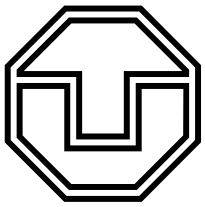

TECHNISCHE UNIVERSITÄT DRESDEN

Department of Mathematics
Institute of Geometry

On the length of group laws

Master's Thesis

To Receive the Degree

Master of Science (M.Sc.)

Author:
Jakob Schneider
(* January 26, 1994, Peine, Germany)

Supervisor:
Prof. Dr. Andreas Thom
(Institute of Geometry)

August 19, 2016
(date of submission)

Statutory Declaration

I declare on oath that I completed this work on my own and that information which has been directly or indirectly taken from other sources has been noted as such. Neither this, nor a similar work, has been published or presented to an examination committee.

Dresden, August 19, 2016 Jakob Schneider

Contents

Introduction

In this thesis, we will concentrate on the content of the two articles [14] and [3] (but mainly on the first). These two articles deal with so-called *group laws* and we want to explain at first what is behind this term.

Definition 1 (Group law)**.** *A* group law *in k letters or a k-*letter group law *for any given group G is a word $w \in \mathbf{F}_k = \langle a_1, \ldots, a_k \rangle$ in the free group of rank k such that for any mapping $\varphi\colon \{a_1, \ldots, a_k\} \to G$ the evaluation of the word $\varphi(w)$ yields the neutral element 1_G in G. Here $\varphi(w)$ denotes the image of w under the homomorphism corresponding to the mapping φ (see Chapter 0, Definition 2 for details). The word $w \in \mathbf{F}_k$ is called* non-trivial *if $w \neq 1_{\mathbf{F}_k}$.*

Example 1. A group G is called *elementary abelian* if it is the direct sum of cyclic groups of order p for a fixed prime p. When G is elementary abelian with respect to the prime p, we have the non-trivial two-letter group laws $aba^{-1}b^{-1}$ and $a^p \in \mathbf{F}_2$ for G.

In this thesis, we will mainly deal with group laws in two letters. However, the first few results will be presented in a more general form. Our main goal is to find a satisfactory answer to the following question.

Question 1. *Let $n \in \mathbb{Z}_+$. What is the minimum length of a non-trivial two-letter law which holds simultaneously for all finite groups G of order at most n?*

Essentially, the answer which we present in the last chapter is the answer presented in [14] (page 7, Theorem 5.1). Its proof is spread over the whole thesis:

Theorem 1 (Main result)**.** *For $n \in \mathbb{Z}_+$ large enough there exists a non-trivial two-letter law $w_n \in \mathbf{F}_2$ holding for all finite groups of order at most n of length*

$$\ell(w_n) \leq C \log^*(n)^2 \frac{n}{\log(n)^2}$$

for some fixed constant $C > 0$ (see Chapter 0, Definition 3 for the definition of the word length $\ell(w_n)$). In the former inequality, $\log^*(n)$ denotes the smallest integer $k \geq 0$ such that

$$\underbrace{\log \cdots \log}_{k}(n) \leq 1.$$

The main idea

Having presented the main theorem as an answer to Question 1, we now present the main idea to prove good bounds for the length of such a word w_n which is a non-trivial group law for each finite group of order at most n.

Take a finite group G and consider a special *short exact sequence* (see Section 1.2 for explanation), namely the following:

$$\mathbf{1} \to \mathbf{S}(G) \to G \to G/\mathbf{S}(G) \to \mathbf{1}.$$

Here $\mathbf{S}(G)$ is the *solvable radical* of G and $G/\mathbf{S}(G)$ is the corresponding *semi-simple* quotient in the sense of Fitting (see Chapter 3, Definition 11, where these notions are defined).

Considering the group G as an extension of $G/\mathbf{S}(G)$ by $\mathbf{S}(G)$ as above, we can combine a non-trivial law $w_{G/\mathbf{S}(G)}$ for $G/\mathbf{S}(G)$ and a non-trivial law $w_{\mathbf{S}(G)}$ for $\mathbf{S}(G)$ to obtain a non-trivial law w_G for G of length at most $\ell(w_{G/\mathbf{S}(G)})\ell(w_{\mathbf{S}(G)})$. The tool we need for the construction of w_G is the so-called extension lemma (Chapter 0, Lemma 4).

Setting $n_1 := |\mathbf{S}(G)|$ and $n_2 := |G/\mathbf{S}(G)|$ and distinguishing several cases, which altogether cover all combinations of $(n_1, n_2) \in \mathbb{Z}_+^2$ such that $n_1 n_2 \leq n$ for some fixed n, we can find a non-trivial law w_n holding for all finite groups G of order at most n. This last process will be described in detail in Chapter 4.

Therefore, for the construction of the word w_n, we will need a short non-trivial law holding for all finite solvable groups of order at most n_1 and a short non-trivial law for all finite semi-simple groups of order at most n_2 (for $n_1, n_2 \in \mathbb{Z}_+$).

Thus, in Chapter 2, we will focus on short non-trivial laws for finite solvable groups up to a certain order. Indeed, at first we will derive bounds for the length of short non-trivial laws for finite nilpotent groups of bounded order. Using these estimates and a result of Mike F. Newman, we will derive similar bounds for short non-trivial laws of finite solvable groups (see Proposition 4).

Thereafter, we will deal with non-trivial laws for finite semi-simple groups in Chapter 3. For this purpose it will become necessary to consider symmetric and finite simple groups. This is due to the fact that finite semi-simple groups (in the sense of Fitting) can be identified as subgroups of products of wreath products of automorphism groups of finite simple groups and symmetric groups (see Section 3.1 for details).

However, we will need the classification of finite simple groups to derive the main result of this thesis (see in Section 3.3 the proof of Proposition 5, and in Section 3.5 the use of Schreier's conjecture).

Acknowledgments

I wish to thank my supervisor Andreas Thom for numerous valuable discussions. Moreover, I thank Andreas Thom and Christiane Zyrus for proofreading.

Chapter 0

Essentials from group theory

In this chapter, we recall some basics from group theory. As a reference for a book in combinatorial group theory, we recommend [10], and as a reference for a book in finite group theory, we suggest [6].

At first, we recall the definition of a free group over a given set S.

Definition 2 (Free group)**.** *The* free group $\mathbf{F}(S)$ *over a given set S is the unique group (up to isomorphism) satisfying the following property. It holds that $S \subseteq \mathbf{F}(S)$ and for any mapping of sets $\varphi\colon S \to G$, where G is an arbitrary group, there is a unique extension $\varphi\colon \mathbf{F}(S) \to G$ being a homomorphism of groups.*

It is clear that $\mathbf{F}(S) \cong \mathbf{F}(T)$ if and only if $|S| = |T|$. Namely, when S and T are finite, we have that

$$\mathrm{Hom}(\mathbf{F}(S), \mathbb{Z}_2) \cong \mathbb{Z}_2^S$$

and thus $|\mathrm{Hom}(\mathbf{F}(S), \mathbb{Z}_2)| = 2^{|S|} \neq |\mathrm{Hom}(\mathbf{F}(T), \mathbb{Z}_2)| = 2^{|T|}$ if and only if $|S| \neq |T|$. In the case that S and T are infinite with $|S| \neq |T|$, we obviously have $|\mathbf{F}(S)| = |S| \neq |\mathbf{F}(T)| = |T|$.

Thus we can write $\mathbf{F}_k$ to denote the free group of rank k, i.e., over a k-element set. It is a well-known fact that the elements of the free group $\mathbf{F}(S)$ can be seen as *reduced words* over the set S, i.e., words of the form $w = s_1 \cdots s_n$, where $s_i \in S \cup S^{-1}$ for $i = 1, \ldots, n$ and $s_i \neq s_{i+1}^{-1}$ for $i = 1, \ldots, n-1$ (here $S^{-1} \coloneqq \{s^{-1} \mid s \in S\}$; see again [10], page 4, Proposition 1.9).

We can measure the *length* of an element of the free group $\mathbf{F}(S)$ by considering its representation as a reduced word:

Definition 3 (Word length)**.** *Let $w \in \mathbf{F}(S)$ be an element of the free group over the set S with reduced representation $s_1 \cdots s_n$ ($s_i \in S \cup S^{-1}$ for $i = 1, \ldots, n$). Then we call n the* word length *or just the* length *of w and define $\ell(w) \coloneqq n$.*

Next we define the *vanishing set* of a word inside a group G.

Definition 4 (Vanishing set)**.** *Let $w \in \mathbf{F}_k$. Set*

$$Z(G, w) \coloneqq \{(g_1, \ldots, g_k) \in G^k \mid w(g_1, \ldots, g_k) = 1_G\}$$

and call this the vanishing set *of w in G. Here 1_G denotes the neutral element of G and $w(g_1, \ldots, g_n)$ denotes the image of $w \in \mathbf{F}_k = \langle a_1, \ldots, a_k \rangle$ under the homomorphism which is induced by the mapping $\varphi\colon a_i \mapsto g_i$ for $i = 1, \ldots, n$ (see Definition 2).*

As stated in the introduction and using the defining property of the free group given in Definition 2, a law $w \in \mathbf{F}_k$ for a group G is a word which is mapped to 1_G under each homomorphism $\varphi\colon \mathbf{F}_k \to G$.

Thus we can describe the set of k-letter laws $\mathcal{L}_k$ as follows

$$\mathcal{L}_k = \bigcap_{\varphi\colon \mathbf{F}_k \to G} \ker(\varphi).$$

Considering this equality, it is immediately clear that $\mathcal{L}_k$ is a *characteristic* subgroup of $\mathbf{F}_k$, i.e., invariant under each automorphism $\alpha\colon \mathbf{F}_k \to \mathbf{F}_k$, as

$$\mathcal{L}_k = \bigcap_{\varphi\colon \mathbf{F}_k \to G} \ker(\varphi) = \bigcap_{\varphi\colon \mathbf{F}_k \to G} \ker(\varphi \circ \alpha).$$

Moreover, when G is finite, $\mathcal{L}_k$ is a subgroup of finite index as the intersection of finitely many subgroups of finite index.

Having defined the set of k-letter laws for G, we can ask for a non-trivial element of minimal length inside this subgroup.

Definition 5 (k-letter girth of a group)**.** *Let*

$$\operatorname{girth}_k(G) \coloneqq \min\{\ell(w) \mid w \in \mathbf{F}_k \setminus \{1\} \text{ is a law for } G\} \cup \{\infty\}$$

denote the k-letter girth of G.

Let $\operatorname{ord}(g)$ for $g \in G$ denote the least positive integer e such that $g^e = 1_G$, i.e., the *order of g*. Then the 1-letter girth of a finite group G is just the least common multiple of all orders $\operatorname{ord}(g)$ of elements of G, which is called the *exponent* of G.

However, as mentioned previously, our main focus will be set on the 2-letter girth of a group. Now we can reformulate the main question of this thesis as follows.

Question 2. *What is the 2-letter girth of the group Γ_n, where Γ_n is the direct product of all finite groups of order at most n?*

Note that a 2-letter word $w \in \mathbf{F}_2$ is a law for all finite groups of order at most n if and only if it is a law for the direct product of all these groups. Thus Question 1 and 2 ask for the same number.

We mention the following fact relating the 2-letter girth of a group with its k-letter girth.

Lemma 1. *There is an embedding of* $\mathbf{F}_{2\cdot 3^k} = \langle a_1, \ldots, a_{2\cdot 3^k} \rangle$ *into* $\mathbf{F}_2 = \langle a, b \rangle$ *such that* $\ell(a_i) = 2k + 1$*, where* $\ell(w)$ *is the length of the word* $w \in \mathbf{F}_2 = \langle a, b \rangle$*.*

For the proof of this lemma we need the following definition.

Definition 6 (Cayley graph)**.** *Let* G *be a group and* $S \subseteq G$ *be a subset of* G*. Then by* $\mathrm{Cay}(G, S)$ *we denote the directed graph with vertices* $V = G$ *and edges* $\{(p, q) \mid p^{-1}q \in S\}$*. This is called the* Cayley graph *of* G *with respect to* S*.*

Now we prove Lemma 1:

Proof. Consider the tree $T_k \leq \mathrm{Cay}(\mathbf{F}_2, \{a, b\})$ consisting of all words of length at most k. Every vertex of this tree which is not a leaf has two incoming edges labeled with a and b and two outgoing edges labeled with a and b. Thus, when we complete T_k arbitrarily to a Schreier graph (i.e., a graph which has two outgoing edges labeled with a and b and two incoming edges labeled with a and b at every vertex), only leaves are joined by new edges. Let S_k be such a completion. Now T_k is a spanning tree for S_k. Consider the fundamental group $\pi_1(S_k)$. As S_k is a graph with spanning tree T_k, this group is generated by circles in S_k which are of the following kind: Choose an edge of S_k, which is not in T_k, and prepend the unique non-returning path (i.e., a path which visits no vertex twice) in T_k starting at $1_{\mathbf{F}_2}$ and ending at the starting point of this edge and append the unique non-returning path in T_k starting at the end point of this edge and ending at $1_{\mathbf{F}_2}$ (the root of the tree T_k). Thus all these circles are of (word) length $2k + 1$. But for each edge of S_k which is not an edge in T_k there is a unique such circle. Thus, as there are $4 \cdot 3^{k-1}$ leaves of T_k, there are $3/2 \cdot 4 \cdot 3^{k-1} = 2 \cdot 3^k$ such circles $a_1, \ldots, a_{2\cdot 3^k}$, so that $\pi_1(S_k) \cong \mathbf{F}_{2\cdot 3^k}$. □

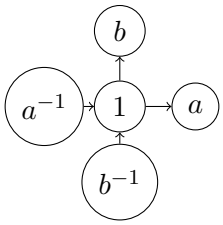

Figure 1: The tree T_1

We can thus draw a connection between the k-letter girth and the 2-letter girth of a group. An immediate consequence of this lemma is the following result.

