *Lecture Notes in Mathematics*. Springer-Verlag Berlin Heidelberg, 1997.

[LGM02] C. R. Leedham-Green and S. McKay. *The structure of groups of prime power order*. Oxford University Press, Oxford, 2002.

[MA96] M.W.Liebeck and A.Shalev. Simple groups, probabilistic methods, and a conjecture of kantor and lubotzky. *J. Algebra*, 184(1):31–57, 1996.

[Man96] A. Mann. Positively finitely generated groups. *Forum Math.*, 8:429–459, 1996.

[MS96] A. Mann and A. Shalev. Simple groups, maximal subgroups, and probabilistic aspects of profinite groups. *Israel J. Math.*, 96(B):449–468, 1996.

[Net82] E. Netto. *Substitutionentheorie and ihre Anwendungen auf die Algebra*. Teubner, Leipzig, 1882. English transl. 1892, second ed., Chelsea, New York, 1964.

[Pin98] R. Pink. Compact subgroups of linear algebraic groups. *J. Algebra*, 206(2):438–504, 1998.

[Sak64] S. Saks. *Theory of the integral*. Dover Publications, Inc., New York, 1964.

[Sha99] A. Shalev. Probabilistic group theory. In C. M. Campbell, E. F. Robertson, N. Ruskuc, and G. C. Smith, editors, *Groups St. Andrews 1997 in Bath*, London Math. Soc. Lecture Note Series 261, pages 648–678. Cambridge University Press, Cambridge, 1999.

[Sha00] A. Shalev. Lie methods in the theory of pro-p groups. In M. du Sautoy, D. Segal, and A. Shalev, editors, *New Horizons in Pro-p Groups*, pages 1–54. Birkhäuser, Basel, 2000.

[Sze05] B. Szegedy. Almost all finitely generated subgroups of the Nottingham group are free. *Bull. London Math. Soc.*, 37(1):75–79, 2005.

[Yor90a] I. O. York. The exponent of certain p-groups. *Proc. Edinburg Math. Soc.*, 33(3):483–490, 1990.

[Yor90b] I. O. York. *The group of formal power series under substitution*. PhD thesis, Nottingham University, UK, 1990.