Corollary 1. *Let* G *be a group and* $k \geq 2$ *be an integer. Then*

$$\mathrm{girth}_k(G) \leq \mathrm{girth}_2(G)$$

and

$$\mathrm{girth}_2(G) \leq (2\lceil \log_3(k/2) \rceil + 1)\,\mathrm{girth}_k(G).$$

Proof. The first inequality is obvious since every two-letter law is a k-letter law (as $k \geq 2$). The second inequality is true as we can embed $\mathbf{F}_{2 \cdot 3^{\lceil \log_3(k/2) \rceil}}$ into $\mathbf{F}_2$ such that none of its generators is longer than $2\lceil \log_3(k/2) \rceil + 1$. But it holds that $\mathbf{F}_k \subseteq \mathbf{F}_{2 \cdot 3^{\lceil \log_3(k/2) \rceil}}$, therefore the inequality follows. □

Chapter 1

The two main tools

The article [8] contains the following key lemma for proving good bounds for the length of non-trivial group laws (page 6, Lemma 2.2 in the article). It enables us to construct a new non-trivial word from m given non-trivial words such that the vanishing set of this word in any group G contains each of the vanishing sets of the m given words. We present it in the following section.

1.1 The commutator lemma

We start this section with a notation convention.

Notation 1. We define the *commutator* by $[x, y] \coloneqq xyx^{-1}y^{-1}$ and the *conjugate* $y^x \coloneqq x^{-1}yx$. Moreover, if $X, Y \subseteq G$ are subgroups of the group G, we define $[X, Y] \coloneqq \langle [x, y] \,|\, x \in X, y \in Y \rangle$.

The first main tool is a lemma which we formulate here in various versions. The first version is just applicable in the special case that we want to combine 2^e words, for some $e \in \mathbb{N}$, to get a new word.

To state it, we need the following definition.

Definition 7. *Let G be a group. An element g is said to be a non-trivial power if there exist an element $h \in G$ and an integer $n > 1$ such that $g = h^n$.*

Lemma 2 (Commutator lemma for powers of two)**.** *Let $k \geq 2$, $e \in \mathbb{N}$ and let non-trivial words $w_1, \ldots, w_m \in \mathbf{F}_k$ be given where $m = 2^e$. Then there exists a word $w \in \mathbf{F}_k$ of length*

$$\ell(w) \leq 2m \left(\sum_{i=1}^{m} \ell(w_i) + m \right)$$

which is not a non-trivial power such that for any group G we have

$$Z(G, w) \supseteq Z(G, w_1) \cup \cdots \cup Z(G, w_m).$$

Proof. We will prove this by induction on $e \in \mathbb{N}$. For $e = 0$ (i.e., $m = 1$) simply take $v = [s, w_1]$ for $s \in S$ such that w_1 is not a power of s. Then the commutator $[s, w_1]$ cannot be a non-trivial power and has length at most $2(\ell(w_1) + 1)$. Moreover, $Z(G, [s, w_1]) \supseteq Z(G, s) \cup Z(G, w_1)$ is immediate for any group G.

For $e \geq 1$ (i.e., $m \geq 2$) we are given $w_1, \ldots, w_{m/2}, w_{m/2+1}, \ldots, w_m$. By induction, there exist words v_1, v_2, which are no non-trivial powers, such that

$$\ell(v_1) \leq m \left(\sum_{i=1}^{m/2} \ell(w_i) + \frac{m}{2} \right)$$
$$\ell(v_1) \leq m \left(\sum_{i=1+m/2}^{m} \ell(w_i) + \frac{m}{2} \right).$$

and

$$Z(G, v_1) \supseteq Z(G, w_1) \cup \cdots \cup Z(G, w_{m/2})$$
$$Z(G, v_2) \supseteq Z(G, w_{m/2+1}) \cup \cdots \cup Z(G, w_m)$$

for any group G.

Since a commutator $[a, b]$ for $a, b \in \mathbf{F}_k \setminus \{1\}$ is trivial if and only if a and b lie a free subgroup of $\mathbf{F}_k$ of rank one, it can only be trivial if $a = c^j$, $b = c^k$ for some $c \in \mathbf{F}_k$ (this is an application of the Nielsen–Schreier theorem, see [10], page 7, Proposition 2.6). Thus the commutator $v \coloneqq [v_1, v_2]$ can only be trivial if $v_1 = v_2^{\pm 1}$ since v_1, v_2 are no non-trivial powers. If this is the case, we have that

$$Z(G, v_1) = Z(G, v_2)$$

and thus we can simply set $w \coloneqq v_1$ or $w \coloneqq v_2$. The restriction on the length of w is then obviously satisfied.

In the opposite case, we take $w \coloneqq v$. Then $v \neq 1_{\mathbf{F}_k}$ and at the same time v is not a non-trivial power (by a result of Schützenberger, see [13]) and

$$\begin{aligned} \ell(w) &\leq 2 \cdot m \left(\sum_{i=1}^{m/2} \ell(w_i) + \frac{m}{2} \right) + 2 \cdot m \left(\sum_{i=1+m/2}^{m} \ell(w_i) + \frac{m}{2} \right) \\ &= 2m \left(\sum_{i=1}^{m} \ell(w_i) + m \right). \end{aligned}$$

This finishes the proof. □

Now we generalize the lemma to cover the case where m is not a power of 2.

Lemma 3 (Commutator lemma, general case)**.** *Let $k \geq 2$ and let non-trivial words*

$w_1, \ldots, w_m \in \mathbf{F}_k$ be given. Then there exists a word $w \in \mathbf{F}_k$ of length

$$\ell(w) \leq 8m \left(\sum_{i=1}^{m} \ell(w_i) + m \right)$$

which is not a non-trivial power such that for any group G we have

$$Z(G, w) \supseteq Z(G, w_1) \cup \ldots \cup Z(G, w_m).$$

Proof. Let 2^e be the smallest power of two which is greater than or equal to m. Then $2^e < 2m$. Set

$$w'_1 \coloneqq w_1, \ldots, w'_m = w_m, w'_{m+1} \coloneqq w_1, \ldots, w'_{2^e} \coloneqq w_{2^e - m}$$

and apply the lemma for powers of two for the w'_i $(i = 1, \ldots, 2^e)$ to finish the proof. □

We can weaken the result a bit in the following corollary to make it more handy.

Corollary 2. *Let $k \geq 2$ and let non-trivial words $w_1, \ldots, w_m \in \mathbf{F}_k$ be given. Then there exists a word $w \in \mathbf{F}_k$ of length*

$$\ell(w) \leq 8m^2 \left(\max_{i \in \{1, \ldots, m\}} \ell(w_i) + 1 \right)$$

which is not a non-trivial power such that for any group G we have

$$Z(G, w) \supseteq Z(G, w_1) \cup \ldots \cup Z(G, w_m).$$

Proof. The result follows immediately from the previous lemma and the straightforward estimate

$$\sum_{i=1}^{m} \ell(w_i) \leq m \max_{i \in \{1, \ldots, m\}} \ell(w_i).$$

□

1.2 The extension lemma

The second important tool is the following simple result from [14] (page 3, Lemma 2.2) which allows us to construct non-trivial laws for group extensions from two given non-trivial group laws. For this purpose recall that a short exact sequence of groups is a chain of homomorphisms

$$\mathbf{1} \to A \xrightarrow{\varphi} B \xrightarrow{\psi} C \to \mathbf{1}$$

such that $\ker(\psi) = \mathrm{im}(\varphi)$, φ is injective and ψ is surjective.

Lemma 4 (Extension lemma)**.** *Let*

$$\mathbf{1} \to N \to G \to G/N \to \mathbf{1}$$

be a short exact sequence of groups and let w_N be a non-trivial law for N and $w_{G/N}$ be a non-trivial law for G/N, both in $\mathbf{F}_k = \langle a_1, \ldots, a_k \rangle = \langle S \rangle$. Then there is a non-trivial law for G in these k letters of length at most $\ell(w_N)\ell(w_{G/N})$. Thus,

$$\operatorname{girth}_k(G) \leq \operatorname{girth}_k(N) \operatorname{girth}_k(G/N).$$

Proof. If $w_{G/N} = s^n$ for some $s \in S$ and $n \in \mathbb{Z} \setminus \{0\}$, then t^n (where $t \in S$) is also a law for G/N and we can take

$$w = w_N(a_1^n, \ldots, a_k^n)$$

as a non-trivial law for G (it is a law for G since the map $g \mapsto g^n$ maps G to N, as s^n is a law for G/N, and w_N is a non-trivial law for N).

In the opposite case, we may assume, w.l.o.g., that $w_{G/N} = a_1 w'_{G/N} a_2$, where $w'_{G/N}$ does neither start with a_1^{-1} nor end with a_2^{-1}. The reason for this is that, if $w_{G/N}$ starts and ends with the same letter (or a letter and its inverse), we may perform a cyclic rotation of the word and take $w^v_{G/N} = sw't$, which will start and end with different letters $s, t \in S \cup S^{-1}$ with $st \neq 1$. Then, if necessary, we apply the appropriate automorphism of $\mathbf{F}_k$ induced by $s \mapsto a_1, a_1 \mapsto s$ and $t \mapsto a_2, a_2 \mapsto t$ and $r \mapsto r$ for all other $r \in S$.

Set

$$w_i \coloneqq w_{G/N}(a_i, \ldots, a_k, a_1 \ldots, a_{i-1}).$$

It is now routine to check that all non-trivial combinations (i.e., the words $w_i w_j$, $w_i^{-1} w_j$, $w_i^{-1} w_j^{-1}$, $w_i w_j^{-1}$, $w_i w_i$, $w_i^{-1} w_i^{-1}$ for all $i, j = 1, \ldots, k, i \neq j$) of the words w_i involve no cancellation.

We can thus take

$$w \coloneqq w_N(w_1, \ldots, w_k)$$

as a non-trivial law for G, since for all $i \in \{1, \ldots, k\}$ the words w_i induce mappings from G to N and w_N is a non-trivial law for N. □

Chapter 2

Nilpotent and solvable groups

In this chapter, we wish to find short non-trivial laws for finite nilpotent and solvable groups only depending on their order. More precisely, we intend to find bounds for the length of the shortest non-trivial law holding in all finite nilpotent (resp. solvable) groups of order at most n. Our main references here are again the articles [14] and [3].

We need to introduce some notation conventions for this and the following chapters.

Notation 2. In the following, we write $a_n = o(1)$ to express that $\lim_{n\to\infty} a_n = 0$.

Notation 3 (Center, centralizer)**.** We write $\mathbf{Z}(G)$ for the *center* of G, i.e.,

$$\mathbf{Z}(G) = \{g \in G \,|\, \forall h \in G : gh = hg\}.$$

Moreover, we write $\mathbf{C}_G(H)$ for the *centralizer* of H in G, i.e.,

$$\mathbf{C}_G(H) = \{g \in G \,|\, \forall h \in H : hg = gh\}.$$

2.1 Definitions and basic properties

At first, we recall the definition of the *lower central series* and a *nilpotent group.*

Definition 8 (Nilpotent group, lower central series, nilpotency class)**.** *A group G is called* nilpotent *if and only if the* lower central series $(\gamma_k(G))_{k\geq 1}$ *terminates with the trivial group, where*

$$\gamma_1(G) \coloneqq G \quad \textit{and} \quad \gamma_{k+1}(G) \coloneqq [\gamma_k(G), G].$$

The smallest k such that $\gamma_{k+1}(G) = \mathbf{1}$ is called the nilpotency class *of G.*

A *central series* is a series $(H_k)_{k\geq 1}$ where $H_1 = H$ and $H_i/H_{i+1} \subseteq \mathbf{Z}(G/H_{i+1})$ or equivalently $[H_i, G] \subseteq H_{i+1}$ (for $i \geq 1$).

Remark 1. The reason why we call $(\gamma_k(G))_{k\geq 1}$ the lower central series is that for any *central series* $(H_k)_{k\geq 1}$ it holds that $H_i \supseteq \gamma_i(G)$ for all i (this can easily be proven by induction: $\gamma_{i+1}(G) = [\gamma_i(G), G] \subseteq [H_i, G] \subseteq H_{i+1} \subseteq H_i$).

Example 2. Considering the group of *unitriangular matrices* $\mathrm{UT}_n(R)$ over some ring R, one notes that this is in fact nilpotent. More precisely, $\mathrm{UT}_n(R)$ consists of all matrices of the form

$$\begin{pmatrix} 1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

where all $*$'s are chosen arbitrarily from R. Its class of nilpotency is $n-1$.

Secondly, we define the *derived series* and a *solvable group.*

Definition 9 (Solvable group, derived series, solvability class)**.** *Let G be a group. Its* derived series *is defined by*

$$G^{(0)} \coloneqq G \quad \textit{and} \quad G^{(k+1)} \coloneqq [G^{(k)}, G^{(k)}].$$

A group is called solvable *if and only if $G^{(k)} = \mathbf{1}$ for some $k \in \mathbb{N}$. The smallest such k is called* solvability class *of G.*

Example 3. Considering the group of *upper triangular matrices* $\mathrm{T}_n(R)$ over some ring R, one notes that this is solvable since $[\mathrm{T}_n(R), \mathrm{T}_n(R)] \subseteq \mathrm{UT}_n(R)$ and $\mathrm{UT}_n(R)$ is nilpotent. More precisely, $\mathrm{T}_n(R)$ is the group of matrices of the form

$$\begin{pmatrix} d_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

where all $*$'s are chosen arbitrarily from R and $d_i \in R^\times$ are units ($i = 1, \ldots, n$).

Next we state some simple lemmas relating the derived series to the lower central series.

Notation 4 (Commutators)**.** We write $[g_1, \ldots, g_n]$ for $[[\cdots[g_1, g_2]\cdots], g_n]$ and analogously for groups $[G_1, \ldots, G_n]$ for $[[\cdots[G_1, G_2]\cdots], G_n]$.

Lemma 5. *It holds that $[\gamma_i(G), \gamma_j(G)] \subseteq \gamma_{i+j}(G)$.*

This lemma basically tells us that there is a unique biggest commutator subgroup written as a commutator of weight n, namely $\gamma_n(G)$.

To prove this result, we need the so-called three subgroups lemma which is a consequence of the so-called *Hall–Witt-identity*.

Lemma 6 (Hall–Witt identity)**.** *Let $x, y, z \in G$ for a group G. Then*

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

Proof. The proof happens by direct computation. □

Lemma 7 (Three subgroups lemma)**.** *Let* $[X, Y, Z] = [Y, Z, X] = \mathbf{1}$*. Then* $[Z, X, Y] = \mathbf{1}$*.*

The following proof is taken from [6] (page 126, Lemma 4.9).

Proof. We want to prove that every element of $[Z, X]$ commutes with every element of Y. For this it suffices to prove this for every generating commutator $[z, x] \in [Z, X]$ and $y \in Y$, i.e., it suffices to prove that $[y, [z, x]] = 1$ for $x \in X, y \in Y, z \in Z$, which is equivalent to the fact that $[y, [z^{-1}, x]] = 1$ for all $x \in X, y \in Y, z \in Z$. But this follows from the Hall–Witt identity since

$$[z, [x^{-1}, y]] = [x^{-1}, y, z]^{-1} = 1 \quad \text{and} \quad [x, [y^{-1}, z]] = [y^{-1}, z, x]^{-1} = 1$$

by the assumptions $[X, Y, Z] = [Y, Z, X] = \mathbf{1}$. Thus we are done. □

The proof of Lemma 5 can now be given.

Proof. We proceed by induction on j. For $j = 1$ we have $[\gamma_i(G), G] = \gamma_{i+1}(G) \subseteq \gamma_{i+1}(G)$. For $j > 1$ it holds that

$$[\gamma_i(G), \gamma_j(G)] = [\gamma_j(G), \gamma_i(G)] = [\gamma_{j-1}(G), G, \gamma_i(G)].$$

Using the three subgroups lemma, it suffices to prove that

$$[G, \gamma_i(G), \gamma_{j-1}(G)] \subseteq \gamma_{i+j}(G) \quad \text{and} \quad [\gamma_i(G), \gamma_{j-1}(G), G] \subseteq \gamma_{i+j}(G).$$

But the first inclusion holds by induction since

$$[G, \gamma_i(G), \gamma_{j-1}(G)] = [\gamma_i(G), G, \gamma_{j-1}(G)] = [\gamma_{i+1}(G), \gamma_{j-1}(G)] \subseteq \gamma_{i+j}(G)$$

and the same is true for the second inclusion

$$[\gamma_i(G), \gamma_{j-1}(G), G] \subseteq [\gamma_{i+j-1}(G), G] = \gamma_{i+j}(G).$$

This finishes the proof. □

A simple consequence of Lemma 5 is the following corollary.

Corollary 3. *It holds that* $G^{(n)} \subseteq \gamma_{2^n}(G)$*.*

Next we focus on short non-trivial words in the derived series of $\mathbf{F}_2$.

2.2 Short non-trivial words in the derived series of $\mathbf{F}_2$

The construction we present here is essentially from [3] (Chapter 3). We set

$$a_0 \coloneqq a \quad \text{and} \quad b_0 \coloneqq b$$

and define the sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \subseteq \mathbf{F}_2$ recursively by

$$a_{n+1} \coloneqq [b_n^{-1}, a_n] \quad \text{and} \quad b_{n+1} \coloneqq [a_n, b_n].$$

The first lemma we need is a bit technical (it is essentially Lemma 3.1 in [3]).

Lemma 8 (No cancellation in products)**.** *For all $n \in \mathbb{N}$ the products $a_n a_n$, $a_n^{-1}a_n^{-1}$, $b_n b_n$, $b_n^{-1}b_n^{-1}$, $a_n^{-1}b_n$, $b_n^{-1}a_n$, $a_n b_n^{-1}$, $b_n a_n^{-1}$, $a_n^{-1}b_n^{-1}$, $b_n a_n$ involve no cancellation.*

Proof. We prove the result by induction on n. For $n = 0$ the statement of the lemma is obvious. Let $n > 0$. The word $a_n a_n$ and its inverse have no cancellation since $a_{n-1}^{-1}b_{n-1}^{-1}$ has no cancellation.

$$a_n a_n = [b_{n-1}^{-1}, a_{n-1}]^2 = b_{n-1}^{-1}a_{n-1}b_{n-1} \underbrace{a_{n-1}^{-1}b_{n-1}^{-1}}_{\text{no cancellation}} a_{n-1}b_{n-1}a_{n-1}^{-1}$$

The word $b_n b_n$ and its inverse have no cancellation since $b_{n-1}^{-1}a_{n-1}$ has no cancellation. The word $a_n^{-1}b_n$ and its inverse admit no cancellation since $b_{n-1}a_{n-1}$ has no cancellation. The word $a_n b_n^{-1}$ and its inverse have no cancellation since $a_{n-1}^{-1}b_{n-1}$ has no cancellation. Finally, $a_n^{-1}b_n^{-1}$ has no cancellation since $b_{n-1}b_{n-1}$ has no cancellation. □

The next result tells us that a_n and b_n are of the same length and gives a first estimate for this (it is essentially Lemma 3.2 from [3]).

Lemma 9. *It holds that $4^n \geq \ell(a_n) = \ell(b_n) \geq 2^n$. Moreover, $a_n, b_n \in \mathbf{F}_2^{(n)}$ for $n \in \mathbb{N}$.*

Proof. Plugging in the definition of b_{n+1} yields

$$\begin{aligned}\ell(b_{n+1}) &= \ell(a_n b_n a_n^{-1} b_n^{-1}) \\ &= \ell(a_n b_n) + \ell(a_n) + \ell(b_n) \\ &= \ell(b_n^{-1} a_n b_n a_n^{-1}) \\ &= \ell(a_{n+1})\end{aligned}$$

Here we use Lemma 8 in the second and the third line. It follows from the second line of the computation that $\ell(b_{n+1}) \geq 2\ell(b_n)$ and thus by induction $\ell(a_n) = \ell(b_n) \geq 2^n$. Moreover, it follows from the third line (without considering cancellation) that $\ell(b_{n+1}) \leq 4\ell(b_n)$ and thus inductively $\ell(b_n) \leq 4^n$.

The last fact follows inductively from $a_0 = a \in \mathbf{F}_2^{(0)}$ and $b_0 = b \in \mathbf{F}_2^{(0)}$ and $a_{n+1} = [b_n^{-1}, a_n] \in [\mathbf{F}_2^{(n)}, \mathbf{F}_2^{(n)}] = \mathbf{F}_2^{(n+1)}$ as well as $b_{n+1} = [a_n, b_n] \in [\mathbf{F}_2^{(n)}, \mathbf{F}_2^{(n)}] = \mathbf{F}_2^{(n+1)}$. □

From now on, we set $c_n \coloneqq \ell(a_n) = \ell(b_n)$. The main result of this section is presented in the subsequent lemma.

Lemma 10. *For all $n \in \mathbb{N}$ we have $c_{n+2} = 3c_{n+1} + 2c_n$. Thus it holds that*

$$\begin{aligned} c_n &= \left(\frac{1}{2} + \frac{5}{2\sqrt{17}}\right)\left(\frac{3+\sqrt{17}}{2}\right)^n + \left(\frac{1}{2} - \frac{5}{2\sqrt{17}}\right)\left(\frac{3-\sqrt{17}}{2}\right)^n \\ &\leq C_1 \iota^n + o(1). \end{aligned}$$

where $\iota \coloneqq (3+\sqrt{17})/2 = 3.5615528\ldots$ and $C_1 \coloneqq \frac{1}{2} + \frac{5}{2\sqrt{17}} = 1.10633906\ldots$.

Proof. The proof happens by the following computation and makes use of Lemma 8.

$$\begin{aligned} c_{n+2} = \ell(b_{n+2}) &= \ell([a_{n+1}, b_{n+1}]) \\ &= \ell([[b_n^{-1}, a_n], [a_n, b_n]]) \\ &= \ell(b_n^{-1} a_n b_n \underbrace{a_n^{-1} a_n}_{\text{cancels}} b_n a_n^{-1} b_n^{-1}) + \ell([a_n, b_n^{-1}]) + \ell([b_n, a_n]) \\ &= \underbrace{\ell(b_n^{-1} a_n b_n)}_{\ell(a_{n+1}) - \ell(a_n^{-1}) = c_{n+1} - c_n} + \underbrace{\ell(b_n) + \ell(a_n^{-1}) + \ell(b_n^{-1})}_{3c_n} + \underbrace{\ell([a_n, b_n^{-1}])}_{\ell(a_{n+1}) = c_{n+1}} + \underbrace{\ell([b_n, a_n])}_{\ell(b_{n+1}) = c_{n+1}} \end{aligned}$$

Thus we obtain

$$c_{n+2} - 3c_{n+1} - 2c_n = 0 \ (n \geq 0), \text{ where } c_0 = 1 \text{ and } c_1 = 4.$$

It follows inductively that for all $n \in \mathbb{N}$ we have

$$c_n = \left(\frac{1}{2} + \frac{5}{2\sqrt{17}}\right)\left(\frac{3+\sqrt{17}}{2}\right)^n + \left(\frac{1}{2} - \frac{5}{2\sqrt{17}}\right)\left(\frac{3-\sqrt{17}}{2}\right)^n$$

as the polynomial $\lambda^2 - 3\lambda - 2$ has the roots

$$\lambda_{1,2} = \frac{3 \pm \sqrt{17}}{2}.$$

This finishes the proof. □

Hence for large k there are non-trivial words in $\mathbf{F}_2^{(k)}$ of length remarkably shorter than 4^k as $\iota = (3+\sqrt{17})/2 = 3.5615528\ldots < 4$ and $a_k, b_k \in \mathbf{F}_2^{(k)}$ (cf. Lemma 9).

The subsequent proposition is an immediate consequence of the preceding lemma.

Proposition 1. *There is a non-trivial two-letter word w of length*

$$\ell(w) \leq C_1 \iota^n + o(1)$$

which is a law for every solvable group of solvability class at most n.

Proof. Let k be the solvability class of G. The previous lemma shows that there is a non-trivial word of length $C_1\iota^n + o(1)$ in $\mathbf{F}_2^{(n)}$. But since G is solvable of class $k \leq n$, every word of $\mathbf{F}_2^{(k)} \supseteq \mathbf{F}_2^{(n)}$ is a law for G. Thus this word, as an element of $\mathbf{F}_2^{(n)} \subseteq \mathbf{F}_2^{(k)}$, is a law for G. □

We proceed by considering short non-trivial words in the lower central series of $\mathbf{F}_2$.

2.3 Short non-trivial words in the lower central series of $\mathbf{F}_2$

Set

$$\kappa \coloneqq \log_2(\iota) = \log_2\left(\frac{3+\sqrt{17}}{2}\right) = 1.832506\ldots < 2.$$

As a simple consequence of Lemma 10 and Corollary 3, we obtain the following corollary. However, it is not optimal as we will see.

Corollary 4. *For $k \in \mathbb{Z}_+$ there is a non-trivial word w of length*

$$\ell(w) \leq C_2 k^{\log_2(\iota)} + o(1) = C_2 k^{\kappa} + o(1)$$

in $\gamma_k(\mathbf{F}_2)$ for some fixed constant $C_2 \coloneqq 2 + 8/\sqrt{17} = 3.940285\ldots$.

Proof. Take 2^e as the smallest power of two which is not less than k, i.e., $e = \lceil \log_2(k) \rceil < \log_2(k) + 1$. Then

$$\mathbf{F}_2^{(e)} \subseteq \gamma_{2^e}(\mathbf{F}_2) \subseteq \gamma_k(\mathbf{F}_2).$$

In $\mathbf{F}_2^{(e)}$ there exists a non-trivial word of length at most $C_1\iota^e + o(1)$ by the previous section. But $\iota^e = 2^{\log_2(\iota)e} < (2k)^{\log_2(\iota)} = (2k)^{\kappa}$. Thus we can take $C_2 \coloneqq \iota C_1 = 2 + 8/\sqrt{17} = 3.940285\ldots$ to finish the proof. □

In fact, one can find a better exponent for k in the last corollary. The idea we will present again stems from [3].

But firstly, note that the previous result implies the following lemma.

Lemma 11. *There is non-trivial word w which is a law for all nilpotent groups G of order at most n of length*

$$\ell(w) \leq C_3 \log(n)^{\kappa} + o(1)$$

for some fixed constant $C_3 \coloneqq 7.712869694\ldots$.

Proof. Since G is nilpotent, it holds that $\gamma_i(G)/\gamma_{i+1}(G)$ has at least two elements when $\gamma_i(G)$ is not trivial. For the class k of nilpotency of G we have then $k \leq \lfloor \log_2|G| \rfloor \leq \log_2(n)$. Thus, setting $C_3 \coloneqq C_2/\log(2)^{\kappa} = 7.712869694\ldots$, the lemma follows from the previous corollary. To see this, take a non-trivial word

$$w \in \gamma_{\lfloor \log_2(n) \rfloor}(\mathbf{F}_2) \subseteq \gamma_k(\mathbf{F}_2)$$

of length at most $C_2 \log_2(n)^\kappa + o(1)$. Then $w(g,h)$ for $g,h \in G$ lies in $\gamma_k(G) = \mathbf{1}$, thus $w(g,h) = 1_G$. □

Now let us explain how to improve the exponent $\log_2(\iota)$ to

$$\log_{1+\sqrt{2}}(\iota) = 1.44115577\ldots$$

following the second part of Chapter 3 of [3]. At first define

$$\gamma(w) \coloneqq \max\{n \in \mathbb{Z}_+ \mid w \in \gamma_n(\mathbf{F}_2)\} \cup \{\infty\}.$$

Lemma 12 (Properties of γ)**.** *The function γ has the properties*

$$\gamma(w_1 w_2) \geq \min\{\gamma(w_1), \gamma(w_2)\}$$

and

$$\gamma([w_1, w_2]) \geq \gamma(w_1) + \gamma(w_2).$$

Moreover, it is true that

$$\gamma(a_n) = \gamma(b_n)$$

and we define

$$\gamma_n \coloneqq \gamma(a_n) = \gamma(b_n).$$

Proof. The first identity holds since, when $w_1 \in \gamma_i(\mathbf{F}_2)$ and $w_2 \in \gamma_j(\mathbf{F}_2)$, then $w_1 w_2 \in \gamma_i(\mathbf{F}_2)\gamma_j(\mathbf{F}_2) = \gamma_{\min\{i,j\}}(\mathbf{F}_2)$, as the sequence $(\gamma_k(\mathbf{F}_2))_{k\geq 1}$ is monotone. The second identity holds by the fact of Lemma 5 that

$$[\gamma_i(\mathbf{F}_2), \gamma_j(\mathbf{F}_2)] \subseteq \gamma_{i+j}(\mathbf{F}_2).$$

Hence, when $w_1 \in \gamma_i(\mathbf{F}_2)$ and $w_2 \in \gamma_j(\mathbf{F}_2)$, then

$$[w_1, w_2] \in \gamma_{i+j}(\mathbf{F}_2).$$

The last claim can be verified by the fact that $a_n = [b_{n-1}^{-1}, a_{n-1}] = [a_{n-1}, b_{n-1}]^{b_{n-1}} = b_n^{b_{n-1}}$ and the fact that all $\gamma_k(\mathbf{F}_2)$ are normal. This ends the proof. □

Now we can present the following lemma (this is essentially Lemma 3.6 from [3]).

Lemma 13. *We have that $\gamma_{n+2} - 2\gamma_{n+1} - \gamma_n \geq 0$ for all $n \geq 0$. Thus, by induction, one obtains*

$$\begin{aligned}\gamma_n &\geq \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)(1+\sqrt{2})^n + \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)(1-\sqrt{2})^n \\ &\geq C_4 \mu^n - o(1),\end{aligned}$$

where $\mu \coloneqq 1 + \sqrt{2}$ and $C_4 \coloneqq \frac{1}{2} + \frac{1}{2\sqrt{2}} = 0.85355339\ldots$.

Proof. The key for the proof of this lemma are the following two identities (see [3], page 5)

$$[[a^{-1}, b], [a, b]] = [[[a^{-1}, b], a], [a, b]] \tag{2.1}$$

$$[[a^{-1}, b], [b, a]] = [[[a^{-1}, b], a], [b, a]] \tag{2.2}$$

We briefly verify the first identity

$$[[a^{-1}, b], a] = a^{-1}bab^{-1}aba^{-1}b^{-1}aa^{-1} = [a^{-1}, b][a, b]$$

and thus

$$\begin{aligned}[[[a^{-1}, b], a], [a, b]] &= [a^{-1}, b][a, b][a, b][a, b]^{-1}[a^{-1}, b]^{-1}[a, b]^{-1} \\ &= [[a^{-1}, b], [a, b]]\end{aligned}$$

Now we compute

$$\begin{aligned}b_{n+2} &= [a_{n+1}, b_{n+1}] \\ &= [[b_n^{-1}, a_n], [a_n, b_n]] \overset{(2.2)}{=} [[[b_n^{-1}, a_n], b_n], [a_n, b_n]] \\ &= [[a_{n+1}, b_n], b_{n+1}],\end{aligned}$$

from which we obtain $\gamma_{n+2} \geq 2\gamma_{n+1} + \gamma_n$ by Lemma 12 and the estimate of the lemma follows inductively from the initial values $\gamma_0 = 1$ and $\gamma_1 = 2$. □

Remark 2. It is an interesting question if each identity which can be expressed only by commutators $[\bullet, \bullet]$ and inverses $\bullet^{-1}$ in the free group $\mathbf{F}_2$ follows from the two identities (2.1) and (2.2), and the identity $[a, b]^{-1} = [b, a]$ or if there are *more* such relations.

We are now able to improve the result of Corollary 4. Set

$$\nu \coloneqq \log_{1+\sqrt{2}}\left(\frac{3+\sqrt{17}}{2}\right) = 1.44115577\ldots.$$

Corollary 5 (Improved exponent for k)**.** *For $k \in \mathbb{Z}_+$ there is a non-trivial word w of length*

$$\ell(w) \leq (C_5 + o(1))k^{\log_{(1+\sqrt{2})}(\iota)} = (C_5 + o(1))k^{\nu}$$

in $\gamma_k(\mathbf{F}_2)$ for some fixed constant $C_5 = 4.95033877\ldots$.

Proof. Let $k \in \mathbb{Z}_+$. Then by the previous lemma (Lemma 13)

$$b^{\left\lceil \log_{1+\sqrt{2}}\left(\frac{k+o(1)}{C_4}\right)\right\rceil}$$

lies in $\gamma_k(\mathbf{F}_2)$ by the choice of the index and its length is bounded by

$$C_1\iota^{\left\lceil \log_{1+\sqrt{2}}\left(\frac{k+o(1)}{C_4}\right)\right\rceil} + o(1) \leq C_1\iota^{1+\log_{1+\sqrt{2}}\left(\frac{k+o(1)}{C_4}\right)} + o(1)$$
$$= \frac{C_1\iota}{C_4^\nu}(k+o(1))^\nu + o(1) = (C_5 + o(1))k^\nu$$

by Lemma 10. Plugging in the values for the constants reveals $C_5 = 4.95033877\ldots$. □

2.4 Laws for finite nilpotent groups

The results concerning nilpotent groups can now be stated as the following proposition (cf. Proposition 3.1 in [14]).

Proposition 2 (Short non-trivial laws for nilpotent groups)**.** *There exists a non-trivial two-letter word w of length*

$$\ell(w) \leq (C_6 + o(1))\log(n)^\nu = (8.395184144\ldots + o(1))\log(n)^{1.44115577\ldots}$$

for some fixed constant $C_6 \coloneqq 8.395184144\ldots$, which is a law for every nilpotent group of order at most n.

Proof. The proof of the fact relies on the fact that the nilpotency class of G can at most be $\lfloor \log_2|G| \rfloor \leq \log_2(n)$ and on Corollary 5 (choose $C_6 = C_5/\log(2)^\nu$). In fact, take $k \coloneqq \lfloor \log_2(n) \rfloor$. Then k is greater than or equal to the class of nilpotency of G. By Corollary 5 there is a non-trivial word of length at most $(C_5 + o(1))k^\nu$ in $\gamma_k(\mathbf{F}_2)$. This word is a law for G since its evaluation $w(g,h)$ lies in $\gamma_k(G) = \mathbf{1}$ for all $g, h \in G$. □

2.5 Laws for finite solvable groups

To prove good bounds for solvable groups, one needs a somehow more complicated machinery. In fact, one needs to consider the so-called *Fitting subgroup* $\mathbf{F}(G)$ of G which is the unique biggest normal nilpotent subgroup of G containing any other nilpotent normal subgroup.

Let $|G| \leq n$ for some $n \in \mathbb{Z}_+$ in the subsequent considerations. The group G acts on $\mathbf{F}(G)$ by conjugation. This gives us the homomorphism

$$\varphi\colon G \to \mathrm{Aut}(\mathbf{F}(G)) = \prod_{p||G|} \mathrm{Aut}(\mathbf{O}_p(G)).$$

Here $\mathbf{O}_p(G)$ denotes the unique characteristic Sylow p-subgroup of $\mathbf{F}(G)$. The map φ will be composed with the map α where $\alpha = \prod_p \alpha_p$ and

$$\alpha_p\colon \mathrm{Aut}(\mathbf{O}_p(G)) \to \mathrm{Aut}(\mathbf{O}_p(G)/\Phi(\mathbf{O}_p(G)).$$

The last group $\mathbf{O}_p(G)/\Phi(\mathbf{O}_p(G))$ is the *Frattini quotient* of the p-group $\mathbf{O}_p(G)$ (which is elementary abelian and thus its automorphism group is a linear group $\mathrm{GL}_{d_p}(p)$ for some dimension d_p). Finally, the homomorphism $\psi \coloneqq \alpha \circ \varphi$ will be used to construct a non-trivial law for G using the extension lemma (Lemma 4).

But now let us explain the previously mentioned thought in greater detail. Recall briefly the following definition of the *Fitting subgroup*.

Lemma 14 (Fitting subgroup)**.** *For a finite group G there exists a unique biggest normal nilpotent subgroup $\mathbf{F}(G)$ containing every other nilpotent normal subgroup.*

Proof. For p a prime dividing the order of G consider the core (the intersection of all its conjugates) of a Sylow p-subgroup $\mathbf{O}_p(G)$ (the so-called *p-core*, i.e., the intersection of all Sylow p-subgroups). We claim that

$$\mathbf{F}(G) \coloneqq \prod_{p||G|} \mathbf{O}_p(G)$$

has the desired properties. In fact let $N \subseteq G$ be a normal subgroup of G which is nilpotent. Then any Sylow p-subgroup P is characteristic in N and hence normal in G. Thus it is contained in every Sylow p-subgroup of G, i.e., $P \subseteq \mathbf{O}_p(G)$. Taking the product over all primes p dividing $|G|$ we obtain

$$N = \prod P \subseteq \prod_{p||G|} \mathbf{O}_p(G) = \mathbf{F}(G),$$

which completes the proof. □

The next fact that we use is the following proposition.

Proposition 3. *Let G be solvable. The map*

$$\varphi \colon G \to \mathrm{Aut}(\mathbf{F}(G))$$

given by conjugation has kernel equal to the center of $\mathbf{F}(G)$, i.e., $\mathbf{C}_G(\mathbf{F}(G)) = \mathbf{Z}(\mathbf{F}(G))$.

The proof we present is taken from [5] (page 218, Section 6.1, Theorem 1.3).

Proof. Assume the contrary. Define $C \coloneqq \mathbf{C}_G(\mathbf{F}(G))$ and $H \coloneqq \mathbf{Z}(\mathbf{F}(G)) = C \cap \mathbf{F}(G)$. Since $H \subseteq \mathbf{F}(G)$, every element of H commutes with every element of C. Therefore $H \subseteq \mathbf{Z}(C)$ and H is normal in C. Now consider the derived series of C/H. Since G is solvable, C/H is solvable as well. Let B be the inverse image of the term in the derived series just before the trivial group. Then $[B, B] \subseteq H$ and $B \subseteq C$. But since H commutes with every element of C we have that

$$[[B, B], B] \subseteq [[B, B], C] \subseteq [H, C] \subseteq 1$$

and hence B is nilpotent of class two.

Now we justify that B is normal in G and obtain a contradiction since B is not contained in $\mathbf{F}(G)$. In fact, $\mathbf{F}(G)$ and $\mathbf{C}_G(\mathbf{F}(G))$ are characteristic and so is their intersection H. B/H is characteristic in C/H as a member of the derived series. Thus B is characteristic in C. Because C is characteristic in G and B is characteristic in C, it is normal in G. This completes the proof, showing that $C/H = \mathbf{1}$ (and thus $C \subseteq \mathbf{F}(G)$) by contradiction. □

We proceed with our considerations for solvable groups. By Proposition 3, we have the short exact sequence

$$\mathbf{1} \to \mathbf{Z}(\mathbf{F}(G)) \to G \to \mathrm{im}(\varphi) \to \mathbf{1}$$

and thus by the extension lemma (Lemma 4) we obtain from a non-trivial law for $\mathrm{im}(\varphi)$ of length k a non-trivial law for G of length $4k$, since we can combine the former law with the law $aba^{-1}b^{-1}$ for the center $\mathbf{Z}(\mathbf{F}(G))$, which is abelian.

Since the Fitting subgroup is nilpotent, it is the direct product of its Sylow p-subgroups, i.e.,

$$\mathbf{F}(G) = \prod_p \mathbf{O}_p(G)$$

as already mentioned in the proof of the existence of the Fitting subgroup. Thus it is clear that

$$\mathrm{Aut}(\mathbf{F}(G)) = \prod_p \mathrm{Aut}(\mathbf{O}_p(G))$$

since for each prime p there is only one Sylow p-group. Hence this must be mapped to itself under every automorphism.

Now we want to consider the *Frattini subgroups* of the p-groups $\mathbf{O}_p(G)$. So we introduce the *Frattini subgroup* of a group G.

Definition 10 (Frattini subgroup)**.** *The* Frattini subgroup $\Phi(G)$ *of a group G is the intersection of all maximal subgroups $M \lessdot G$ of G.*

This group can be characterized by the following lemma.

Lemma 15. *The Frattini subgroup $\Phi(G)$ of a finite group G is precisely the set of non-generators, i.e., $g \in \Phi(G)$ if and only if for each set $U \subseteq G$ such that $\langle g, U\rangle = G$ it holds that already $\langle U\rangle = G$.*

Proof. Let us prove both directions.

$\Rightarrow$: Assume U would be a subset such that $\langle g, U\rangle = G$, $\langle U\rangle < G$, then there exists a maximal subgroup M such that $\langle U\rangle \subseteq M$ and thus it follows that $\langle g, U\rangle \subseteq M$ since $g \in \Phi(G) \subseteq M$, a contradiction.

$\Leftarrow$: Assume that there would be a maximal subgroup $M \lessdot G$ such that $g \notin M$. Then $\langle g, M\rangle = G$ by maximality of M and $\langle M\rangle = M$, contradicting the assumption on g to be a *non-generator*. □

The next lemma characterizes the Frattini subgroup in the case that the original group G is a p-group.

Lemma 16. *Let P be a finite p-group. Then $\Phi(P) = P^p[P,P]$ is the smallest normal subgroup such that the corresponding quotient is elementary abelian. Here P^p is the group generated by all g^p for $g \in P$ and $[P,P]$ the derived subgroup.*

Proof. We prove both inclusions.

$\supseteq$: Let $M \lessdot P$ be maximal. By nilpotency of P, it follows that $M \trianglelefteq P$ and P/M is a cyclic group of order p by maximality of M. Since P/M is abelian, we have $[P,P] \subseteq M$ and since P/M as cyclic of order p we have that $g^p \in M$ for all $g \in P$. Thus $P^p[P,P] \subseteq M$ and hence taking the intersection of all maximal subgroups gives $P^p[P,P] \subseteq \Phi(P)$.

$\subseteq$: Conversely, assume $N \trianglelefteq P$ and that P/N to be elementary abelian. Then P/N is a vector space and can be seen as d-fold direct product of cyclic groups of order p, where d is the dimension of the vector space. The product of all but one of these factors has index p and thus is maximal in P/N. The intersection of these subgroups is trivial. Taking the preimages under the canonical mapping $P \to P/N$ yields maximal subgroups whose intersection is N. Thus $\Phi(P) \subseteq N$. Choosing $N \coloneqq P^p[P,P]$ yields the desired inclusion $\Phi(P) \subseteq P^p[P,P]$. □

Returning to our considerations, let us denote the Frattini quotients $\mathbf{O}_p(G)/\Phi(\mathbf{O}_p(G))$ by V_p. Then the dimensions d_p of these vector spaces are bounded by

$$d_p \coloneqq \log_p|\mathbf{O}_p(G)/\Phi(\mathbf{O}_p(G))| \leq \log_p|\mathbf{O}_p(G)| \leq \log_p|G| \leq \log_2(n),$$

where $|G| \leq n$ (as postulated at the beginning of the section).

Next we consider the homomorphisms

$$\alpha_p\colon \operatorname{Aut}(\mathbf{O}_p(G)) \to \operatorname{Aut}(V_p) = \mathrm{GL}_{d_p}(p).$$

Note that the homomorphism α_p is well-defined only since $\Phi(\mathbf{O}_p(G))$ is characteristic in $\mathbf{O}_p(G)$. The image of G under

$$\alpha_p \circ \pi_p \circ \varphi\colon G \to \operatorname{Aut}(\mathbf{F}(G)) = \prod_p \operatorname{Aut}(\mathbf{O}_p(G)) \to \operatorname{Aut}(\mathbf{O}_p(G)) \to \mathrm{GL}_{d_p}(p)$$

in $\mathrm{GL}_{d_p}(p)$ is solvable since G is solvable. We next wish to bound the solvability class of this image in terms of the dimension.

The subsequent result is due to Mike F. Newman (see [12], page 1, Theorem $\mathrm{A_S}$) and we leave it as a black box here.

Lemma 17 (Mike F. Newman)**.** *A solvable subgroup N of $\mathrm{GL}_d(p)$ has solvability class at most*

$$5\log_9(d) + C_7.$$

for a constant $C_7 = 10 - \frac{15\log(2)}{2\log(3)} = 5.268026848\ldots > 0$ and $d \geq 66$.

Thus, when $\pi_p\colon \mathrm{Aut}(\mathbf{F}(G)) \to \mathrm{Aut}(\mathbf{O}_p(G))$ is the projection, the image $(\alpha_p \circ \pi_p \circ \varphi)(G)$ of G in $\mathrm{GL}_{d_p}(p)$ has solvability class at most

$$5\log_9(\log_2(n)) + C_7 = \frac{5\log(\log(n))}{2\log(3)} + C_8$$

where $n \geq |G|$ is large and $C_8 = C_7 - \frac{5\log(\log(2))}{2\log(3)} = 10 - \frac{15\log(2)+5\log(\log(2))}{2\log(3)} = 6.102062942\ldots$.

Next we need a result of Burnside.

Lemma 18. *Let P be a finite p-group. Then the kernel of the homomorphism*

$$\mathrm{Aut}(P) \to \mathrm{Aut}(P/\Phi(P))$$

is itself a p-group.

Remark 3. As we mentioned already, the homomorphism $\mathrm{Aut}(P) \to \mathrm{Aut}(P/\Phi(P))$ is well-defined only since $\Phi(P)$ is *characteristic* in P. Thus each automorphism α of P permutes the cosets of $\Phi(P)$ by sending $g\Phi(P)$ to $\alpha(g)\Phi(P)$.

The proof is taken from [5] (Theorem 1.4, Chapter 5).

Proof. Define $\overline{P} \coloneqq P/\Phi(P)$ and set $p^a \coloneqq |\overline{P}|$, $p^b \coloneqq \Phi(P)$. It is known as a property of $\Phi(P)$ that for any subset $S \subseteq P$ we have $\langle S, \Phi(P)\rangle = P$ if and only if $\langle S\rangle = P$ (see Lemma 15). Thus S generates P if and only if the image $\overline{S} \subseteq \overline{P}$ generates $\overline{P}$. Hence, as $\overline{P}$ is elementary abelian of dimension a (see Lemma 16), it can be generated by no fewer than a elements and the same holds for P.

Assume $(s_i)_{i=1}^a$ is an ordered minimal generating system of P. Then if we pick $\varphi_i \in \Phi(P)$, we have that $\varphi_i s_i$ is as well a minimal generating set for P. On the other hand, it is clear that any ordered minimal generating system of P whose images in $\overline{P}$ are $(\overline{s}_i)_{i=1}^a$ is of this form. Thus there are exactly $(p^b)^a = p^{ab}$ minimal ordered generating sets of P whose images are the ordered set $(\overline{s}_i)_{i=1}^a$. Let M be the set of all these tuples.

Let α be an automorphism of P of p'-order (i.e., of order prime to p) which induces the identity on $\overline{P}$. It is then easy to see that for a tuple $(s_i)_{i=1}^a \in M$, we have $(\alpha(s_i))_{i=1}^a \in M$, since α fixes the cosets of $\Phi(P)$. Thus α permutes M.

Now clearly the number of elements in a cycle of this permutation action is a divisor of the order $t \coloneqq \mathrm{ord}(\alpha)$. Assume some cycle had order $s < t$. Then set $\alpha' \coloneqq \alpha^s$. Note that α' fixes each element of this cycle and thus fixes some ordered minimal generating set $(s_i')_{i=1}^a$ of P. But this set generates P and thus α' must be the identity contradicting $s < t = \mathrm{ord}(\alpha)$. Thus the permutation of M decomposes into a product of cycles of length t. Hence t must divide $p^{ab} = |M|$. On the other hand, t was prime to p, so $t = 1$ and $\alpha = \mathrm{id}_P$ is the identity on P. Hence there are no p'-elements in the kernel of the map $\mathrm{Aut}(P) \to \mathrm{Aut}(P/\Phi(P))$. Thus this kernel is a p-group. □

From the previous lemma (Lemma 18) it follows that the kernel of the map

$$\alpha_p \colon \operatorname{Aut}(\mathbf{O}_p(G)) \to \mathrm{GL}_{d_p}(p)$$

is a p-group. Therefore the kernel of the map

$$\alpha = \prod_p \alpha_p \colon \prod_p \operatorname{Aut}(\mathbf{O}_p(G)) = \operatorname{Aut}(\mathbf{F}(G)) \to \prod_p \mathrm{GL}_{d_p}(p)$$

is a direct product of p-groups (running over all primes p dividing $|G|$) and hence nilpotent. Restricting the map α to $\operatorname{im}(\varphi) = \varphi(G)$, the kernel remains nilpotent and the image has solvability class at most

$$\frac{5\log(\log(n))}{2\log(3)} + C_8$$

since every component of $(\alpha_p \circ \pi_p \circ \varphi)(G)$ in the product has this property. We now apply the extension lemma (Lemma 4) to the map

$$\alpha|_{\operatorname{im}(\varphi)} \colon \operatorname{im}(\varphi) \to (\alpha \circ \varphi)(G) \subseteq \prod_p \mathrm{GL}_{d_p}(p).$$

Consider the shortest non-trivial word which is a law for every nilpotent group of order at most n. Then by Proposition 2, we can bound the length of this word by

$$(C_6 + o(1)) \log(n)^{\nu}.$$

This is a law for the kernel of $\alpha|_{\operatorname{im}(\varphi)}$.

At the same time, consider the shortest non-trivial word which is a law for all groups of solvability class at most

$$\frac{5\log(\log(n))}{2\log(3)} + C_8.$$

This is a law for the image $(\alpha\circ\varphi)(G)$. By Proposition 1, the length of this word is bounded by

$$C_1 \iota^{\frac{5\log(\log(n))}{2\log(3)}+C_8} + o(1) = C_9 \log(n)^{\frac{5\log(\iota)}{2\log(3)}} + o(1) = 2,570.55273\ldots \log(n)^{2.890457\ldots} + o(1).$$

Combining these two words using the extension lemma, again we obtain a law for $\operatorname{im}(\varphi) = \varphi(G)$. This gives a non-trivial word whose length is bounded by

$$(21,580.26355\ldots + o(1)) \log(n)^{4.331612776\ldots}.$$

Combining this again with the word $aba^{-1}b^{-1}$, which is a law for $\ker(\varphi) = \mathbf{Z}(\mathbf{F}(G))$, gives a non-trivial law for all solvable groups G of order at most n whose length is bounded by

$$(86,321.05422\ldots + o(1)) \log(n)^{4.331612776\ldots}.$$

We thus have proven the following fact (cf. Proposition 3.2 in [14]).

Proposition 4 (Short non-trivial laws for solvable groups)**.** *For $n \in \mathbb{Z}_+$ there exists a non-trivial two-letter word w of length*

$$\ell(w) \leq (C_{10} + o(1)) \log(n)^{\lambda}$$

which is a law for all solvable groups of order at most n, where $C_{10} \coloneqq 86,321.05422\ldots$ and $\lambda \coloneqq 4.331612776\ldots$.

This ends this chapter.

Chapter 3

Semi-simple groups

In this chapter, we study laws for so-called *semi-simple groups*. This becomes necessary since we want to use the following short exact sequence

$$\mathbf{1} \to \mathbf{S}(G) \to G \to G/\mathbf{S}(G) \to \mathbf{1}$$

to get group laws for the group G from laws for its *solvable radical* $\mathbf{S}(G)$ and the *semi-simple quotient* $G/\mathbf{S}(G)$ using the extension lemma (Lemma 4) as explained in the introduction. We thus start with some definitions.

3.1 Definitions and basic facts

At first we recall the notion of a *semi-simple group* and the *solvable radical* in the sense of Fitting.

Definition 11 (Solvable radical, semi-simple group)**.** *Let G be a finite group. Then we call the unique largest normal solvable subgroup of G the* solvable radical $\mathbf{S}(G)$*. If $\mathbf{S}(G) = \mathbf{1}$ we call G* semi-simple*.*

Remark 4. The solvable radical $\mathbf{S}(G)$ is well-defined for a finite group G, since when N and M are normal and solvable in G, then MN is normal and solvable.

Proof. To see this, consider the short exact sequence

$$\mathbf{1} \to M \to MN \to MN/M \cong N/(N \cap M) \to \mathbf{1}.$$

Since N is solvable, there is $n_N \in \mathbb{N}$ such that $N^{(n_N)} = \mathbf{1}$ and thus $(N/(N \cap M))^{(n_N)} = (MN/M)^{(n_N)} = \mathbf{1}$. Hence $(MN)^{(n_N)} \subseteq M$. Moreover, as M is solvable, we have an integer n_M such that $M^{(n_M)} = \mathbf{1}$. From this we deduce that $(MN)^{(n_M+n_N)} = \mathbf{1}$, so MN is solvable. □

Remark 5. Semi-simple groups can thus be characterized as groups without non-trivial normal abelian subgroups.

Proof. On the one hand side, if G is semi-simple, it cannot have a non-trivial abelian normal subgroup, since this would be a solvable normal subgroup. On the other hand side, if G is not semi-simple, it has a non-trivial solvable normal subgroup S. Let N be the subgroup of S which occurs as the group of the derived series of S just before the trivial group $\mathbf{1}$. Then N is characteristic in S and $[N, N] = \mathbf{1}$. Therefore N is abelian and normal in G. □

In order to describe the structure of semi-simple groups, we need to introduce the notion of the *wreath product*.

Definition 12 (Wreath product)**.** *Let G and H be groups such that G acts on a set Ω via the homomorphism $\alpha\colon G \to \mathrm{Sym}(\Omega)$. Then we define the* wreath product $H \wr G$ *as the semi-direct product $H^\Omega \rtimes G$ with respect to the induced action $\alpha'\colon G \to \mathrm{Aut}(H^\Omega)$. More visually, one can think of this group as the subgroup of the linear group $\mathrm{GL}(\Omega, \mathbb{Z}[H])$ which is generated by the subgroup of permutation matrices from $\alpha(G)$ and the subgroup of diagonal matrices with entries in H.*

Indeed, one can show that semi-simple groups are always groups of the following kind.

Theorem 2. *Let G be a semi-simple group. Then there exist different non-abelian simple groups H_i and integers $k_i \geq 1$ (for $i = 1, \ldots, l$) such that*

$$\prod_{i=1}^{l} H_i^{k_i} \leq G \leq \mathrm{Aut}\left(\prod_{i=1}^{l} H_i^{k_i}\right) \cong \prod_{i=1}^{l} \mathrm{Aut}(H_i) \wr \mathrm{S}_{k_i}.$$

From this theorem it is clear that we need to understand the automorphism groups of simple groups and the symmetric groups (i.e., we need to find short non-trivial laws for these). We prove the preceding theorem in four steps (which means we deduce it from the following four lemmas).

Definition 13. *Let G be a group. The subgroup which is generated by all minimal normal subgroups of G is called the* socle $\mathrm{soc}(G)$ *of G.*

Remark 6. The socle $\mathrm{soc}(G)$ is clearly characteristic in G since every automorphism of G permutes the minimal normal subgroups of G.

The first step is done by proving the following lemma whose statement is essentially the statement of Theorem 2. We are only left to explain the form of $\mathrm{soc}(G)$ and of its automorphism group $\mathrm{Aut}(\mathrm{soc}(G))$.

Lemma 19. *Let G be a semi-simple group. Then* $\mathrm{soc}(G) \subseteq G \leq \mathrm{Aut}(\mathrm{soc}(G))$.

Proof. Let $\mathrm{soc}(G)$ be the characteristic subgroup of G which is generated by all minimal normal subgroups of G. Then clearly $\mathrm{soc}(G) \subseteq G$ and G acts on $\mathrm{soc}(G)$ via conjugation. We need to show that the corresponding map $G \to \mathrm{Aut}(\mathrm{soc}(G))$ is actually injective, i.e., it has trivial kernel $\mathbf{C}_G(\mathrm{soc}(G))$.

So let us assume that the centralizer $\mathbf{C}_G(\mathrm{soc}(G))$ is non-trivial. Then $\mathbf{C}_G(\mathrm{soc}(G))$ is normal (since $\mathrm{soc}(G)$ is normal) and so there is a minimal normal subgroup $M \subseteq \mathbf{C}_G(\mathrm{soc}(G))$ of G. Thus it holds that $[M, M] \subseteq [\mathrm{soc}(G), \mathbf{C}_G(\mathrm{soc}(G))] = \mathbf{1}$. But this contradicts the fact that G has no non-trivial normal solvable subgroup, as M is abelian and thus solvable. Hence it follows that $\mathbf{C}_G(\mathrm{soc}(G)) = \mathbf{1}$. Thus G embeds into $\mathrm{Aut}(\mathrm{soc}(G))$. This ends the proof. □

In the following lemma, we identify $\mathrm{soc}(G)$ as the internal direct product of *characteristically simple* subgroups of G, i.e., it is isomorphic to the abstract direct product of characteristically simple groups, which is embedded into G.

Lemma 20. *Let G be a group. Each minimal normal subgroup is characteristically simple, i.e., it has no non-trivial characteristic subgroups.*

Proof. Assume the opposite, namely that M is minimal normal in G with characteristic subgroup N. Then each element $g \in G$ induces an automorphism of M which fixes N, as N is characteristic in M. Thus N is also normal in G and by minimality of M it follows that $N = \mathbf{1}$ or $N = M$. □

The third step in the proof of Theorem 2 is the subsequent lemma which characterizes the finite characteristically simple groups as direct powers of finite simple groups.

Lemma 21. *A finite characteristically simple group G is the direct power of a finite simple group. The converse is also true.*

The idea of proof is taken from [1] and [2].

Proof. Take a minimal normal subgroup $\mathbf{1} \subset M \trianglelefteq G$ and consider its image $\alpha(M)$ under all automorphisms $\alpha \in \mathrm{Aut}(G)$. Clearly the groups $M^{\mathrm{Aut}(G)}$ generate a characteristic subgroup of G so the whole group G. Moreover, if $\alpha(M)$ and M are not the same subgroup, they must intersect trivially since M was minimal normal. Thus G is a direct power of the group M. Now let $N \trianglelefteq M$ be normal in M. Then, since G is a direct product of copies $\alpha(M)$ (for $\alpha \in \mathrm{Aut}(G)$), these copies centralize M and so N. Thus N is normal in G, forcing that $N = \mathbf{1}$ or $N = M$, so M is simple.

Conversely, assume that $G = M^k$ for a simple group M and $k \in \mathbb{Z}_+$. In the abelian case, G is an $\mathbb{F}_p$-vectors space. In particular, there are automorphisms carrying each subspace of dimension d to another subspace of the same dimension. Hence there are no non-trivial characteristic subgroups.

In the opposite case, each factor M is a perfect simple group (*perfect* means $[M, M] = M$). For any normal subgroup N of G, N must be a subdirect product of some subset of the factors $M_1, \ldots, M_k \cong M$. This holds since the surjective image of N under the projections $\pi_i \colon G \to M_i$ must again be normal and thus either M_i or $\mathbf{1}$. But then, since all M_i are perfect, it follows that N must be the full internal direct product of the corresponding M_i where $\pi_i(N) = M_i$.

Let us briefly prove this fact. It suffices to prove it for a product of two groups. So let K and L be perfect groups such that the group $J \subseteq K \times L$ is a subdirect product and normal in $K \times L$, i.e., $\pi_K(J) = K$ and $\pi_L(J) = L$, where $\pi_K \colon K \times L \to K$ and $\pi_L \colon K \times L \to L$ are the projection maps. Then we can naturally embed K and L into $K \times L$ via the inclusions $K \times \{1_L\}, \{1_K\} \times L \subseteq K \times L$. We then have that $[J, K \times \{1_L\}] \subseteq K \times \{1_L\}$ since $K \times \{1_L\}$ is normal in $K \times L$. Furthermore, $\pi_K([J, K \times \{1_L\}]) = [\pi_K(J), K] = [K, K] = K$. Thus $[J, K \times \{1_L\}] = K \times \{1_L\}$ and similarly $[J, \{1_K\} \times L] = \{1_K\} \times L$. Hence $K \times \{1_L\}, \{1_K\} \times L \subseteq J$ as J is normal, so $J = K \times L$ is the full product.

Returning to the proof of our lemma, we see that the normal subgroups of $G = \prod_{i=1}^{k} M_i$ are precisely the subgroups $G = \prod_{i \in I} M_i$ for $I \subseteq \{1, \ldots, n\}$. But we may only take $I = \emptyset$ or $I = \{1, \ldots, n\}$ to get a characteristic subgroup, since S_n operates on the product naturally. □

We can complete the proof of Theorem 2 by identifying the socle $\mathrm{soc}(G)$ of G as the internal direct product of minimal normal subgroups, which are direct powers of simple groups by the preceding two lemmas. The last step is to prove the statement about the shape of the automorphism group as a subgroup of a wreath product of simple groups and symmetric groups.

Lemma 22. *Let H_i be pairwise non-isomorphic finite non-abelian simple groups and $k_i \geq 1$ integers ($i \in \{1, \ldots, l\}$, $l \in \mathbb{Z}_+$). Then*

$$\mathrm{Aut}\left(\prod_{i=1}^{l} H_i^{k_i}\right) = \prod_{i=1}^{l} \mathrm{Aut}(H_i) \wr \mathrm{S}_{k_i}.$$

Proof. Let α be an automorphism from the left side of the equation. Then consider the image of the lth component $H_{i,l}$ of the group H_i under α ($1 \leq l \leq k_i$). This component is embedded normally into the above product as

$$H_{i,l} \times \mathbf{1} \trianglelefteq \prod_{i=1}^{l} H_i^{k_i}.$$

Let $\pi_{j,m}$ be the projection onto the factor $H_{j,m}$ ($1 \leq m \leq k_j$). Then $\pi_{j,m}(\alpha(H_{i,l}))$ is normal in $H_{j,m}$ as a surjective image of a normal subgroup of $\prod_{i=1}^{l} H_i^{k_i}$ and hence either $\mathbf{1}$ or $H_{j,m}$ as $H_{j,m}$ is simple. The latter can only happen if $i = j$ since $H_i \ncong H_j$ for $i \neq j$. Moreover, $\pi_{i,m}(\alpha(H_{i,l})) = H_{i,m}$ can only happen for precisely one m, since otherwise $\alpha(H_{i,l})$ would be a subdirect product of more than one copies of H_i. Thus (since H_i is perfect), according to the proof of the previous lemma, it would be the full direct product of more than one copy of H_i, a contradiction due to cardinality reasons. From this we can immediately deduce the theorem. □

Reconsidering Theorem 2, we see that, if we want to find short non-trivial laws for all semi-simple groups of order at most n, we need to find short non-trivial laws for groups

of the form

$$\prod_{i=1}^{l} \mathrm{Aut}(H_i) \wr \mathrm{S}_{k_i}.$$

In order to find these laws, at first, we need to study the wreath products $\mathrm{Aut}(H) \wr \mathrm{S}_k$ (where H is simple; $k \geq 1$) which fit into the short exact sequence

$$\mathbf{1} \to \mathrm{Aut}(H)^k \to \mathrm{Aut}(H) \wr \mathrm{S}_k \to \mathrm{S}_k \to \mathbf{1}.$$

Thus, in view of the extension lemma (Lemma 4), we need to find short non-trivial laws for the automorphism groups of simple groups and symmetric groups.

Hence in the next sections, we focus on symmetric and simple groups, respectively. We will see how non-trivial laws for simple groups can be used (using the characterization of finite simple groups and Schreier's conjecture stating that for each non-abelian simple group G the outer automorphism group $\mathrm{Out}(G)$ is solvable of class at most three) to find non-trivial laws for their automorphism group.

3.2 Laws for the symmetric group S_n

In this section, we present the content of the article [11]. We have the following result due to Landau (see [9]).

Theorem 3. *The element of maximal order in S_n has order at most*

$$\exp\left((1+o(1))\sqrt{n\log(n)}\right).$$

To establish this result, we make the following definitions.

Definition 14. *Set $g\colon \mathbb{Z}_+ \to \mathbb{Z}_+$ where $g(n)$ is the maximum order of an element of S_n.*

Definition 15. *Set $s\colon \mathbb{Z}_+ \to \mathbb{N}$ to*

$$s(n) = s(p_1^{e_1} \cdots p_m^{e_m}) = p_1^{e_1} + \cdots + p_m^{e_m},$$

where $p_1^{e_1} \ldots p_m^{e_m}$ is the prime factorization of n.

Lemma 23. *Let $a_1, \ldots, a_n \in \mathbb{Z}_+$ be positive integers and a be its least common multiple. Then we have*

$$s(a) \leq \sum_{i=1}^{n} a_i.$$

Proof. Suppose $a_1, \ldots, a_n$ is a counterexample with $\sum_{i=1}^{n} a_i$ as small as possible.

At first, we note that $1 \notin \{a_1, \ldots, a_n\}$, since otherwise we could delete the 1's to get a smaller sum, but the least common multiple of the a_i would be the same. The next thing to notice is that each a_i must be a prime power ($\neq 1$). If not, we would have $a_i = xy$ for

some i and $x > 1$ and $y > 1$ relatively prime. But then we can replace a_i by x and y in the sequence $(a_1, \ldots, a_n)$ to obtain a smaller sum as (w.l.o.g., $x < y$)

$$x + y \leq x + (x-1)y = xy + x - y < xy = a_i.$$

Lastly, all the prime powers a_i must be distinct since otherwise, if $a_i = p^{e_1}$ and $a_j = p^{e_2}$ with $e_1 \leq e_2$, we can delete a_i to obtain a smaller sum and the same least common multiple. But in this case it is easy to verify that the sequence $(a_i)_{i=1}^n$ is not a counterexample since then $s(a) = \sum_{i=1}^n a_i$. This ends the proof. □

From this lemma we now obtain the following corollary.

Corollary 6. *For $m \in \mathbb{Z}_+$ there is a permutation of order m in S_n if and only if $s(m) \leq n$.*

Proof. If

$$m = \prod_{i=1}^{l} p_i^{e_i}$$

and thus

$$s(m) = \sum_{i=1}^{l} p_i^{e_i} \leq n$$

we can simply take a permutation with cycles of length $p_i^{e_i}$ and if necessary cycles of length one. Conversely, let $a_1, \ldots, a_l \neq 1$ be the length of the cycles of a permutation $\sigma \in \mathrm{S}_n$ having order m. Then the sum of the cycle length is n and their least common multiple is m. Thus $s(m) \leq \sum_{i=1}^{l} a_i \leq n$ by the previous lemma. □

Corollary 7. *There is a permutation $\sigma \in \mathrm{S}_n$ having order $g(n)$ whose cycle length are all prime powers.*

Proof. The construction is carried out in the proof of the previous lemma (if $(a_i)_{i=1}^l$ is the sequence of cycle length, then we can split each cycle $a_i = xy$ consisting of two coprime non-trivial factors to obtain the same least common multiple and a smaller sum of the new cycle lengths $a_1, \ldots, a_{i-1}, x, y, a_{i+1}, \ldots, a_l$). □

Corollary 8. *It holds that*

$$g(n) = \max_{s(m) \leq n} m.$$

Proof. It holds that $s(g(n)) \leq n$, thus $g(n)$ does not exceed the maximum on the right. Conversely, by Corollary 6, if $s(m) \leq n$ there is a permutation in S_n of order m. Thus the opposite inequality holds as well. □

Now we introduce a new function $f \colon \mathbb{Z}_+ \to \mathbb{Z}_+$ and a notation convention:

Notation 5. We write $a(n) \sim b(n)$ if and only if

$$\lim_{n \to \infty} \frac{a(n)}{b(n)} = 1.$$

The definition of the function f is somehow greedy.

Definition 16. *Let n be a positive integer. Start with an empty list $()$ of primes. Then take the smallest prime p which is not already contained in the list $(p_i)_{i=1}^{l}$ and check if*

$$\sum_{i=1}^{l} p_i + p \leq n.$$

If this holds, append p to the list, i.e., $p_{l+1} \coloneqq p$. Otherwise, the list is complete and f is defined as

$$f(n) = \prod_{i=1}^{l} p_i.$$

Example 4. For $n = 8$ we obtain a list $(2, 3)$ and thus $f(8) = 6$ and obviously this is far from being the maximal order in S_8 which is 15.

We will prove the following result.

Theorem 4. *We have that* $\log(f(n)) \sim \log(g(n))$.

For the proof, we will need two lemmas. We already have that $f(n) \leq g(n)$ by the definition of $g(n)$. Thus we only need to bound $g(n)$ by some appropriate expression in $f(n)$.

Lemma 24 (Shah)**.** *Let $q_1 < \cdots < q_l$ be the primes dividing $g(n)$ and p be the biggest prime such that the sum of all primes smaller than p is less than or equal to n. Then it holds that*

$$\sum_{i=1}^{l} \log(q_j) < 2 + \log(f(n)) + \log(p)$$

Proof. At first, observe that $n \mapsto \log(n)/n$ is decreasing for $n \geq 3$ since its derivative

$$n \mapsto \frac{1 - \log(n)}{n^2}$$

is negative. Thus for $3 \leq x \leq y$ we have the two inequalities

$$\frac{x}{\log(x)} \log(y) \leq y \text{ and } x \leq \frac{y}{\log(y)} \log(x).$$

In the case $n = 1$, it holds that $p = 2$ and the conclusion of the lemma is obvious. So assume $p \geq 3$. Set $q_1, \ldots, q_s$ to be the primes dividing $g(n)$ which do not exceed p and $p_1, \ldots, p_t$ to be the odd primes not exceeding p which do not divide $g(n)$. The list $q_1, \ldots, q_s, p_1, \ldots, p_t$ contains every prime not exceeding p exactly once except for 2 (which could be omitted). We have that

$$\sum_{i=1}^{l} q_i \leq s(g(n)) \leq n \leq \sum_{p' \leq p \text{ prime}} p'.$$

Canceling the q_i for $i = 1, \ldots, s$ on both sides of the last inequality gives

$$\sum_{i=s+1}^{l} q_i \leq 2 + \sum_{j=1}^{t} p_j.$$

We now use that $3 \leq p < q_i$ for $i = s+1, \ldots, l$ and $3 \leq p_j \leq p$ $(j = 1, \ldots, t)$ and the first observation in the proof (namely that $\log(q_i) \leq (\log(p)/p)q_i$ and $(\log(p)/p)p_i \leq \log p_i$) to get

$$\begin{aligned}\sum_{i=s+1}^{l} \log(q_i) &< \frac{\log(p)}{p} \sum_{i=s+1}^{l} q_i \\ &\leq 2\frac{\log(p)}{p} + \frac{\log(p)}{p} \sum_{j=1}^{t} p_j \\ &\leq 2 + \sum_{j=1}^{t} \log(p_j).\end{aligned}$$

Now we add the terms $\log(q_i)$ for $i = 1, \ldots, s$ to both sides to get the desired result (since $q_1, \ldots, q_s, p_1 \ldots, p_t$ is the list of primes not exceeding p except for 2 which could be omitted). □

We need a further lemma.

Lemma 25. *Let q be a prime and $e \in \mathbb{Z}_{>1}$ and let p be as in the previous lemma. If $q^e | g(n)$, then $q^e \leq 2p$, in particular $q \leq \sqrt{2p}$.*

Proof. Let Q be the smallest prime not dividing $g(n)$, then all primes less than Q divide $g(n)$, so their sum is at most $s(g(n))$. Thus, since the sum of primes not exceeding p is greater than n, we find that $Q \leq p$. It is therefore sufficient to show that $q^e \leq 2Q$. Thus assume the contrary, namely that $q^e > 2Q$, and define $N \in \mathbb{Z}_+$ by $q < Q^N < qQ$ (equality cannot occur since $q \mid g(n)$ while Q does not divide $g(n)$). Now set

$$m = (Q^N/q)g(n) > g(n).$$

Then we have

$$s(m) = s(g(n)) + (Q^N - q^e + q^{e-1}).$$

Now we claim that the last expression in parenthesis is negative. If $q < Q$, we have $N = 1$ and thus obtain

$$-q^e + q^{e-1} \leq -\frac{q^e}{2} < -\frac{2Q}{2} = -Q$$

from the assumption $q^e > 2Q$. On the other hand, if $q > Q$ (i.e., $q - 1 \geq Q$) we obtain similarly as $Q^N < qQ$ and $e > 1$

$$Q^N - q^e + q^{e-1} < qQ - q(q-1) \leq qQ - qQ = 0.$$

However, this is a contradiction since $m > g(n)$ and we can find a new partition into cycles (by deleting the cycle of length q^e and replacing it by two cycles of length q^{e-1} and Q^N, respectively) such that the corresponding new element of S_n has a bigger order. □

Now we are able to prove Theorem 4.

Proof. Assume that $\prod_{i=1}^{l} p_i^{e_i}$ is the prime factorization of $g(n)$. We can write

$$\log(g(n)) = \sum_{i,e_i=1} \log(p_i) + \sum_{j,e_j>1} e_j \log(p_j).$$

This can be estimated using the preceding two lemmas by

$$\log(g(n)) = \underbrace{\sum_{i,e_i=1} \log(p_i)}_{\leq 2+\log(f(n))+\log(p)} + \underbrace{\sum_{j,e_j>1} e_j \log(p_j)}_{\leq \sqrt{2p}(\log(2p))}.$$

The second estimate is correct since there are at most $\sqrt{2p}$ of the q_i having exponent $e_i > 1$ and for these we have by the previous lemma $\log(q_i^{e_i}) \leq \log(2p)$. During the analysis of $f(n)$, we will find that $\log(f(n)) > cp$ for some constant $c > 0$. Thus dividing both sides by $\log(f(n))$ gives the desired result of the theorem. □

Properties of $f(n)$

Now we can use the prime number theorem to prove that

$$\log(f(n)) \sim \sqrt{n \log(n)}.$$

The prime number theorem gives us two statements. Set

$$h(x) \coloneqq \sum_{p \leq x} p \text{ and } \theta(x) \coloneqq \sum_{p \leq x} \log(p).$$

Then the prime number theorem gives us

$$h(x) \sim \frac{x^2}{2\log(x)} \text{ and } \theta(x) \sim x.$$

To get an expression for $\log(f(n))$, we must solve the inequality

$$h(p-1) \leq n < h(p)$$

and compute $\theta(p-1)$. Fortunately, we can find an almost inverse to the function $H\colon x \mapsto \frac{x^2}{2\log(x)}$, namely $G\colon y \mapsto \sqrt{y\log(y)}$. It is easily verified that $(G \circ H)(x) \sim (H \circ G)(x) \sim x$. Thus a solution to the former inequality for p must be $p \sim \sqrt{n\log(n)}$ and $\theta(p-1) \sim \theta(p) \sim \log f(n) \sim \sqrt{n\log(n)}$.

All in all, we are now ready to derive the following theorem.

Theorem 5. *We have that for all $n \in \mathbb{Z}_+$*

$$\max_{\sigma \in \mathrm{S}_n} \mathrm{ord}(\sigma) = \exp\left((1+o(1))\sqrt{n\log(n)}\right).$$

In particular, there is a non-trivial law of length at most

$$8\exp\left(3(1+o(1))\sqrt{n\log(n)}\right)$$

holding simultaneously for S_k (where $k = 1, \ldots, n$).

Proof. The first statement is clear from the preceding considerations in this section. It follows from the fact that $\log(g(n)) \sim \log(f(n)) \sim \sqrt{n\log(n)}$. The second fact is just the application of Corollary 2 to the words $a, \ldots, a^{g(n)}$ (where a is a generator of a free group). □

3.3 Laws for simple groups

In this section, we deal with laws for simple groups. However, we will see that the estimates which one gets from the classification of finite simple groups for non-trivial laws of simple groups are much worse than the estimates for finite solvable or nilpotent groups. The main result of this section is the following proposition (cf. Proposition 4.1 in [14]).

Proposition 5. *For $n \in \mathbb{Z}_+$ there is a non-trivial word $w \in \mathbf{F}_2$ of length*

$$\ell(w) \leq (C_{11} + o(1))\frac{n}{\log(n)^2}$$

which is a law for all finite non-abelian simple groups of order at most n.

Proof. The proof makes use of classification of finite simple groups by considering each family of simple groups separately and using the commutator lemma (Lemma 3) or Corollary 2. The worst case (i.e., the family of groups where we have the longest non-trivial laws of minimal length compared to the size) is $\mathrm{PSL}_2(q)$ where $q = p^e$ is some prime power. Let x be an element of $\mathrm{PSL}_2(q)$. Then there are three cases to consider when we want to find the order of x.

In the first case, x comes from a diagonalizable matrix

$$x \equiv \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$$

then $\mathrm{ord}(x) \mid q-1$ since $\alpha^{q-1} = \beta^{q-1} = 1$ as $\alpha, \beta \in \mathbb{F}_q^\times$. In the second case, x comes from an unipotent matrix

$$x \equiv \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} = 1 + N$$

where 1 is the identity and N the nilpotent part of x. Taking the pth power gives $(1+N)^p = 1 + pN + \cdots = 1$. Thus $\mathrm{ord}(x) = p$ in this case (when $N \neq 0$; otherwise $x = 1$).

The last case is the irreducible case. Here we can diagonalize the matrix of x in $\mathbb{F}_{q^2}$ and its eigenvalues are conjugate by the automorphism $y \mapsto y^q$.

$$x \equiv \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^q \end{pmatrix}$$

Thus we have $\alpha\alpha^q = \alpha^{q+1} = 1$ and so $\mathrm{ord}(x) = q + 1$. Using Lemma 3, we obtain a non-trivial law of length less then $24(2q+p+3) \leq 72(q+1)$ for $\mathrm{PSL}_2(q)$. The last step to obtain a non-trivial law for all $\mathrm{PSL}_2(q)$ where $|\mathrm{PSL}_2(q)| \leq n$ is to use Corollary 2 again. The size of $\mathrm{PSL}_2(q)$ is precisely

$$\frac{1}{2(q-1)}(q^2-1)(q^2-q) = \frac{1}{2}(q^2-1)q$$

when q is odd (otherwise we need not to divide by two). It holds that

$$\frac{1}{2}(q^2-1)q \leq n \Rightarrow q \leq \sqrt[3]{(1+o(1))2n}.$$

Combining all the laws for prime powers less than or equal to $\sqrt[3]{(1+o(1))2n}$, we obtain, by applying Corollary 2, a non-trivial law of length at most

$$8 \cdot \left(\frac{3\sqrt[3]{(1+o(1))2n}}{\log((1+o(1))2n)} \right)^2 \cdot 72\sqrt[3]{(1+o(1))2n} \leq 10,368(1+o(1))\frac{n}{\log(n)^2}.$$

This proves the claim for the simple groups $\mathrm{PSL}_2(q)$. For all other simple groups, explicit considerations show that the order of an element is bounded by $D\sqrt[4]{n}$ for some explicit constant $D > 0$.

In fact, we have shown that the maximal order of an element from the alternating group A_n is

$$\exp\left((1+o(1))\sqrt{n\log(n)}\right)$$

which is much smaller than $\sqrt[4]{n!/2}$. For the other finite simple groups we can cite the results from [7] or [14], respectively. The following table contains a lower bound for the size of each classical Chevalley and Steinberg group and an upper bound for the maximal element order (again up to multiplicative constants). Each entry is just up to asymptotic equivalence as $q \to \infty$.

	$\mathrm{A}_d(q)$	${}^2\mathrm{A}_d(q^2)$	$\mathrm{B}_d(q)$	$\mathrm{C}_d(q)$	$\mathrm{D}_d(q)$	${}^2\mathrm{D}_d(q^2)$
restriction		$d > 1$	$d > 1$	$d > 2$	$d > 3$	$d > 3$
size	q^{d^2+2d}/d	q^{d^2+2d}/d	q^{2d^2+d}	q^{2d^2+d}	q^{2d^2-d}	q^{2d^2-d}
meo	q^d	q^d	q^d	q^d	q^d	q^d

Another table presents the families of bounded rank containing the twisted forms of classical Chevalley groups (which only exist for small rank), exceptional Chevalley groups and Suzuki–Ree groups.

	${}^2\mathrm{B}_2(q)$	${}^3\mathrm{D}_4(q^3)$	$\mathrm{F}_4(q)$	${}^2\mathrm{F}_4(q)$	$\mathrm{E}_6(q)$	${}^2\mathrm{E}_6(q^2)$	$\mathrm{E}_7(q)$	$\mathrm{E}_8(q)$	$\mathrm{G}_2(q)$	${}^2\mathrm{G}_2(q)$
size	q^5	q^{28}	q^{52}	q^{26}	q^{78}	q^{78}	q^{133}	q^{248}	q^{14}	q^7
meo	q	q^4	q^4	q^2	q^6	q^6	q^7	q^8	q^2	q

Sporadic groups and Tits' group are not interesting for us, since we are interested only in asymptotic bounds. From these tables we see that the only case for which the maximal element order cannot be bounded by $D\sqrt[4]{n}$ is $\mathrm{A}_1(q) = \mathrm{PSL}_2(q)$. Thus applying the Corollary 2 to the words $a^1, \ldots, a^{\lceil D\sqrt[4]{n}\rceil}$ yields a non-trivial word of length $En^{3/4}$ which is a law for all other simple groups of order less than n (for large n). Combining this word with the law for the groups $\mathrm{PSL}_2(q)$, using the same corollary again, yields a non-trivial word of length at most

$$8 \cdot 2^2 \cdot 10,368(1+o(1))\frac{n}{\log(n)^2} = (331,776+o(1))\frac{n}{\log(n)^2} = (C_{11}+o(1))\frac{n}{\log(n)^2}.$$

□

Actually, one can prove that for $\mathrm{PSL}_2(p)$ there is no non-trivial law shorter than p. This result is taken from [4] (page 9, Lemma 12).

Lemma 26. *The length of the shortest non-trivial law* $\mathrm{girth}_2(\mathrm{PSL}_2(p))$ *is at least* p.

Proof. We argue by contradiction, i.e., we take a non-trivial word of length less than p and show that it is not a law for $\mathrm{PSL}_2(p)$. W.l.o.g., we need to consider only two cases (up to 'cyclic rotation').

The first case is that our word is conjugate to a^n or b^n in $\mathbf{F}_2 = \langle a, b\rangle$ $(n < p)$. Then we plug in for a the matrix

$$a = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

to get

$$w(a,b) = a^n = \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} \neq \mathrm{id}\,.$$

In the second case, our word is equivalent via conjugation to some word

$$w(a,b) = a^{k_1}b^{l_1}\cdots a^{k_n}b^{l_n},$$

where $k_i, l_i \in \mathbb{Z}\setminus\{0\}$ and $\sum_{i=1}^{n}(|k_i|+|l_i|) < p$. We plug in for a and b the following matrices

$$a = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$$

and

$$b = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

We prove inductively that the evaluation $w(a,b)$ has the following form

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} + k_1 l_1 \cdots k_n l_n x^{2n} \end{pmatrix},$$

where the f_{ij} are polynomials of degree at most $2n-1$. Indeed,

$$\begin{aligned} &\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}^{k_0} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}^{l_0} \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} + k_1 l_1 \cdots k_n l_n x^{2n} \end{pmatrix} \\ &= \begin{pmatrix} 1 & l_0 x \\ k_0 x & 1 + k_0 l_0 x^2 \end{pmatrix} \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} + k_1 l_1 \cdots k_n l_n x^{2n} \end{pmatrix} \\ &= \begin{pmatrix} f_{11} + l_0 x f_{21} & f_{12} + l_0 x (f_{22} + k_1 l_1 \cdots k_n l_n x^{2n}) \\ k_0 x f_{11} + (1 + k_0 l_0 x^2) f_{21} & k_0 x f_{12} + (1 + k_0 l_0 x^2)(f_{22} + k_1 l_1 \cdots k_n l_n x^{2n}) \end{pmatrix}, \end{aligned}$$

so by induction, f_{11} has degree at most $2n-2$ and f_{21}, f_{12}, f_{22} have degree at most $2n-1$. In particular, the degree of the polynomial $-f_{11} + f_{22} + k_1 l_1 \cdots k_n l_n x^{2n}$, which is the difference of the diagonal entries, lies between zero and p (since $2n < p$ by assumption). Thus we can find an $x \in \mathbb{F}_p$ such that $w(a,b)$ is not a scalar matrix.

Here we use the following fact: The polynomials $\mathbb{F}_p[X]$ of degree less than p are in one-to-one correspondence with the functions $\mathbb{F}_p \to \mathbb{F}_p$ via the evaluation map $\mathrm{ev}\colon f \mapsto (x \mapsto f(x))$. This can be shown via Lagrange interpolation.

Therefore any polynomial in $f \in \mathbb{F}_p[X]$ which evaluates to the zero map lies in the ideal $(X^p - X)$, so it is either zero or has degree at least p. □

We end this section by a question.

Question 3. *How long must a non-trivial two-letter word be which is a law for all groups* $\mathrm{PSL}_2(p)$ *of order at most* n*?*

Regrettably, we have not yet found a satisfactory answer to this question.

3.4 Laws for finite linear groups

Although we already discussed this case (indeed we just cited [14]), we present a theorem from the article [8] (page 17, Theorem 6.1).

Theorem 6. *Let* G *be a finite linear group of rank* d *over the field* $\mathbb{F}_q$*. There is a non-trivial law* w *for* G *with*

$$\ell(w) \leq 8d^{2\sqrt{2d}} q^d p d.$$

Proof. Consider a matrix $M \in \mathrm{GL}_d(q)$. By the Jordan decomposition, we can write $M = M_s M_u$, where M_s is semi-simple and M_u is unipotent. Let the characteristic polynomial χ have distinct irreducible factors of degrees d_i, where each degree is counted only once. When we set

$$e = \prod_{i=1}^{l} (q^{d_i} - 1)$$

we obtain that $M_s^e = 1$, since we have an isomorphism from the algebra $\mathbb{F}_q[M_s]$ into an algebra which is a sum of the fields $\mathbb{F}_{q^{d_i}}$ with the corresponding multiplicities. Now we want to count the different possibilities for the exponent e, i.e., the different possibilities for the degrees d_i. We have that

$$\frac{l^2}{2} \leq \frac{l(l+1)}{2} = \sum_{i=1}^{l} i \leq \sum_{i=1}^{l} d_i \leq d.$$

Thus we can choose at most $\sqrt{2d}$ times a degree d_i. Thus there are at most $d^{\sqrt{2d}}$ possibilities for e. Moreover, note that we have

$$e \leq q^d - 1 < q^d$$

Now since $M_s^e = 1$, we must have that $(M_s M_u)^e = M'$ is unipotent. Thus we need an additional exponent of $p^{\lceil \log_p(d) \rceil} < pd$ to trivialize the matrix M'. This exponent trivializes every unipotent matrix $U = 1 + N \in \mathrm{GL}_d(q)$ where N is the nilpotent part. Indeed, as the Frobenius map $x \mapsto x^p$ is an automorphism of $\mathbb{F}_q$, we have

$$(1+N)^{p^{\lceil \log_p(d) \rceil}} = 1 + N^{p^{\lceil \log_p(d) \rceil}} = 1.$$

Altogether, we can now apply Corollary 2 and obtain a non-trivial law w such that

$$\ell(w) \leq 8d^{2\sqrt{2d}} q^d p^{\lceil \log_p(d) \rceil} \leq 8d^{2\sqrt{2d}} q^d pd.$$

This finishes the proof. □

3.5 Returning to semi-simple groups

At first, we study laws for *isotypic* semi-simple groups, which are semi-simple groups G that fit in the following chain of inclusions, where H is non-abelian simple.

$$H^k \leq G \leq \mathrm{Aut}(H^k) = \mathrm{Aut}(H) \wr \mathrm{S}_k$$

Now recall that for the group $\mathrm{Aut}(H^k)$ we have a short exact sequence

$$\mathbf{1} \to \mathrm{Aut}(H)^k \to \mathrm{Aut}(H) \wr \mathrm{S}_k \to \mathrm{S}_k \to \mathbf{1} \tag{3.1}$$

and we can use the results of the previous section (namely the fact that there exists a non-trivial law of length at most $8\exp\left(3(1+o(1))\sqrt{k\log(k)}\right)$ for S_j (for $j \leq k$); in fact there is a much better bound due to Kozma and Thom in [8], but we will be satisfied with the weaker bound here).

Since H is non-abelian simple, we have another short exact sequence, namely

$$\mathbf{1} \to H = \mathrm{Inn}(H) \to \mathrm{Aut}(H) \to \mathrm{Out}(H) \to \mathbf{1}.$$

By Schreier's conjecture, which holds only due to the classification of finite simple groups, $\mathrm{Out}(H)$ is solvable of length at most three. Thus, using the extension lemma (Lemma 4), we obtain a non-trivial law β for $\mathrm{Aut}(H)$ from a non-trivial law α for H such that

$$\ell(\beta) = 50\ell(\alpha),$$

since we have constructed a non-trivial word $w = b_3$ of length $c_3 = 50$ which is a law for every solvable group of solvability class at most three (and thus for $\mathrm{Out}(H)$) in the section about nilpotent and solvable groups using Lemma 10.

Return to Equation (3.1) and note that, obviously, by definition a word is a law for $\mathrm{Aut}(H)^k$ if and only if it is a law for $\mathrm{Aut}(H)$. Again we use the extension lemma (Lemma 4) to combine the law which we have constructed for all automorphism groups of simple groups of order at most m with the universal law for all symmetric groups of order at most $k!$ to get a non-trivial law of length at most

$$8\exp\left(3(1+o(1))\sqrt{k\log(k)}\right) \cdot 50\frac{(331,776+o(1))m}{\log(m)^2}.$$

We distinguish two cases which cover all isotypic semi-simple groups of order at most n (where the only restriction to m and k is $m^k \leq n$). The first case is $k = 1$ and $m = n$ giving us a contribution of

$$8 \cdot 50\frac{(331,776+o(1))n}{\log(n)^2}.$$

The second case is $2 \leq m \leq \sqrt{n}$ where $2 \leq k \leq \log_2(n)$. Here we get a contribution of

$$8\exp\left(3(1+o(1))\sqrt{\log_2(n)\log(\log_2(n))}\right) \cdot 50\frac{4(331,776+o(1))\sqrt{n}}{\log(n)^2}.$$

Combining both words to get a non-trivial law which holds for every isotypic semi-simple group (using Corollary 2), we note that the first contribution is greater than the second. Thus we have a non-trivial law of length

$$8\cdot 2^2\cdot 8\cdot 50(331,776+o(1))\frac{n}{\log(n)^2} = (4,246,732,800+o(1))\frac{n}{\log(n)^2} = (C_{12}+o(1))\frac{n}{\log(n)^2}.$$

for all isotypic semi-simple groups of order at most n. But in fact, this law is already a law for all semi-simple groups of order at most n. To see this, let H_i be different non-abelian

simple groups and $k_i \in \mathbb{Z}_+$. A general semi-simple group G is then a group such that there exist H_i and k_i such that

$$\prod_{i=1}^{l} H_i^{k_i} \leq G \leq \mathrm{Aut}\left(\prod_{i=1}^{l} H_i^{k_i}\right) = \prod_{i=1}^{l} \mathrm{Aut}(H_i) \wr \mathrm{S}_{k_i}.$$

Thus, if n is the size of G and m_i the size of H_i, we have that $m_i^{k_i} \leq n$ and thus we have already constructed a universal non-trivial law for all groups $\mathrm{Aut}(H_i) \wr \mathrm{S}_{k_i}$ with the above property of length $(C_{12} + o(1))n/\log(n)^2$ which must also be a law for the product of these groups and thus for G. Hence we have derived a universal non-trivial law for all semi-simple groups of order at most n. This completes this section.

Chapter 4

The final conclusion

Finally, we have considered the cases of solvable and semi-simple groups, so that we can draw our conclusion using the extension lemma (Lemma 4).

Keeping in mind that (as we have proven in Proposition 4) there is a universal non-trivial law for all solvable groups of order at most n of length at most $(C_{10}+o(1))\log(n)^{\lambda}$ and another non-trivial law for all semi-simple groups of order at most n of length at most $(C_{12}+o(1))n/\log(n)^2$ (cf. Section 3.5), we can compose these two cases to get a non-trivial law for every group of order at most n.

For this purpose, consider the solvable radical $\mathbf{S}(G)$ of the finite group G and the short exact sequence

$$\mathbf{1} \to \mathbf{S}(G) \to G \to G/\mathbf{S}(G) \to \mathbf{1}$$

to which we want to apply the extension lemma (Lemma 4). Note that by Definition 11, the quotient $G/\mathbf{S}(G)$ is semi-simple. Let $n_1 \coloneqq |\mathbf{S}(G)|$ and $n_2 \coloneqq |G/\mathbf{S}(G)|$.

Let k be a some fixed positive integer. Now we distinguish k cases which together cover all finite groups of order at most n. In the ith case, it holds that

$$\alpha_{i-1}(n) \leq n_1 \leq \alpha_i(n) \text{ and } n_2 \leq \frac{n}{\alpha_{i-1}(n)} \ (i = 1, \ldots, k).$$

Then we plug in

$$\alpha_0 \coloneqq 1 \text{ and } \alpha_i(n) = \underbrace{\log \cdots \log}_{k-i \text{ times}}(n)^{\lambda}.$$

Considering the ith case ($i > 1$), we combine the two laws for $\mathbf{S}(G)$ and $G/\mathbf{S}(G)$ (with the given restrictions to n_1 and n_2 in that case) to get a non-trivial law of length at most

$$\begin{aligned}
&(C_{10}+o(1))\lambda^{\lambda}\underbrace{\log \cdots \log}_{k-i+1}(n)^{\lambda}(C_{12}+o(1))\frac{n/\overbrace{\log \cdots \log}^{k-(i-1)}(n)^{\lambda}}{\log(n/\log \cdots \log(n)^{\lambda})^2} \\
&\leq (C_{13}+o(1))\frac{n}{\log\left(n/\log(n)^{\lambda}\right)^2} \leq (C_{13}+o(1))\frac{n}{\log(n)^2}.
\end{aligned}$$

The only remaining case is $i = 1$, for which we obtain

$$\begin{aligned}&(C_{10} + o(1))\lambda^{\lambda} \underbrace{\log \cdots \log}_{k}(n)^{\lambda}(C_{12} + o(1))\frac{n}{\log(n)^2}\\ &\leq (C_{13} + o(1))\frac{n}{\log(n)^2}.\end{aligned}$$

Here the last inequality holds for an appropriate choice of k, i.e., choose $k(n)$ minimal such that

$$\underbrace{\log \cdots \log}_{k(n)}(n) \leq 1.$$

This function $k(n)$ is called *iterated logarithm* $\log^*(n)$ in computer science. Moreover, $C_{13} = C_{10}C_{12}\lambda^{\lambda} = 209.8414729\ldots \cdot 10^{15}$. Using this notation, we can combine the k cases to obtain the following theorem (using Corollary 2).

The main theorem

Finally, we can deduce the following theorem from this last fact.

Theorem 7. *For $n \in \mathbb{Z}_+$ large enough there exists a non-trivial two-letter law w_n holding for all finite groups of order at most n of length*

$$\ell(w_n) \leq C \log^*(n)^2 \frac{n}{\log(n)^2}$$

for some fixed constant $C > 0$. E.g., one can take $C > 209.8414729\ldots \cdot 10^{15}$.

We leave it here as an open question if the factor $\log^*(n)$ can be removed by some additional argument.

Index

Bibliography

[1] Equivalence of definitions of finite characteristically simple group. `http://groupprops.subwiki.org/wiki/Equivalence_of_definitions_of_finite_characteristically_simple_group`, August 2016.

[2] Normal subdirect product of perfect groups equals direct product. `http://groupprops.subwiki.org/wiki/Normal_subdirect_product_of_perfect_groups_equals_direct_product`, August 2016.

[3] Abdelrhman Elkasapy and Andreas Thom. On the length of the shortest non-trivial element in the derived and the lower central series. *Journal of Group Theory*, 18(5):793–804, 2015.

[4] Alexander Gamburd, Shlomo Hoory, Mehrdad Shahshahani, Aner Shalev, and Balint Virág. On the girth of random cayley graphs. *Random Structures & Algorithms*, 35(1):100–117, 2009.

[5] Daniel Gorenstein. *Finite groups*, volume 301. American Mathematical Society, 2007.

[6] I. Martin Isaacs. *Finite group theory*, volume 92. American Mathematical Society, 2008.

[7] William Kantor and Ákos Seress. Large element orders and the characteristic of Lie-type simple groups. *Journal of Algebra*, 322(3):802–832, 2009.

[8] Gady Kozma and Andreas Thom. Divisibility and laws in finite simple groups. *Mathematische Annalen*, 364(1–2):79–95, 2016.

[9] Edmund Landau. Über die Maximalordnung der Permutation gegebenen Grades. *Archiv der Mathematik und Physik*, 3(5):92–103, 1903.

[10] Roger Lyndon and Paul Schupp. *Combinatorial group theory*. Springer Science & Business Media, 2015.

[11] William Miller. The maximum order of an element of a finite symmetric group. *The American Mathematical Monthly*, 94(6):497–506, 1987.

[12] Mike F. Newman. The soluble length of soluble linear groups. *Mathematische Zeitschrift*, 126(1):59–70, 1972.

[13] Marcel-Paul Schützenberger. Sur l'equation $a^{2+n} = b^{2+m}c^{2+p}$ dans un groupe libre. *Comptes rendus de l'Académie des Sciences Paris, Série I Mathémathique*, 248:2435–2436, 1959.

[14] Andreas Thom. About the length of laws for finite groups. *Israel Journal of Mathematics*, 219(1):469–478, 2017.